DATA SCIENCE HOMEWORK 5

KUSHAL CHANDANI (kc07535)

Kc07535@st.habib.edu.pk

1. **Queries for creating tables for both of the CSV files:**

- **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

);

For checking:

select * from EmployeeAttrition1;

- **Creating table for Employeeattrition2:**

  CREATE TABLE EmployeeAttrition2 (

  Employee_Number INT,

  Over_18 VARCHAR(64),

  Over_Time VARCHAR(64),

  Attrition VARCHAR(64)

  );

For checking:

select * from EmployeeAttrition2;

2. **The data in these tables was imported manually through the CSV files through pgAdmin4's direct access.**

Q1 Solution:

select count(*) as Total_Number_Of_Records from Employeeattrition1

	total_number_of_records bigint
1	1470

Q2 Solution:

select Job_Role as Job_Role, count(Job_Role) as Total_Count_JobRole from employeeattrition1

group by Job_Role order by count(Job_Role) desc;

	job_role character varying (255)	total_count_jobrole 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

Q3 Solution:

select job_role as Job_Role, round(avg(monthly_Income), 3) as Avg_Monthly_Income, round(avg(percent_salary_hike), 3) as Salary_Hike

from Employeeattrition1

group by job_Role order by job_role asc;

	job_role character varying (255)	avg_monthly_income numeric	salary_hike numeric
1	Healthcare Representative	7528.763	15.450
2	Human Resources	4235.750	14.808
3	Laboratory Technician	3237.170	15.046
4	Manager	17181.676	15.137
5	Manufacturing Director	7295.138	15.593
6	Research Director	16033.550	14.950
7	Research Scientist	3239.973	15.449
8	Sales Executive	6924.279	14.890
9	Sales Representative	2626.000	15.675

Q4 Solution:

```
select round(avg(Job_Satisfaction),3) as Average_Job_Satis, Gender, Marital_Status
from EmployeeAttrition1
group by gender, Marital_Status;
```

	average_job_satis numeric	gender character varying (255)	marital_status character varying (255)
1	2.530	Female	Divorced
2	2.684	Female	Married
3	2.764	Male	Single
4	2.790	Male	Divorced
5	2.738	Male	Married
6	2.774	Female	Single

Q5 Solution:

```
select job_role as Job_Role, max(age) as Maximum_Age, min(age) as Minimum_Age,
max(Hourly_Rate) as Maximum_HRate, min(Hourly_Rate) as Minimum_HRate
from EmployeeAttrition1
group by job_role;
```

	job_role character varying (255)	maximum_age integer	minimum_age integer	maximum_hrate integer	minimum_hrate integer
1	Manager	60	30	99	30
2	Research Scientist	59	18	100	30
3	Healthcare Representative	60	24	100	30
4	Human Resources	59	19	100	31
5	Laboratory Technician	59	18	100	30
6	Manufacturing Director	59	22	100	30
7	Sales Representative	53	18	100	30
8	Sales Executive	60	24	100	30
9	Research Director	58	27	99	30

Q6 Solution:

```
Select
   Emp1.Employee_Number as EmployeeNumber,
   Emp1.Age as Age,
   Emp1.Gender as Gender,
   Emp1.Job_Role as JobRole,
   Emp2.Over_Time as OverTime,
   Emp2.Attrition as Attrition
from EmployeeAttrition1 Emp1
Inner join EmployeeAttrition2 Emp2
On Emp1.Employee_Number = Emp2.Employee_Number
Limit 20;
```

	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	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