**HR ANALYTICS QUERIES**

* See all the data imported:

SELECT \*FROM HR\_Analytics

**DATA CLEANING**:

1. **REMOVE UNNECESSARY COLUMN**:

Removing unnecessary columns such as **YearsWithCurrManager, StockOptionLevel, and RelationshipSatisfaction** helps improve data quality and efficiency in analysis. These fields may contain less relevant, or unused information that does not contribute to the objectives of the HR analytics study.

ALTER TABLE HR\_Analytics

DROP COLUMN YearsWithCurrManager, StockOptionLevel, RelationshipSatisfaction;

1. **CHECK DUPLICATE DATA IN COLUMN:**

Checking for duplicate values in the **EmpID** column is essential to maintain data integrity and accuracy. Since EmpID represents a unique identifier for each employee, any duplicate entries may indicate data entry errors in the dataset.

SELECT EmpID, COUNT (\*) AS DuplicateCount FROM HR\_Analytics

GROUP BY EmpID

HAVING COUNT (\*) > 1;

1. **REMOVE DUPLICATES IN COLUMNS:**

WITH CTE AS (

SELECT \*,

ROW\_NUMBER () OVER (PARTITION BY EmpID ORDER BY EmpID) AS row\_num

FROM HR\_Analytics

)

DELETE FROM CTE WHERE row\_num > 1;

1. **UPDATE INCONSISTENT VALUES IN BUSINESSTRAVEL COLUMN:**

Updating inconsistent values in the **BusinessTravel** column ensures consistency and prevents duplicate categories. Standardizing entries like “TravelRarely” to “Travel\_Rarely” improves data quality, making reporting and analysis more accurate.

UPDATE HR\_Analytics

SET BusinessTravel = 'Travel\_Rarely'

WHERE BusinessTravel IN ('TravelRarely');

1. **UPDATE OVER18 COLUMN:**

Updating the **Over18** column ensures consistent values by replacing “Y” with “Yes”. This standardization improves clarity and avoids confusion in reporting or filtering.

UPDATE HR\_Analytics

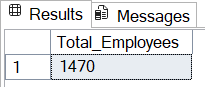
SET Over18 = 'Yes'

WHERE Over18 = 'Y';

**QUERY PROBLEMS:**

1. **Total Employees**

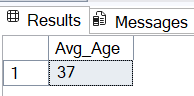
SELECT COUNT (EmpID) AS Total\_Employees

FROM HR\_Analytics;

1. **Avg Age of Employees**

SELECT ROUND (AVG (CAST (Age AS FLOAT)), 0) AS Avg\_Age

FROM HR\_Analytics;



1. **Avg Year of Employees at Company**.

SELECT AVG(YearsAtCompany) AS Avg\_Years\_At\_Company

FROM HR\_Analytics;

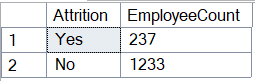


1. **Overall Attrition Employees Count**

SELECT Attrition, COUNT (\*) AS EmployeeCount

FROM HR\_Analytics

GROUP BY Attrition;



1. **Avg Salary of Employees**

SELECT (CONCAT (AVG (MonthlyIncome),' K')) AS Avg\_Salary

FROM HR\_Analytics;



1. **Overall Attrition Rate**

SELECT

CAST (

(SUM (CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) \* 100.0) / COUNT (\*) AS decimal (5, 2)) AS Attrition\_Rate

FROM HR\_Analytics;



1. **Attrition by Department**

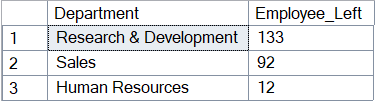
SELECT Department, COUNT (\*) AS Employee\_Left

FROM HR\_Analytics

WHERE Attrition ='Yes'

GROUP BY Department

ORDER BY Employee\_Left desc;

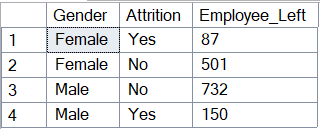


1. **Attrition by Gender (Left and Stayed)**

SELECT Gender, Attrition, COUNT (\*) AS Employee Count

FROM HR\_Analytics

GROUP BY Gender, Attrition;



1. **Attrition Percentage by Department**

SELECT

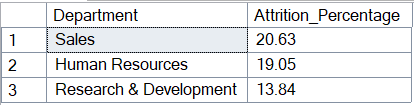
Department,

CAST (SUM (CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) \* 100.0 / COUNT (\*) AS DECIMAL (5, 2)) AS Attrition\_Percentage

FROM HR\_Analytics

GROUP BY Department

ORDER BY Attrition\_Percentage DESC;



1. **Attrition by Job Role**

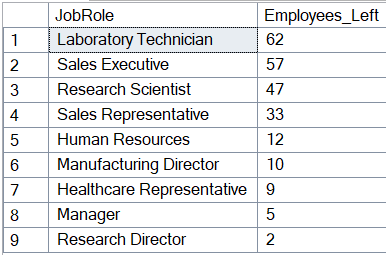
SELECT JobRole, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY JobRole

ORDER BY Employees\_Left DESC;



1. **Total Employees by Departments**

SELECT

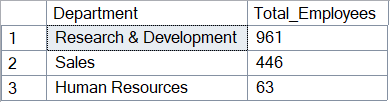
Department,

COUNT (\*) AS Total\_Employees

FROM HR\_Analytics

GROUP BY Department

ORDER BY Total\_Employees DESC;



1. **Total Employees by Job Role**

SELECT

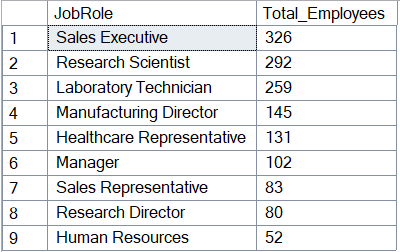
JobRole,

COUNT (\*) AS Total\_Employees

FROM HR\_Analytics

GROUP BY JobRole

ORDER BY Total\_Employees DESC;



1. **Attrition by Age Group**

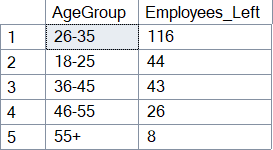
SELECT AgeGroup, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY AgeGroup

ORDER BY Employees\_Left DESC;



1. **Attrition by Salary Slab.**

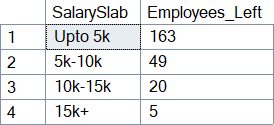
SELECT SalarySlab, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY SalarySlab

ORDER BY Employees\_Left DESC;



1. **Attrition by Gender**

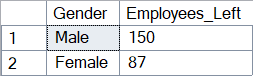
SELECT Gender, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY Gender

ORDER BY Employees\_Left DESC;



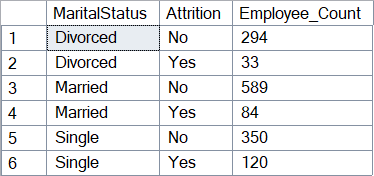
1. **Attrition by Marital Status(Left and Stayed)**

SELECT MaritalStatus, Attrition, COUNT (\*) As Employee\_Count

FROM HR\_Analytics

GROUP BY MaritalStatus, Attrition

ORDER BY MaritalStatus, Attrition;



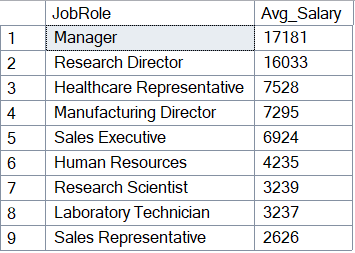
1. **Average Salary by Job Role**

SELECT JobRole, AVG (MonthlyIncome) AS Avg\_Salary

FROM HR\_Analytics

GROUP BY JobRole

ORDER BY Avg\_Salary DESC;



1. **Attrition by Years at Company**

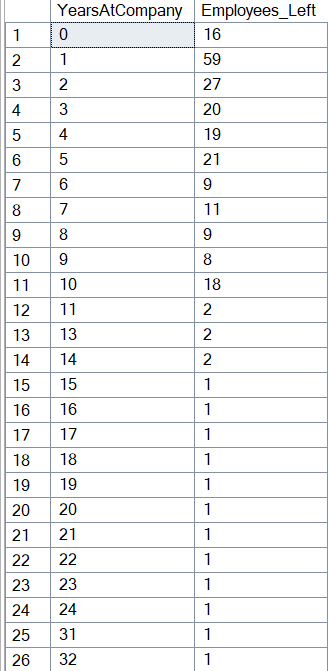
SELECT YearsAtCompany, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY YearsAtCompany

ORDER BY YearsAtCompany;





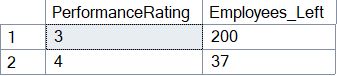
1. **Performance Rating vs Attrition**

SELECT PerformanceRating, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY PerformanceRating;



1. **Attrition by Business Travel**

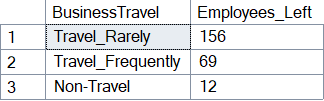
SELECT BusinessTravel, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY BusinessTravel

ORDER BY Employees\_Left DESC;



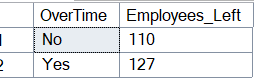
1. **Attrition by Overtime**

SELECT OverTime, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY OverTime;

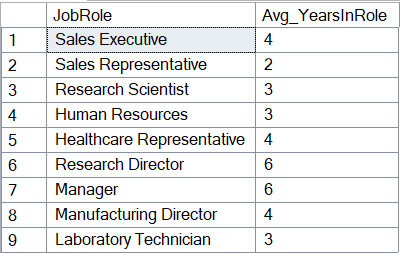


1. **Average Years in Current Role by Job Role**

SELECT JobRole, AVG (YearsInCurrentRole) AS Avg\_YearsInRole

FROM HR\_Analytics

GROUP BY JobRole;



1. **Attrition by Years Since Last Promotion**

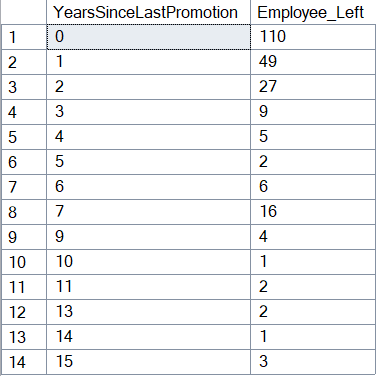
SELECT YearsSinceLastPromotion, COUNT (\*) AS Employee\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY YearsSinceLastPromotion

ORDER BY YearsSinceLastPromotion;



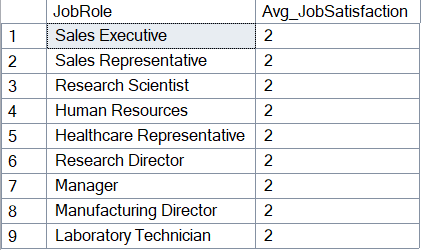
1. **Job Satisfaction by Job Role**

SELECT JobRole, AVG (JobSatisfaction) AS Avg\_JobSatisfaction

FROM HR\_Analytics

GROUP BY JobRole

ORDER BY Avg\_JobSatisfaction DESC;



1. **Attrition by Work-Life Balance Rating**

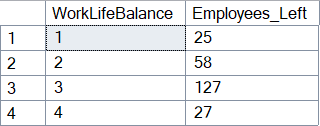
SELECT WorkLifeBalance, COUNT (\*) AS Employees\_Left

FROM HR\_Analytics

WHERE Attrition = 'Yes'

GROUP BY WorkLifeBalance

ORDER BY WorkLifeBalance;



\*\*\*\*\*\*\*END\*\*\*\*\*\*\*