Question: Why are those best and most experienced employees leaving prematurely?

Report by: Qi Jing Qian (Eugene)

## Problem scoping and identification

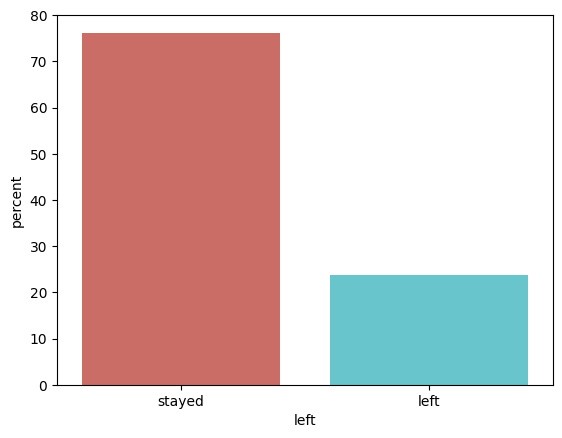


Fig 1. Proportion of all employees who stayed or left

As a whole, the company had about 3,571 employees (23.8%) leaving the company, as seen from Fig 1. However, we are looking at the best and most experienced employees. Hence, we need to define and zoom in on the portion of the company that is the best and most experienced. The chart below shows the descriptive statistics of the data.

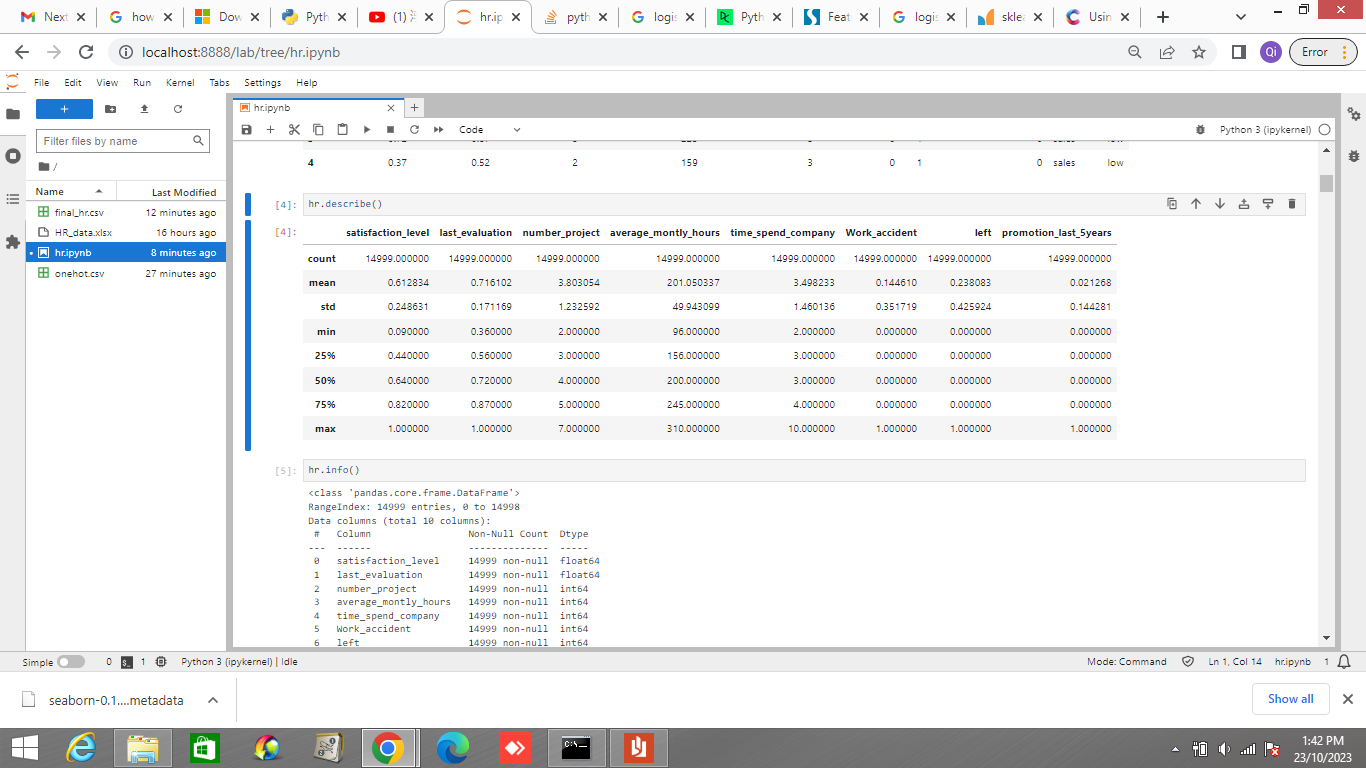


Fig 2. Table showing the descriptive statistics of the whole employee dataset

From the descriptive statistics in Fig 2, we can define the best through the last\_evaluation, which gives us a number from 0 - 1, giving us an indication of performance, where the higher the score, the better the performance. We can also define the most experienced as the employees that have done the most number of projects. Using the top 25% figure as a benchmark, we can infer that the best and most experienced personnel are those that have done 5 or more projects and have an last evaluation score of 0.87 and above.

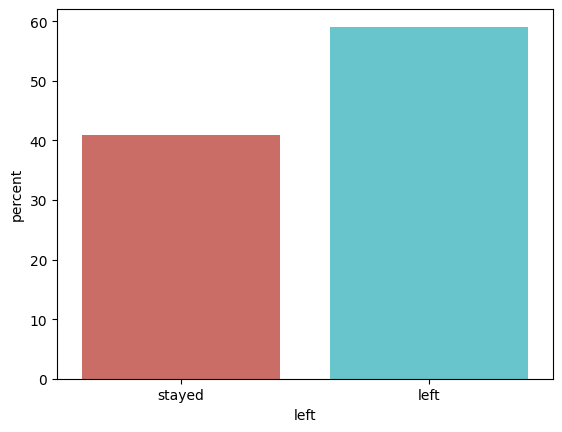


Fig 3. Proportion of the best and most experienced employees who stayed or left

As we can see in Fig 3, 906 of the 1534 best and most experienced employees (59.0%) have left the company. This percentage (59.0%) is about 2.5 more as compared to that of the overall population (23.8%). This proves that the talent churn problem among the best and the most experienced is a huge problem.

## Possible Reasons for Employee Churn

There could be a few possible reasons for this churn.

1. Lack of opportunities as seen by the near negligible promotions over the last 5 years

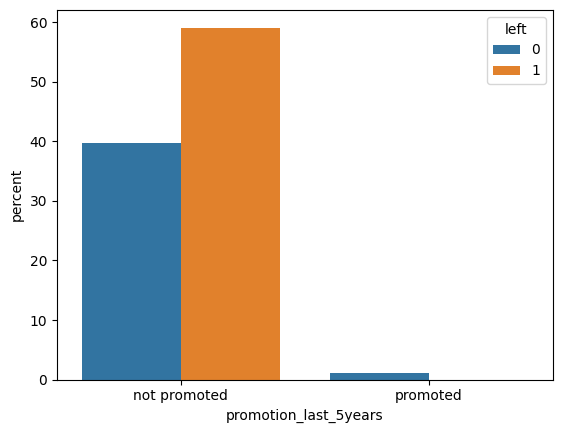


Fig 5. Proportion of the best and most experienced employees who are promoted and whether they stayed or left

From the Fig. 5, we can see that more than 95% of the target employees (best and experienced) have not been promoted over the past 5 years, despite having stellar evaluation scores and deep expertise as shown by the number of projects they’ve been involved in. Of those that have yet to be promoted, close to 60% of them left the company and everyone who got promoted stayed in the company, which empirically shows a correlation between the lack of promotion and employee leaving.

|  | Stayed in company | Left company |
| --- | --- | --- |
| Not Promoted in last 5 years | 610 | 906 |
| Promoted in last 5 years | 18 | 0 |

Fig 6. Table showing proportion of the best and most experienced employees who are promoted and whether they stayed or left

Moreover, from Fig 6., the number of promotions in general may not be encouraging. Over the past 5 years, only 18 of 1534 exceptional employees (1.1%) have been promoted. Due to the lack of opportunities to move upwards in the corporate ladder in such a long period of time, this might push employees to leave and pursue other opportunities.

1. Overworked employees that work a high number of hours on average would leave the company.

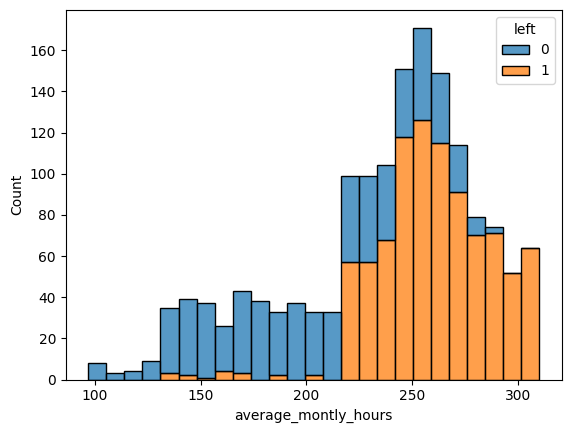


Fig 7. Graph showing count of the best and most experienced employees and their average monthly hours, grouped by whether they stayed or left

From Fig 7., we can tell that employee churn is high for employees working above 220 hrs, and increases as the average monthly hours increase. This is very logical as if we assume 22 working days in a month, these employees have been working for more than an average of 10 hours every working day. From the graph, we can infer that for employees working more than about 220 hours in a month on average, as the monthly average hours the employee works, the more likely the employee is to leave the company as well.

1. Exceptional employees may leave the company after a certain number of years and having sufficient experiences

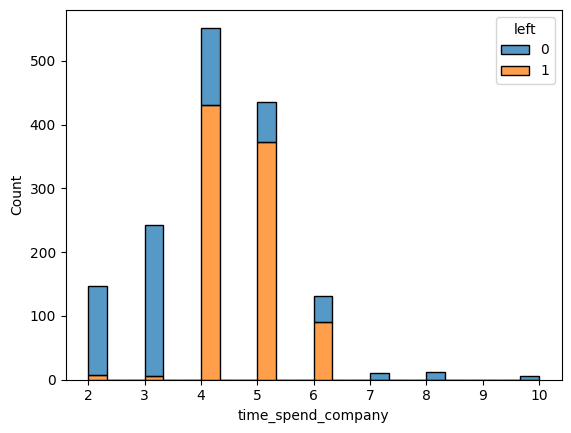


Fig 8. Graph showing count of the best and most experienced employees and their time spent in company, grouped by whether they stayed or left

Exceptional employees may leave the company after a certain number of years/ having enough experiences, possibly realizing the lack of opportunities and to look for other opportunities after staying for a certain number of years. From the graph above, we can observe that there is a tendency for employees working for mid-level employees (4 - 6 years) to leave the company. This could also explain why for exceptional employees, there tends to be very few senior employees (those working for more than 7 years), as employees might have less with 4 - 6 years of experience in the company.

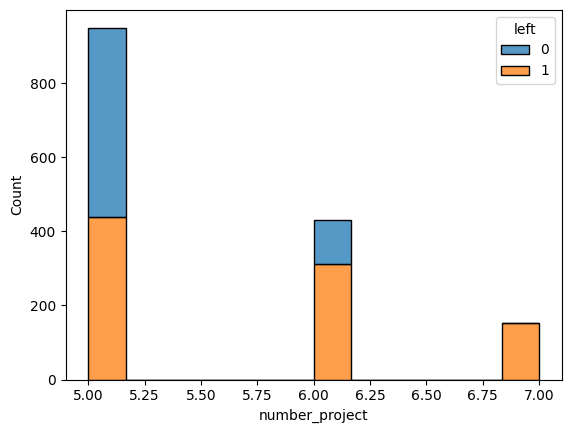


Fig 9. Graph showing count of the best and most experienced employees and their project count, grouped by whether they stayed or left

This is further supported as we see that for experienced employees, the trend seems to be that as more projects are done, employees with more project experiences would leave, with nearly all employees who have completed 7 projects leaving the company.

## Feature Comparison using logistic regression

In order to verify the reasons as listed above, I created a machine learning model to compare the importance of different features in the dataset for the target employees. This can be compared to the baseline model of predicting that all exceptional employees would leave, which would give the baseline model at 59.0%. With minor tweaking, I achieved a 92.1% accuracy score and 90.7% ROC AUC score, which shows that my logistic regression model works well in predicting employee churn.

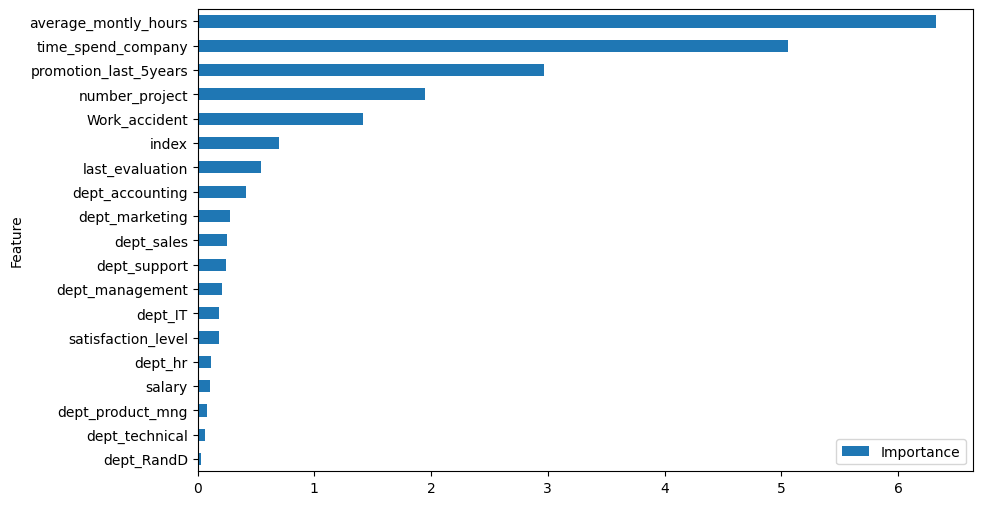


Fig 10. Graph showing features ranked by their importance in predicting employee churn

Extracting the important features, we realize that almost all the features that are the most significant proved to be important in the model, including:

1. the average monthly hours spent;
2. time spent in the company;
3. whether the employee was promoted over last 5 years;
4. the number of projects the employee was involved in.

The ‘index’ feature is random indexes assigned to the employees in the excel, hence we can assume that any features that rank lower than the ‘index’ feature are not significant as they have less predicting power than random. From the model, the ‘Work accident’ feature is notably an important feature for prediction as well, albeit the least important feature for prediction from the graph above.