Data Science

Homework 5

Muhammad . (mm06369)

Part 1

Create Table Queries

```
CREATE TABLE EmployeeAttrition1 (
    EmployeeNumber INT,
    Age INT,
    BusinessTravel TEXT,
    DailyRate INT,
    Department TEXT,
    DistanceFromHome INT,
    Education INT,
    EducationField TEXT,
    EnvironmentSatisfaction INT,
    Gender TEXT,
    HourlyRate INT,
    JobInvolvement INT,
    JobLevel INT,
    JobRole TEXT,
    JobSatisfaction INT,
    MaritalStatus TEXT,
    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
);
```

```
CREATE TABLE EmployeeAttrition2 (
    EmployeeNumber INT,
    Over18 CHAR(1) CHECK (Over18 IN ('Y', 'N')),
    OverTime CHAR(3) CHECK (OverTime IN ('Yes', 'No')),
    Attrition CHAR(3) CHECK (Attrition IN ('Yes', 'No'))
);

select * from employeeattrition2
```

Data Output | Messages | Notifications

employeenumber integer	over18 character	overtime character	attrition character

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

```
50  select count(*) from employeeattrition1
51
```

Data Output | Messages | Notifications

	count bigint
1	1470

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

```
52  select jobrole, count(employeenumber) from employeeattrition1 group by jobrole
53  order by count(employeenumber) desc
54
```

Data Output | Messages | Notifications

	jobrole text	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

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

```
84  SELECT JobRole, AVG(MonthlyIncome) AS AverageMonthlyIncome,
85  AVG(PercentSalaryHike) AS AveragePercentSalaryHike
86  FROM employeeattrition1
87  GROUP BY JobRole
88  ORDER BY JobRole ASC;
89
```

Data Output | Messages | Notifications

	jobrole text	averagemonthlyincome numeric	averagepercentsalaryhike numeric
1	Healthcare Representative	7528.7633587786259542	15.4503816793893130
2	Human Resources	4235.7500000000000000	14.8076923076923077
3	Laboratory Technician	3237.1698841698841699	15.0463320463320463
4	Manager	17181.676470588235	15.1372549019607843
5	Manufacturing Director	7295.1379310344827586	15.5931034482758621
6	Research Director	16033.550000000000	14.9500000000000000
7	Research Scientist	3239.9726027397260274	15.4486301369863014
8	Sales Executive	6924.2791411042944785	14.8895705521472393
9	Sales Representative	2626.0000000000000000	15.6746987951807229

4. the average JobSatisfaction for each Gender and MaritalStatus

```
select gender, maritalstatus, avg(jobsatisfaction) as AverageJobSatisfaction from employeeattrition1
group by gender, maritalstatus
```

Data Output | Messages | Notifications

	gender text	maritalstatus text	averagejobsatisfaction numeric
1	Female	Divorced	2.5299145299145299
2	Female	Married	2.6838235294117647
3	Male	Single	2.7638376383763838
4	Male	Divorced	2.7904761904761905
5	Male	Married	2.7381546134663342
6	Female	Single	2.7738693467336683

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

```
SELECT
    JobRole,
    MIN(Age) AS MinAge,
    MAX(Age) AS MaxAge,
    MIN(HourlyRate) AS MinHourlyRate,
    MAX(HourlyRate) AS MaxHourlyRate
FROM EmployeeAttrition1
GROUP BY JobRole;
```

Data Output | Messages | Notifications

	jobrole text	minage integer	maxage integer	minhourlyrate integer	maxhourlyrate 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

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

```
select e1.employeenumber, age, gender, jobrole, e2.overtime, e2.attrition
from employeeattrition1 e1 join employeeattrition2 e2
on e1.employeenumber = e2.employeenumber limit 20;
```

Data Output | Messages | Notifications

	employeenumber integer	age integer	gender text	jobrole text	overtime character	attrition character
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