

IBM Employee Attrition and Performance Analysis Using MySQL



Submitted By:
JISA VARGHESE

TABLE OF CONTENTS

- Introduction
- Objectives
- Dataset Overview
- Methodology
- Detailed Analysis
- Key Findings and Observations
- Surprising Factors
- Recommendations
- Conclusion
- References

INTRODUCTION

Employee attrition is a critical concern for organizations striving to retain top talent and maintain operational efficiency. High turnover rates can lead to increased costs, loss of institutional knowledge, and disruption in workplace culture. Analyzing attrition trends and performance data can help organizations proactively address the underlying factors contributing to employee turnover.

This report leverages **SQL** to perform an in-depth analysis of the **IBM HR Analytics Employee Attrition & Performance** dataset. Through this report, we aim to uncover actionable insights that can help HR teams implement strategies to reduce attrition, improve employee satisfaction, and optimize workforce performance.

OBJECTIVES

The primary objectives of this Pandas-based analysis are:

1. Understand Employee Attrition Patterns

- Analyze attrition trends by age, gender, marital status, and department.
- Identify which groups are most vulnerable to attrition.

2. Evaluate Income and Performance Relationships

- Explore the correlation between monthly income, performance ratings, and attrition.
- Assess whether financial incentives could help reduce turnover.

3. Analyze the Impact of Tenure and Distance

- Investigate how factors like years with the company, years under the current manager, and commuting distance influence attrition.

4. Provide Actionable Insights

- Use findings to propose strategic interventions to improve retention and enhance employee satisfaction. This analysis aims to empower HR teams with data-driven recommendations to address workforce challenges effectively.

DATASET OVERVIEW

The dataset consists of **1,470** records and **35** features, including demographic details, professional characteristics, satisfaction levels, and attrition status.

ATTRIBUTES:

- **Age:** Age of the employee in years.
- **Attrition:** Whether the employee left the company (Yes/No).
- **BusinessTravel:** Frequency of business travel (e.g., Rarely, Often).
- **DailyRate:** Daily income of the employee in USD.
- **Department:** Department of the employee (e.g., Sales, R&D).
- **DistanceFromHome:** Distance from home to the workplace (in miles).
- **Education:** Education level (1: Below College, 5: Doctorate).
- **EducationField:** Field of education (e.g., Life Sciences, Technical Degree, Marketing).
- **EmployeeCount:** Always 1
- **EmployeeNumber:** Unique identifier for each employee (not useful for analysis).

- **EnvironmentSatisfaction:** Employee's satisfaction with the work environment (1-4).
- **Gender:** Gender of the employee (Male/Female).
- **HourlyRate:** Hourly rate of the employee's pay.
- **JobInvolvement:** Employee's involvement level in their job (1-4).
- **JobLevel:** Level of the job (1-5).
- **JobRole:** Designation of the employee (e.g., Manager, Analyst).
- **JobSatisfaction:** Employee's job satisfaction (1: Low, 4: High).
- **MaritalStatus:** Marital status of the employee (Single/Married/Divorced).
- **MonthlyIncome:** Monthly salary in USD.
- **MonthlyRate:** Monthly rate of the employee's pay.
- **NumCompaniesWorked:** Number of companies the employee has worked for.
- **Over18:** whether an employee is above the age of 18(yes/no).
- **OverTime:** Whether the employee works overtime (Yes/No).
- **PercentSalaryHike:** Percentage increase in salary.
- **PerformanceRating:** Performance rating (1: Low, 4: Excellent).
- **RelationshipSatisfaction:** Satisfaction level with workplace relationships (1-4).
- **StandardHours:** Standard hours (always 80, irrelevant for analysis).

- **StockOptionLevel:** Level of stock options granted to the employee (0-3).
- **TotalWorkingYears:** Total years of professional experience.
- **TrainingTimesLastYear:** Number of training sessions attended last year.
- **WorkLifeBalance:** Work-life balance rating (1-4).
- **YearsAtCompany:** Total years spent at the company.
- **YearsInCurrentRole:** Number of years in the current role.
- **YearsSinceLastPromotion:** Years since the last promotion.
- **YearsWithCurrManager:** Years working with the current manager.

EDA STEP BY STEP PROCESS

1. Understanding the Dataset
2. Data Cleaning
3. Data type Conversion
4. Feature Engineering
5. Data Analysis
 - Univariate analysis
 - Bivariate analysis
 - Multivariate analysis
6. Insights Extraction
7. Conclusion and Recommendations

DATA PREPROCESSING

1.UNDERSTANDING THE DATASET

DESCRIBE ibmemployees;

2.HANDLING MISSING AND DUPLICATE VALUES:

No missing or duplicate values were found in the dataset after checking the dataset.

SELECT COLUMN_NAME, COUNT() AS Missing_Values FROM IBMEmployee WHERE COLUMN_NAME IS NULL GROUP BY COLUMN_NAME;*

*SELECT *, COUNT(*) FROM IBMEmployee GROUP BY EmployeeNumber, Age, Department, MonthlyIncome HAVING COUNT(*) > 1;*

3.FEATURE ENGINEERING:

Created new features, such as **AgeGroup**, **RecentlyPromoted**, **IncomeLevel**, based on domain knowledge.

```
ALTER TABLE ibmemployee
ADD Age_Category VARCHAR(10),
ADD Income_Category VARCHAR(10),
ADD RecentlyPromoted VARCHAR(3);
UPDATE ibmemployee
SET Age_Category = CASE
    WHEN Age < 30 THEN 'Young'
    WHEN Age < 50 THEN 'Adult'
    ELSE 'Senior'
END,
Income_Category = CASE
    WHEN MonthlyIncome < 4000 THEN 'Low'
    WHEN MonthlyIncome < 8000 THEN 'Medium'
    WHEN MonthlyIncome < 12000 THEN 'High'
    ELSE 'Very High'
END,
RecentlyPromoted = CASE
    WHEN YearsSinceLastPromotion = 0 THEN 'Yes'
    ELSE 'No'
END;
```

4.COLUMNS NOT USEFUL FOR ANALYSIS:

- **EmployeeCount:** All values are 1. This column does not provide any meaningful variation for analysis.
- **StandardHours:** All values are 80. This column is uniform and not useful for identifying trends or insights.
- **EmployeeNumber:** This is likely an identifier and not meaningful for analysis or modeling.
- **Over18:** Only contains the value 'Yes' for all rows

```
ALTER TABLE IBMEmployee  
DROP COLUMN EmployeeCount,  
DROP COLUMN Over18,  
DROP COLUMN StandardHours;
```

DATA ANALYSIS

1.OVERALL ATTRITION RATE

The attrition rate is **16%** of total employees, indicating that approximately 1 in 6 employees have left the organization..

SELECT

COUNT() AS Total_Employees,*

*SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,*

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee;

Total_Employees	Attrition_Count	Attrition_Percentage
1470	237	16.12

2. ATTRITION BY DEPARTMENT:

High attrition in **Sales** and **HR** department (20%).

The high attrition rate in Sales and HR departments is likely due to high job stress, performance pressure in Sales, and emotional burnout or limited career growth opportunities in HR.

```
SELECT
    Department,
    COUNT(*) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,
    ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage
FROM ibmemployee
GROUP BY Department
ORDER BY Attrition_Percentage DESC;
```

Department	Total_Employees	Attrition_Count	Attrition_Percentage
Sales	446	92	20.63
Human Resources	63	12	19.05
Research & Development	961	133	13.84

3.ATTRITION BY GENDER

Males (17%) have a slightly higher attrition rate than **females (15%)**. This may be due to greater pursuit of external career opportunities or dissatisfaction with current job roles and compensation.

```
SELECT
    Gender,
    COUNT(*) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,
    ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100.0
/ COUNT(*), 2) AS Attrition_Percentage
FROM ibmemployee
GROUP BY Gender
ORDER BY Attrition_Percentage DESC;
```

Gender	Total_Employees	Attrition_Count	Attrition_Percentage
Male	882	150	17.01
Female	588	87	14.80

4.ATTRITION BY AGE GROUP:

Employees in the **20–30 age group** exhibit the highest attrition rate at **28%**, likely due to career exploration and unmet expectations. In contrast, employees aged **40 and above** have a significantly lower attrition rate of **12.5%**, indicating greater job stability and loyalty. This suggests that younger employees are more likely to leave in search of better opportunities.

```
SELECT
    Age_Category,
    COUNT(*) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,
    ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage
FROM ibmemployee
GROUP BY Age_Category
ORDER BY Attrition_Percentage DESC;
```

Age_Category	Total_Employees	Attrition_Count	Attrition_Percentage
Young	326	91	27.91
Senior	173	23	13.29
Adult	971	123	12.67

5.ATTRITION BY MARITAL STATUS

Single employees have a higher attrition rate(25.5%). This may be due to greater flexibility and fewer personal commitments, allowing them to pursue new opportunities more freely.

```
SELECT
    MaritalStatus,
    COUNT(*) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,
    ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage
FROM ibmemployee
GROUP BY MaritalStatus
ORDER BY Attrition_Percentage DESC;
```

MaritalStatus	Total_Employees	Attrition_Count	Attrition_Percentage
Single	470	120	25.53
Married	673	84	12.48
Divorced	327	33	10.09

6.ATTRITION BY EDUCATION LEVEL

Employees with higher education (such as Master's or Doctorate) are more likely to leave the company, possibly due to higher expectations for career advancement.

SELECT

Education,

COUNT() AS Total_Employees,*

SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS

Attrition_Count,

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100.0
/ COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY Education

ORDER BY Attrition_Percentage DESC;

Education	Total_Employees	Attrition_Count	Attrition_Percentage
1	170	31	18.24
3	572	99	17.31
2	282	44	15.60
4	398	58	14.57
5	48	5	10.42

7.ATTRITION BY EDUCATION FIELD

Employees in Human Resources, Technical Degree, and Marketing roles may leave due to stress, lack of growth, or skill mismatches. High demand for technical skills and pressure in marketing jobs can also push them to look for better opportunities.

SELECT

EducationField,

COUNT() AS Total_Employees,*

*SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,*

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY EducationField

ORDER BY Attrition_Percentage DESC;

EducationField	Total_Employees	Attrition_Count	Attrition_Percentage
Human Resources	27	7	25.93
Technical Degree	132	32	24.24
Marketing	159	35	22.01
Life Sciences	606	89	14.69
Medical	464	63	13.58
Other	82	11	13.41

8. ATTRITION VS. MONTHLY INCOME

Lower-income employees (earning below **\$4,000**) have a significantly higher attrition rate(25%), while those earning above **\$8,000** are less likely to leave(10%),suggesting dissatisfaction with financial compensation as a key factor.

SELECT

Income_Category,

COUNT() AS Total_Employees,*

SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY Income_Category

ORDER BY Attrition_Percentage DESC;

Income_Category	Total_Employees	Attrition_Count	Attrition_Percentage
Low	542	137	25.28
High	186	29	15.59
Medium	547	60	10.97
Very High	195	11	5.64

9.ATTRITION VS. JOB SATISFACTION

Employees with **low job satisfaction (1 or 2)** are far more likely to leave compared to those with high satisfaction (**3 or 4**).

SELECT

JobSatisfaction,

COUNT() AS Total_Employees,*

*SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,*

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY JobSatisfaction

ORDER BY Attrition_Percentage DESC;

JobSatisfaction	Total_Employees	Attrition_Count	Attrition_Percentage
1	289	66	22.84
3	442	73	16.52
2	280	46	16.43
4	459	52	11.33

10.ATTRITION VS. WORK-LIFE BALANCE

Employees with **poor work-life balance (rating of 1)** are more likely to leave compared to those with higher ratings (**3–4**).

SELECT

WorkLifeBalance,

COUNT() AS Total_Employees,*

*SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,*

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY WorkLifeBalance

ORDER BY Attrition_Percentage DESC;

WorkLifeBalance	Total_Employees	Attrition_Count	Attrition_Percentage
1	80	25	31.25
4	153	27	17.65
2	344	58	16.86
3	893	127	14.22

11. ATTRITION BY JOB ROLE AND DEPARTMENT:

Specific roles such as **Sales Representative, Lab Technician, HR** have the highest percentage.

SELECT

Department,

JobRole,

COUNT() AS Total_Employees,*

SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY Department, JobRole

ORDER BY Department, Attrition_Percentage DESC;

Department	JobRole	Total_Employees	Attrition_Count	Attrition_Percentage
Human Resources	Human Resources	52	12	23.08
Human Resources	Manager	11	0	0.00
Research & Development	Laboratory Technician	259	62	23.94
Research & Development	Research Scientist	292	47	16.10
Research & Development	Manufacturing Director	145	10	6.90
Research & Development	Healthcare Representative	131	9	6.87
Research & Development	Manager	54	3	5.56
Research & Development	Research Director	80	2	2.50
Sales	Sales Representative	83	33	39.76
Sales	Sales Executive	326	57	17.48
Sales	Manager	37	2	5.41

12.ATTRITION VS. BUSINESS TRAVEL

The workers who **travel a lot** are more likely to quit than other employees (**25%**). This may be due to work-life balance challenges, increased stress, and the desire for more stability or local opportunities.

SELECT

BusinessTravel,

COUNT() AS Total_Employees,*

*SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,*

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY BusinessTravel

ORDER BY Attrition_Percentage DESC;

BusinessTravel	Total_Employees	Attrition_Count	Attrition_Percentage
Travel_Frequently	277	69	24.91
Travel_Rarely	1043	156	14.96
Non-Travel	150	12	8.00

13.ATTRITION VS. DISTANCE FROM HOME

Low attrition for moderate commute distances (10-15 km) may indicate a balance between convenience and job satisfaction.

Employees with **very short (under 5 km)** or **very long commutes (over 20 km)** experience higher attrition, potentially due to factors such as **commute fatigue, work-life balance issues, or lower job satisfaction.**

The results suggest that employees with **extremely short or long commutes** may face challenges in balancing work and personal life, which could influence their decision to leave.

SELECT

CASE

WHEN DistanceFromHome BETWEEN 1 AND 5 THEN '1-5'

WHEN DistanceFromHome BETWEEN 6 AND 10 THEN '6-10'

WHEN DistanceFromHome BETWEEN 11 AND 15 THEN '11-15'

WHEN DistanceFromHome BETWEEN 16 AND 20 THEN '16-20'

ELSE '>21'

END AS Distance_Category,

COUNT() AS Total_Employees,*


```

SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,

ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage

FROM

ibmemployee

GROUP BY

Distance_Category

ORDER BY

Attrition_Percentage DESC;

```

Distance_Category	Total_Employees	Attrition_Count	Attrition_Percentage
>21	204	45	22.06
11-15	115	25	21.74
16-20	125	23	18.40
6-10	394	57	14.47
1-5	632	87	13.77

14.ATTRITION VS. OVER TIME

Employees who work overtime are **2.5 times** more likely (**30%**) to leave compared to those who don't. This indicates that work-life balance plays a significant role in attrition rates.

SELECT

OverTime,

COUNT() AS Total_Employees,*

SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY OverTime

ORDER BY Attrition_Percentage DESC;

OverTime	Total_Employees	Attrition_Count	Attrition_Percentage
Yes	416	127	30.53
No	1054	110	10.44

15.ATTRITION VS. PROMOTION

Employees **who have not been promoted in the last 5 years** show significantly higher attrition rates. This is due to stagnation, lack of career growth, and frustration with limited advancement opportunities.

SELECT

CASE

WHEN YearsSinceLastPromotion BETWEEN 1 AND 5 THEN '1-5'

WHEN YearsSinceLastPromotion BETWEEN 6 AND 10 THEN '6-10'

WHEN YearsSinceLastPromotion BETWEEN 11 AND 15 THEN '11-15'

WHEN YearsSinceLastPromotion BETWEEN 16 AND 20 THEN '16-20'

ELSE '>20'

END AS yearsincelastpromotion_Category,

COUNT() AS Total_Employees,*

SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM

ibmemployee

GROUP BY

yearsincelastpromotion_Category

ORDER BY

Attrition_Percentage DESC;

<i>yearsincelastpromotion_Category</i>	<i>Total_Employees</i>	<i>Attrition_Count</i>	<i>Attrition_Percentage</i>
>20	581	110	18.93
6-10	149	27	18.12
1-5	674	92	13.65
11-15	66	8	12.12

16. ATTRITION VS. JOB INVOLVEMENT

Employees with **low job involvement** are more likely to leave (**33%**). Higher involvement might indicate stronger engagement with work, which leads to higher retention rates.

SELECT

JobInvolvement,

COUNT() AS Total_Employees,*

*SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,*

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM ibmemployee

GROUP BY JobInvolvement

ORDER BY Attrition_Percentage DESC;

JobInvolvement	Total_Employees	Attrition_Count	Attrition_Percentage
1	83	28	33.73
2	375	71	18.93
3	868	125	14.40
4	144	13	9.03

17. ATTRITION VS. NUMBER OF COMPANIES WORKED

Employees who have worked at **fewer companies** (15.23%) tend to stay longer, while those who have worked at **many companies** (21.48%) are more likely to leave. This may suggest that employees who have stayed in one place for a longer time are more likely to be loyal.

SELECT

CASE

WHEN NumCompaniesWorked BETWEEN 1 AND 4 THEN '1-4'

WHEN NumCompaniesWorked BETWEEN 5 AND 8 THEN '5-8'

ELSE '>9'

END AS NumCompanies_Category,

COUNT() AS Total_Employees,*

*SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,*

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM

ibmemployee

GROUP BY

NumCompanies_Category

ORDER BY

Attrition_Percentage DESC;

NumCompanies_Category	Total_Employees	Attrition_Count	Attrition_Percentage
5-8	256	55	21.48
1-4	965	147	15.23
>9	249	35	14.06

18.ATTRITION VS. TOTAL WORKING YEARS

Employees who have a total working years <5 have a high attrition rate (28%) than those with **total working years >10(9%)**. This is because they are often in the early stages of their careers, more likely to explore new opportunities, and may feel less committed to the organization.

SELECT

CASE

WHEN TotalWorkingYears BETWEEN 1 AND 5 THEN '1-5'

WHEN TotalWorkingYears BETWEEN 6 AND 10 THEN '6-10'

WHEN TotalWorkingYears BETWEEN 11 AND 15 THEN '11-15'

WHEN TotalWorkingYears BETWEEN 16 AND 20 THEN '16-20'

ELSE '>20'

END AS years_Category,

COUNT() AS Total_Employees,*

SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM

ibmemployee

GROUP BY

years_Category

ORDER BY

Attrition_Percentage DESC;

<i>years_Category</i>	<i>Total_Employees</i>	<i>Attrition_Count</i>	<i>Attrition_Percentage</i>
1-5	305	86	28.20
6-10	607	91	14.99
11-15	191	24	12.57
16-20	149	15	10.07
>20	218	21	9.63

19.ATTRITION VS. TRAINING HOURS

Employees who have attended <2 (18%) training sessions last year show high attrition rate than those who have attended >6(9%). This may be because they may feel undervalued, lack opportunities for skill development, and see limited career growth within the company.

SELECT

CASE

WHEN TrainingTimesLastYear BETWEEN 0 AND 2 THEN '<2'

WHEN TrainingTimesLastYear BETWEEN 3 AND 5 THEN '3-5'

WHEN TrainingTimesLastYear BETWEEN 6 AND 8 THEN '6-8'

ELSE '>9'

END AS trainingtime_Category,

COUNT() AS Total_Employees,*

*SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS
Attrition_Count,*

*ROUND(SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0
END) * 100.0 / COUNT(*), 2) AS Attrition_Percentage*

FROM

ibmemployee

GROUP BY

trainingtime_Category

ORDER BY

Attrition_Percentage DESC;

trainingtime_Category	Total_Employees	Attrition_Count	Attrition_Percentage
<2	672	122	18.15
3-5	733	109	14.87
6-8	65	6	9.23

20.ATTRITION VS. YEARS AT COMPANY

Employees who have been with the company for less than 2 years are likely to have higher attrition rates, which is common in many organizations as early-stage employees often leave for better opportunities or due to dissatisfaction.

SELECT

CASE

WHEN YearsAtCompany < 2 THEN 'Less than 2 years'

WHEN YearsAtCompany BETWEEN 2 AND 5 THEN '2-5 years'

WHEN YearsAtCompany BETWEEN 6 AND 10 THEN '6-10 years'

ELSE 'More than 10 years'

END AS TenureRange,

*AVG(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END)*100 AS
AttritionRate*

FROM ibmemployee

GROUP BY TenureRange;

TenureRange	AttritionRate
6-10 years	12.2768
Less than 2 years	34.8837
2-5 years	15.5080
More than 10 years	8.1301

21.ATTRITION VS. STOCK OPTION LEVEL

The highest attrition rate (24.41%) is seen in employees who don't receive stock options. However, Level 3 employees, despite receiving the highest stock options, have a relatively higher attrition rate (17.64%), indicating that factors beyond compensation might influence retention.

SELECT

StockOptionLevel,

COUNT() AS TotalEmployees,*

*AVG(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END)*100 AS
AttritionRate*

FROM ibmemployee

GROUP BY StockOptionLevel

ORDER BY StockOptionLevel;

StockOptionLevel	TotalEmployees	AttritionRate
3	85	17.6471
2	158	7.5949
1	596	9.3960
0	631	24.4057

22.ATTRITION VS. PERCENT SALARY HIKE

Employees with smaller salary hikes (less than 15%) tend to have a higher attrition rate, indicating that low salary increases might contribute to higher turnover. High salary hikes (more than 20%) lead to a surprisingly higher attrition rate.

SELECT

CASE

WHEN PercentSalaryHike < 15 THEN 'Less than 15'

WHEN PercentSalaryHike BETWEEN 16 AND 20 THEN '16-20'

ELSE 'More than 20'

END AS Percent_Salary_Hike,

COUNT() AS TotalEmployees,*

*AVG(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END)*100 AS
AttritionRate*

FROM ibmemployee

GROUP BY Percent_Salary_Hike

ORDER BY Percent_Salary_Hike;

Percent_Salary_Hike	TotalEmployees	AttritionRate
16-20	380	15.0000
Less than 15	818	16.1369
More than 20	272	17.6471

KEY FINDINGS AND INSIGHTS

1. Demographics:

- **Age:** Younger employees (under 30) and employees with less than 2 years of tenure have higher attrition rates.
- **Marital Status:** Single employees are more prone to leaving, while married employees exhibit greater stability.
- **Education Field:** Employees with education in the "Human Resources" field have slightly higher attrition rates compared to other fields.

2. Job-Related Factors:

- **Job Satisfaction:** Employees dissatisfied with their jobs are significantly more likely to leave.
- **Department:** Higher attrition rates were observed in the **Sales** and **R&D** departments compared to others.
- **Overtime:** Employees frequently working overtime are at a much higher risk of attrition.
- **Promotions:** Employees who **haven't received a promotion in 5+ years** exhibit a significantly higher attrition rate compared to those recently promoted.

3. Compensation and Benefits:

- **Income:** Lower-income employees (e.g., earning below the median income) are at a significantly higher risk of leaving.

4. Work-Life Balance:

- Employees with poor work-life balance (indicated by high overtime) and long commuting distances face higher attrition risks.
- Training time doesn't have a significant influence on attrition but maintaining development programs still supports retention.

SURPRISING TRENDS

Gender Neutrality:

- Gender was not a significant factor in attrition. Both males and females displayed similar attrition patterns, contradicting common assumptions.

Distance from Home:

- Employees living closer to their workplace (<5 km) were surprisingly at slightly higher risk of leaving compared to those with moderate commuting distances.

Job Role Differences:

- Employees in job roles such as Sales Executives and Laboratory Technicians had significantly higher attrition rates, even when controlling for compensation and satisfaction levels.

Overtime's Impact:

- A staggering 60% attrition rate was observed among employees working overtime, far higher than other factors.

RECOMMENDATIONS

Improve Job Satisfaction:

- Regular surveys to gauge employee satisfaction and take corrective actions based on feedback.
- Provide clear career growth opportunities and structured paths for promotions.

Work-Life Balance:

- Limit overtime hours to avoid employee burnout.
- Implement flexible work arrangements or hybrid working options for employees facing commuting difficulties.

Employee Retention Programs:

- Increase engagement and retention efforts, especially for younger and single employees, through tailored mentorship and career development programs.
- Enhance rewards such as performance-based bonuses, stock options, or recognition awards for high-performing employees.

Focus on High-Risk Groups:

- Prioritize efforts in high-attrition departments like **Sales** and **R&D**.
- Introduce additional training and development initiatives for roles such as **Sales Executives** and **Technicians** to boost morale and skills.

Promotion Framework:

- Develop clear policies for timely promotions and communicate them transparently to employees.
- Offer alternatives like lateral moves or increased responsibilities to employees who may not immediately qualify for promotion.

Fair Compensation:

- Conduct market comparisons to ensure competitive salaries, especially for lower-income groups.
- Provide annual raises to keep compensation aligned with employee expectations.

CONCLUSION

The analysis highlights that attrition in the organization is influenced by a combination of demographic, job-related, and compensation factors. Key contributors to attrition include poor job satisfaction, frequent overtime, lack of career advancement opportunities, and inadequate compensation. Surprisingly, gender and training time were found to have minimal impact on attrition.

By implementing the recommendations above, the organization can improve retention rates, reduce turnover costs, and foster a more satisfied and engaged workforce. Specifically, addressing issues related to promotions, overtime, and employee recognition can have a profound impact on improving employee loyalty and overall organizational performance.

REFERENCES

Thathvar, Pavan Subhash. "IBM HR Analytics Employee Attrition Dataset." *Kaggle*, <https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset>.