DATA SCIENCE INTERNSHIP

SQL Project – Task 1: HR Analytics

PRASHANT SINGH CHAUHAN

LinkedIn: https://www.linkedin.com/in/prashant-singh-chauhan-1072298a/
Email: chauhanp72@gmail.com

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

```
select count(employeeid) as Total_Number_of_Employees from general_data
```

	Total_Number_of_Employees
1	4410

2. List all unique job roles in the dataset.

```
select distinct(JobRole) from general_data
```

	JobRole
1	Sales Representative
2	Manager
3	Healthcare Representative
4	Laboratory Technician
5	Sales Executive
6	Manufacturing Director
7	Human Resources
8	Research Director

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

```
SELECT [Emp name] as Employee_Name,AGE FROM GENERAL_DATA WHERE yearsatcompany<5;
```

	Employee_Name	AGE
3	JENNIFER MATHE...	54
4	KEVIN CHOCKER	24
5	ANGELA CALVILLO	44
6	GAIL NORMAN	37
7	MARYANNE LOW...	32
8	LOUISE DIMATTIO	20
9	JOHN RUGGEIRO	37
10	JOSEPH SALLADY	59

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

```
SELECT Department, COUNT(*) as Number FROM GENERAL_DATA GROUP BY DEPARTMENT;
```

	Department	Number
1	Sales	1338
2	Research & Development	2883
3	Human Resources	189

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

```
SELECT EMPLOYEEID,[Emp Name] FROM GENERAL_DATA where joblevel=3;
```

	EMPLOYEEID	Emp Name
1	1333	CORNELIUS JOHNSON
2	1337	BETH GOUDREAU
3	1351	CHARLES RICARDO
4	1357	KEVIN BYRNE
5	1358	KAREN ACOSTA
6	1360	PETER DAILEY
7	1364	THERESA FOGARTY
8	1368	KEVIN MURRAY

7. Find the highest Monthly Income in the dataset.
8. List employees who have 'Travel_Rarely' as their BusinessTravel type.
9. Retrieve the distinct MaritalStatus categories in the dataset.
10. Get a list of employees with more than 2 years of work experience but less than 4 years in their current role.

```
7.SELECT MAX(MONTHLYINCOME) as Max_Monthly_Income FROM GENERAL_DATA;
8.SELECT EMPLOYEEID,[emp name] as Employees_who_Travel_Rarely FROM GENERAL_DATA WHERE BusinessTravel
IN('TRAVEL_RARELY');
9.SELECT DISTINCT(MaritalStatus) as MaritalStatus_Categories FROM GENERAL_DATA;
10.SELECT EMPLOYEEID,[emp name] as Employee_with_years_between_2_and_4_at_Company FROM GENERAL_DATA
WHERE TotalWorkingYears>2 AND
YEARSATCOMPANY<4;
```

	Max_Monthly_Income
1	199990

	EMPLOYEEID	Employees_who_Travel_Rarely
1	1299	ANDRE KURY
2	1300	CHARLES ROWLES
3	1301	ROSE-ELLEN FAIRGRIEVE
4	1302	ALEX TAKAOKA
5	1304	DAVID HAMILTON
6	1305	RANDALL SOO-HOO
7	1307	GERALD SCULLION
8	1308	JENNIFER MATHEWS

	MaritalStatus_Categories
1	Single
2	Divorced
3	Married

	EMPLOYEEID	Employee_with_years_between_2_and_4_at_Company
1	1299	ANDRE KURY
2	1309	KEVIN CHOCKER
3	1312	ANGELA CALVILLO

Query executed successfully. CRYPTOGRAPHIC-A\SQLEXPRESS ... CRYPTOGRAPHIC-A\Crypto... master 00:00:01 654 rows

11. List employees who have changed their job roles within the company (JobLevel and JobRole differ from their previous job).
12. Find the average distance from home for employees in each department.
13. Retrieve the top 5 employees with the highest MonthlyIncome.
14. Calculate the percentage of employees who have had a promotion in the last year.

```
11.SELECT
    EmployeeID,
    [Emp Name],
    CurrentJobRole,
    PreviousJobRole,
    CurrentJobLevel,
    PreviousJobLevel
FROM (
    SELECT
        EmployeeID,
        [Emp Name],
        JobRole AS CurrentJobRole,
        JobLevel AS CurrentJobLevel,
        LAG(JobRole) OVER (PARTITION BY EmployeeID ORDER BY YearsAtCompany) AS PreviousJobRole,
        LAG(JobLevel) OVER (PARTITION BY EmployeeID ORDER BY YearsAtCompany) AS PreviousJobLevel
    FROM general_data
) AS JobChanges
WHERE (CurrentJobRole <> PreviousJobRole)
    OR (CurrentJobLevel <> PreviousJobLevel)

12.SELECT Department,AVG(DistanceFromHome) as AVG_Distance_from_Home_to_Office FROM GENERAL_DATA GROUP BY DEPARTMENT;

13.SELECT top 5 EmployeeID, [Emp Name], MonthlyIncome FROM general_data ORDER BY MonthlyIncome DESC

14.SELECT
    COUNT(CASE WHEN YearsSinceLastPromotion <= 1 THEN 1 END) AS EmployeesWithPromotionLastYear,
    COUNT(*) AS TotalEmployees,
    (COUNT(CASE WHEN YearsSinceLastPromotion <= 1 THEN 1 END) * 100.0 / COUNT(*)) AS PercentagePromotedLastYear
FROM
    general_data;
```

	EmployeeID	Emp Name	CurrentJobRole	PreviousJobRole	CurrentJobLevel	PreviousJobLevel

	Department	AVG_Distance_from_Home_to_Office
1	Sales	9.23094170403587
2	Research & Development	9.23621227887617
3	Human Resources	8.25396825396825

	EmployeeID	Emp Name	MonthlyIncome
1	1856	DAVID KUCIA	199990
2	3326	LAWRENCE LAU	199990
3	386	KEVIN LABANO...	199990
4	3882	SHANNON STA...	199730
5	942	KEVIN MCNAU...	199730

	EmployeesWithPromotionLastYear	TotalEmployees	PercentagePromotedLastYear
1	2814	4410	63.809523809523

15. List the employees with the highest and lowest EnvironmentSatisfaction.
16. Find the employees who have the same JobRole and MaritalStatus.

```
15.SELECT a.EmployeeID, a.[emp name],b.environmentsatisfaction FROM general_data a JOIN employee_survey_data b ON a.EmployeeID = b.EmployeeID
WHERE b.EnvironmentSatisfaction IN (
    SELECT MAX(EnvironmentSatisfaction)
    FROM employee_survey_data
    UNION
    SELECT MIN(EnvironmentSatisfaction)
    FROM employee_survey_data)


16.SELECT EmployeeID, JobRole, MaritalStatus FROM general_Data e1
WHERE EXISTS (
    SELECT 1
    FROM general_Data e2
    WHERE e1.EmployeeID <> e2.EmployeeID
    AND e1.JobRole = e2.JobRole
    AND e1.MaritalStatus = e2.MaritalStatus
)
ORDER BY JobRole, MaritalStatus, EmployeeID;
```

	EmployeeID	emp name	environmentsatisfaction
1	1300	CHARLES ROWLES	4
2	1308	JENNIFER MATHEWS	1
3	1309	KEVIN CHOCKER	1
4	1311	LISA MOY	4
5	1314	MARYANNE LOWMAN	4
6	1315	LOUISE DIMATTIO	4
7	1316	JAMES FEWELL	1
8	1319	JULIE RUSSELL	1

	EmployeeID	JobRole	MaritalStatus
1	64	Healthcare Representative	Divorced
2	83	Healthcare Representative	Divorced
3	147	Healthcare Representative	Divorced
4	152	Healthcare Representative	Divorced
5	161	Healthcare Representative	Divorced
6	203	Healthcare Representative	Divorced
7	234	Healthcare Representative	Divorced
8	262	Healthcare Representative	Divorced

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

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

```
17.SELECT a.EmployeeID, a.[Emp name]
FROM general_data a
JOIN manager_survey_data b ON a.EmployeeID = b.EmployeeID WHERE b.performancerating = 4
AND a.totalworkingyears = (
    SELECT MAX(totalworkingyears)
    FROM general_data
    WHERE EmployeeID IN (
      SELECT EmployeeID
        FROM manager_survey_data
        WHERE performancerating = 4
    )
);
```

```
18.SELECT AVG(age) as Average_Age,AVG(JobSatisfaction) as JobSatisfaction_AVG
FROM GENERAL_DATA a  join employee_survey_data b on a.EmployeeID=b.EmployeeID
GROUP BY BusinessTravel;
```

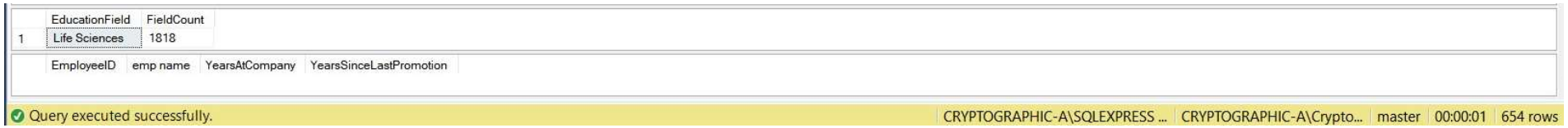
	EmployeeID	Emp name
1	786	STEVEN SETO
2	2256	SHARON LEGENZA
3	3726	EMILY MURASE

	Average_Age	JobSatisfaction_AVG
1	36.62	2.78794642857143
2	37.0920421860019	2.70369181380417
3	36.4548736462094	2.78839177750907

19. Retrieve the most common EducationField among employees.
20. List the employees who have worked for the company the longest but haven't had a promotion.

```
19.SELECT TOP 1 EducationField, COUNT(*) AS FieldCount
FROM general_Data
GROUP BY EducationField
ORDER BY COUNT(*) DESC;


20.SELECT EmployeeID, [emp name], YearsAtCompany, YearsSinceLastPromotion
FROM general_data
WHERE YearsAtCompany = (
    SELECT MAX(YearsAtCompany)
    FROM general_data
)
AND YearsSinceLastPromotion = 0;
```



	EducationField	FieldCount
1	Life Sciences	1818

EmployeeID	emp name	YearsAtCompany	YearsSinceLastPromotion

Query executed successfully. CRYPTOGRAPHIC-A\SQLEXPRESS ... CRYPTOGRAPHIC-A\Crypto... master 00:00:01 654 rows