



# **HR ANALYSIS**

(EMPLOYEE ATTRITION)

# **PROJECT**

# OVERVIEW



This project showcases the use of MySQL in MySQL Workbench to analyze current employee and attrite employee data. MySQL queries are written to answer important questions required by human resource managers.

# GOAL



The goal is to provide insights and identify areas for improvement for HR.

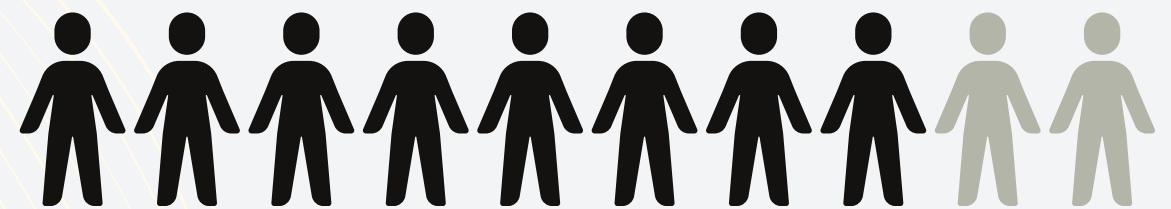


LET'S QUICKLY CHECK THE TOTAL NUMBER OF EMPLOYEES PRESENT IN THE TABLE.

## Query

```
select count(EmpID) as total_employees  
from `hr_analytics - csv`
```

**TOTAL EMPLOYEES:**  
**1413**

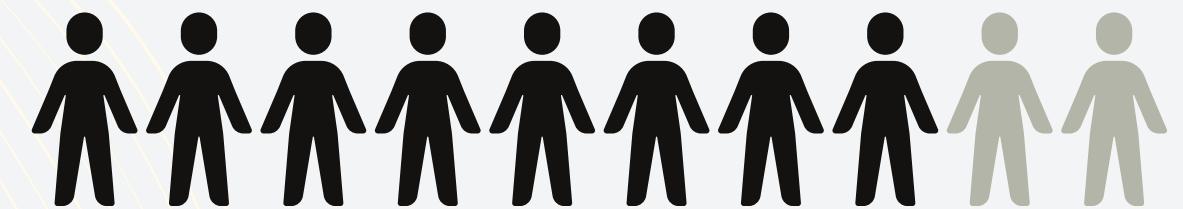


How many male and female employees were there in the company before Attrition?

## Query

```
select Gender, count(*)  
from `hr_analytics - csv`  
group by Gender
```

**847 Male and 566 Female employees  
were there in the company.**



# HOW MANY ARE THE CURRENT EMPLOYEES AND HOW MANY HAVE LEFT THE COMPANY?

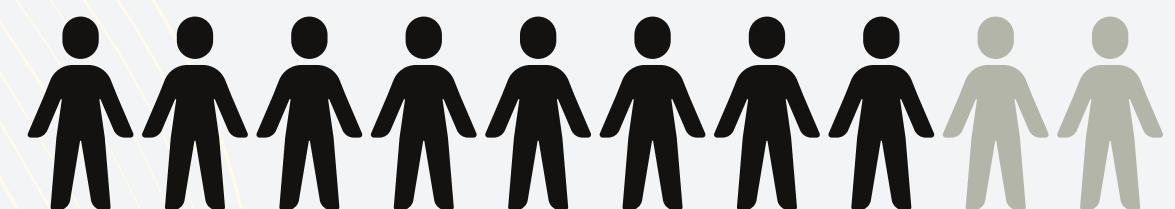
## Query

```
select Attrition, count(Attrition) as count  
from `hr_analytics - csv`  
group by Attrition
```

Insights:

**Current employees: 1184**

**left the company: 229**



# CURRENT EMPLOYEE ANALYSIS

WHO ARE THE TOP 5 CURRENT EMPLOYEES HAVING THE HIGHEST SALARY?

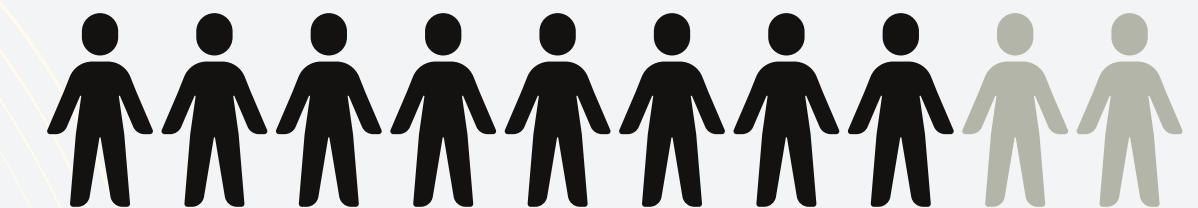
# Query

```
select EmpID, Department, MonthlyIncome  
from `hr_analytics - csv`  
where Attrition='No'  
order by MonthlyIncome desc  
limit 5
```

# Output

| EmpID | Department             | MonthlyIncome |
|-------|------------------------|---------------|
| RM191 | Research & Development | 19999         |
| RM747 | Research & Development | 19973         |
| RM852 | Research & Development | 19943         |
| RM166 | Research & Development | 19926         |
| RM919 | Sales                  | 19847         |

Top four employees are from “R&D” department and fifth one is from “Sales” department.



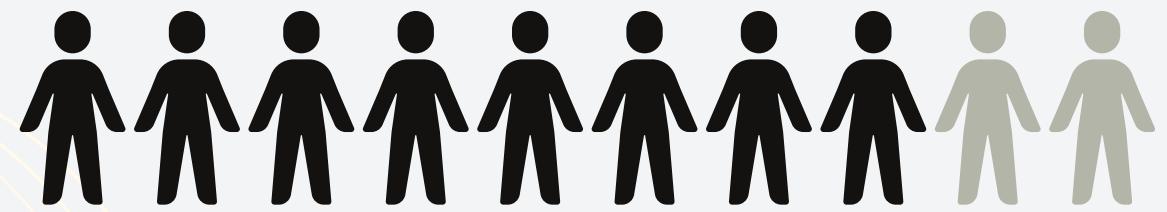
WHAT IS THE LARGEST DIFFERENCE BETWEEN MINIMUM AND MAXIMUM  
MONTHLY SALARY VALUES OF EACH DEPARTMENT FOR ALL CURRENT  
EMPLOYEES ?

## Query

```
select department, max(MonthlyIncome)-min(MonthlyIncome) as income_difference from `hr_analytics - csv`  
where Attrition='No'  
group by Department
```

## Insights

| department             | income_difference |
|------------------------|-------------------|
| Research & Development | 18948             |
| Sales                  | 18795             |
| Human Resources        | 17653             |

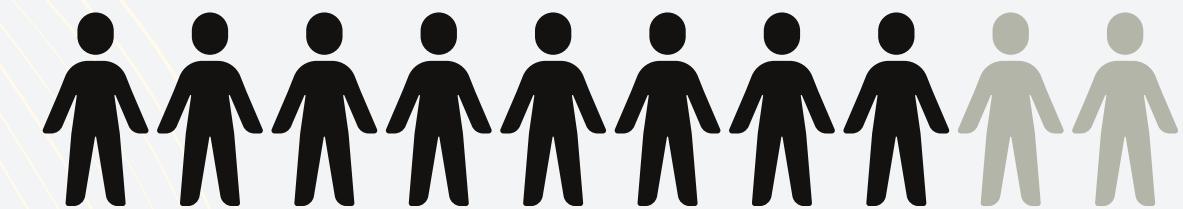


# WHICH DEPARTMENT HAS THE LEAST NUMBER OF CURRENT EMPLOYEES?

```
select department, count(EmpID) as count_
from `hr_analytics - csv`
where attrition ='No'
group by Department
order by count_
```

| department             | count_ |
|------------------------|--------|
| Human Resources        | 50     |
| Sales                  | 343    |
| Research & Development | 791    |

Human Resources has least  
number of current employees  
i.e. 50s

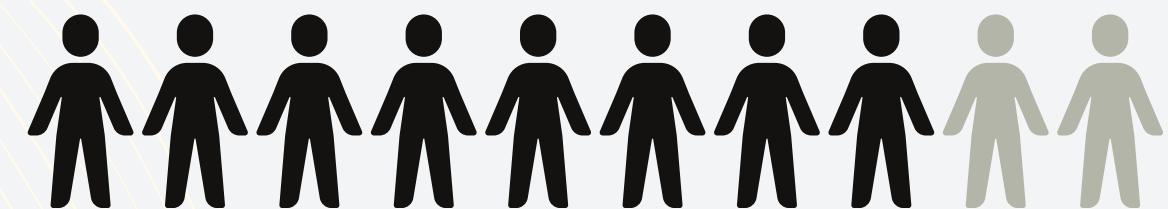


# WHICH DEPARTMENT HAS THE HIGHEST AVERAGE SALARY FOR FEMALE EMPLOYEES?

## Query

```
select Department, round(avg(MonthlyIncome),2) as average_income  
from `hr_analytics - csv`  
where Attrition='No' and Gender='Female'  
group by Department  
order by average_income desc  
limit 1
```

**Human Resources has the highest average monthly income for female employees i.e. 8761.14**



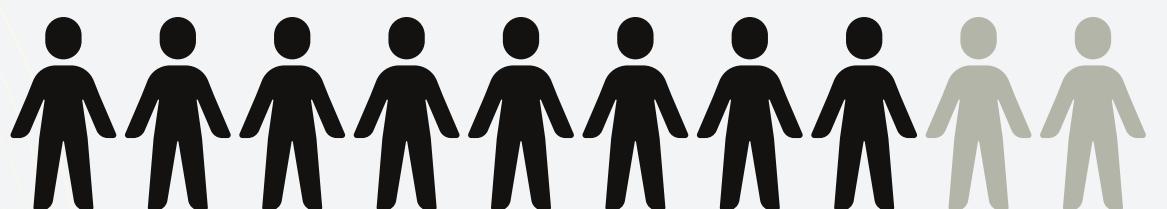
For each department, how is the Work life balance for the outstanding performing and highest salaried employees ?

```
with table1 as (
select EmpID, Department, n.WorkLifeBalance, MonthlyIncome,
row_number() over(partition by Department order by h.PerformanceRating desc , MonthlyIncome desc ) as rank_
from `hr_analytics - csv` h join `num- csv` n on h.WorkLifeBalance=n.Num
where Attrition='No')

select EmpID, Department, MonthlyIncome, WorkLifeBalance from table1
where rank_ in (1,2)
```

| EmpID  | Department             | MonthlyIncome | WorkLifeBalance |
|--------|------------------------|---------------|-----------------|
| RM106  | Human Resources        | 18844         | Better          |
| RM1097 | Human Resources        | 16437         | Better          |
| RM747  | Research & Development | 19973         | Better          |
| RM1010 | Research & Development | 19701         | Better          |
| RM919  | Sales                  | 19847         | Good            |
| RM1117 | Sales                  | 19586         | Better          |

The work life balance of the employees is neither Bad nor Best. It is Better.



# What are the departments where employees are highly satisfied with their job ?

```
select Department,  
       (sum(case when JobSatisfaction=4 or JobSatisfaction=3 then 1 else 0 end)/count(*)*100) as percent_satisfied  
  from `hr_analytics - csv`  
 where Attrition='No'  
 group by Department
```

| Department             | percent_satisfied |
|------------------------|-------------------|
| Research & Development | 63.2111           |
| Sales                  | 63.5569           |
| Human Resources        | 52.0000           |



In which departments, the highest percentage of male and female employees are doing “Overtime“ ?

## Query

```
select Department,  
(sum(case when Overtime='Yes' then 1 else 0 end)/count(*)*100) as percent_overtime  
from `hr_analytics - csv`  
where Gender='Male' and Attrition='No'  
group by Department  
order by percent_overtime desc
```

## Male Employees Percentage

| Department             | percent_overtime |
|------------------------|------------------|
| Research & Development | 22.5053          |
| Human Resources        | 22.2222          |
| Sales                  | 19.4872          |



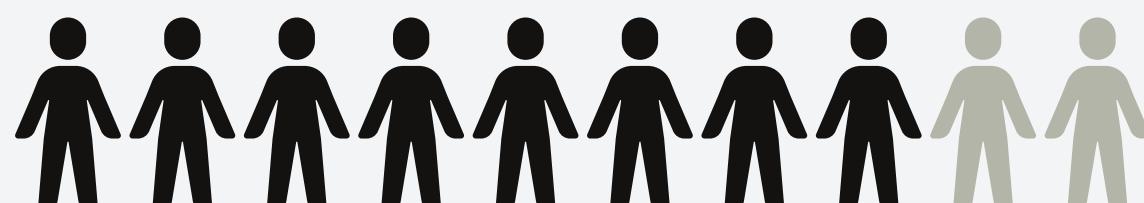
In which departments, the highest percentage of male and female employees are doing “Overtime“ ?

## Query

```
select Department,  
(sum(case when Overtime='Yes' then 1 else 0 end)/count(*)*100) as percent_overtime  
from `hr_analytics - csv`  
where Gender='Female' and Attrition='No'  
group by Department  
order by percent_overtime desc
```

## Female Employees Percentage

| Department             | percent_overtime |
|------------------------|------------------|
| Sales                  | 28.3784          |
| Research & Development | 25.9375          |
| Human Resources        | 21.4286          |





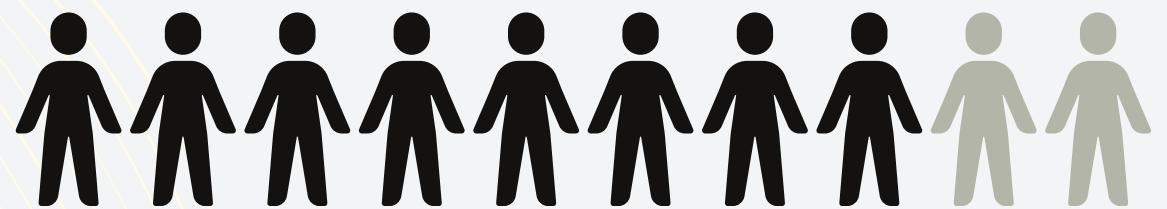
# **ATTRITE EMPLOYEE ANALYSIS**

How many male and female employees have left the company ?

## Query

```
select Gender, count(*)  
from `hr_analytics - csv`  
where Attrition='Yes'  
group by Gender
```

**145 Male and 84 Female employees  
have left the company.**



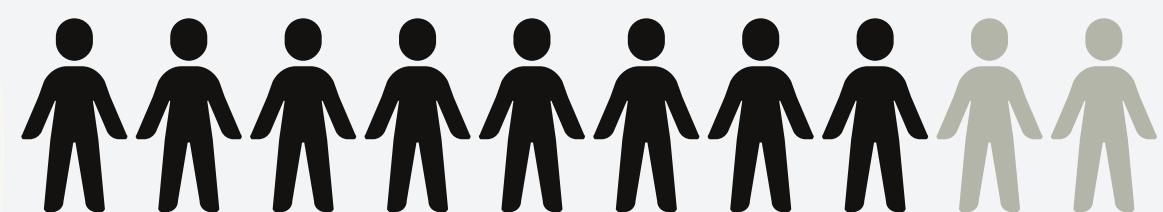
Department wise, How many employees left the company ?

## Query

```
select Department, count(*) as number_of_employees_left  
from `hr_analytics - csv`  
where Attrition='Yes'  
group by Department
```

## Output

| Department             | number_of_employees_left |
|------------------------|--------------------------|
| Research & Development | 127                      |
| Sales                  | 90                       |
| Human Resources        | 12                       |



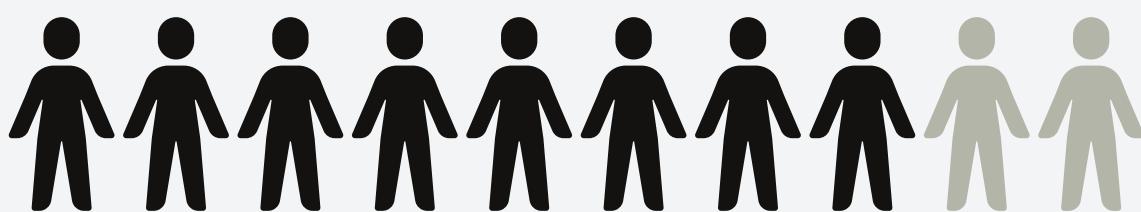
What was the average salary for each male and female employee who has left the company ?

## Query

```
select Gender, avg(MonthlyIncome) as average_income  
from `hr_analytics - csv`  
where attrition = 'Yes'  
group by Gender
```

## Output

| Gender | average_income |
|--------|----------------|
| Female | 4752.4881      |
| Male   | 4749.9793      |

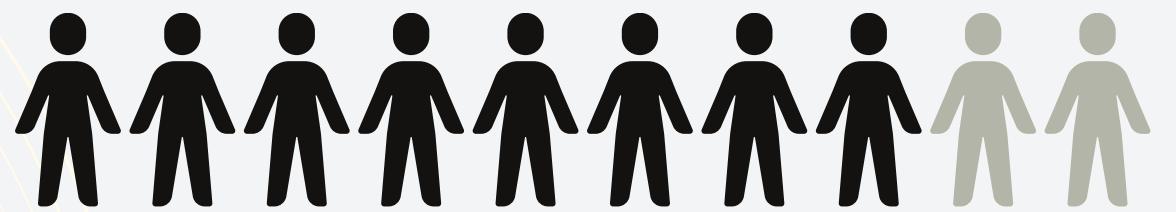


What was the average total company tenure for each churn employee just before they left?

## Query

```
select round(avg(YearsAtCompany),2) as average_company_tenure  
from `hr_analytics - csv`  
where attrition ='Yes'
```

The average company tenure was  
5.10 years



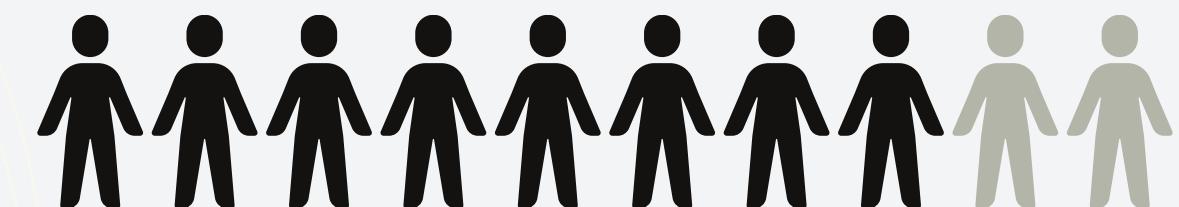
How many employees of each age group have left the company ?

## Query

```
select AgeGroup, count(*) as count_
from `hr_analytics - csv`
where Attrition='Yes'
group by AgeGroup
order by count_ desc
```

## Output

| AgeGroup | count_ |
|----------|--------|
| 26-35    | 111    |
| 18-25    | 43     |
| 36-45    | 41     |
| 46-55    | 26     |
| 55+      | 8      |



What was the percentage of employees for each level of job satisfaction ?

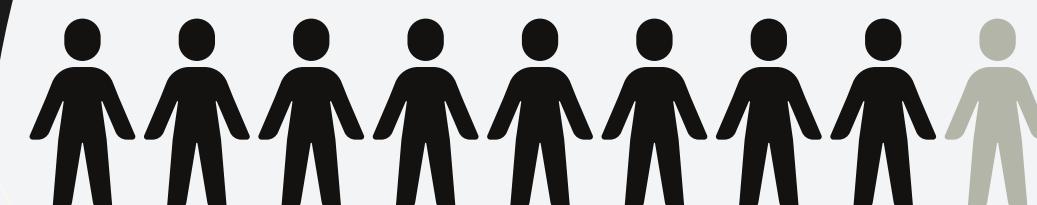
## Query

```
with cte as (
  select h.JobSatisfaction as h_,n.JobSatisfaction,count(*) as count_
  from `hr_analytics - csv` h join `num- csv` n on h.JobSatisfaction=n.Num
  where Attrition='Yes'
  group by h.JobSatisfaction
  order by count_ desc)

select c.JobSatisfaction, count(EmpID) as total_employees,
count_ as Employees_left, (count_/count(EmpID)*100) as percentage_employees_left
from cte c join `hr_analytics - csv` h1 on h1.JobSatisfaction=c.h_
group by h1.JobSatisfaction
order by percentage_employees_left desc
```

## Output

| JobSatisfaction | total_employees | Employees_left | percentage_employees_left |
|-----------------|-----------------|----------------|---------------------------|
| Low             | 274             | 62             | 22.6277                   |
| Medium          | 274             | 46             | 16.7883                   |
| High            | 426             | 71             | 16.6667                   |
| Very High       | 439             | 50             | 11.3895                   |



What was the percentage of employees for each level of environment satisfaction ?

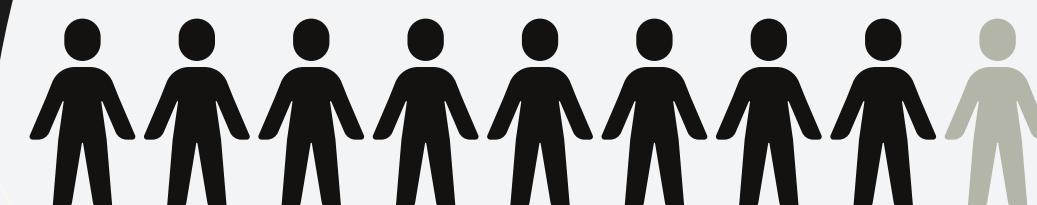
## Query

```
with cte as(
select h.EnvironmentSatisfaction as h_,n.EnvironmentSatisfaction,count(*) as count_
from `hr_analytics - csv` h join `num- csv` n on h.EnvironmentSatisfaction=n.Num
where Attrition='Yes'
group by h.EnvironmentSatisfaction
order by count_ desc)

select c.EnvironmentSatisfaction, count(EmpID) as total_employees,
count_ as Employees_left, (count_/count(EmpID)*100) as percentage_employees_left
from cte c join `hr_analytics - csv` h1 on h1.EnvironmentSatisfaction=c.h_
group by h1.EnvironmentSatisfaction
order by percentage_employees_left desc
```

## Output

| EnvironmentSatisfaction | total_employees | Employees_left | percentage_employees_left |
|-------------------------|-----------------|----------------|---------------------------|
| Low                     | 270             | 69             | 25.5556                   |
| Medium                  | 278             | 40             | 14.3885                   |
| High                    | 437             | 62             | 14.1876                   |
| Very High               | 428             | 58             | 13.5514                   |



What was the average salary of employees for each role in the company ?

## Query

```
select JobRole, count(*) as count_, avg(MonthlyIncome) as average_income  
from `hr_analytics - csv`  
where Attrition='Yes'  
group by JobRole  
order by average_income
```

## Output

| JobRole                   | count_ | average_income |
|---------------------------|--------|----------------|
| Sales Representative      | 33     | 2364.73        |
| Research Scientist        | 44     | 2736.70        |
| Laboratory Technician     | 60     | 2879.38        |
| Human Resources           | 12     | 3715.75        |
| Manufacturing Director    | 9      | 7067.44        |
| Sales Executive           | 55     | 7433.35        |
| Healthcare Representative | 9      | 8548.22        |
| Manager                   | 5      | 16797.40       |
| Research Director         | 2      | 19395.50       |

