# Addressing High Turnover

## Recommendations

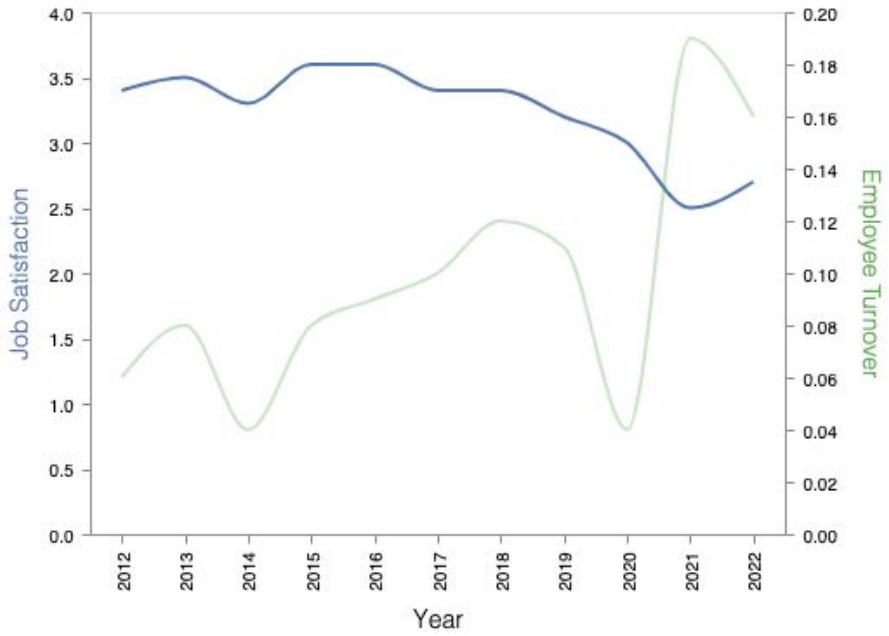
Sarah Amiraslani
December 13th, 2022

Acme Aroma

# Introduction

## Table of Contents

# Problem Statement
# Addressing High Employee Turnover




## Problem

Acme Aroma is facing a **high attrition rate of 16%**, an increase of 12% since 2020. Compounding the issue, there is **less interest from job seekers**. In 2022 candidates submitted an average of 8 job applications per opening, nearly 17 applications less than the running average.

Because our ideal employees have experience with our internal operations, maintaining our talent is a priority. Moreover, training and **recruiting new candidates is costly**, costing an average of ₹30,000 per position.

## Opportunity

Understanding why individuals leave and predicting who is likely to allow us to strategically invest our efforts for the greatest payoff.

Possible solutions are targeted interventions for at-risk staff or workforce planning to minimize the inevitable costs.

| **16%** | **24%** | **2.39 out of 4** | **15** | **₹21,330,000** |
|---|---|---|---|---|
| attrition in 2022, down by over 6% from the average to date | of employees who left in 2022 held mid to senior level roles | average job satisfaction of employees who left in 2022, 0.83 less than employees who stayed | fewer applications for new positions in 2022 compared to the average to date | is required to replace all the employees lost in 2022, assuming the average acquisition cost |

# Approach
# Root Association and Predictive Model

## Methods

Using the records of **4,400 employees** sourced from the human resources information system, we take two approaches to address turnover.

1. **Root Association Analyses** to understand the driving factors of attrition and identify a single action to address the challenge.

2. **Predictive modeling** to identify potentially at-risk employees using logistic regression so that management can intervene and support those individuals.

While the root analyses will not be an ongoing development, a good predictive model can run against changes to the human resource system to monitor risk continuously.

## Evaluation and Trade Offs

We can optimize our model against two types of errors:

- **Type I Error** occurs when we falsely classify an employee as at risk of attrition.
- **Type II Error** occurs when we fail to classify an at-risk employee as at-risk of attrition.

Because recruiting and training are more costly than remediation, we can afford to falsely reach out to employees who are not at risk of leaving. Given this bias, we will vet models against the following evaluation metrics:

- **Precision** is the proportion of employees labeled as at-risk who are at-risk.
- **Sensitivity** is the proportion of at-risk employees identified as at-risk.

The 44 predictor variables can be described as falling under **four dimensions of work life** described below.

### Quality of Life

- Work life balance
- Distance from home
- Business travel
- Job and environment satisfaction

### Individual Differences

- Personal attributes (e.g., age and gender)
- Education
- Career path and length
- Job role and level

### Performance

- Performance rating
- Years since promotion
- Job involvement

### Financial

- Income
- salary hike
- Stock option levels

**Outcome Variable: Attrition (Yes/No)**
Did the individual leave the company?
Attrition occurs in 16% of outcomes

# Results
# Root Association Analyses

Our analyses reveal that the following characteristics of Acme employees' jobs are the highest contributors to attrition.

## #1: Job Seniority

Senior employees with a work level of four are 3.5 times more likely to leave than their junior and executive counterparts.

Employees are 1.99 times more likely to leave for roughly every six years they work at Acme beyond the sample mean of seven years.

Employees are 1.60 times more likely to leave for roughly every three other companies they have worked at beyond the sample mean of three.

While we cannot design HR interventions to combat seniority, the knowledge that seniority is a risk factor for attrition paints a concerning picture: employees most likely to leave are those we would like to keep the most.

## #2: Work Related Travel

Employees who travel frequently are 2.02 times more likely to leave than their counterparts.

Employees are 2.09 times more likely to leave for roughly every seven additional kilometers beyond the sample mean of eight kilometers they live from the workplace.

This insight offers an actionable work effort for HR. By providing work-from-home options to individuals who live beyond eight kilometers from the office, we can significantly improve the employee experience.

## #3: Performance

Employees are 1.3 times more likely to leave for roughly every three passing years beyond the sample mean of two years during which they are not promoted.

This insight offers an actionable insight, suggesting that HR should consider offering regular promotions to employees who meet their duties. Promotion reviews should occur at least every two years.

# Results
## Final Model

The final model contained 41 variables, falling under the four dimensions of work-life described previously: quality of life, individual differences, performance, and financial incentives.

The choice of variables to include in the model was optimized to maximize the number of at risk-individuals we identify (i.e., precision) while also minimizing model complexity.

We removed highly correlated variables and variables encoding the same information from the model. Continuous encodings are chosen when given the option between a continuous or categorical encoding of the same data.

All continuous variables were standardized to avoid adding significance to variables with higher variances.

l2 penalty was used to avoid overfitting. We did not use stronger penalties as the number of features was small.

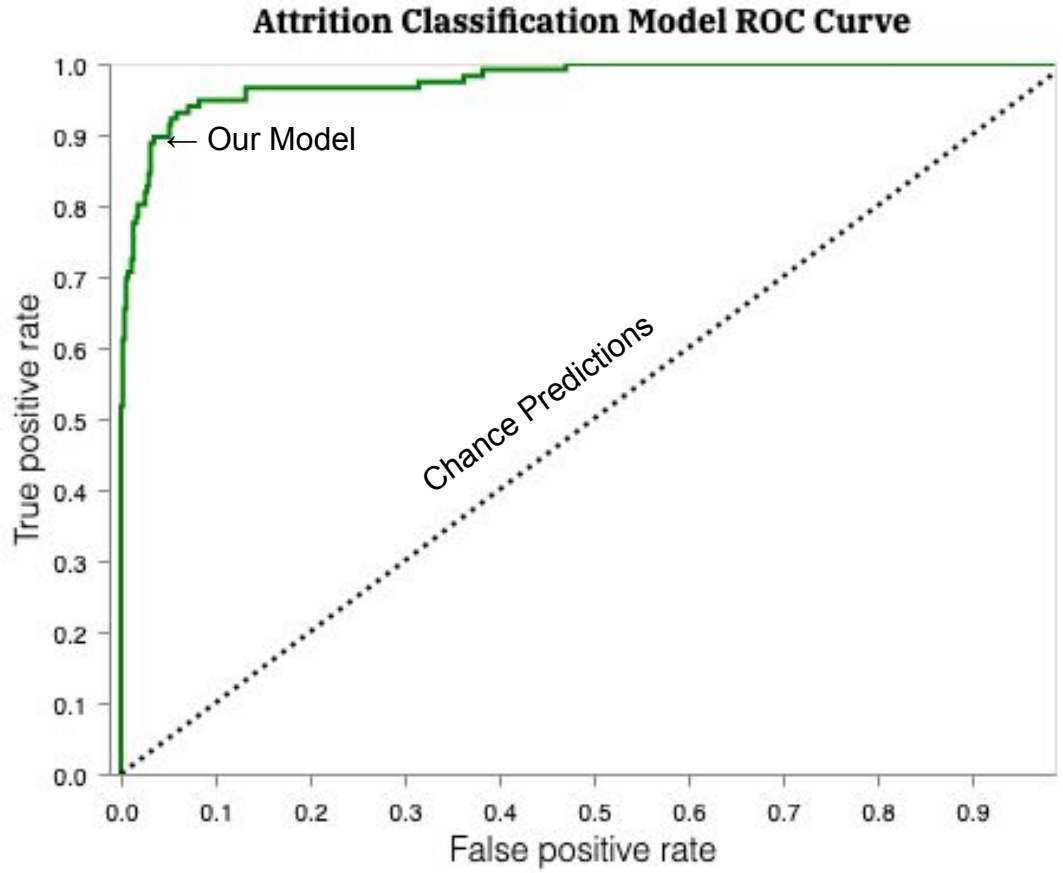
We used the liblinear solver as we have a small dataset and because there are binary outcomes.

# Results

## Summary of Model Performance

The model performs well on new data and is able to correctly identify 89% of individuals who will leave Acme as at risk.

| | |
|---|---|
| **Correct Rejection**<br>*True Negative* | **Type I error**<br>*False Positive* |
| 96% of individuals who stayed with Acme were correctly classified as having a low likelihood of attrition.<br><br>They would aptly not receive remediation. | 4% of of individuals who did not leave the company were falsely identified as at risk.<br><br>They would receive unnecessary remediation. |
| **Type II error**<br>*False Negative* | **Hit**<br>*True Positive* |
| 10% of individuals who would leave were incorrectly classified as having low risk of attrition.<br><br>They would have benefited from remediation as they left the company, but would not receive it. | 90% of individuals who would leave the company were classified as at risk of attrition.<br><br>They would receive necessary remediation. |




Better models arch strongly towards high true positive rates, as does our model. For comparison, randomly classified attrition events would result in the dotted line.

# Conclusion

## Recommendation

We **recommend implementing the turnover model** because its performance meets the needs of the business while maintaining simplicity, which means that it will require minimal engineering effort once deployed.

| When given new data, the model... |
| --- |
| does **+94% better** than always predicting the most common outcome meaning that our model has strong predictive efficiency. |
| **correctly classified 83%** of individuals who plan to leave as at risk. |
| When given novel data, the model **misclassified only 3%** of individuals who plan to stay and labeled them as at risk of attrition. |
| scores an **F1 of .86**, where the best possible performance is 1. |



### **Model in Practice**: a possible execution

Run quarterly batch jobs against HR database.

Managers of at risk individuals will be notified.

Begin workforce planning and remediation.

### **HR Recommendation**: limit work-related travel

While job seniority variables were the strongest predictors of attrition, this information does not lend well to action. Therefore, as work-related travel was the second greatest contributor to attrition **Acme should limit** business trips and allow the possibility to work from home when possible.

Our model projects that by offering the option to work from home, we would **retain 300 extra employees** (an improvement of 6%) in 2023, resulting in a projected savings of **₹90,000,000** of acquisition costs plus savings from limiting office space rentals.

# Limitations

## Root Association and Predictive Model

### 01 Data and wrangling

**Attrition is a rare event** in our dataset (16% of sample) meaning that our model can be biased against attrition and underpredict attrition. A technical correction was applied to mitigate this limitations.

Roughly **0.4% of rows contained missing data** and were filled with column means without consideration for skew. This practice decreases variance in our predictors which can limit our model.

### 02 Assumptions

The model **assumes that variables are not correlated**, which is not the case (e.g., years at the company and years since an employee's last promotion are strongly correlated).

We can mitigate this risk by removing some highly correlated variables or creating new linear combinations of the correlated variables to include in the model.

### 03 Model

Logistic Regression **creates linear boundaries** and between the predictors and the log odds of attrition which may not the true state.

While there is little we can do to correct for fighting a linear boundary to non-linear data, model performance metrics allow us to glean if our model is valid.

# Executive Summary
# Addressing High Employee Turnover

| Problem | Approach | Results | Conclusion |
|---|---|---|---|
| Acme Aroma is facing a high attrition rate of 16%, an increase of 12% since 2020. Because recruiting new candidates is costly and our ideal employees have a few years of experience at Acme, addressing turnover is a priority.<br><br>Goal: limit attrition in 2023 while minimizing costs. | We conducted root association analyses to identify employee attributes that strongly influence attrition.<br><br>We also built a model to run against the HR database and identify individuals likely to leave the organization. HR can then offer remediation to at-risk employees. | Job seniority and the need for work-related travel, including commutes to work, were the highest predictors of attrition.<br><br>Senior employees are 3.5 times more likely to leave than their junior counterparts. Likewise, the risk of attrition doubles for every seven kilometers of commuting beyond a baseline of eight kilometers.<br><br>When tested against new data, the model correctly identified 90% of individuals who left as in need of remediation. | While work level doesn't lend well to intervention by HR, limiting work-related travel is cost-effective in that it doesn't require expenditure and will result in considerable savings in rental fees.<br><br>Our model projects that by offering the option to work from home, we would retain 300 extra employees (an improvement of 6%) in 2023, resulting in a projected savings of ₹90,000,000 of acquisition costs plus savings from limiting office space rentals. |

We recommend

- Offering to work from home, especially if they live more than 8 km from the nearest office (top action item)
- Offer additional support to senior employees especially those in job level 4
- HR runs the logistic regression model quarterly against new data to identify and plan upcoming remediation efforts and necessary recruiting.