Diksha
diksharajsonth90@gmail.

HR DATA ASSESSMENT QUESTIONS

[1]
```
%load_ext sql
%%capture
%%bash
pip install -sqlalchemy-access
```
```
The sql extension is already loaded. To reload it, use:
  %reload_ext sql
```

1. Retrieve the total number of employees in the dataset.

[2]
```
%sql select count(*) as Total_employees from general_data;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
1 rows affected.
```

[2]

Total_employees

4401

2. Count all unique job roles in the dataset.

[3]
```
%sql select distinct(JobRole) from general_data ;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
9 rows affected.
```

[3]

JobRole
Healthcare Representative
Research Scientist
Sales Executive
Human Resources
Research Director
Laboratory Technician
Manufacturing Director
Sales Representative
Manager

3. Find the average age of employees.

[4]
```
%sql select avg(Age) as average_age from general_data;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
1 rows affected.
```

[4]

average_age
36.92047259189126

4. Retrieve the names and ages of employees who have worked at the company for more than 5 years.

[5]
```
# limit has been applied only to reduce the number of rows displayed in the result.
%%bash select Emp_name, Age, YearsWithCompany from general_data where YearsWithCompany> 5 limit 10 ;
prettyPrint
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
10 rows affected.
+-----------------+-----+-----------------+
|    Emp_name     | Age | YearsWithCompany |
+-----------------+-----+-----------------+
| RENEE MARQUART  |  18 |               6 |
|  HARVEY EVANS   |  32 |               6 |
|  DENNIS HERRERA |  28 |               6 |
|  NATHAN HARDY   |  31 |               9 |
| SUSAN BUCHBINDER |  25 |               6 |
|  DONALD BRYANT  |  29 |               9 |
|  ERNEST SUTTON  |  36 |              23 |
|   JOHN BROWN    |  55 |              16 |
| KATHRYN BALLOU  |  47 |              10 |
|   BRAD GREEN    |  35 |               7 |
+-----------------+-----+-----------------+
```

5. Get a count of employees grouped by their department.

[6]
```
%%bash select Department, count(*) as No_of_Employees from general_data
group by Department;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
3 rows affected.
```

[6]

Department	No_of_Employees
Sales	1336
Research & Development	2877
Human Resources	188

6. List employees who have 'High' Job Satisfaction.

[7]
```
%%bash select * from general_data inner join employee_survey_data on general_data.EmployeeID= employee_survey_data.EmployeeID
where JobSatisfaction= 3 limit 10;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
10 rows affected.
```

[7]

EmployeeID	Emp_name	Age	Attrition	business_travel	Department	DistanceFromHome	Education	EducationField	EmployeeCount	Gender	JobLevel	JobRole
7	DENNIS HERRERA	28	Yes	Travel_Rarely	Research & Development	11	2	Medical	1	Male	2	Sales Executive
17	GEORGE FOURAS	31	No	Travel_Rarely	Research & Development	2	3	Life Sciences	1	Male	1	Laboratory Technician
23	ELLIE BANKS	53	No	Travel_Rarely	Research & Development	11	4	Life Sciences	1	Female	2	Research Scientist
29	CHRISTINA HAZEN	54	Yes	Travel_Frequently	Research & Development	1	2	Medical	1	Male	2	Research Scientist
39	CRUCE CASCIATO	18	Yes	Travel_Rarely	Research & Development	1	4	Life Sciences	1	Male	1	Sales Executive
44	GERARDO PINTO	18	No	Travel_Rarely	Sales	7	3	Life Sciences	1	Male	1	Research Scientist
49	JUAN HARVEY	32	No	Travel_Rarely	Sales	2	3	Marketing	1	Male	2	Laboratory Technician
51	BARBARA SCHLEYNES	58	No	Non-Travel	Research & Development	2	3	Life Sciences	1	Male	2	Manufacturing Director
59	BARRY BROOM	49	No	Travel_Rarely	Research & Development	1	3	Life Sciences	1	Male	3	Manufacturing Director
63	ARNOLD CASTRO	27	No	Travel_Rarely	Research & Development	13	2	Medical	1	Female	4	Research Director

In the above cell limit has only been applied to reduce the number of the rows to be displayed

7. Find the highest Monthly Income in the dataset

[8]
```
%sql select max(MonthlyIncome) as highest_income from general_data;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
1 rows affected.
```

[8]

highest_income
19999

8. List employees who have 'Travel_Rarely' as their BusinessTravel type.

[9]
```
%%sql select * from general_data
where business_travel="Travel_Rarely" limit 10;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
10 rows affected.
```

[9]

EmployeeID	Emp_name	Age	Attrition	business_travel	Department	DistanceFromHome	Education	EducationField	EmployeeCount	Gender	JobLevel	JobRole
1	ALBERTO PEDRALLI	51	No	Travel_Rarely	Sales	6	2	Life Sciences	1	Female	1	Healthcare Representative
5	HARVEY EVANS	32	No	Travel_Rarely	Research & Development	10	1	Medical	1	Male	1	Sales Executive
6	LEON WHITE	46	No	Travel_Rarely	Research & Development	8	3	Life Sciences	1	Female	4	Research Scientist
7	DENNIS HERRERA	28	Yes	Travel_Rarely	Research & Development	11	2	Medical	1	Male	2	Sales Executive
8	DONALD BRYANT	29	No	Travel_Rarely	Research & Development	18	3	Life Sciences	1	Male	2	Sales Executive
11	GEORGE FOURAS	31	No	Travel_Rarely	Research & Development	2	3	Life Sciences	1	Male	1	Laboratory Technician
12	CHRISTOPHER HAGEN	44	Yes	Travel_Frequently	Research & Development	1	2	Medical	1	Male	2	Research Scientist
13	PETER PRANGLOU	38	No	Travel_Rarely	Sales	2	3	Marketing	1	Female	1	Manager
14	ROSEANN SQUIRRO	26	Yes	Travel_Rarely	Research & Development	4	3	Medical	1	Male	3	Research Scientist
15	COLLEEN RILEY	49	No	Travel_Frequently	Research & Development	1	1	Medical	1	Female	2	Research Scientist
16	TIMOTHY SULLIVAN	36	No	Travel_Frequently	Research & Development	8	3	Other	1	Female	3	Sales Executive
17	STEPHEN TACHIAN	31	No	Travel_Rarely	Sales	7	4	Marketing	1	Male	2	Sales Executive
18	GREGORY NATHANIAN	31	No	Travel_Rarely	Sales	16	4	Marketing	1	Female	2	Research Scientist

In the above cell limit has only been applied to reduce the number of the rows to be displayed

9. Retrieve the distinct MaritalStatus categories in the dataset.

[10]
```
%sql select distinct(MaritalStatus) from general_data;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
3 rows affected.
```

[10]

MaritalStatus
Married
Single
Divorced

10. Get a list of employees with more than 2 years of work experience but less than 4 years in their current role.

[11]
```
%%sql select * from general_data
where YearsSinceLastPromotion> 2 and YearsInCompany= 4 limit 10;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
10 rows affected.
```

[11]

EmployeeID	Emp_name	Age	Attrition	business_travel	Department	DistanceFromHome	Education	EducationField	EmployeeCount	Gender	JobLevel	JobRole
7	DENNIS HERRERA	28	Yes	Travel_Rarely	Research & Development	11	2	Medical	1	Male	2	Sales Executive
8	DONALD BRYANT	29	No	Travel_Rarely	Research & Development	18	3	Life Sciences	1	Male	2	Sales Executive
11	GEORGE FOURAS	31	No	Travel_Rarely	Research & Development	2	3	Life Sciences	1	Male	1	Laboratory Technician
29	CHRISTOPHER HAGEN	44	Yes	Travel_Frequently	Research & Development	1	2	Medical	1	Male	2	Research Scientist
30	PETER PRANGLOU	38	No	Travel_Rarely	Sales	2	3	Marketing	1	Female	1	Manager
31	ROSEANN SQUIRRO	26	Yes	Travel_Rarely	Research & Development	4	3	Medical	1	Male	3	Research Scientist
33	COLLEEN RILEY	49	No	Travel_Frequently	Research & Development	1	1	Medical	1	Female	2	Research Scientist
41	TIMOTHY SULLIVAN	36	No	Travel_Frequently	Research & Development	8	3	Other	1	Female	3	Sales Executive
53	STEPHEN TACHIAN	31	No	Travel_Rarely	Sales	7	4	Marketing	1	Male	2	Sales Executive
55	GREGORY NATHANIAN	31	No	Travel_Rarely	Sales	16	4	Marketing	1	Female	2	Research Scientist

In the above cell limit has only been applied to reduce the number of the rows to be displayed

11. Retrieve the top 5 employees with the highest MonthlyIncome

[12]
```
%%bash select * from general_data
order by MonthlyIncome desc limit 5;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
5 rows affected.
```

[12]

EmployeeID	Emp_name	Age	Attrition	business_travel	Department	DistanceFromHome	Education	EducationField	EmployeeCount	Gender	JobLevel	JobRole
3326	LAWRENCE LAU	42	No	Travel_Rarely	Research & Development	18	1	Technical Degree	1	Female	2	Research Scientist
386	KEVIN LABARGNONI	42	No	Travel_Rarely	Research & Development	18	1	Technical Degree	1	Male	2	Research Scientist
1856	DAVID KUCIA	42	No	Travel_Rarely	Research & Development	18	1	Technical Degree	1	Female	2	Research Scientist
2412	NATHAN DEUTSU	36	No	Travel_Rarely	Research & Development	2	2	Technical Degree	1	Male	3	Sales Executive
942	KEVIN MCNAUGHTON	36	No	Travel_Rarely	Research & Development	2	2	Technical Degree	1	Male	3	Sales Executive

14. Calculate the percentage of employees who have had a promotion in the last year.

[13]
```
%%bash select round(select count(EmployeeID) from general_data where YearsSinceLastPromotion>1)/count(EmployeeID) *100,1)
as Percentage from general_data;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
1 rows affected.
```

[13]

Percentage
24.27

15. List the employees with the highest and lowest EnvironmentSatisfaction.

[14]
```
%%bash select general_data.Emp_name, general_data.JobRole,employee_survey_data.JobSatisfaction, employee_survey_data.EnvironmentSatisfaction
from general_data inner Join employee_survey_data
on general_data.EmployeeID= employee_survey_data.EmployeeID
where employee_survey_data.EnvironmentSatisfaction= (select min(EnvironmentSatisfaction) from employee_survey_data) or
employee_survey_data.EnvironmentSatisfaction= (select max(EnvironmentSatisfaction) from employee_survey_data) limit 10 ;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
10 rows affected.
```

[14]

Emp_name	EmployeeID	JobSatisfaction	EnvironmentSatisfaction
RENEE MARQUART	4	4	4
HARVEY EVANS	5	1	4
DENNIS HERRERA	7	3	1
DONALD BRYANT	8	2	1
JOHN BROWN	10	1	4
KATHRYN BALLOU	14	2	1
DONALD FILDS	15	4	4
GEORGE FOURAS	17	3	4
MARTIN LALOR JR	18	4	1
DARCY MILLER	20	1	1

In the above cell limit has only been applied to reduce the number of the rows to be displayed

16. Find the employees who have the same JobRole and MaritalStatus.

[15]
```
%%sql SELECT * FROM general_data
WHERE (JobRole, MaritalStatus) IN (
    SELECT JobRole, MaritalStatus
    FROM general_data
    GROUP BY JobRole, MaritalStatus
    HAVING COUNT(*) > 1) LIMIT 10;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
10 rows affected.
```

[15]

EmployeeID	Emp_name	Age	Attrition	business_travel	Department	DistanceFromHome	Education	EducationField	EmployeeCount	Gender	JobLevel	JobRole
1	ALBERTO PEDRALLI	51	No	Travel_Rarely	Sales	6	2	Life Sciences	1	Female	1	Healthcare Representative
2	LAWRENCE LEE	31	No	Travel_Frequently	Research & Development	10	1	Life Sciences	1	Female	1	Research Scientist
3	DWAYNE CURRY	32	No	Travel_Frequently	Research & Development	17	4	Other	1	Male	4	Sales Executive
4	RENEE MARQUART	18	No	Travel_Rarely	Research & Development	2	5	Life Sciences	1	Male	3	Human Resources
5	HARVEY EVANS	32	No	Travel_Rarely	Research & Development	10	1	Medical	1	Male	1	Sales Executive
6	LEON WHITE	46	No	Travel_Rarely	Research & Development	8	3	Life Sciences	1	Female	4	Research Director
7	DENNIS HERRERA	28	Yes	Travel_Rarely	Research & Development	11	2	Medical	1	Male	2	Sales Executive
8	DONALD BRYANT	29	No	Travel_Rarely	Research & Development	18	3	Life Sciences	1	Male	2	Sales Executive
9	NATHAN HARDY	31	No	Travel_Rarely	Research & Development	1	3	Life Sciences	1	Male	3	Laboratory Technician
10	SUSAN BUCHBINDER	25	No	Non-Travel	Research & Development	7	4	Medical	1	Female	4	Laboratory Technician

In the above cell limit has only been applied to reduce the number of the rows to be displayed

17. List the employees with the highest TotalWorkingYears who also have a PerformanceRating of 4

[16]
```
%%sql select * from general_data inner Join employee_survey_data
on general_data.EmployeeID= employee_survey_data.EmployeeID
where general_data.TotalWorkingYears= (select max(TotalWorkingYears) from general_data)
and manager_survey_data.PerformanceRating=4;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
0 rows affected.
```

[16]

EmployeeID	Emp_name	Age	Attrition	business_travel	Department	DistanceFromHome	Education	EducationField	EmployeeCount	Gender	JobLevel	JobRole	Marital

Note:
There is no such employee with the highest TotalWorkingYears who also have a performance rating of 4

18. Calculate the average Age and JobSatisfaction for each BusinessTravel type.

[17]
```
%%sql select general_data.business_travel,round( avg(general_data.Age)) as Average_age,round(avg(employee_survey_data.JobSatisfaction),2) as Average_JobSat
from general_data inner Join employee_survey_data on general_data.EmployeeID= employee_survey_data.EmployeeID
group by general_data.business_travel;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
3 rows affected.
```

[17]

business_travel	Average_age	Average_JobSatisfaction
Travel_Rarely	37.0	2.7
Travel_Frequently	36.0	2.79
Non-Travel	37.0	2.78

19. Retrieve the most common EducationField among employees.

[18]
```
%%sql select EducationField,count(EducationField) as count from general_data
group by EducationField
order by count desc limit 1 ;
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
1 rows affected.
```

[18]

EducationField	count
Life Sciences	1812

20. List the employees who have worked for the company the longest but haven't had a promotion

[19]
```
%%sql select * from general_data
where YearsSinceLastPromotion= (select max(YearsSinceLastPromotion)
```
```
 * mysql+mysqlconnector://root:***@localhost/hr_analysis
0 rows affected.
```

[19]

EmployeeID	Emp_name	Age	Attrition	business_travel	Department	DistanceFromHome	Education	EducationField	EmployeeCount	Gender	JobLevel	JobRole	Marital

Note: There is no such employee.