IBM HR Analytics Employee Attrition & Performance

IBM HR Analytics Employee Attrition & Performance

Context

The company's success is heavily dependent on the performance and engagement of its employees. Therefore, it is critical for every company to understand the factors that lead to employee attrition and develop effective retention strategies. These are some of the several disadvantages of attrition for businesses:
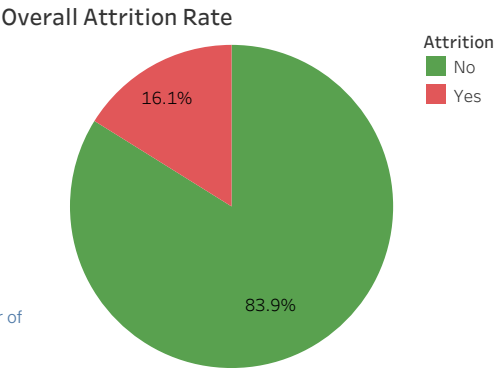
1. **Cost**: Hiring and training new employees can be expensive.
2. **Loss of productivity**: New employees need time to adjust to their role and learn about the company's culture, processes, and expectations.
3. **Reduced morale**: High turnover rates can negatively affect employee morale and engagement.
4. **Knowledge loss**: When experienced employees leave, they take their knowledge, skills, and institutional memory with them.

Problem Statement

1. What is IBM's employee attrition rate in this dataset?
2. Which business department has an attrition issue?
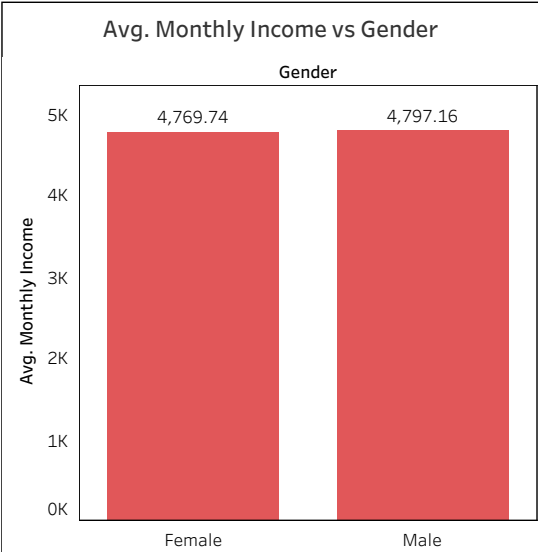3. What are the factors that contribute to employee attrition?

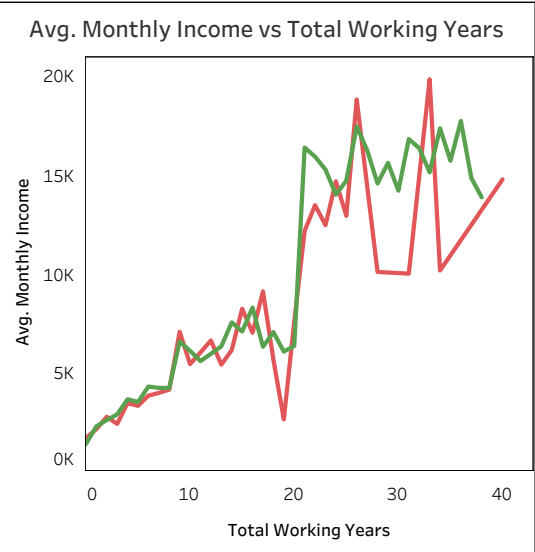The overall attrition rate at IBM is 16.1%.

Overall Attrition Rate

Attrition
- 🟩 No
- 🟥 Yes

16.1%

83.9%

Attrition rate is calculated as the count of employees who left within a specific category divided by the total number of employees in that category.

IBM HR Analytics Employee Attrition & Performance

Avg. Monthly Income vs Total Working Years

Avg. Monthly Income vs Gender

Gender

Female: 4,769.74   Male: 4,797.16

Attrition
- No (green)
- Yes (red)

Attrition Rate by Percent Salary Hike

- 19.52%
- 16.67%
- 16.27%
- 11.94%
- 17.82%
- 17.95%
- 17.07%
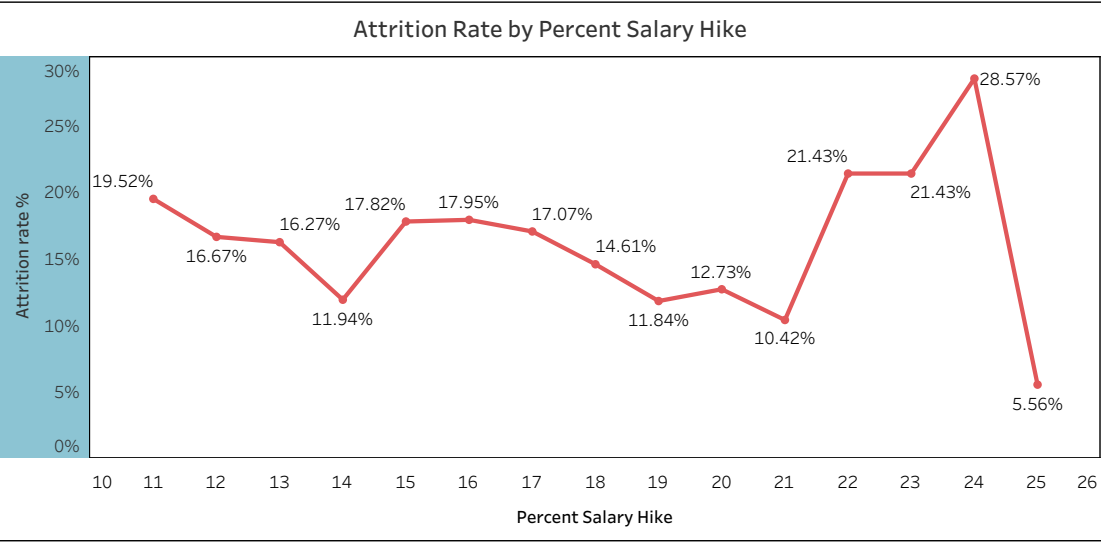- 14.61%
- 11.84%
- 12.73%
- 10.42%
- 21.43%
- 21.43%
- 28.57%
- 5.56%

· **Average monthly income** tends to **increase** with total working years, **with a higher rate for employees without attrition** compared to those with attrition.
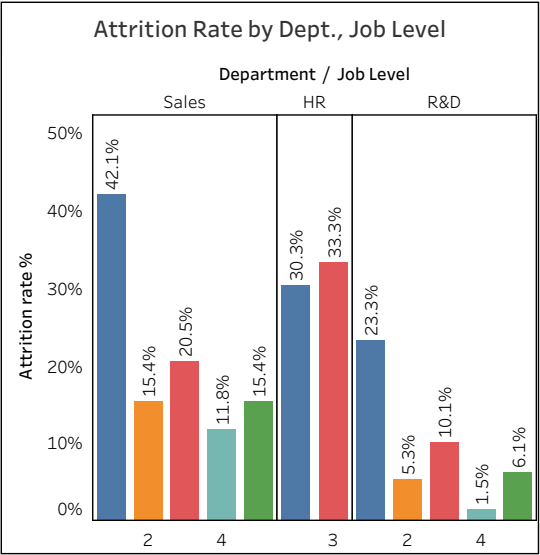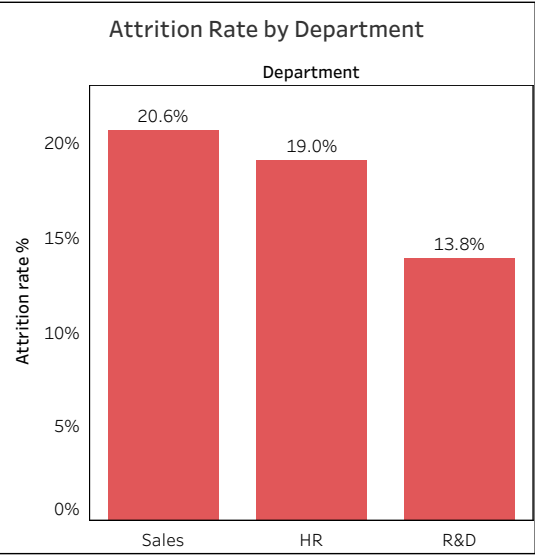
· There is **no gender-based salary difference** among attrition employees, as both male and female employees have similar average monthly income.

· Generally, as the percent salary hike increases, the percentage of employees who left the company decreases with few exceptions percent salary hikes (such as 22%, 23%, and 24%).

IBM HR Analytics Employee Attrition & Performance

Attrition Rate by Department

Department

- Sales: 20.6%
- HR: 19.0%
- R&D: 13.8%

Attrition Rate by Dept., Job Level

Department / Job Level

Job Level
- 1
- 2
- 3
- 4
- 5

Attrition
- Yes

Sales:
- 42.1%
- 15.4%
- 20.5%
- 11.8%
- 15.4%

HR:
- 30.3%
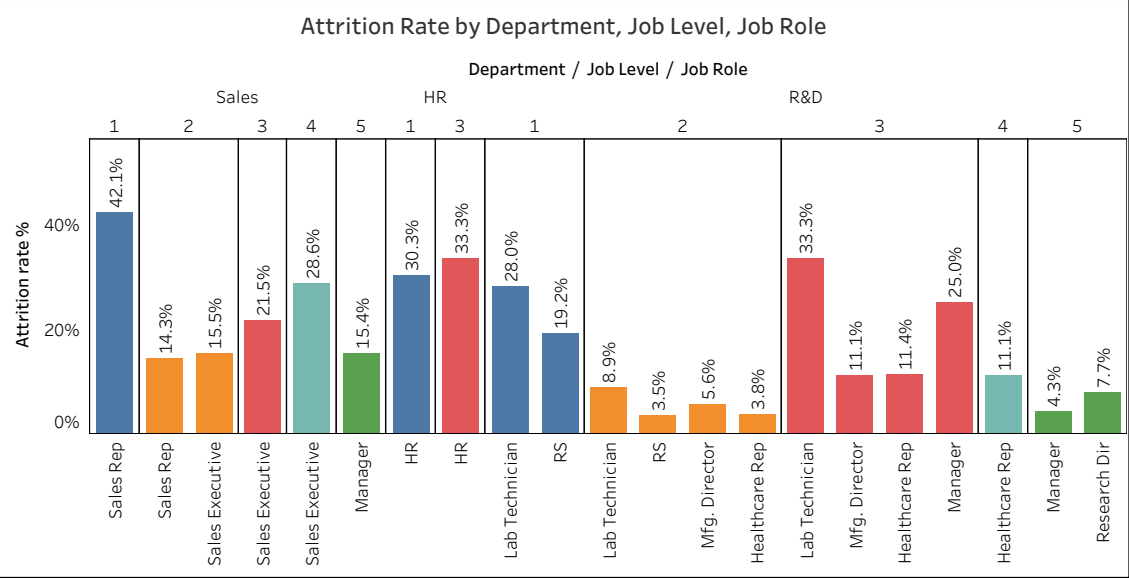- 33.3%

R&D:
- 23.3%
- 5.3%
- 10.1%
- 1.5%
- 6.1%

Note: The numerical scale ranges from 1, indicating the most junior positions, to 5, representing the most senior positions.

· **Sales** department has **highest attrition rate** while **Research & Development** has **lowest attrition rate**.
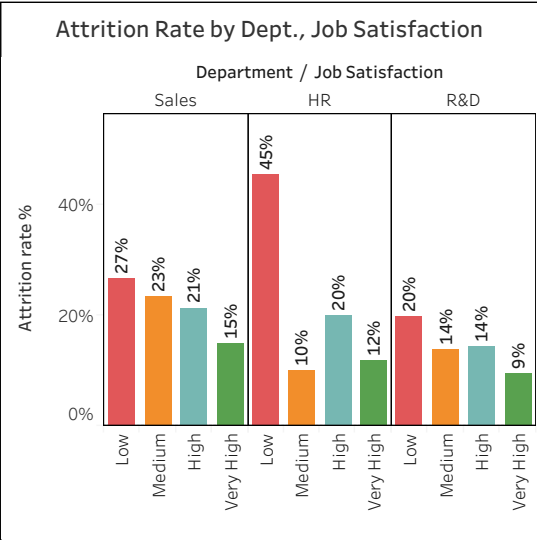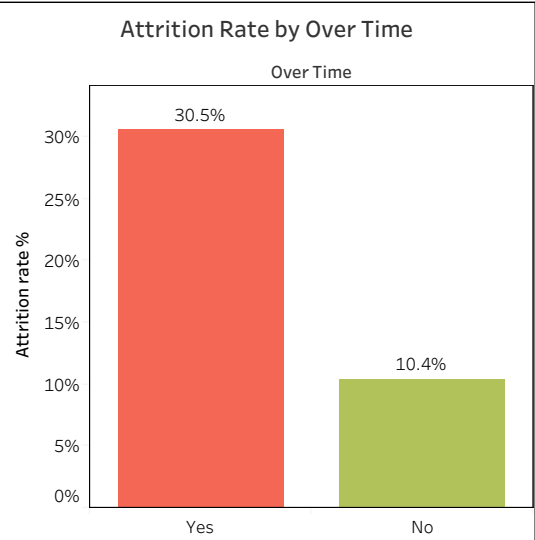
· **Junior most position** has the **highest attrition rate** in Sales and Research & Development departments.

· Junior positions such as **Sales Representative, Lab Technician, and Research Scientist** contribute to high attrition rate of company.

Attrition Rate by Department, Job Level, Job Role

Department / Job Level / Job Role

Sales
- 1 Sales Rep: 42.1%
- 2 Sales Rep: 14.3%
- 3 Sales Executive: 15.5%
- 3 Sales Executive: 21.5%
- 4 Sales Executive: 28.6%
- 5 Manager: 15.4%

HR
- 1 HR: 30.3%
- 3 HR: 33.3%

R&D
- 1 Lab Technician: 28.0%
- 1 RS: 19.2%
- 2 Lab Technician: 8.9%
- 2 RS: 3.5%
- 2 Mfg. Director: 5.6%
- 2 Healthcare Rep: 3.8%
- 3 Lab Technician: 33.3%
- 3 Mfg. Director: 11.1%
- 3 Healthcare Rep: 11.4%
- 3 Manager: 25.0%
- 4 Healthcare Rep: 11.1%
- 4 Manager: 4.3%
- 5 Research Dir: 7.7%

IBM HR Analytics Employee Attrition & Performance

Attrition Rate by Over Time

Over Time

- 30.5% (Yes)
- 10.4% (No)

Attrition Rate by Dept., Job Satisfaction

Department / Job Satisfaction

Sales				HR				R&D			
Low	Medium	High	Very High	Low	Medium	High	Very High	Low	Medium	High	Very High
27%	23%	21%	15%	45%	10%	20%	12%	20%	14%	14%	9%

Over Time
- Yes
- No

Job Involvement
Job Satisfaction
Performance Rating
- Low
- Medium
- High
- Very High

Attrition Rate by Dept., Job Involvement

Department / Job Involvement

Sales				HR				R&D			
Low	Medium	High	Very High	Low	Medium	High	Very High	Low	Medium	High	Very High
51.7%	23.6%	16.2%	20.0%	33.3%	29.4%	13.9%	14.3%	23.5%	16.1%	13.6%	4.9%

Attrition Rate by Performance Rating

Department / Performance Rating

HR		R&D		Sales	
High	Very High	High	Very High	High	Very High
20.4%	11.1%	13.3%	16.7%	21.3%	16.4%

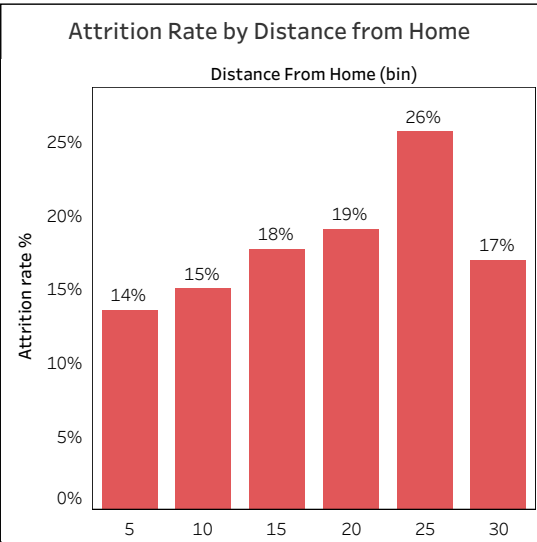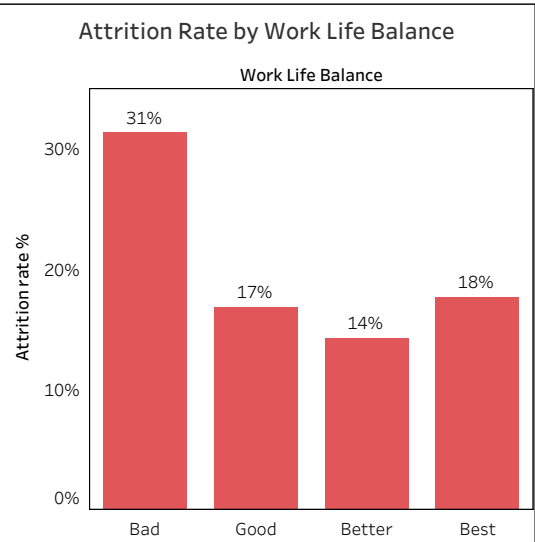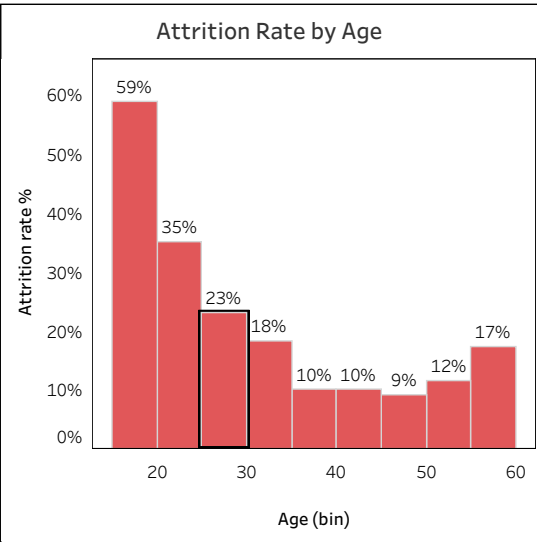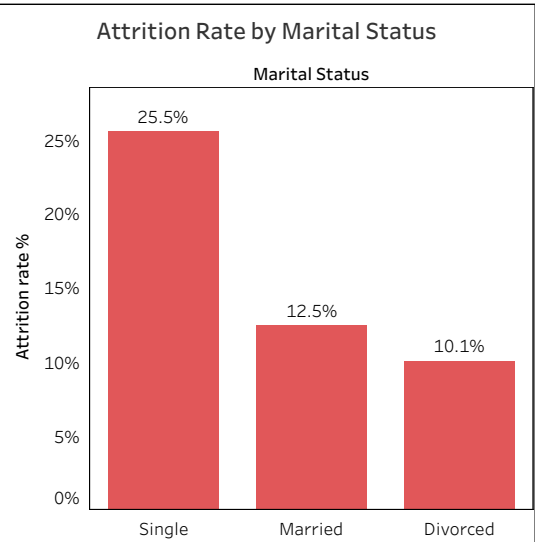· **Overtime workers** have a **3 times higher attrition rate** than non-overtime workers.
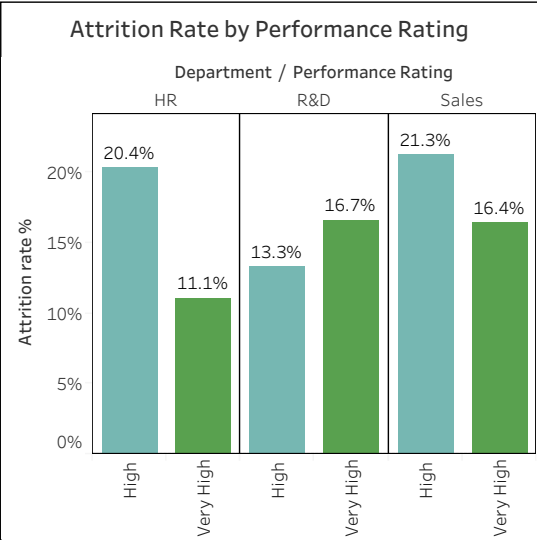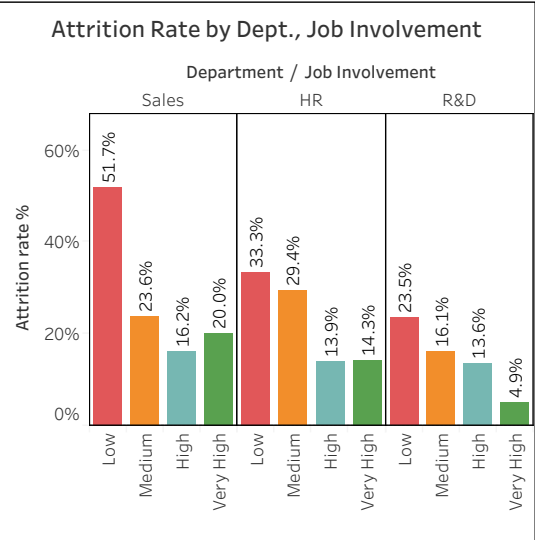
· **Low job satisfaction and low employee involvement** is associated with **higher attrition** rates within departments.

· **Performance rating does not impact attrition rate** as attrition employees had High and Very High rating.

IBM HR Analytics Employee Attrition & Performance

Attrition Rate by Marital Status

Marital Status

- Single: 25.5%
- Married: 12.5%
- Divorced: 10.1%

Attrition Rate by Age

Age (bin)

- 20: 59%, 35%
- 30: 23%, 18%
- 40: 10%, 10%
- 50: 9%, 12%
- 60: 17%

Attrition ■ Yes

· **Single employees** exhibit a **higher attrition rate** compared to married and divorced employees.

· Attrition is **highest** among employees **under 20 years old**, decreases with age until 50, and slightly rises for employees in their 50s.

· As an employee's **work-life balance improves**, the **attrition rate decreases**.

· The **attrition rate** exhibits a **slight increase with distance from home** up to 25 units, followed by a subsequent drop.

Attrition Rate by Work Life Balance

Work Life Balance

- Bad: 31%
- Good: 17%
- Better: 14%
- Best: 18%

Attrition Rate by Distance from Home

Distance From Home (bin)

- 5: 14%
- 10: 15%
- 15: 18%
- 20: 19%
- 25: 26%
- 30: 17%

IBM HR Analytics Employee Attrition & Performance

Model Results

Algorithm	Area Under the Precision-Recall Curve (PR AUC)	Macro Avg Recall	Accuracy
Logistic Regression	0.62	0.69	87.07%
Random Forest	0.53	0.57	85.49%
Random Forest with hyperparameter tuning	0.57	0.56	85.71%
XGBoost	0.56	0.63	85.49%
XGBoost with hyperparameter tuning	0.62	0.63	87.30%

The XGBoost model far outperformed the other approaches on the test and validation sets.

I tried above five models and evaluated them using **Precision-Recall Area Under the Curve (PR AUC) and Accuracy**.

I opted for the PR AUC instead ROC AUC because for imbalanced datasets PR AUC provides a more accurate representation of the classifier's performance, particularly in situations where correctly identifying the positive instances is more critical than minimizing false positives.

The other important metric used was **Macro Avg recall.**
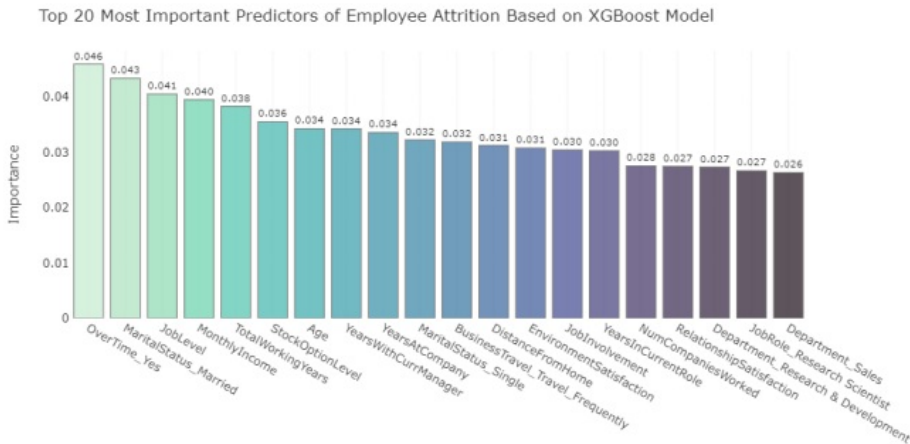
Top 20 Most Important Predictors of Employee Attrition Based on XGBoost Model

IBM HR Analytics Employee Attrition & Performance

Recommended Business Plan

Based on the features importance derived from the model, here is a suggested employee attrition plan:

1. **Manage Overtime**: Implement strategies to reduce or manage overtime for employees. This can include workload distribution, better resource planning, and promoting work-life balance to reduce the likelihood of attrition among employees who work overtime.

2. **Support Marital Relationships**: Provide support and resources to married employees to help them maintain a healthy work-life balance. This can include flexible work arrangements, employee assistance programs, and communication channels to address any concerns or challenges they may face.

3. **Career Development and Promotions:** Create opportunities for career growth and development within the organization. Offer training programs, mentoring, and clear paths for advancement to encourage employees to stay and progress within the company. Focus on providing a clear career trajectory and recognizing and rewarding employees for their achievements.

4. **Competitive Compensation:** Ensure that the company offers competitive compensation packages, including higher monthly income or wages for employees. Regularly review and benchmark salaries to ensure they are aligned with industry standards and provide financial incentives to retain valuable employees.

5. **Employee Retention Programs**: Develop employee retention programs that specifically target employees who have reached a certain threshold of total working years (e.g., after year 3). These programs can include additional benefits, recognition, and opportunities for growth to motivate and retain experienced employees.