

IBM HR Analytics – Employee Attrition Project

Project under Unified Mentor Data Science Internship

1. Project Overview

The main goal of this project is to understand **why employees leave a company** by analyzing various HR factors.

The dataset — “*WA_Fn-UseC_-HR-Employee-Attrition.csv*” — contains information about employees such as their age, job role, salary, distance from home, satisfaction levels, and more.

The project uses data analysis and visualization to identify patterns that contribute to **employee attrition (resignation)**.

Objective:

- Find out which factors most influence attrition.
- Help HR teams take proactive steps to retain employees.

2. Data Cleaning and Preparation

The first step was to clean the data.

- There were **no missing or duplicate values**, which made the work easier.
- Some unnecessary columns like *EmployeeNumber*, *EmployeeCount*, and *Over18* were removed.
- New columns were created for better understanding, such as:
 - **Total yearly salary** ($\text{MonthlyIncome} \times 12$)
 - **Overtime flag** (Yes = 1, No = 0)
 - **Age groups** (<25, 25–35, 35–45, 45+)
 - **Distance category** (0–5 km, 6–10 km, etc.)
 - **Tenure band** (<2 yrs, 2–5 yrs, etc.)

Insight:

After cleaning and adding new features, the data became much clearer and ready for analysis.

3. Exploring the Data (EDA)

Here's what the analysis showed:

- **Age:** Younger employees (under 35) are more likely to leave. They often look for better growth or new experiences.
- **Overtime:** Employees who work overtime often feel stressed and are more likely to resign.

- **Distance from Home:** People living far from the office tend to leave more often.
- **Salary:** Employees with lower pay are more likely to quit.
- **Job Role:** Some roles like Sales or Lab Technicians had higher turnover rates.
- **Satisfaction:** Low job or environment satisfaction directly increases attrition.

Insight:

The main reasons people leave are **stress (too much overtime), low pay, long distance, and lack of satisfaction.**

4. Key Factors Behind Attrition

When checking correlations, the factors most linked to employee attrition were:

1. **Overtime**
2. **Monthly income**
3. **Job satisfaction**
4. **Years at the company**
5. **Distance from home**

Insight:

Workload, pay, and job satisfaction are the biggest reasons employees quit.

5. Business Insights and Recommendations

Here's what the company can do to reduce employee turnover:

Factor	Impact on Attrition	Suggested Action
Overtime	High overtime increases attrition	Implement workload balance policies
Salary	Low income leads to attrition	Review compensation structure
Age	Younger employees leave more often	Offer career development and mentorship
Distance	Long commute increases attrition	Allow flexible or remote work options
Satisfaction	Low satisfaction drives attrition	Conduct regular feedback and engagement programs

Insight:

Improving work-life balance, fair pay, and employee satisfaction can greatly reduce attrition.

6. Conclusion

The IBM HR Analytics project clearly shows that **attrition is not random** — it's caused by specific factors like **overtime, pay, satisfaction, and distance**.

By understanding these issues, companies can take better decisions to **retain employees, improve happiness, and build a stronger workplace**.