HR DATA ANALYSIS

Presented By: Mansi Tyagi

Using MySQL Workbench

ABOUT DATASET

```
USE hr_db;
SHOW TABLES;
```

Result Grid | Filter Rows:

Tables_in_hr_db
employee_survey_data
general_data
manager_survey_data

Table: general_data

Columns:

EmpName	text
Age	int
Attrition	text
BusinessTravel	text
Department	text
DistanceFromHome	int
Education	int
EducationField	text
EmployeeCount	int
EmployeeID	int
Gender	text
JobLevel	int
JobRole	text
MaritalStatus	text
MonthlyIncome	int
NumCompaniesWorked	int
Over18	text
PercentSalaryHike	int
StandardHours	int
StockOptionLevel	int
TotalWorkingYears	int
TrainingTimesLastYear	int
YearsAtCompany	int
YearsSinceLastPromotion	int
YearsWithCurrManager	int

Table: employee_survey_data

Columns:

EmployeeID	int
EnvironmentSatisfaction	int
JobSatisfaction	int
WorkLifeBalance	int

Table: manager_survey_data

Columns:

EmployeeID	int
JobInvolvement	int
PerformanceRating	int

****

1. RETRIEVE THE TOTAL NUMBER OF EMPLOYEES IN THE DATASET.

total number of employees

```sql
SELECT count(EmployeeID) AS Total_employees
FROM general_data;
```

Total_employees
4382

2. LIST ALL UNIQUE JOB ROLES IN THE DATASET.

different job roles

```
SELECT DISTINCT JobRole
FROM general_data;
```

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.

average age of employees

```
SELECT round(avg(Age),2) AS Avg_employee_age
FROM general_data;
```

Avg_employee_age
36.93

The company has a young workforce with an average of 36.93

4. RETRIEVE THE NAMES AND AGES OF EMPLOYEES WHO HAVE WORKED AT THE COMPANY FOR MORE THAN 5 YEARS.

employees worked at the company (>5 years)

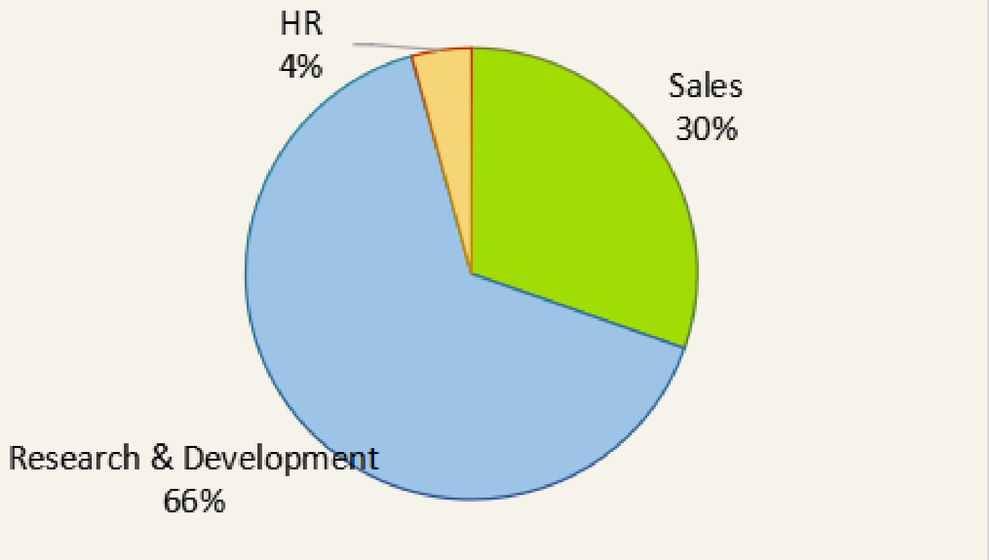
```sql
SELECT EmpName, Age FROM general_data
WHERE YearsAtCompany > 5
ORDER BY EmpName;
```

It is a positive sign that 47.2% worked at the company for more than 5 years which indicates good retention.

5. GET A COUNT OF EMPLOYEES GROUPED BY THEIR DEPARTMENT.

no. of employees in each department

```sql
SELECT Department, count(EmployeeID) AS No_of_Employees
FROM general_data
GROUP BY Department;
```



6. LIST EMPLOYEES WHO HAVE 'HIGH' JOB SATISFACTION.

employees having high job satisfaction

```
SELECT employee_survey_data.EmployeeID, general_data.EmpName
FROM employee_survey_data
INNER JOIN general_data
ON employee_survey_data.EmployeeID = general_data.EmployeeID
WHERE JobSatisfaction = 4;
```

It's encouraging that 1000 employees report 'High' job satisfaction.

7. FIND THE HIGHEST MONTHLY INCOME IN THE DATASET.

highest monthly income

```sql
SELECT EmployeeID, EmpName, MonthlyIncome AS HighestMonthlyIncome
FROM general_data
WHERE MonthlyIncome = (SELECT max(MonthlyIncome) FROM general_data);
```

EmployeeID	EmpName	HighestMonthlyIncome
386	KEVIN LABANOWSKI	199990
1856	DAVID KUCIA	199990
3326	LAWRENCE LAU	199990

8. LIST EMPLOYEES WHO HAVE 'TRAVEL_RARELY' AS THEIR BUSINESS TRAVEL TYPE.

employees who travel rarely

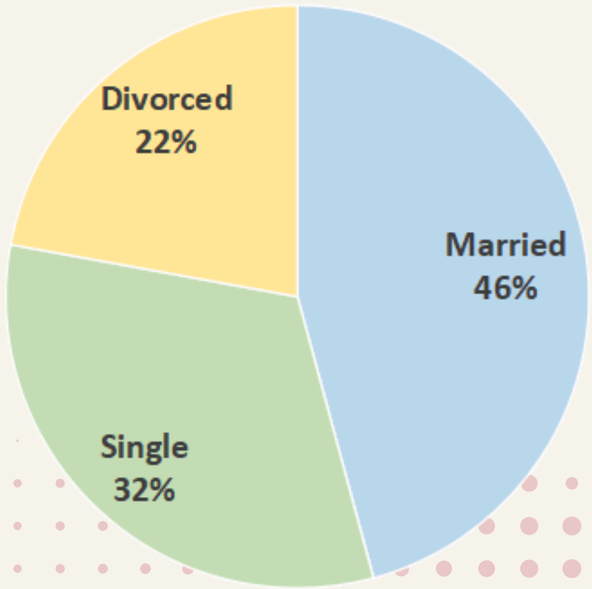
```
SELECT EmpName FROM general_data
WHERE BusinessTravel = "Travel_Rarely";
```

There are 3109 employees who travel rarely for business purpose.

9. RETRIEVE THE DISTINCT MARITAL STATUS CATEGORIES IN THE DATASET.

different marital status

```sql
SELECT DISTINCT MaritalStatus, count(EmployeeID) AS No_of_Employees
FROM general_data
GROUP BY MaritalStatus;
```



There is a good mix of marital status with the majority being married.

10. GET A LIST OF EMPLOYEES WITH MORE THAN 2 YEARS OF WORK EXPERIENCE BUT LESS THAN 4 YEARS IN THEIR CURRENT ROLE.

employees who worked for 3 years

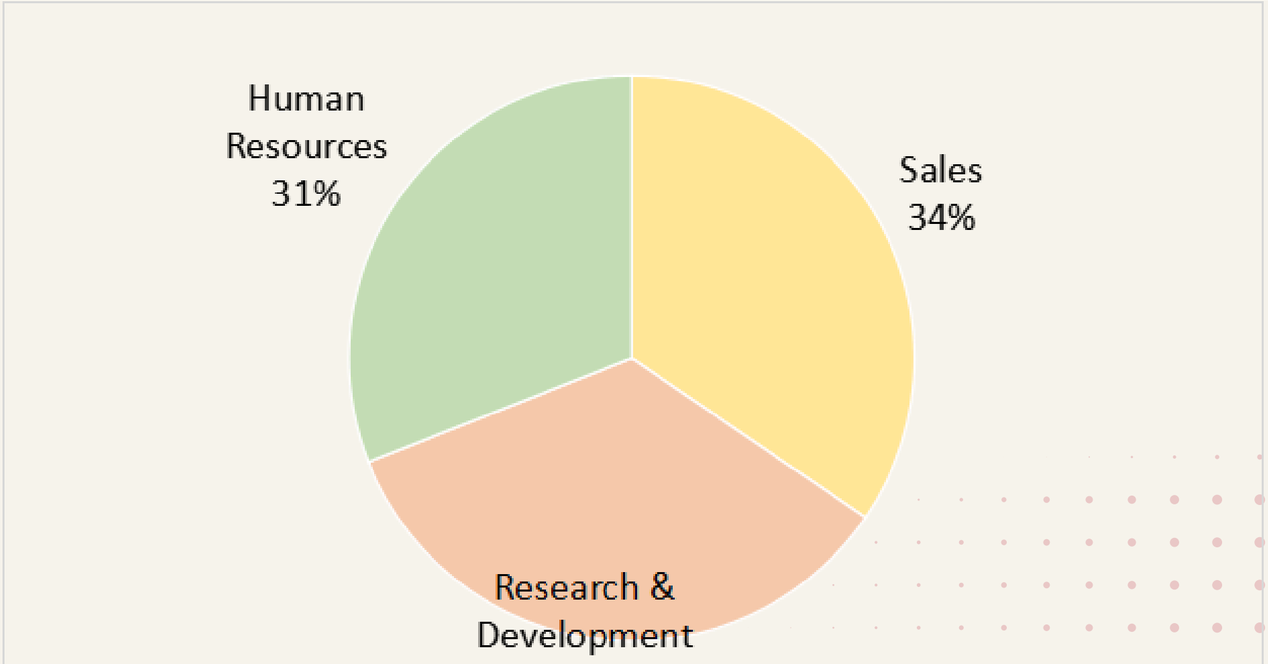
```sql
SELECT EmployeeID, EmpName FROM general_data
WHERE TotalWorkingYears = 3;
```

There are a total of 126 employees who worked for 3 years.

11. FIND THE AVERAGE DISTANCE FROM HOME FOR EMPLOYEES IN EACH DEPARTMENT.

avg distance from home by each department

```sql
SELECT Department, round(avg(DistanceFromHome),2) AS Avg_DistanceFromHome
FROM general_data
GROUP BY Department;
```



12. RETRIEVE THE TOP 5 EMPLOYEES WITH THE HIGHEST MONTHLY INCOME.

employee having highest monthly income

```sql
SELECT EmployeeID, EmpName, MonthlyIncome FROM general_data
ORDER BY MonthlyIncome DESC
LIMIT 5;
```

EmployeeID	EmpName	MonthlyIncome
386	KEVIN LABANOWSKI	199990
1856	DAVID KUCIA	199990
3326	LAWRENCE LAU	199990
942	KEVIN MCNAUGHTON	199730
3882	SHANNON STABILE	199730

13. CALCULATE THE PERCENTAGE OF EMPLOYEES WHO HAVE HAD A PROMOTION IN THE LAST YEAR.

percentage of employees who had a promotion in the last year

```sql
SELECT COUNT(*) FROM general_data;    -- Total number of employees

SELECT COUNT(*) FROM general_data
WHERE YearsSinceLastPromotion = 0;    -- no. of employees who have had a promotion in the last
year

SELECT                                --This Outer SELECT statement is used to retrieve the
(SELECT COUNT(*) FROM general_data     --calculated % and labeeled it as %PromotedLastYear
WHERE YearsSinceLastPromotion = 0) * 100.0 / (SELECT COUNT(*) FROM general_data) AS
PercentagePromotedLastYear;
```

PercentagePromotedLastYear
39.47969

It is a positive sign that about 40% of employees received promotions in the last year.

14. LIST THE EMPLOYEES WITH THE HIGHEST AND LOWEST ENVIRONMENT SATISFACTION.

highest and lowest Environment Satisfaction

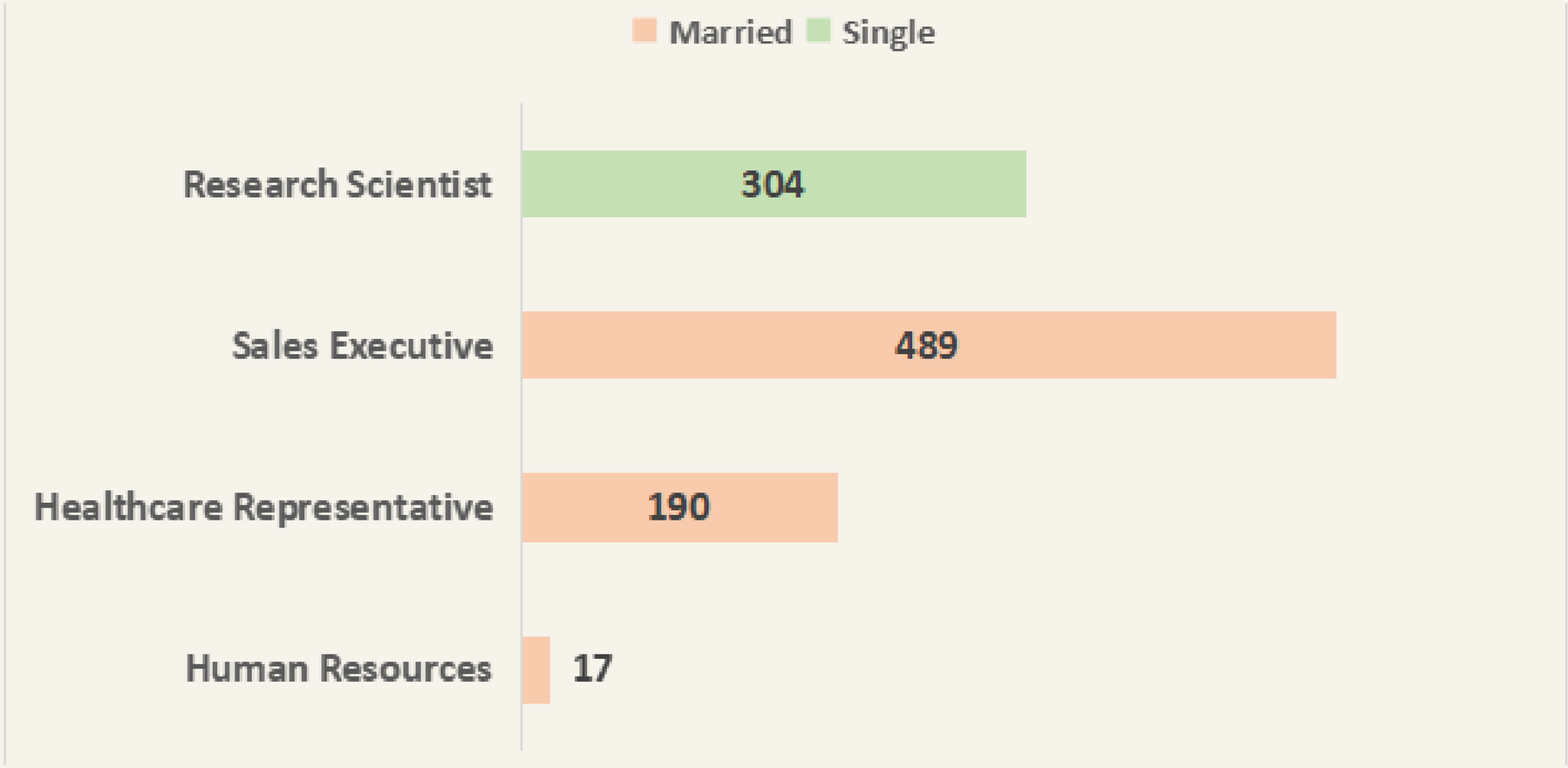
```sql
SELECT g.EmpName, e.EnvironmentSatisfaction
FROM general_data AS g
INNER JOIN employee_survey_data AS e
ON g.EmployeeID = e.EmployeeID
WHERE e.EnvironmentSatisfaction = (SELECT max(EnvironmentSatisfaction) FROM
employee_survey_data)
OR
e.EnvironmentSatisfaction = (SELECT min(EnvironmentSatisfaction) FROM employee_survey_data);
```

Highest = 609
Lowest = 391

15. FIND THE EMPLOYEES WHO HAVE THE SAME JOB ROLE AND MARITAL STATUS.

same job role and marital status

```sql
SELECT e1.EmpName, e1.JobRole, e1.MaritalStatus
FROM general_data e1, general_data e2
WHERE e1.JobRole = e2.JobRole
AND e1.MaritalStatus = e2.MaritalStatus
AND e1.EmployeeID != e2.EmployeeID;
```


Married
Single
Research Scientist
304
Sales Executive
489
Healthcare Representative
190
Human Resources
17

16. LIST THE EMPLOYEES WITH THE HIGHEST TOTAL WORKING YEARS WHO ALSO HAVE A PERFORMANCE RATING OF 4.

highest total working years and performance rating

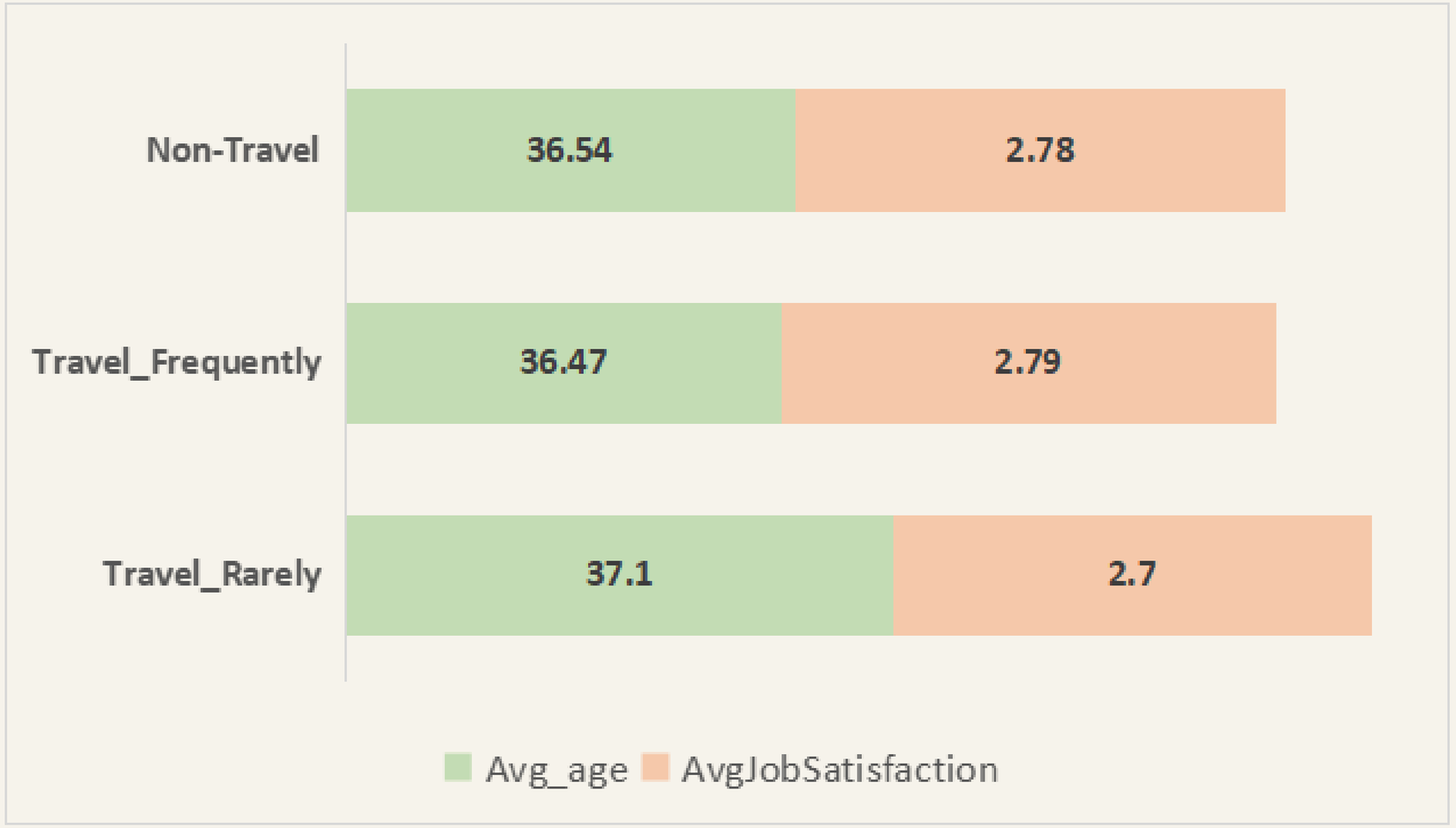
```sql
SELECT g.EmpName, g.TotalWorkingYears, m.PerformanceRating
FROM general_data AS g
INNER JOIN manager_survey_data AS m
ON g.EmployeeID = m.EmployeeID
WHERE m.PerformanceRating = 4
AND g.TotalWorkingYears = (
    SELECT MAX(g.TotalWorkingYears)
    FROM general_data AS g
    INNER JOIN manager_survey_data AS m
    ON g.EmployeeID = m.EmployeeID
    WHERE m.PerformanceRating = 4
);   -- ensures that the maximum Total Working Years is calculated only for employees with a Performance Rating of 4
```

EmpName	TotalWorkingYears	PerformanceRating
STEVEN SETO	35	4
SHARON LEGENZA	35	4
EMILY MURASE	35	4

17. CALCULATE THE AVERAGE AGE AND JOBSATISFACTION FOR EACH BUSINESSTRAVEL TYPE.

average age and job satisfaction

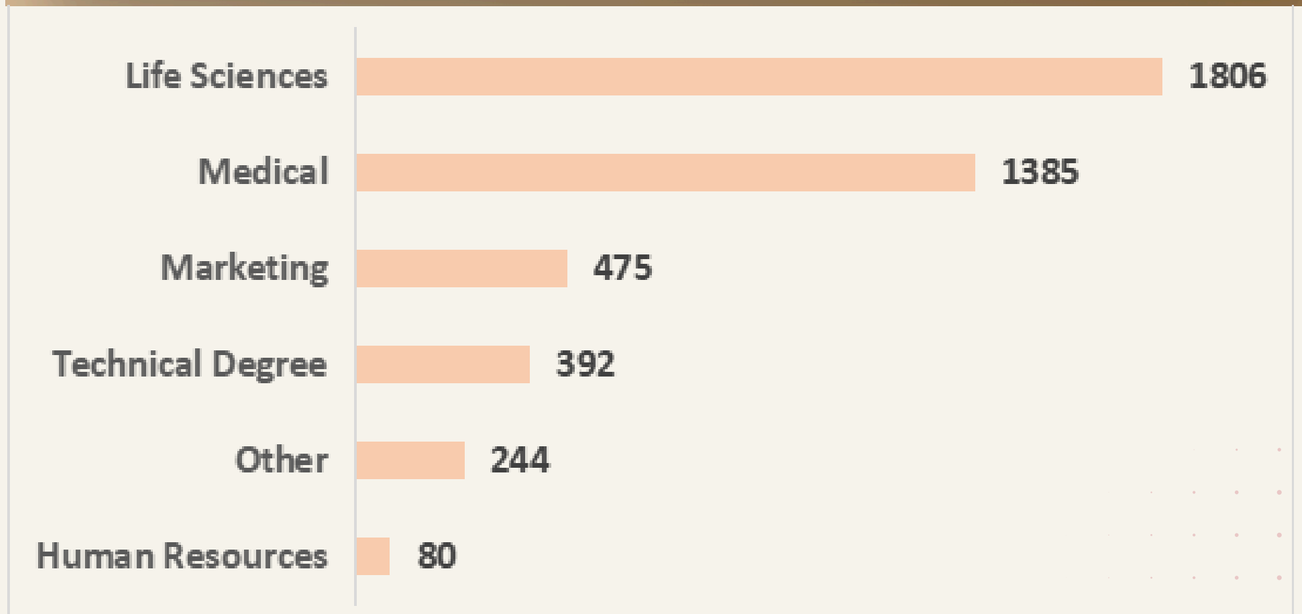
```sql
SELECT g.BusinessTravel, round(avg(g.Age),2) AS Avg_age,
round(avg(e.JobSatisfaction),2) AS AvgJobSatisfaction

FROM general_data AS g
INNER JOIN employee_survey_data AS e
ON g.EmployeeID = e.EmployeeID
GROUP BY BusinessTravel;
```


Non-Travel
36.54
2.78
Travel_Frequently
36.47
2.79
Travel_Rarely
37.1
2.7
Avg_age
AvgJobSatisfaction

18. RETRIEVE THE MOST COMMON EDUCATION FIELD AMONG EMPLOYEES.

most common education field

```sql
SELECT EducationField, count(*) AS Count
FROM general_data
GROUP BY EducationField
ORDER BY count(*) > 1 desc, EducationField;
```



Life Sciences is the dominant field of education with 41.2% of employees.

19. LIST THE EMPLOYEES WHO HAVE WORKED FOR THE COMPANY THE LONGEST BUT HAVEN'T HAD A PROMOTION.

longest working employee but had not promoted

```sql
SELECT EmpName, TotalWorkingYears, YearsSinceLastPromotion
FROM general_data
WHERE YearsSinceLastPromotion = 0
AND TotalWorkingYears =
(
    SELECT MAX(TotalWorkingYears)
    FROM general_data
);
```

There is no such employees who have worked longest for the company but have not promoted.

RECOMMENDATION

Recommendation 1

Analyze factors that influence job satisfaction and environment satisfaction. This can help to identity the areas of improvement.

Recommendation 2

Analyze factors that influence promotions (department, experience etc.) to understand if there is any pattern.

****

Recommendation 3

Analyze the reason behind the top 5 employees who have high monthly income. Is it performance, specialization or seniority?

****

THANK YOU