ACME Aroma Project

Table Of Contents

1. Problem
2. Approach
3. Insights (Data Exploration)
4. Insights (Model Quality)
5. Insights (Results)
6. Recommendations
7. Limitations

Patrick Coady

12/11/2023

ACME Aroma Project
Problem

□2X Turnover

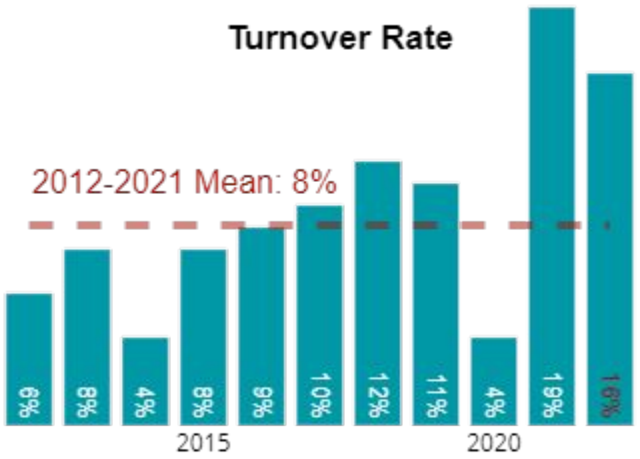
₹19 Million Replacement Cost

□15% Job Satisfaction

Turnover

Acme Aroma has recently ran into issues maintaining it's workforce, particularly in the post-covid world.

In this last year, 2022 we had a turnover rate of 16%, double that of the average turnover rate from the previous 10 years (2012-2021) of 8%.



Cost

The current cost of employee acquisition is 30,000 Rupees, this is a 2x increase from the cost in 2021. That means at our current turnover rate it is costing our company:

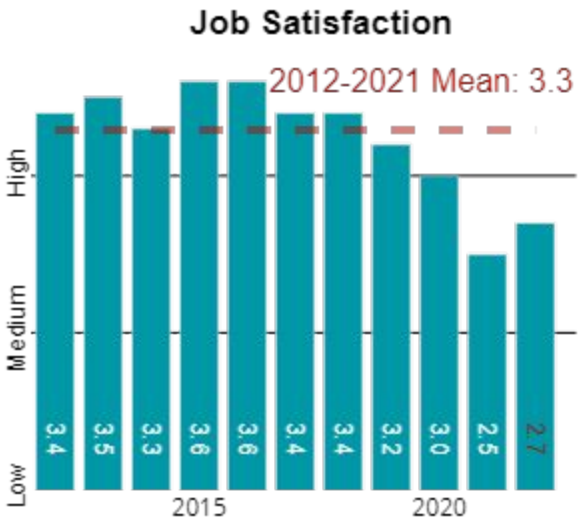
3965(Employees)×.16(Turnover)×30k(Cost)

= ₹19 Million

This cost should be seen as the minimum expense for our turnover because it doesn't include the loss of knowledge, and expertise that these employees take with them when they leave the company. It also doesn't capture that additional complexity to strategic planning as it's difficult to estimate the amount of work a team can finish, if so many members are either leaving, or spending their time training new employees.

Job Satisfaction

Employee Job Satisfaction has also decreased in this same time period. In the previous 10 years we had an average job satisfaction of **3.3/4**, but since 2021 this has dropped significantly currently sitting at 2.7/4. This represents a drop in employee job satisfaction by 15%.



ACME Aroma Project
Approach

Problem | Approach | Insights | Recommendations | Limitations

Dataset

Our dataset came from the HR team's information system. It includes objective measures like education, salary, and distance from work, along with subjective measures from employee surveys like Job Satisfaction, Performance Rating, and Work-Life Balance.

The model employs these variables, among others to predict employee attrition.

Goals

Our model predicts employee turnover, addressing two key issues:

1.Identifying reasons for departures and implementing solutions.

2. Enhancing capacity forecasting for better planning.

Using logistic regression we predict if an employee will leave and to identify the main factors from the dataset that contribute to attrition.

Evaluation

The model will be judged using the following metrics

False Positive: We predict that the employee will leave, but they stay instead.

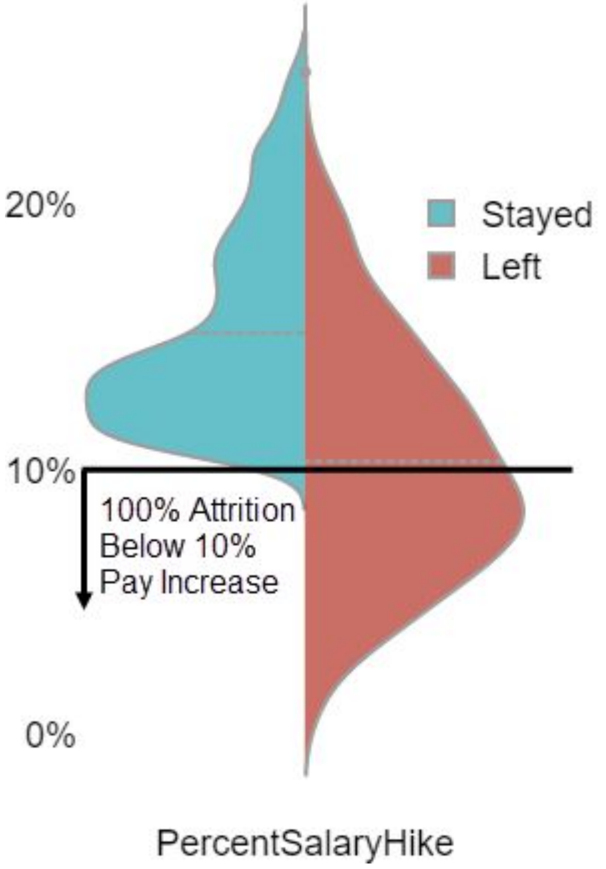
False Negative: We predict that the employee will stay, but they end up leaving.

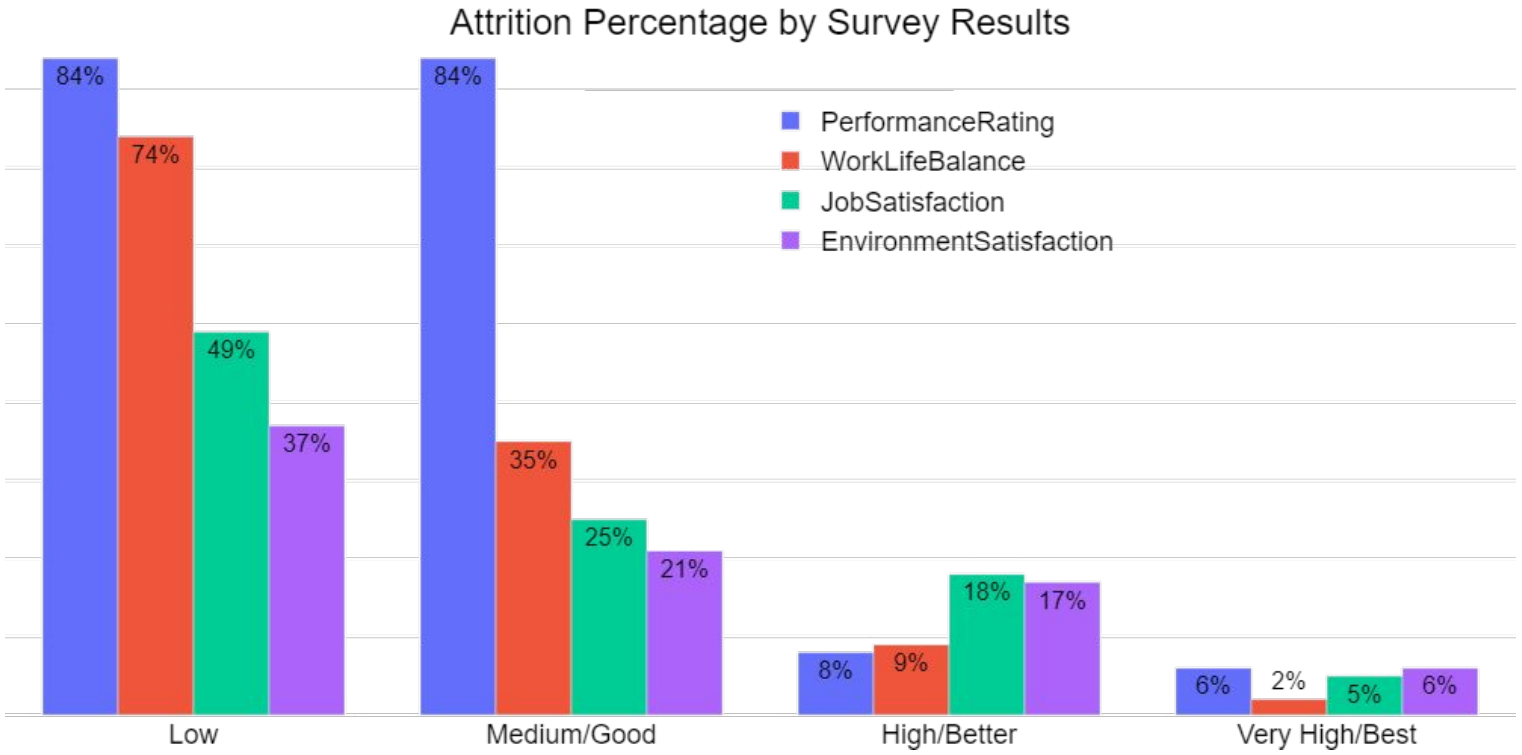
Recall: For everyone who left the company, how many of those people did the model correctly identify?

Precision: For everyone the model predicted would leave the company, how many actually did.

ACME Aroma Project
Insights (Explore Data)

We discovered that **every single employee leaves the company if their pay increase was less than 10%**. Inflation in India 2022 was 6.7%, that means the impact of our pay raises was damped significantly, as the cost of living for our of our employees increased.





In the bar chart above we can see the relationship between the various survey questions and what percentage of the respondents left given their response. Moving from left to right the survey answers get more positive. A couple of standouts Performance rating which jumps from 84% of respondents leaving the company if they got a ratings of Good OR Low, but drops significantly at the high level. This may be because of a limitation in the survey which I talk about in the limitations section.

Another Standout is Work Life Balance, which has a steady decline in attrition until it reaches the remarkable rate of 2% for those who scored work life balance a “Best” score.

ACME Aroma Project
Insights (Model Quality)

Quality

In the approach section we talked about different measure of quality for a model. Specifically, Recall and Precision.

From the table on the right, Recall is "For everyone who left the company, how many of those people did the model correctly identify?" That would be the number of people in the bottom right quadrant divided by the number of people in the bottom row, the closer to 100% the better. Our score was:

Recall: 81%

The other metric we used was precision which we defined as: "For everyone the model predicted would leave the company, how many actually did."

Looking at the same table, this means our precision is the number of people in the bottom right divided by the total number of people in the right row. Again the closer to 100% the better, our precision score was:

Precision: 63%

This model tends to classify too many people as leaving, when they actually stay, but this can be tuned as discussed in the next section

Choice

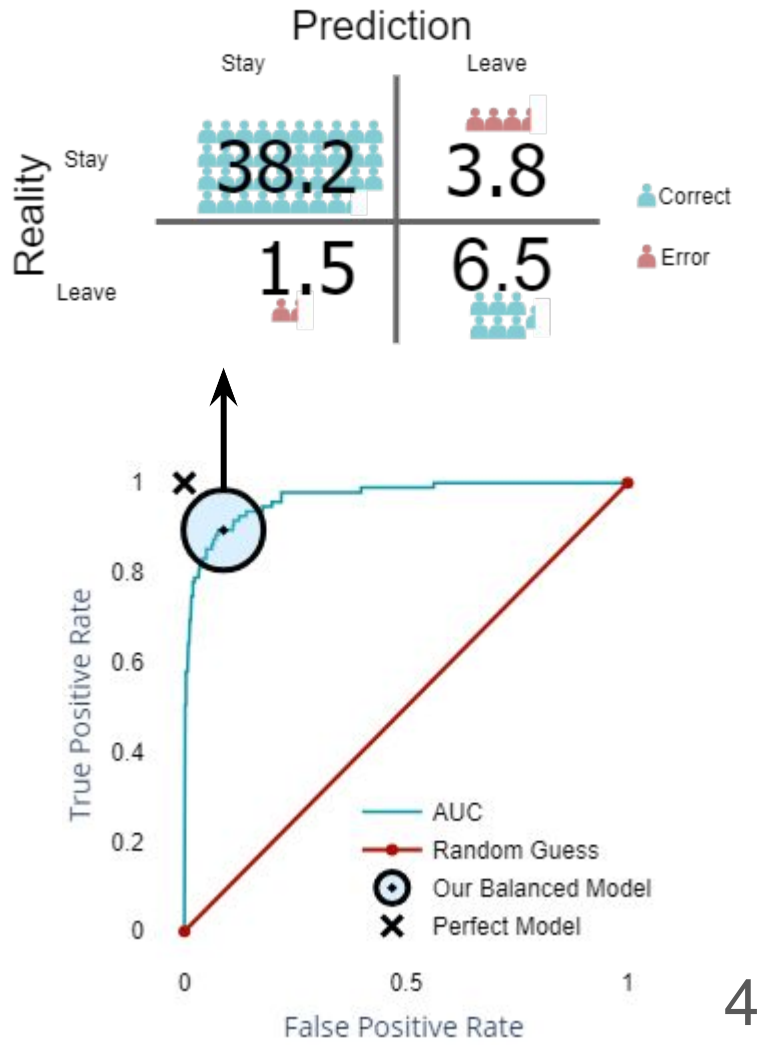
This Logistic Model uses multiple variables and their relationship to attrition to score each person on how likely they are to leave the company. We can then set a threshold, such that, anyone beyond it will be marked as "Predicted to Leave" and below it "Predicted to Stay".

There are two kinds of mistakes we can make, we can either predict someone will stay when they actually leave, or we can predict someone will leave when they actually stay.

Using the threshold we can tune the model to prefer to make one error over the other while still keeping the overall model quality high. On the bottom right we have a visualization of this choice, with the table above showing a single point on that graph.

As we move left to right on the x-axis we will increase the number of people we incorrectly assume will leave the company, the means the top right quadrant will get larger As we move up the y-axis the number of people we accurately mark as leaving increase so the bottom right quadrant will get larger.

For our purposes each kind of error is equally costly, so we have going to minimize error in general. However, if this model is used in the future different choices may need to be made.



ACME Aroma Project

Insights (Results)

Variables Effects on Attrition

Variable	% Increase of Chance to Stay
Percent Salary Hike	**1.5X** per 1% increase in pay
Environment Satisfaction	**2X** per 1 point increase in environment satisfaction survey
Training Time Since Last Year	**1.4X** per 1 point increase in training
Work Life Balance	**7.1X** per 1 point increase in work life balance
Job Satisfaction	**2.9X** per 1 point increase in job satisfaction survey
Distance From Work	**1.08X** per 1 KM closer to the office

The table to the right shows how much more likely an employee is to stay at this company when the corresponding metric increases. THe biggest standout here is Work Life Balance. Our employees are 7 times more likely to keep working for us for every point increase in work life balance.

This is obviously something our employees care deeply about, and an area where we can improve a lot. Our average work life balance score is 2.7/4 which is somewhere between “good” and “better” so there is a lot of room for growth. Even beyond our specific recommendations, this seems like an area that require more investigation on how we can improve employees work life balance.

There are two other variables that need to be called out, because they don’t compare well to the others. Percent Salary Hike, and Distance from work. 1.5 times more likely to leave the office for each 1% increase in pay doesn’t sound like a lot, but considering our average pay increase was 13% this also represents an easy, although costly way to keep our employees.

Similarly for distance from work. Our farthest employee lives 35 KM away. That means he is 37 times more likely to quit then the average employee. In fact every single employee who lived more than 25 km away left our company. With work from home becoming more popular this is another avenue we can pursue to keep our employees longer

ACME Aroma Project
Recommendations

Problem | Approach | Insights | Recommendations | Limitations

Initiative Effects on Attrition

Initiative	Effect on indicator	Employees who stay instead of leaving	Savings (Millions of ₹)
Pay Base Increase	Increase pay by 7500 Average of 11.7% pay increase	364	15.4 Cost(29.7)
More Training	Increase Training TIme Since Last Year by .5	36	1.5
Workplace Flexibility	**Increase Environment Satisfaction by .75**	**111**	**4.7**
Limit Business Travel	Increase Work Life Balance by .3	121	5.1
Employee Appreciation	Increase Job Satisfaction by .5	110	4.6

Initiative Effects on Attrition

Our Recommendation is to institute Workplace flexibility, allowing our employees to work from home if their role and duties allow.

This may be a surprise given the table to the left showing 3 other indicators that result in less attrition. However, given what we discovered about work life balance, and distance from office and their effects on attrition. .

Pay increase would by far have the largest effect in reducing attrition, and inflation should be considered when giving raises. However, the cost of increasing everyone's pay by ₹7500 eclipses the savings we would gain by retaining more people.

Between Workplace Flexibility, Limit Business Travel, and Employee Appreciation the results are all so close the difference in prediction could be error in the model (which is expected). But working from home would also decrease distance to work, although working from home could be categorically different than working from the office and this model has not captured that relationship for us to calculate the size of the effect.

Also I believe that working from home *should* increase work life balance and suggest we do a study to confirm or reject that belief. Finally, working from home will allows us to capture a larger labor market enabling the acquisition of higher quality employees.

ACME Aroma Project
Limitations

Data Limitations

The survey questions have a positivity bias is the available answers. For instance, the options for “Work Life Balance” are:

- Bad

- Good
- Better
- Best

There is is very little to distinguish someone with ok work life balance and a very bad work life balance. The same pattern holds for all of the other rankings. In the future we suggest using a five point scale with an equal balance of positive and negative responses

For Example:

- Very Bad
- Bad

- Neutral

- Good
- Very Good

Model Limitations

One limitation for almost any model we could choose, is that correlation is not causation. Just because one variable influences the attrition rate doesn’t means that’s what’s causing the increase.

Instead, we should see these relationships as being clues to a larger mystery. We then use these clues to run experiments, and guide our future efforts in collecting more clues.

Another limitation is with Logistic Regression itself. It does a great job when as performance rating increases attrition decreases, and more performance rating is always better.

But the model struggles to capture more complex relationships. For example, if both a low and high performance rating leads to attrition, but a middle rating didn’t our model would fail to capture this relationship.

Impact of Limitations

Because of these limitations we will need to follow two general design principles, experimentation and iterative design.

For every recommendation, we will design an experiment, evaluate the results, then using this information we will update our beliefs and create a new experiment.

We will follow this basic development loop until we gotten the attrition rate down to an acceptable level.