

Acme Aroma Employee Turnover Analysis & Recommendation

Ryan McCurry

December 8th, 2025

Table of Contents

1. Problem Statement & Goal (page 3)

2. Approach to the Analysis (page 4)

3. Insights (pages 5-7)

4. Recommendation (page 8)

5. Project Limitations (page 9)

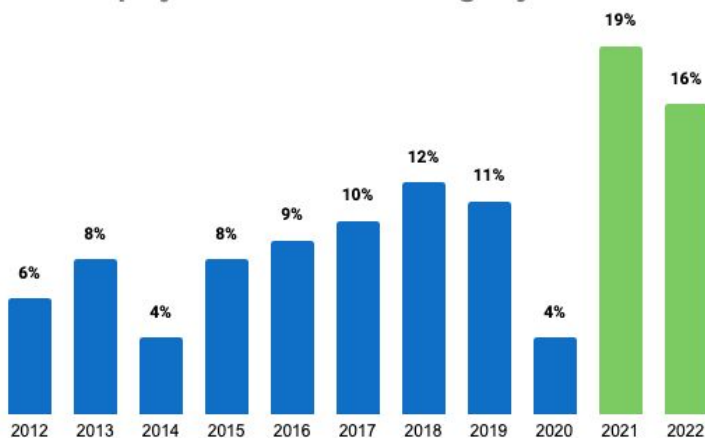
Problem Statement & Goal

Employee Retention Challenge

Acme Aroma is facing a significant employee retention challenge. Attrition rates have climbed from 6% in 2012 to a **peak of 19% in 2021**, representing **more than a 3x increase**. Excluding the outlier year of 2020, rates have consistently risen throughout the decade.

The situation escalated dramatically with a **sharp spike from 11% in 2019 to 19% in 2021**, before slightly declining to 16% in 2022. Despite this recent decline, the 2022 rate remains considerably higher compared to past performance.

Employee Turnover Percentage by Year



Financial Impact of Employee Turnover

These elevated turnover rates have created substantial financial costs for Acme Aroma. In 2021, 721 employees¹ departed at a cost of **₹15,000 per replacement**. In 2022, 634 departures cost **₹30,000 per new hire, doubling the per-employee expense, year-over-year**.

Employee Acquisition Cost

Per Employee vs. Total Cost



The 2021 and 2022 departures generated nearly **₹29.8 million in recruitment expenses** (₹10.8 million and ₹19 million respectively). This marks the **first time this decade that the company has incurred any acquisition costs**, representing an unprecedented financial burden.

Goals of the Project

This project aims to address Acme's employee attrition issue through a few interconnected goals. First, a **root association analysis** will be conducted to enable Acme to understand exactly what is happening within the company. This analysis will identify the **specific factors and conditions driving higher attrition rates** at Acme Aroma.

Leveraging these insights, a **predictive employee turnover model** will be developed to proactively **identify at-risk employees**, allowing Acme to intervene before attrition occurs. Based on the model's findings, an **actionable recommendation** will be provided to strategically reduce employee turnover going forward.

¹Employee Departure Calculation:

$$\text{Turnover Rate} \times \text{Workforce} = \text{Departures}$$

$$\text{2021: } 19\% \times 3,794 \text{ employees} = 721$$

$$\text{2022: } 16\% \times 3,965 \text{ employees} = 634$$

Approach to the Analysis

Data Source & Structure

The analysis draws from Acme Aroma's internal HRIS, **Pegasus**. The dataset's **response variable is attrition**, indicating whether an employee leaves the company. Predictor variables span three categories: personal characteristics (e.g. gender, age, education level), professional history (e.g. income, years at company, years with current manager), and recent employee engagement scores captured through internal surveys.

Initial Data Exploration

Initial exploration shows several variables with notable correlations with attrition. The **strongest relationships are work life balance (-0.5), percent salary hike (-0.4), and job satisfaction (-0.3)**, meaning lower values correspond to higher attrition. Conversely, distance from home (0.2) and marital status (0.2) show positive correlations. These results suggest that **employee engagement factors, particularly work-life balance, compensation, and job satisfaction, are likely primary indicators** for employee turnover at Acme Aroma.

Modeling Approach & Key Metrics

The analysis uses **logistic regression** as the primary modeling technique. Logistic regression is well suited because it meets both predictive and inferential needs: it accurately identifies at-risk employees while providing interpretable coefficients that reveal which factors drive attrition, enabling Acme to both predict and address turnover.

The model's performance is measured by how well it handles different types of mistakes. **Type 1 errors (false positives)** happen when the model predicts an employee will leave but they actually stay, potentially leading to unnecessary retention efforts. **Type 2 errors (false negatives)** occur when the model predicts an employee will stay but they actually leave, resulting in unexpected turnover and lost talent. For Acme, **Type 2 errors are particularly costly** because they allow at-risk employees to leave the company without warning, continuing the attrition problem. The **F1 score** provides a single measure that balances both concerns: making accurate predictions while catching as many actual departures as possible.

Evaluation Process

The evaluation process uses multiple metrics to assess how well the model performs. While overall accuracy measures correct predictions, it can be misleading when most employees stay within the company. Therefore, the evaluation focuses on the **trade-off between Type 1 and Type 2 errors**, understanding that improving one type of error often worsens the other.

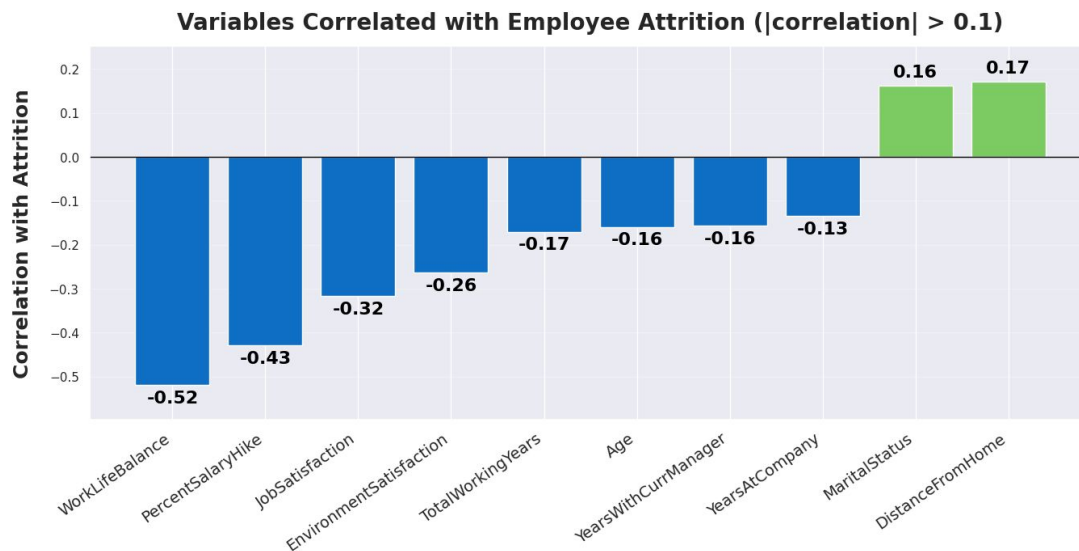
Given Acme's business priorities, the evaluation emphasizes **catching employees who are actually at risk of leaving**. This means the model may occasionally flag employees who weren't actually planning to leave, but this is preferable to missing those who actually depart. The F1 score helps ensure the balance between these main concerns. The model will be adjusted based on how much Acme is willing to spend on preventive retention efforts versus the higher cost of losing employees unexpectedly. Performance will be tested on separate employee data to ensure that the predictions for future employees remain reliable.

Exploratory Insights

Understanding Employee Turnover Risk

The exploratory analysis identifies which factors relate to employee attrition to better understand what drives turnover and predict which employees are at risk of leaving. Initial exploration shows several factors with notable connections to attrition. The **strongest relationships are work life balance (-0.5), percent salary hike (-0.4), and job satisfaction (-0.3)**, meaning lower values in these areas correspond to higher turnover.

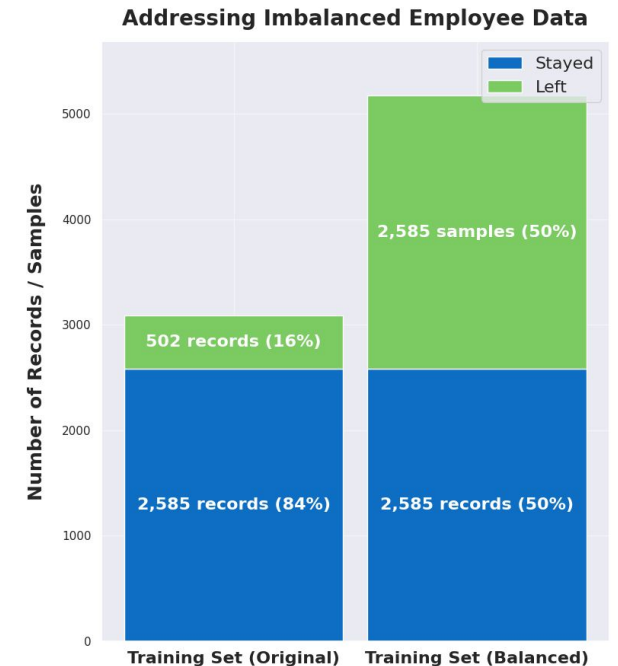
The results suggest that **employee engagement factors, particularly work-life balance, compensation, and job satisfaction, are likely the main warning signs** for employee turnover at Acme Aroma. The data shows that how employees feel about their jobs and work environment matters more than demographic factors alone.



Addressing Data Imbalance

A significant challenge in the dataset is imbalance: **only 16% of employees left the company, while 84% remained**. This imbalance can cause predictive models to simply predict that all employees will stay (looking accurate while missing employees actually at risk of leaving).

To address this, the analysis **creates artificial records of employees who left** to balance the data used for training, ensuring the model learns patterns from both groups equally.



Selecting Variables for the Model

Based on the strength of their relationships and to avoid using too many similar measures, five key variables were selected for the predictive model:

Work-Life Balance (-0.5), Percent Salary Hike (-0.4), Job Satisfaction (-0.3), Age (-0.2), and Distance from Home (+0.2).

These variables provide the strongest and most varied signals of turnover risk, combining pay factors, employee engagement, personal circumstances, and basic demographics.

Model Performance

The Final Model

The final logistic regression model uses **five key predictors** to identify employees at risk of leaving Acme Aroma. To ensure fair comparisons across different scales (such as age in years vs. satisfaction ratings on a 4-point scale), all variables were **standardized**, meaning they were adjusted to a common scale before being used in the model. This allows us to directly compare the relative importance of each factor.

All five variables are **statistically significant: Work-Life Balance (strongest predictor), Percent Salary Hike, Job Satisfaction, Age, and Distance from Home**. This means that each contributes to the overall understanding of employee turnover.

The model shows that **employee engagement factors (work-life balance, compensation, and job satisfaction) have the strongest impact on turnover**. Distance from home is the only factor that increases attrition, while all other factors protect against it.

Model Performance Results

The model performs well at identifying which employees are likely to leave. Most importantly, the model **catches 91% of employees who actually leave** (recall), successfully locating the vast majority of at-risk employees before they depart. When the model flags an employee as at-risk, it's correct about **59% of the time** (precision), meaning roughly 6 out of every 10 predictions are accurate.

Overall, the **model correctly classifies 88% of all employees** in the validation dataset. These results demonstrate that the model can reliably predict employee turnover and provide actionable insights for Acme's retention efforts.

Understanding Errors in the Model

Like any predictive model, this model makes two types of mistakes. **Type 1 errors (false alarms)** occur when the model predicts an employee will leave, but they actually stay. This means that Acme might invest retention resources on employees who weren't actually planning to leave anyway. The **model produces Type 1 errors at a rate of 12.1%**, meaning roughly 1 in 8 employees flagged for retention weren't actually at risk.

Type 2 errors (missed departures) occur when the model predicts an employee will stay, but they actually leave. This is far more costly for Acme because it means losing employees without warning. The **model produces Type 2 errors at only 9.4%**, meaning it successfully catches over 90% of actual departures.

Given Acme's business priorities, the model strikes an excellent balance. It minimizes the more costly Type 2 errors while keeping Type 1 errors at a manageable level. While some retention investments will target employees who weren't planning to leave, this is an acceptable trade-off to ensure Acme doesn't miss the costly departures that drive the attrition problem.

Model Prediction Results (Validation Set)

Actually Stayed	487 Correct (Stayed)	67 Type 1 Error (False Alarm)
Actually Left	10 Type 2 Error (Missed Departure)	97 Correct (Caught Departure)
	Predicted to Stay	Predicted to Leave

Root Association Analysis

Key Factors Driving Employee Turnover

The analysis reveals **five variables with strong associations to employee attrition**. The strongest predictors of turnover are employee engagement factors: **Work-Life Balance, Percent Salary Hike, and Job Satisfaction** show the most powerful relationships with whether or not employees will leave the company. These three factors consistently emerge as critical predictors of turnover risk.

Age and **Distance from Home** also significantly influence turnover decisions, though to a lesser degree. Notably, **distance from home is the only factor that increases attrition risk**. All other factors protect against turnover when they improve.

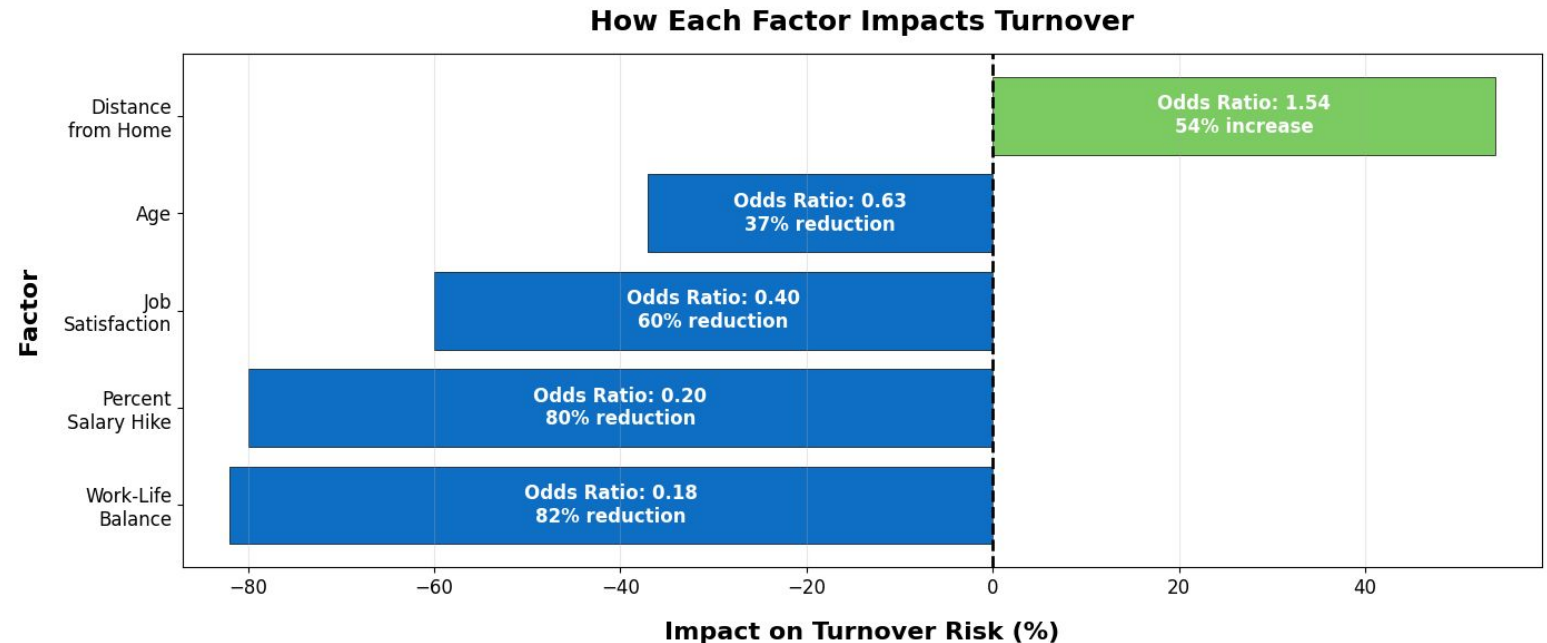
Understanding the Magnitude of Impact

To quantify how much each factor actually matters, **odds ratios** are used which show how much the likelihood of leaving changes when a factor increases. An **odds ratio below 1 means that the factor decreases turnover risk, while an odds ratio above 1 means it increases risk**. The further the distance from 1, the greater the effect on attrition.

Factor Specific Impacts

The three employment factors demonstrate the most powerful effects on employee turnover risk. **Work-life Balance (Odds Ratio: 0.18)** reduces turnover odds by 82%, making it the most protective factor. **Percent Salary Hike (Odds Ratio: 0.20)** follows closely with an 80% reduction, and **Job Satisfaction (Odds Ratio: 0.40)** provides a 60% reduction in the likelihood of an employee staying at Acme Aroma.

Age (Odds Ratio: 0.63) offers moderate protection with a 37% reduction in turnover risk. **Distance from home (Odds Ratio: 1.54)** is the only factor that increases attrition, raising the likelihood of staying at the company by 54%.



Recommendation

Model Implementation: Strong Yes

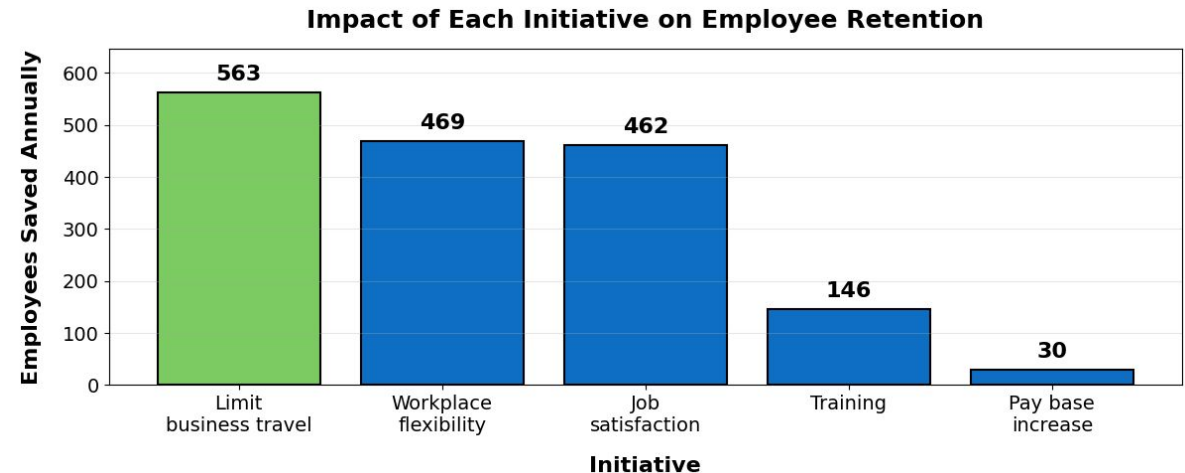
Acme Aroma should implement the employee turnover prediction model. The model demonstrates strong performance with **91% recall rate**, meaning it **successfully identifies 9 out of every 10 employees at risk of leaving the company**.

The model's Type 2 error rate of only 9.4% ensures that the model won't miss critical employee departures. While the model may occasionally flag some employees to leave when they aren't actually at risk (Type 1 errors at 12.1%), this is an acceptable trade-off due to the high cost of unexpected turnover and subsequently onboarding new employees.

The model provides a proactive tool to identify at-risk employees before leaving the company, enabling targeted retention strategies.

Recommended Initiative: *Limit Business Travel*

Based on the analysis of five possible initiatives, **limiting business travel** offers the greatest impact on reducing employee turnover. This initiative increases work-life balance by 0.3 points and would **reduce Acme's attrition probability from 49.6% to 35.4%**, a reduction of 14.2 percentage points. This translates to **saving approximately 563 employees from leaving the company annually**.



Economic Impact

If Acme implements the recommended initiative to limit business travel, the company stands to save approximately **₹318 million annually** in recruitment, training, and onboarding costs. This calculation assumes the industry standard estimate that replacing an employee costs approximately 75% of a full-time annual salary. By comparison:

- Allowing work-from-home would save ₹265 million
- Various employee appreciation initiatives would save ₹261 million
- Providing additional professional development would save ₹82 million
- Increasing base pay for all employees would save ₹17 million

Limiting business travel not only has the strongest impact on retention but also directly addresses the most powerful driver identified in the model: Work-Life Balance. By reducing travel requirements, Acme directly improves the factor with the greatest influence on whether employees will stay or leave at the company.

Project Limitations

Data Limitations

The dataset contains minimal missing data across key variables. Out of 4,410 employee records, there are 19 missing values for the number of companies an employee has worked for, 25 for environment satisfaction, 20 for job satisfaction, and 22 for work-life balance. These **missing values represent less than 1% of observations** and were addressed by replacing empty values with the average of each column. While this approach is standard practice, it slightly reduces variance in these variables.

A more significant challenge is the **imbalance in the data: only 16% of the dataset consists of employees who left the company**. This can cause the model to predict everyone will stay while missing employees who are actually at risk of leaving. To address this, the analysis **creates additional examples of employees who left the company** to balance the training data. Additionally, text-based variables (e.g. gender, department, job role) were converted to numbers, and all variables were adjusted to the same scale to ensure fair comparisons.

Modeling Limitations

The analysis combines two different tools that each serve specific purposes. One tool provides detailed statistical information about which factors matter most and by how much, while the other tool excels at measuring how well the model actually performs at making predictions. Using both together gives a more complete picture than either could provide alone.

A key limitation is that the model **relies on historical patterns**. It learns from how past employees behaved and assumes that these same patterns will continue. If Acme's environment changes significantly (such as gaining new leadership or enacting major policy changes), the model's predictions may become less reliable. There is also a **risk that the model becomes too tailored to past employees** and doesn't work as well for predicting future cases. This is addressed by testing performance on a separate set of employee data and tracking multiple measures of success rather than relying on a single metric.

The model **cannot predict unprecedented events or major disruptions** (such as recessions, industry-wide changes, or competitive shocks) that could fundamentally shift employee behavior, as it only identifies patterns from existing data.

Impact on Recommendations

Despite these limitations, the **constraints are manageable and do not change the core recommendations**. The data quality issues are minor and have been addressed through standard techniques. The data imbalance is directly handled by creating artificial records of employees who left the company. The modeling limitations around historical patterns and unprecedented events are inherent to any predictive approach and don't diminish the value of understanding current attrition drivers.

The recommendations focus on addressing the root causes of employee turnover which remains a valid strategy even accounting for model limitations. As with any analytical project, results should be monitored over time as new data becomes available and business conditions change.