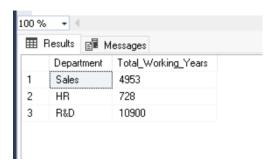
Case Study: Employee Attrition Analysis

Create database HREmployeeDB.

a) Return the shape of the table

b) Calculate the cumulative sum of total working years for each department

```
SELECT Department, SUM(TOTAL_Working_Years) Total_Working_Years
FROM employees
GROUP BY Department
```



c) Which gender have higher strength as workforce in each department

```
☐SELECT Department, Gender, Gender_count, Ranking from

  (SELECT Department, Gender, Gender_count,
  dense rank() Over(partition by Department order by Gender count desc) as Ranking
  from
      (SELECT Department, Gender, count(*) Gender_count
      FROM employees
      Group by Department, Gender) as gcount)
      as sorted_gcount
  where Ranking = 1
    Department
             Gender
                   Gender_count
                             Ranking
   HR
             Male
                   43
                              1
    R&D
             Male
3
    Sales
             Male
                   257
```

d) Create a new column AGE_BAND and Show Distribution of Employee's Age band group (Below 25, 25-34, 35-44, 45-55. ABOVE 55).

```
select emp_no,Age,

case

when Age<25 then 'Below 25'

when Age between 25 and 34 then '25-34'

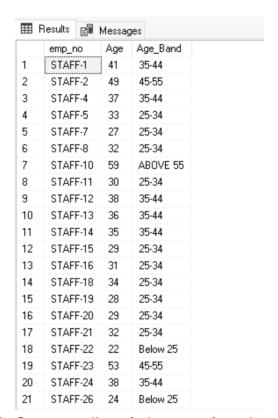
when Age between 35 and 44 then '35-44'

when Age between 45 and 55 then '45-55'

when Age > 55 then 'ABOVE 55'

end as Age_Band

from employees
```



e) Compare all marital status of employee and find the most frequent marital status

```
☐ select Top(1) Marital_Status,count(*) as Total_count
| from employees
| group by Marital_Status
| order by total_count desc
```



f) Show the Job Role with Highest Attrition Rate (Percentage)

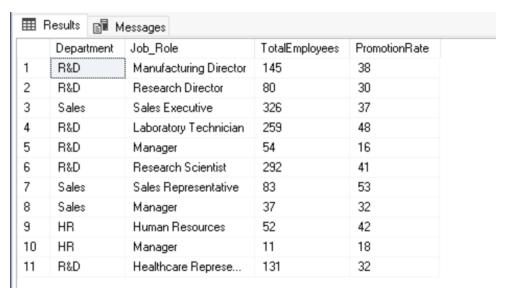
```
    □select job role,

  round(cast(sum(attr) as float)/count(attr)*100,2) as 'Attrition_rate (%)'
  from
  (select job_role,
       when Attrition = 'Yes' then 1
       when Attrition = 'No' then 0
  end as attr
  from employees) as abcd
  group by job_role
  order by 'Attrition_rate (%)' desc
🖽 Results 📳 Messages
    job_role
                          Attrition_rate (%)
   Sales Representative
                          39.76
2
                          23.94
   Laboratory Technician
3
    Human Resources
                          23.08
    Sales Executive
                          17.48
5
    Research Scientist
                          16.1
6
                          6.9
    Manufacturing Director
7
    Healthcare Representative 6.87
8
    Manager
                          4.9
     Research Director
                          2.5
```

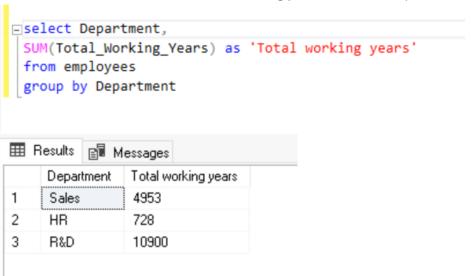
g) Show distribution of Employee's Promotion, Find the maximum chances of employee getting promoted.

```
DWITH PROMO AS (
SELECT Department, Job_Role,
COUNT(*) AS TotalEmployees,
SUM(CASE WHEN Years_Since_Last_Promotion = 0 THEN 1 ELSE 0 END)
AS Promoted,
AVG(Years_Since_Last_Promotion) AS AvgYearsSinceLastPromotion
FROM Employees
GROUP BY Department, Job_Role
)

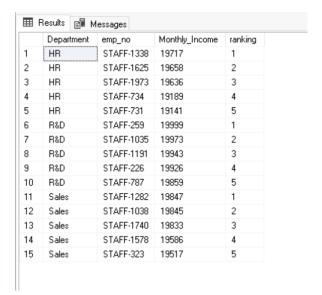
SELECT Department, Job_Role, TotalEmployees,
(Promoted*100/TotalEmployees) as PromotionRate
FROM PROMO
```



h) Show the cumulative sum of total working years for each department.



Find the rank of employees within each department based on their monthly income



j) Calculate the running total of 'Total Working Years' for each employee within each department and age band.

```
Eselect Employee_Number,Department,

CF_age_Band,

SUM(Total_working_years) as 'Total working years'

from employees

group by Department,Employee_Number,CF_age_Band

order by Employee_Number
```

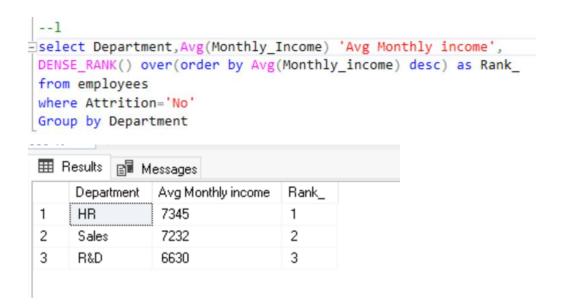
⊞ F	Results 📳 Message	es		
	Employee_Number	Department	CF_age_Band	Total working years
1	1	Sales	35 - 44	8
2	2	R&D	45 - 54	10
3	4	R&D	35 - 44	7
4	5	R&D	25 - 34	8
5	7	R&D	25 - 34	6
6	8	R&D	25 - 34	8
7	10	R&D	Over 55	12
8	11	R&D	25 - 34	1
9	12	R&D	35 - 44	10
10	13	R&D	35 - 44	17
11	14	R&D	35 - 44	6
12	15	R&D	25 - 34	10
13	16	R&D	25 - 34	5
14	18	R&D	25 - 34	3
15	19	R&D	25 - 34	6
16	20	R&D	25 - 34	10

k) Foreach employee who left, calculate the number of years they worked before leaving and compare it with the average years worked by employees in the same department.

```
DWITH avgyears AS (
SELECT Department, AVG(Years_At_Company) AS yr
FROM Employees
GROUP BY Department
)
SELECT
    e.Employee_Number,
    e.Department,
    e.Years_At_Company,
    e.Years_At_Company - a.yr AS Difference_with_dept_avg
FROM Employees e
JOIN avgyears a ON e.Department = a.Department
WHERE e.Attrition = 'No';
```

!!!	Results 📳 Messa	ges		
	Employee_Number	Department	Years_At_Company	Difference_with_dept_avg
1	2	R&D	10	4
2	5	R&D	8	2
3	7	R&D	2	-4
4	8	R&D	7	1
5	10	R&D	1	-5
6	11	R&D	1	-5
7	12	R&D	9	3
8	13	R&D	7	1
9	14	R&D	5	-1
10	15	R&D	9	3
11	16	R&D	5	-1
12	18	R&D	2	-4
13	20	R&D	10	4
14	21	R&D	6	0
15	22	R&D	1	-5
16	23	Sales	25	18
17	24	R&D	3	-3
18	26	R&D	4	-2
19	28	R&D	12	6
20	30	R&D	0	-6
21	32	R&D	14	8

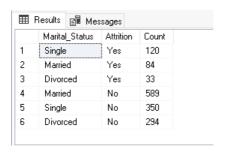
l) Rank the departments by the average monthly income of employees who have left.



m) Find the if there is any relation between Attrition Rate and Marital Status of Employee.

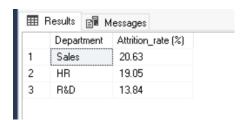
```
□ select Marital_Status, Attrition,

count(Attrition) 'Count' from employees
group by Marital_Status, Attrition
order by Attrition desc, 'Count' desc
```

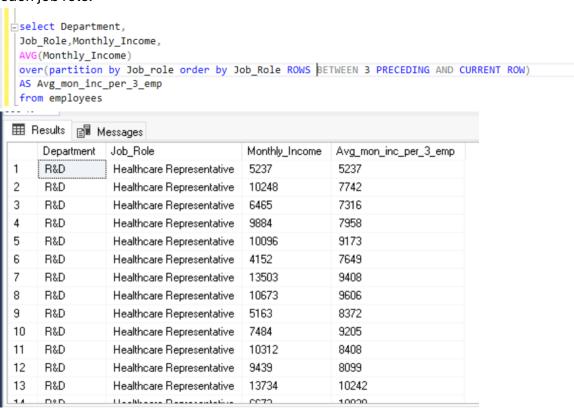


n) Show the Department with Highest Attrition Rate (Percentage)

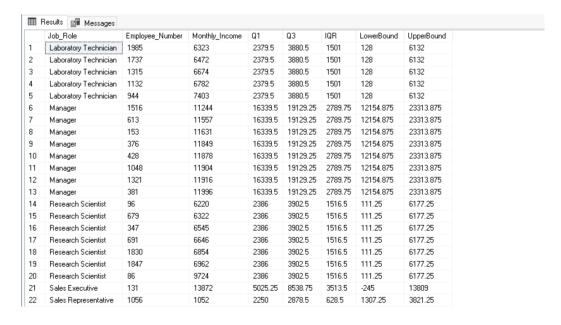
```
select Department,
round(cast(sum(attr) as float)/count(attr)*100,2) as 'Attrition_rate (%)'
from
(select Department,
case
    when Attrition = 'Yes' then 1
    when Attrition = 'No' then 0
end as attr
from employees) as abcd
group by Department
order by 'Attrition_rate (%)' desc
```



o) Calculate the moving average of monthly income over the past 3 employees for each job role.



p) Identify employees with outliers in monthly income within each job role. [
Condition: Monthly_Income < Q1 - (Q3 - Q1) * 1.5 OR Monthly_Income > Q3 + (Q3 - Q1)]



 q) Gender distribution within each job role, show each job role with its gender domination. [Male_Domination or Female_Domination]

```
--q
=select Job_role,Gender as Gender_Dominated,
Gender_count
from
(select Job_Role,
count(gender) as Gender_count,Gender,
DENSE_RANK() over(partition by Job_Role order by count(gender))
As Gender_domination
from employees
group by Job_role,Gender) as sub
```

⊞ F	Results 📳 Messages		
	Job_role	Gender_Dominated	Gender_count
1	Manager	Male	55
2	Laboratory Technician	Female	85
3	Manager	Female	47
4	Healthcare Representative	Female	51
5	Healthcare Representative	Male	80
6	Laboratory Technician	Male	174
7	Research Scientist	Female	114
8	Manufacturing Director	Male	73
9	Research Scientist	Male	178
10	Sales Executive	Male	194
11	Human Resources	Male	36
12	Research Director	Male	47
13	Human Resources	Female	16
14	Research Director	Female	33
15	Sales Representative	Male	45
16	Sales Executive	Female	132
17	Manufacturing Director	Female	72
18	Sales Representative	Female	38

r) Percent rank of employees based on training times last year

```
select Employee_Number,Department,Education_Field,

Job_Role,Training_Times_Last_Year,

round(PERCENT_RANK() OVER(ORDER BY Training_Times_Last_Year)*100,4)

AS Percentile_Rank

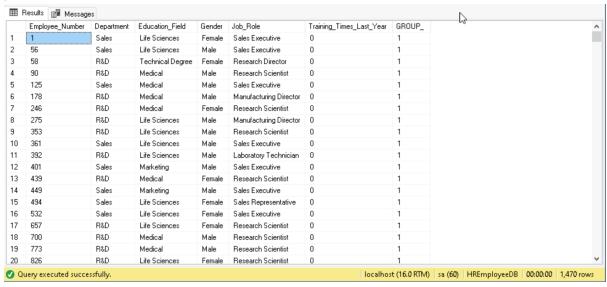
from employees

ORDER BY Employee_Number
```

	Employee_Number	Department	Education_Field	Job_Role	Training_Times_Last_Year	Percentile_Rank	_	
13	16	R&D	Life Sciences	Research Scientist	1	3.676	ν _ζ	
14	18	R&D	Medical	Laboratory Technician	2	8.5092		
15	19	R&D	Life Sciences	Laboratory Technician	4	79.1695		
16	20	R&D	Life Sciences	Manufacturing Director	1	3.676		
17	21	R&D	Life Sciences	Research Scientist	5	87.5425		
8	22	R&D	Medical	Laboratory Technician	2	8.5092		
19	23	Sales	Life Sciences	Manager	3	45.7454		
20	24	R&D	Life Sciences	Research Scientist	3	45.7454		
21	26	R&D	Other	Manufacturing Director	5	87.5425		
22	27	Sales	Life Sciences	Sales Representative	4	79.1695		
23	28	R&D	Life Sciences	Research Director	4	79.1695		
24	30	R&D	Life Sciences	Research Scientist	6	95.6433		
25	31	R&D	Medical	Research Scientist	2	8.5092		
26	32	R&D	Other	Manager	3	45.7454		

s) Divide employees into 5 groups based on training times last year [Use NTILE ()]

```
□select Employee_Number,Department,Education_Field,
Gender,Job_Role,Training_Times_Last_Year,
NTILE(5) OVER(ORDER BY Training_Times_Last_Year) AS GROUP_
from employees
```



t) Categorize employees based on training times last year as - Frequent Trainee, Moderate Trainee, Infrequent Trainee.

```
select emp_no 'Employee', job_role 'Job_Role', Training_Times_Last_Year,

case

when Training_Times_Last_Year <=2 then 'Infrequent Trainee'

when Training_Times_Last_Year <=4 then 'Moderate Trainee'

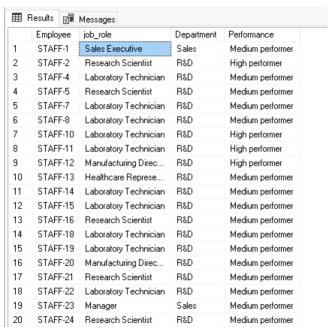
when Training_Times_Last_Year >4 then 'Frequent Trainee'

end 'Training_Frequency'

from employees
```



u) Categorize employees as 'High', 'Medium', or 'Low' performers based on their performance rating, using a CASE WHEN statement.



v) Use a CASE WHEN statement to categorize employees into 'Poor', 'Fair', 'Good', or 'Excellent' work-life balance based on their work-life balance score.

```
□ select emp_no Employee, Department,

case

when Work_Life_Balance = 1 then 'Poor'
when Work_Life_Balance = 2 then 'Fair'
when Work_Life_Balance = 3 then 'Good'
when Work_Life_Balance = 4 then 'Excellent'
end Work_life_balance
from employees
```



w) Group employees into 3 groups based on their stock option level using the [NTILE] function.

```
--W
=select Employee_Number,Department

Job_Role,Stock_option_level, T

NTILE(3) over(ORDER BY Stock_option_level desc) as Stock_op_grp

from employees
```

⊞ Results 🛍 Messages						
	Employee_Number	Job_Role	Stock_option_level	Stock_op_grp		
106	1682	R&D	2	1		
107	2012	R&D	2	1		
108	1803	R&D	2	1		
109	1596	R&D	2	1		
110	1514	R&D	2	1		
111	1558	R&D	2	1		
112	1787	Sales	2	1		
113	1908	Sales	2	1		
114	1995	Sales	2	1		
115	1479	Sales	2	1		
116	2019	Sales	2	1		
117	1482	R&D	2	1		
118	1449	R&D	2	1		
110	1771	D.D	1	1		

x) Find key reasons for Attrition in Company

```
⊟WITH WLB AS (
        SELECT
            Department, Job role,
             'Work_Life_Balance' AS Factor,
            COUNT(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
            (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100 / COUNT(*)) AS AttritionRate
        FROM Employees
        GROUP BY Work Life Balance, Department, Job role
     JS AS (
        SELECT
            Department, Job_role,
             'Job_Satisfaction' AS Factor,
            COUNT(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100 / COUNT(*)) AS AttritionRate
        FROM Employees
        GROUP BY Job_Satisfaction, Department, Job_role
ES AS (
   SELECT
       Department, Job role, 'Environment Satisfaction' AS Factor,
       COUNT(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
       (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100 / COUNT(*)) AS AttritionRate
   FROM Employees
   GROUP BY Environment_Satisfaction, Department, Job_role
YSLP AS (
   SELECT.
       Department, Job_role,
        'Years_since_last_promotion' AS Factor,
       COUNT(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
       (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100 / COUNT(*)) AS AttritionRate
   FROM Employees
   GROUP BY Years_since_last_promotion, Department, Job_role
 SELECT
       Department, Job_role, Factor, Attrition_Count, AttritionRate
 FROM WLB
 UNION ALL
 SELECT
       Department, Job_role, Factor, Attrition_Count, AttritionRate
 FROM JS
 UNION ALL
 SELECT
       Department, Job_role, Factor, Attrition_Count, AttritionRate
 FROM ES
 UNION ALL
 SELECT
       Department, Job_role, Factor, Attrition_Count, AttritionRate
 FROM YSLP
 ORDER BY AttritionRate DESC;
```

	Department	Job_role	Factor	Attrition_Count	AttritionRate
1	Sales	Sales Representative	Years_since_last_promotion	1	100
2	Sales	Sales Executive	Years_since_last_promotion	1	100
3	Sales	Sales Executive	Years_since_last_promotion	1	100
4	R&D	Laboratory Technician	Years_since_last_promotion	1	100
5	Sales	Sales Representative	Work_Life_Balance	9	88
6	R&D	Laboratory Technician	Work_Life_Balance	20	70
7	Sales	Sales Representative	Job_Satisfaction	12	58
8	Sales	Sales Representative	Years_since_last_promotion	44	50
9	HR	Human Resources	Job_Satisfaction	10	50
10	Sales	Sales Executive	Work_Life_Balance	12	50
11	R&D	Research Director	Years_since_last_promotion	2	50
12	R&D	Healthcare Represe	Years_since_last_promotion	2	50
13	R&D	Research Scientist	Years_since_last_promotion	2	50
14	Sales	Manager	Years_since_last_promotion	2	50
15	Sales	Sales Representative	Job_Satisfaction	21	47
16	Sales	Sales Representative	Environment_Satisfaction	11	45
17	Sales	Sales Executive	Years_since_last_promotion	9	44
18	Sales	Sales Representative	Years_since_last_promotion	16	43
19	R&D	Laboratory Technician	Environment_Satisfaction	53	41
20	Sales	Sales Representative	Environment_Satisfaction	22	40
21	HR	Human Resources	Environment_Satisfaction	10	40