# Data Science Assignment 5 Dua Batool db07098@st.habib.edu.pk

#### *Part 1:*

#### Creating Table EmployeeAttrition1:

```
CREATE TABLE employeeattrition1 (
    EmployeeNumber INT,
    Age INT,
    BusinessTravel VARCHAR (50),
    DailyRate INT,
    Department VARCHAR (50),
    DistanceFromHome INT,
    Education INT,
    EducationField VARCHAR(50),
    EnvironmentSatisfaction INT,
    Gender VARCHAR (10),
    HourlyRate INT,
    JobInvolvement INT,
    JobLevel INT,
    JobRole VARCHAR (50),
    JobSatisfaction INT,
    MaritalStatus VARCHAR (10),
    MonthlyIncome INT,
    MonthlyRate INT,
    NumCompaniesWorked INT,
    PercentSalaryHike INT,
    PerformanceRating INT,
    RelationshipSatisfaction INT,
    StandardHours INT,
    StockOptionLevel INT,
    TotalWorkingYears INT,
    TrainingTimesLastYear INT,
    WorkLifeBalance INT,
    YearsAtCompany INT,
    YearsInCurrentRole INT,
    YearsSinceLastPromotion INT,
    YearsWithCurrManager INT
);
```

# Creating Table EmployeeAttrition2:

```
CREATE TABLE employeeattrition2(
EmployeeNumber INT,
Over18 CHAR,
OverTime VARCHAR (4),
Attrition VARCHAR (4)
);
```

# Showing that Data has been imported into employeeattrition1 from the csv file:

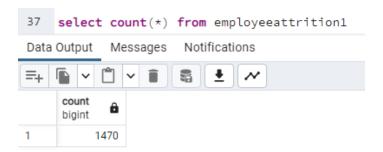
	employeenumber integer	age integer	â	businesstravel character varying (50)	dailyrate integer	department character varying (50)	distancefromhome integer	education integer	educationfield character varying (50)	enviro intege
1	1		41	Travel_Rarely	1102	Sales	1	2	Life Sciences	
2	2		49	Travel_Frequently	279	Research & Development	8	1	Life Sciences	
3	4		37	Travel_Rarely	1373	Research & Development	2	2	Other	
4	5		33	Travel_Frequently	1392	Research & Development	3	4	Life Sciences	
5	7		27	Travel_Rarely	591	Research & Development	2	1	Medical	
6	8		32	Travel_Frequently	1005	Research & Development	2	2	Life Sciences	
7	10		59	Travel_Rarely	1324	Research & Development	3	3	Medical	
8	11		30	Travel_Rarely	1358	Research & Development	24	1	Life Sciences	
9	12		38	Travel_Frequently	216	Research & Development	23	3	Life Sciences	
10	13		36	Travel_Rarely	1299	Research & Development	27	3	Medical	
11	14		35	Travel_Rarely	809	Research & Development	16	3	Medical	
12	15		29	Travel_Rarely	153	Research & Development	15	2	Life Sciences	
13	16		31	Travel_Rarely	670	Research & Development	26	1	Life Sciences	
14	18		34	Travel_Rarely	1346	Research & Development	19	2	Medical	
15	19		28	Travel_Rarely	103	Research & Development	24	3	Life Sciences	
16	20		29	Travel_Rarely	1389	Research & Development	21	4	Life Sciences	
17	21		32	Travel_Rarely	334	Research & Development	5	2	Life Sciences	
18	22		22	Non-Travel	1123	Research & Development	16	2	Medical	
19	23		53	Travel_Rarely	1219	Sales	2	4	Life Sciences	
20	24		38	Travel_Rarely	371	Research & Development	2	3	Life Sciences	
21	26		24	Non-Travel	673	Research & Development	11	2	Other	
22	27		36	Travel_Rarely	1218	Sales	9	4	Life Sciences	
22	20		2.4	Traval Paraly	410	Pacarch & Davalanment	7	А	Life Sciences	

Showing that Data has been imported into employeeattrition2 from the csv file:

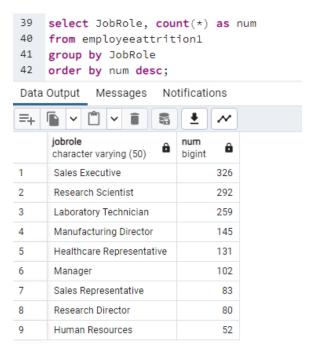
	employeenumber integer	over18 character	overtime character varying (4)	attrition character varying (4)
1	1	Υ	Yes	Yes
2	2	Υ	No	No
3	4	Υ	Yes	Yes
4	5	Y	Yes	No
5	7	Y	No	No
6	8	Υ	No	No
7	10	Υ	Yes	No
8	11	Υ	No	No
9	12	Υ	No	No
10	13	Υ	No	No
11	14	Y	No	No
12	15	Υ	Yes	No
13	16	Υ	No	No
14	18	Υ	No	No
15	19	Υ	Yes	Yes
16	20	Υ	No	No
17	21	Υ	Yes	No
18	22	Υ	Yes	No
19	23	Υ	No	No
20	24	Υ	Yes	No
21	26	Υ	No	No
22	27	Υ	No	Yes
23	28	Υ	No	No

# Interpreting results of EmployeeAttrition1

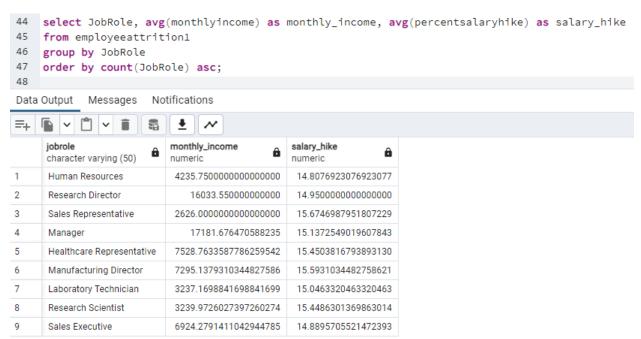
The count of total number of records in the table:



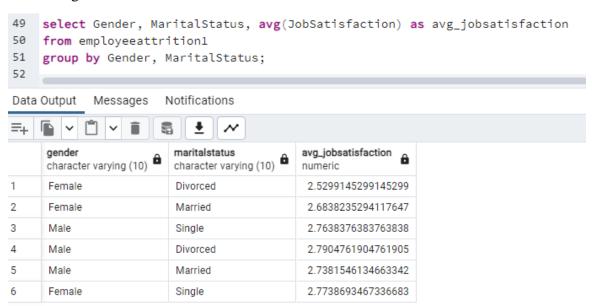
The count of records for each JobRole in descending order of count:



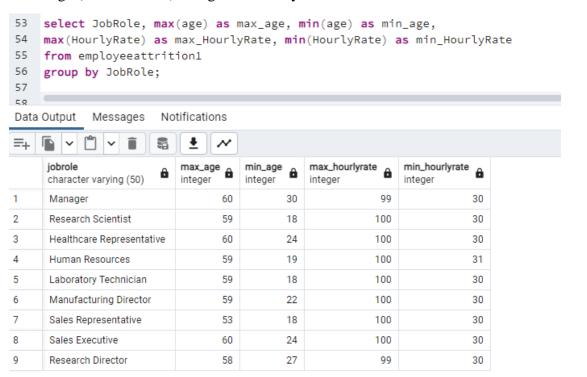
The average MonthlyIncome and PercentSalaryHike for each JobRole in ascending order of JobRole:



#### The average JobSatisfaction for each Gender and MaritalStatus:



#### The range (Min and Max) of Age and HourlyRate for each JobRole:



# Join two tables for EmployeeAttrition1.csv and EmployeeAttrition2.csv and display 20 records:

```
select e1.EmployeeNumber, e1.Age, e1.Gender, e1.JobRole, e2.OverTime, e2.Attrition
from employeeattrition1 e1, employeeattrition2 e2
where e1.EmployeeNumber = e2.EmployeeNumber
limit 20;
```

	employeenumber integer	age integer	gender character varying (10)	jobrole character varying (50)	overtime character varying (4)	attrition character varying (4)
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	35	Male	Laboratory Technician	No	No
12	15	29	Female	Laboratory Technician	Yes	No
13	16	31	Male	Research Scientist	No	No
14	18	34	Male	Laboratory Technician	No	No
15	19	28	Male	Laboratory Technician	Yes	Yes
16	20	29	Female	Manufacturing Director	No	No
17	21	32	Male	Research Scientist	Yes	No
18	22	22	Male	Laboratory Technician	Yes	No
19	23	53	Female	Manager	No	No
20	24	38	Male	Research Scientist	Yes	No