Employee attrition problem

23.11.2022

Ayshik Neogi (MSc II)
Roll No: 2102
Department of Statistics,
Savitribai Phule Pune University

Saurav Jadhav (MSc II)
Roll No: 2120
Department of Statistics,
Savitribai Phule Pune University

Introduction

A large company named ABC, employs, at any given point of time, around 4000 employees. However, every year, around 15% of its employees leave the company and the company needs to replace them with the talent pool available in the job market.

The management believes that this level of attrition (employees leaving, either on their own or because they got fired) is bad for the company, because of the following reasons:

1. The former employees' projects get delayed, making it difficult to meet timelines, resulting in a loss of reputation among clients and partners.

2. More often than not, the new employees have to be trained for the job and/or given time to acclimatize themselves to the company.

The purpose of the analysis is to find what factors share a strong relationship with attrition and use them to decide what changes to be made in the workplace to retain employees.

Before we dive into the analysis of the data. It is first important to think about this problem instinctively.

1. How would a firm reduce its attrition rate?
2. And, more importantly, what causes an increase in the firm's attrition rate?
3. Or, what factors make an employee stay in the company?

Heuristic-based answer to the problem

From the above three questions, we will be specifically looking at the third question:

What factors make an employee stay in the company?

After looking at the question, the first thing which can immediately come to mind is the good pay scale.

1. Pay scale above the industry standards seems like a good reason for an employee to stay in the company.

As this is a business problem, increasing the pay scale of the employee for the sake of reducing the attrition rate is not an economical solution. Also, this may not be the sole reason due to why an employee might stay in the company.

2. Appreciation to the employees for their good work.
3. Promotion for deserving candidates.
4. Friendly environment: communication apart from professional work.
5. Non-toxic work culture. Asking employees to work on weekends, more than 8 hrs per day, and so on may promote some sort of toxicness in the work.

The above solutions, if imparted might reduce the attrition rate in the company.

Data-driven solution to the problem

The data consists of 6 different sheets:

1. Data Dictionary: Consists of metadata
2. Employee Survey Data: This consists of employee survey data such as job satisfaction, work-life balance, and so on.
3. General data: The data consists of age, employee ID, gender, and other general information about the employee.
4. In time: The data consists of the login time of the employee
5. Out Time: The data consists of the logout time of the employee.
6. Manager Survey Data: The survey data consists of manager ratings of the employee

All the above data is considered for analysis.

General Information

Number of employees in the company: 4410
Number of employees who left the company the previous year: 711
Percentage of employees leaving the company: 16.12%

In the data it is verified that, if there is an NA in the employee log-in and log-out time, then an employee was on a leave. And for an NA in log-in time, we have an NA in log-out time for that particular date.

The event under which an employee leaves the company has to be significant. Under normal conditions, observing an employee leave the company may be unlikely. Observing the extreme cases will be more informative and can help understand the cause of an employee leaving the company. For example, companies paying employees below industry standards will have a higher attrition rate. In our case, the base attrition rate of the company is around 16%. But with the employees who have a low pay scale, it is around 18%.

Three features are constructed:

1. avg_time: This feature indicates the average time an employee spends in the company per day.
2. avg_time_In: This feature indicates the average time at which an employee enters the company daily.
3. Employee Leaves: This feature indicates the number of leaves taken by an employee apart from holidays throughout the year.

Tabular interpretations

By looking at the below tables, we observe the attrition rate for very low and very high values of the features.

For the below table, 5-percentile and 95-percentile values of the attrition rate are considered.

The 5-percentile value indicates the value under which the 5 percent of the data lies.
The 95-percentile value indicates the value below which the 95 percent of the data lies.

The highlighted values indicate the significant differences from the base attrition rate of the company that is, 16%.

Attrition rate for the extreme case in the numerical features.

Table 3.1.1: Attrition rate for 5 percentile and 95 percentile feature values

Features	5 Percentile	95 Percentile
MonthlyIncome	17.56	13.51
JobLevel	15.46	14.85
NumCompaniesWorked	14.33	**19.14**
PercentSalaryHike	14.29	18.70
StockOptionLevel	16.80	15.29
TotalWorkingYears	15.64	16.45
TrainingTimesLastYear	16.00	11.41
YearsAtCompany	**34.88**	**9.68**
YearsSinceLastPromotion	18.93	14.61
YearsWithCurrManager	**32.32**	**6.00**
Age	**39.18**	12.64
DistanceFromHome	15.38	13.79
EnvironmentSatisfaction	15.74	15.67
JobSatisfaction	14.65	15.73
WorkLifeBalance	15.90	14.98
JobInvolvement	**21.69**	18.06
PerformanceRating	15.76	18.14
avg_time	**9.50**	**26.24**
avg_time_ln	**8.14**	**21.27**
Employee_leaves	**21.46**	**9.70**

It is important to observe the higher attrition rate for the particular features. But not only this, but we also need to observe lower attrition rates in order to get a picture of why employees stay.

The numbers highlighted in the above table indicate the significant high or significant lower values of the attrition rate from the baseline, that is, 16% which is the attrition rate of the company.

From table 3.1.1 we observe the following:

1. The employees who worked with more companies are more likely to switch than average.
2. Employees who have not spent much time within the company are more likely to make a switch.
3. The employees who stayed for a long time within the company are very less likely to make a switch.
4. The employees who have not spent much time with the current manager are more likely to make a switch.
5. The employees who spent a lot of time with their current manager are less likely to make a switch.
6. Younger employees are more likely to make a switch.
7. The employees who are rated low on a job involvement scale are more likely to make a switch. This is also an indication that their manager may not be satisfied with their performance.
8. The employees who spent lower daily time in the company are less likely to switch the company. This may also be an indication of not having a workload, making an employee satisfied.
9. The employees who spent more daily time in the company are more likely to switch the company. This may indicate a higher workload.
10. The employees who tend to arrive early are less likely to leave the job. Maybe this is an indication of interest in the work.
11. The employees who tend to arrive late at the company are more likely to leave the job. It may be an indication of having a casual attitude toward the job.
12. Employees with lower leaves are more likely to leave the job.
13. The employees with more leaves are less likely to make a switch.

Attrition rate within the different categories of the nominal features.

The highlighted values indicate the significant differences from the base attrition rate of the company that is, 16%.

Table 3.2.1: Attrition rate different categories of Gender

Category	Proportion of employees leaving the job
Female	15.31
Male	16.67

Table 3.2.2: Attrition rate different categories of Job role

Category	Proportion of employees leaving the job
Healthcare Representative	14.50
Research Scientist	18.15
Sales Executive	16.87
Human Resources	13.46
Research Director	**23.75**
Laboratory Technician	16.22
Manufacturing Director	11.03
Sales Representative	14.46
Manager	13.73

The table 3.2.2 compares the attrition rate against the different job categories.

From table 3.2.2, we observe the following:

Research directors are more likely to switch the company.

Table 3.2.3: Attrition rate different categories of Marital Status

Category	Proportion of employees leaving the job
Married	12.48
Single	**25.53**
Divorced	10.09

Table 3.2.3 compares the attrition rate against the marital status.

From Table 3.2.3, we observe the following:

Single employees are more likely to leave the job. The reason might be to explore better opportunities in the initial stages.

Table 3.2.4: Attrition rate different categories of Business Travel

Category	Proportion of employees leaving the job
Travel_Rarely	14.96
Travel_Frequently	**24.91**
Non-Travel	**8.00**

Table 3.2.4 compares the attrition rate against the different business travel categories.

From table 3.2.4, we observe the following:

1. The employees who travel frequently are more likely to leave the job.
2. The employees who do not travel are less likely to leave the job.

Table 3.2.5: Attrition rate different categories of Department

Category	Proportion of employees leaving the job
Sales	15.02
Research & Development	15.71
Human Resources	**30.16**

Table 3.2.5 compares the attrition rate against the different department categories.

From table 3.2.5, we observe the following:

Human resource employees are more likely to leave the job.

Table 3.2.6: Attrition rate for different levels of Education Field

Category	Proportion of employees leaving the job
Life Sciences	16.67
Other	12.20
Medical	16.16
Marketing	15.72
Technical Degree	11.36
Human Resources	**40.74**

Table 3.2.6 compares the attrition rate against the different education departments.

From table 3.2.6, we observe the following:

The employees who come from a human resource background are more likely to make a switch.

Table 3.2.7: Attrition rate for different levels of Education Levels

Category	Proportion of employees leaving the job
Below College	15.29
College	18.79
Bachelor	15.56
Master	15.58
Doctor	14.58

From table 3.2.7, students with a college level of education are more likely to make a switch. Although the attrition rate is not significant.

While studying the data, it is also important to understand the missing values. In employee survey data, suppose there is a missing value in environment satisfaction, then it can be an indication of an employee not being satisfied with the environment and does not want to say it directly. Hence below we understand the attrition rates for different features where the NAs are present.

Below attrition rates are only calculated for those features where there are missing values. And the attrition rates are calculated only considering the missing values and no other values.

Table 3.3.1: Attrition rate for different features only with NAs

Features	Attrition rate
EnvioronmentSatisfaction	**20**
JobSatisfaction	5
WorkLifeBalance	10.53
TotalWorkingYears	**22.22**
NumCompaniesWorked	**21.05**

The above table indicates a higher attrition rate for NAs for the features, Environment Satisfaction, Total Working Years, Number of Companies Worked.

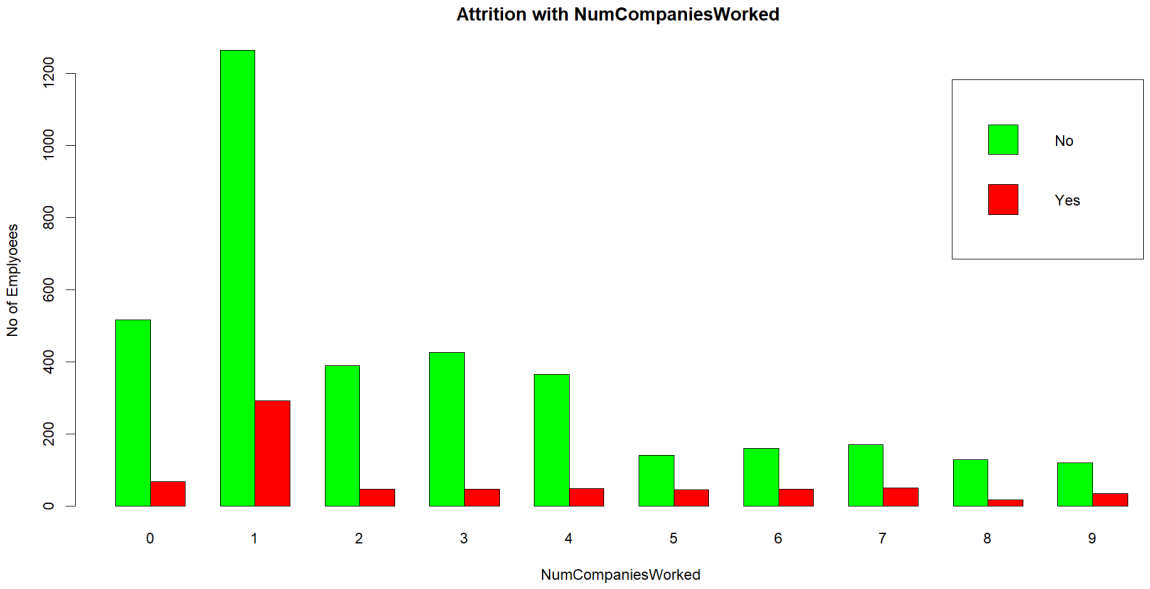
Missing values for Total Working Years and Number of Companies worked may be an indication of frequent switches in companies for that employee, hence a higher attrition rate.

But the conclusion cannot be drawn due to low NA size.

Graphical interpretations

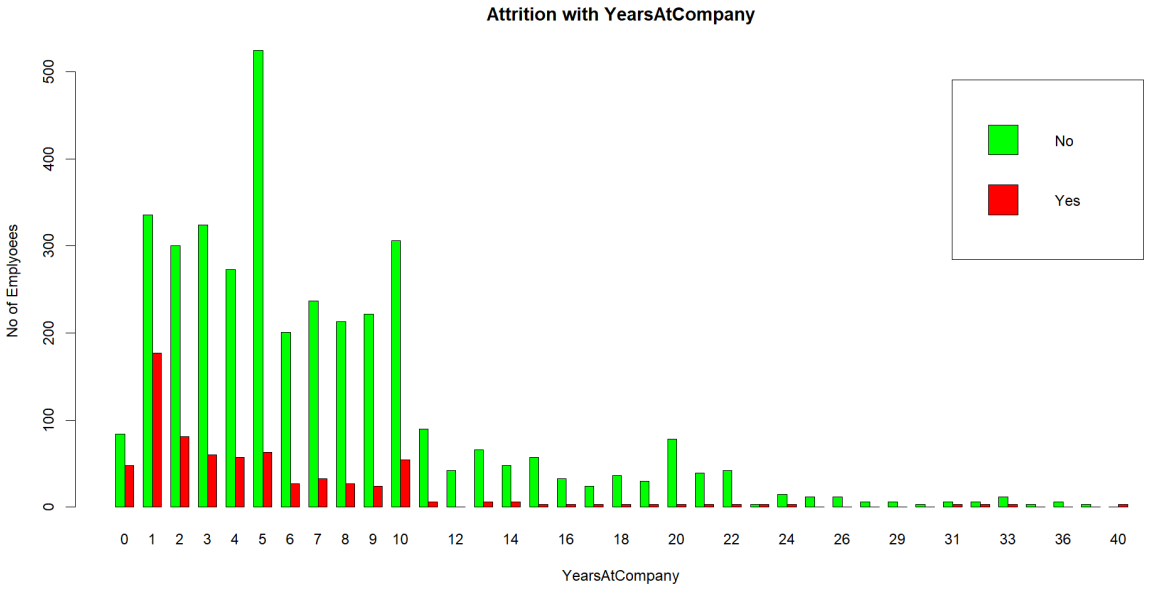
The tables gave us a picture of the attrition for different features. Graphs will help us easily understand the overall picture of attrition for different employee features.

The bars in red indicate the count of employees who left the company with respect to a specific feature. For example, count of employees who left the company with respect to the number of years at the company.
The bars in green indicate the count of employees who is still working in the company for that specific feature.



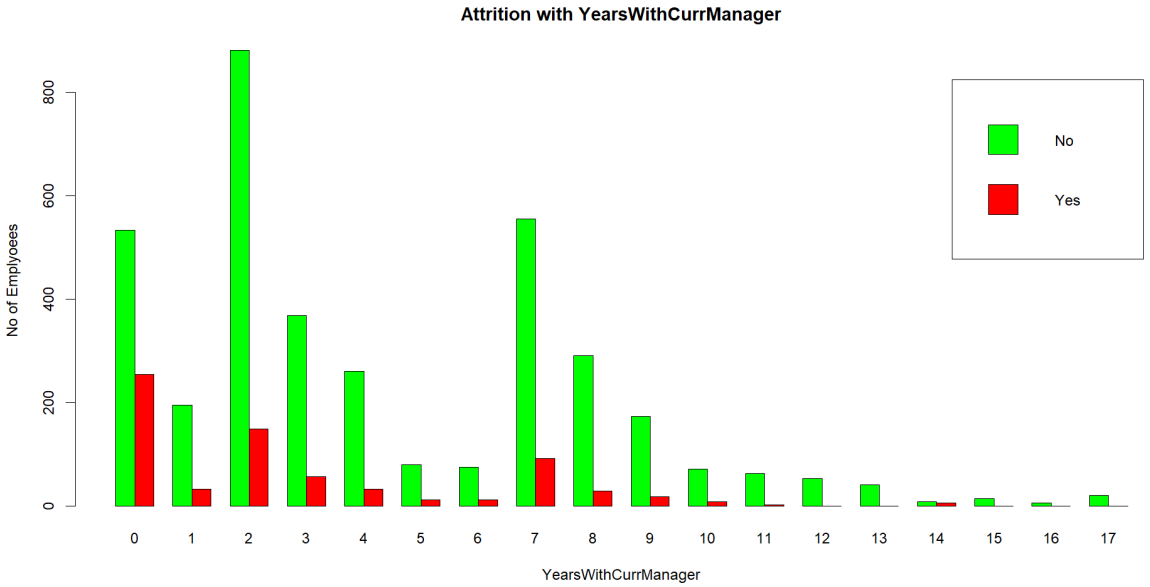
Graph 4.1.1

Graph 4.1.1 indicates the count of the employees with respect to the number of companies they have worked for before.



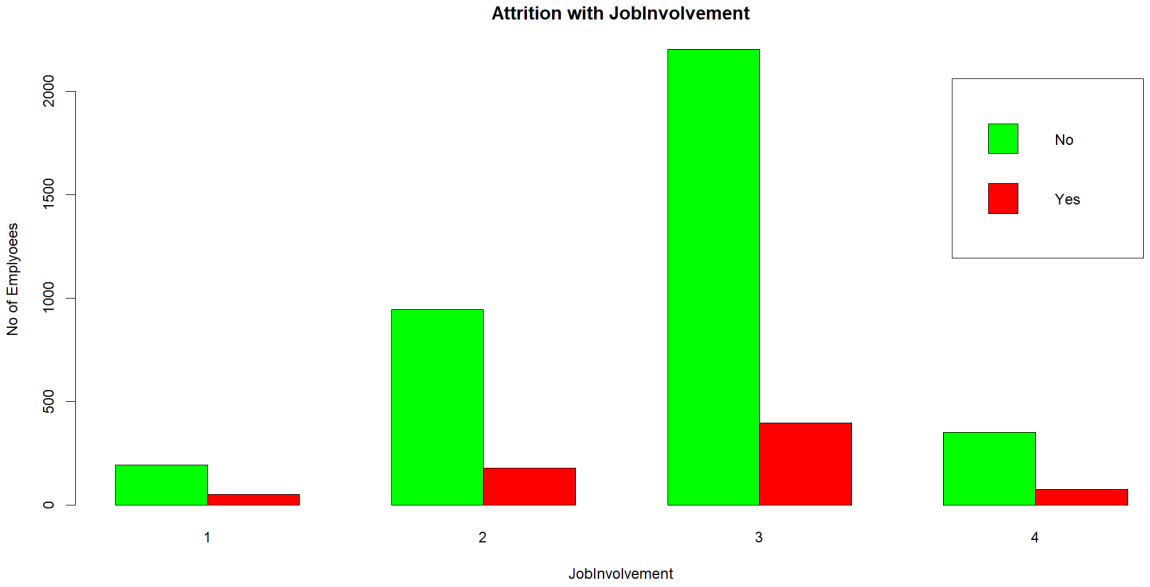
Graph 4.1.2

Graph 4.1.2 indicates the count of the employees with respect to the years at the company.



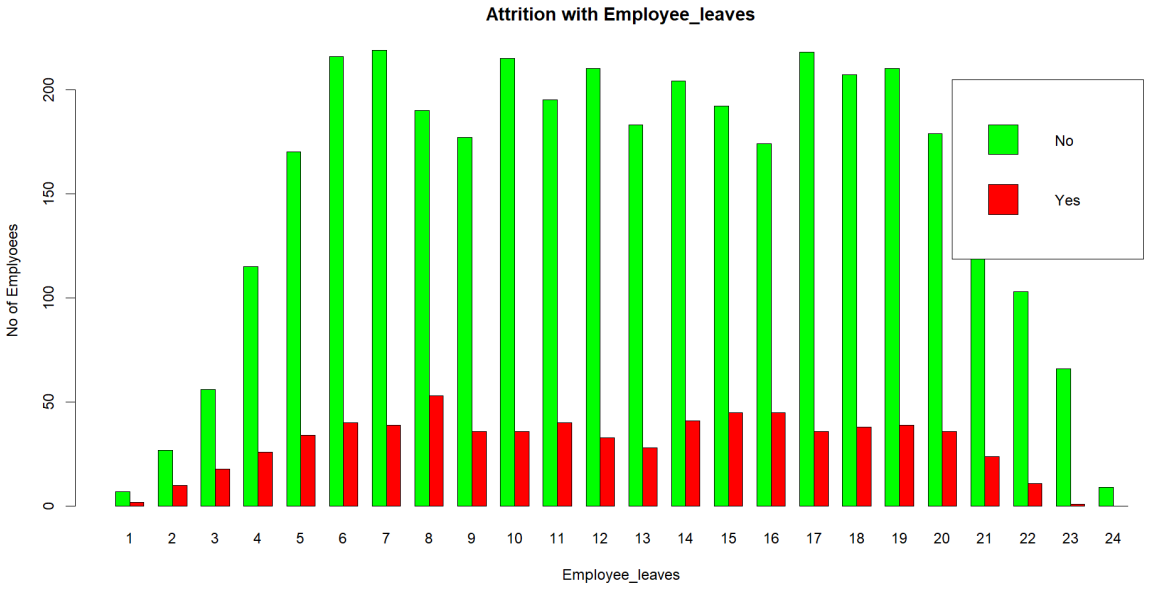
Graph 4.1.3

Graph 4.1.3 indicates the count for the employees with respect to the years with the current manager.



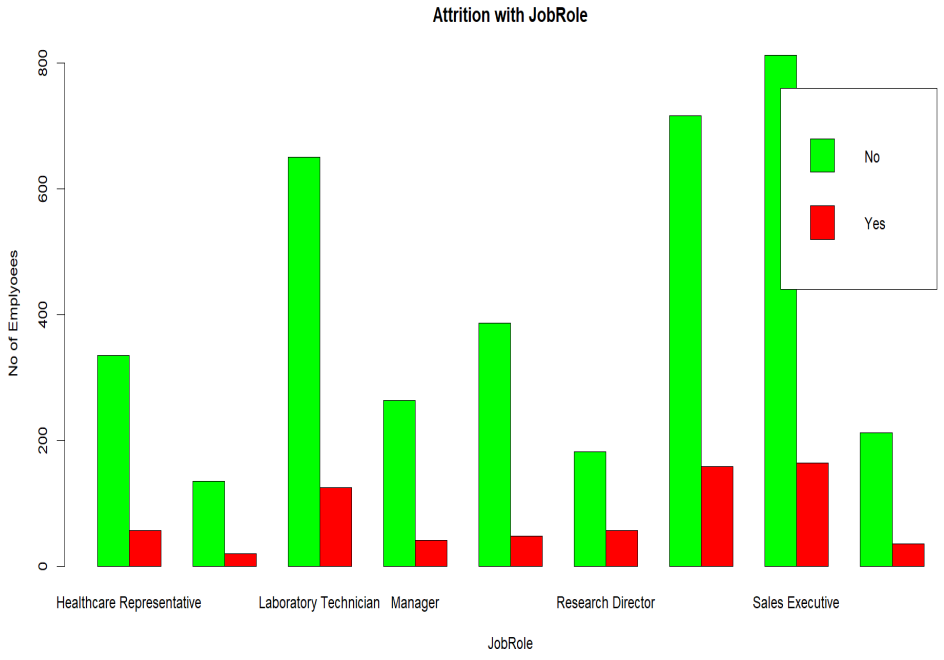
Graph 4.1.4

Graph 4.1.4 indicates the count of the employees with respect to their job involvement.



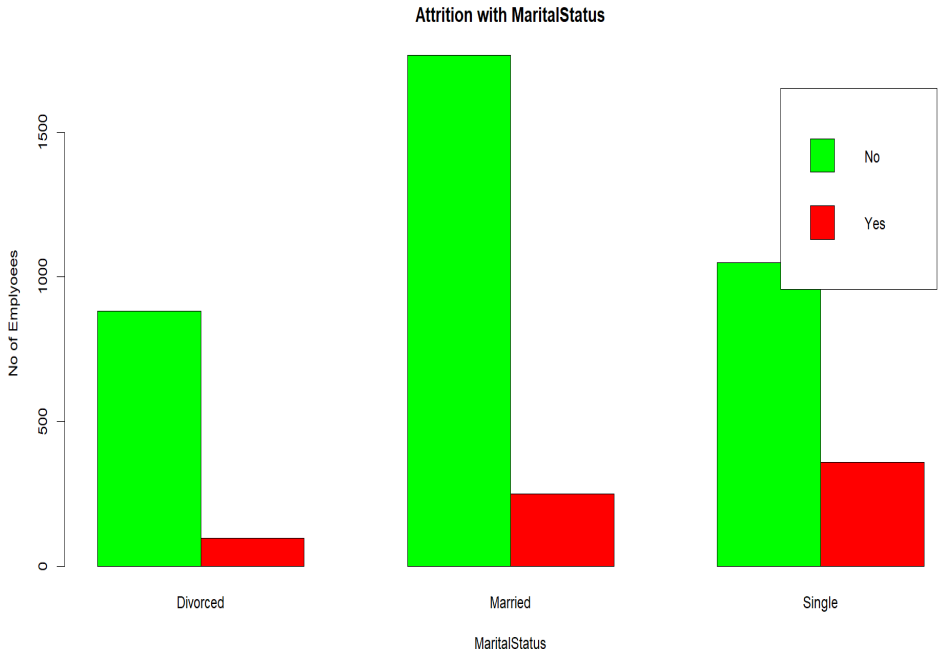
Graph 4.1.5

Graph 4.1.5 indicates the count of the employees with respect to the number of leaves they have taken.



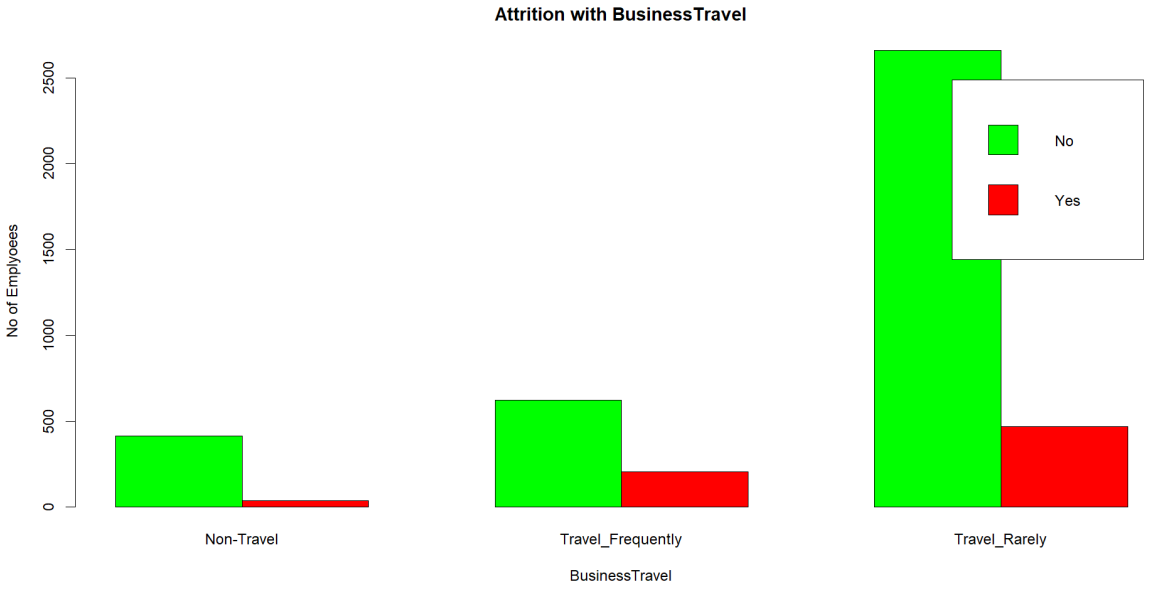
Graph 4.1.6

Graph 4.1.6 indicates the count of the employees with respect to different job roles.



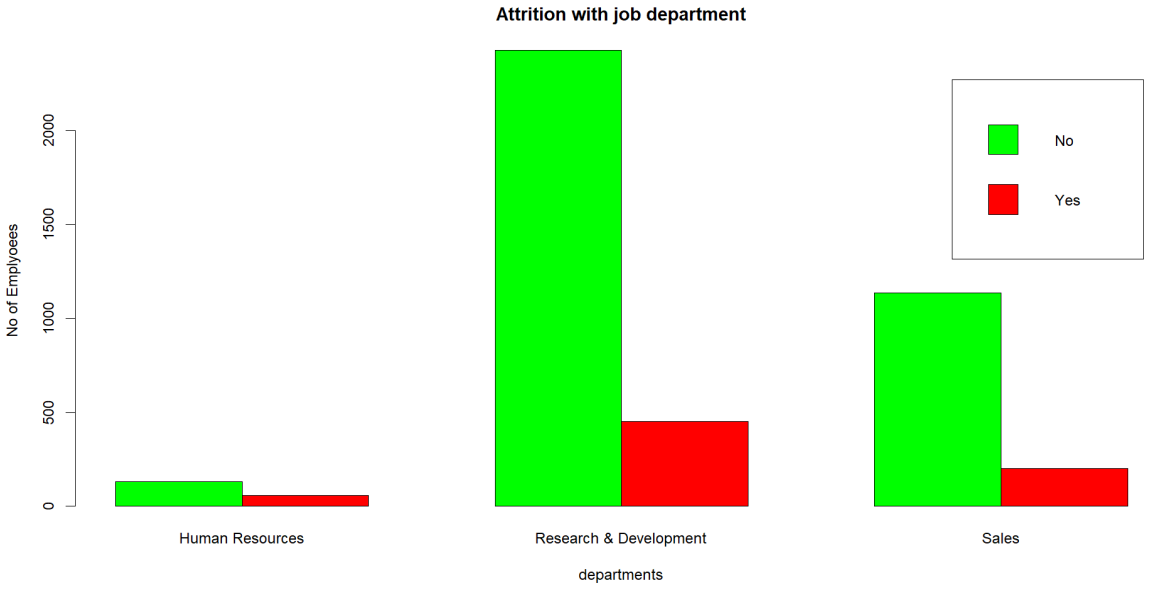
Graph 4.1.7

Graph 4.1.7 indicates the count of the employees with respect to the employee's marital status.



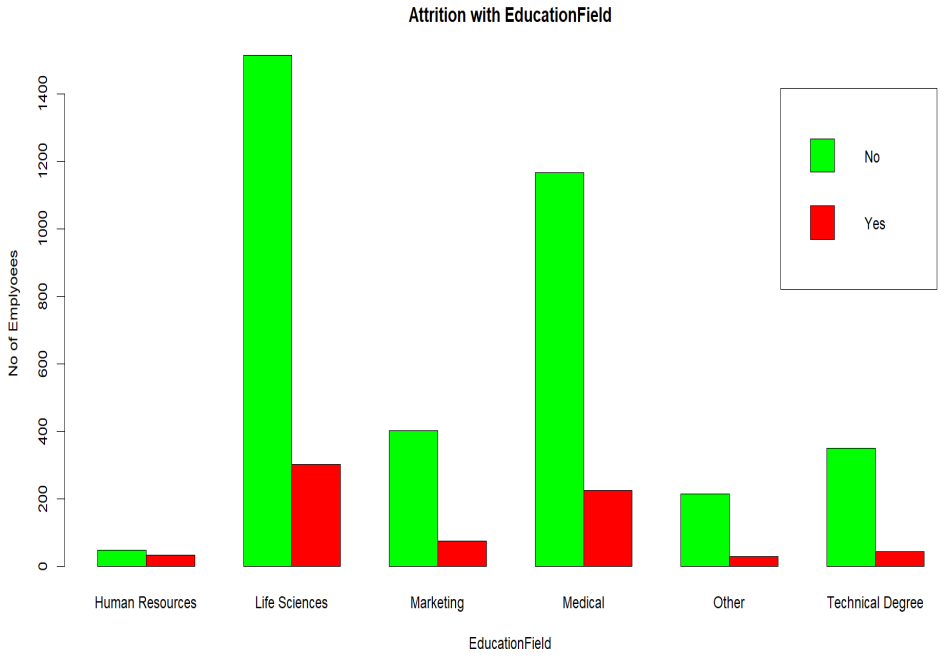
Graph 4.1.8

Graph 4.1.8 indicates the count for the employees with respect to the employee's frequency of business travel.



Graph 4.1.9

Graph 4.1.9 indicates the count for the employees with respect to the employee's job department.



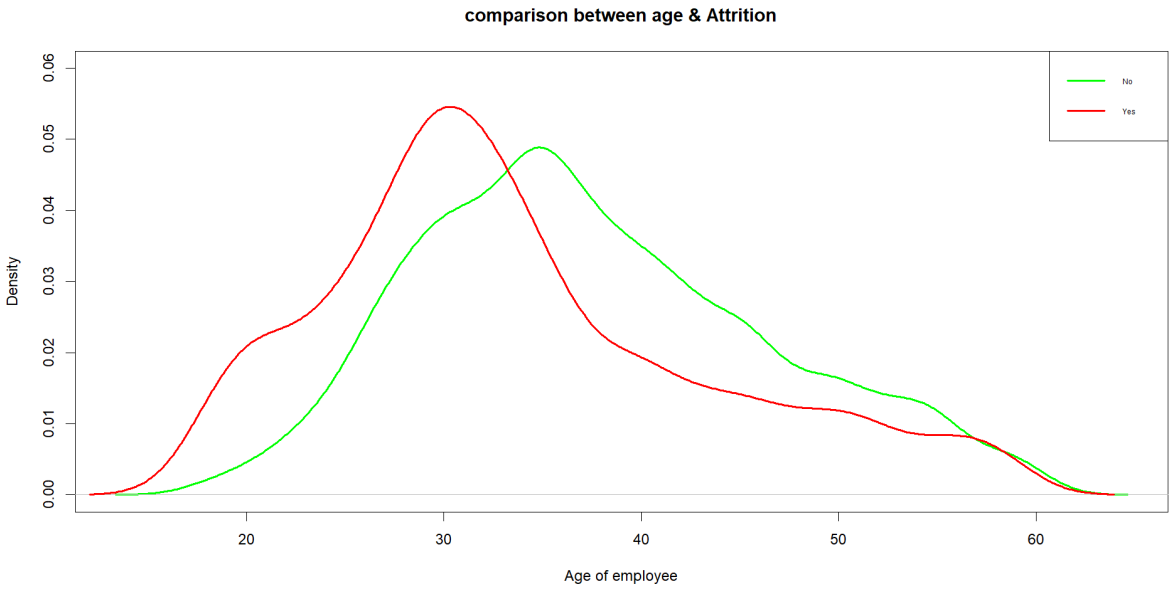
Graph 4.1.10

Graph 4.1.10 indicates the count for the employees with respect to the employee's education field.

Curve Interpretations

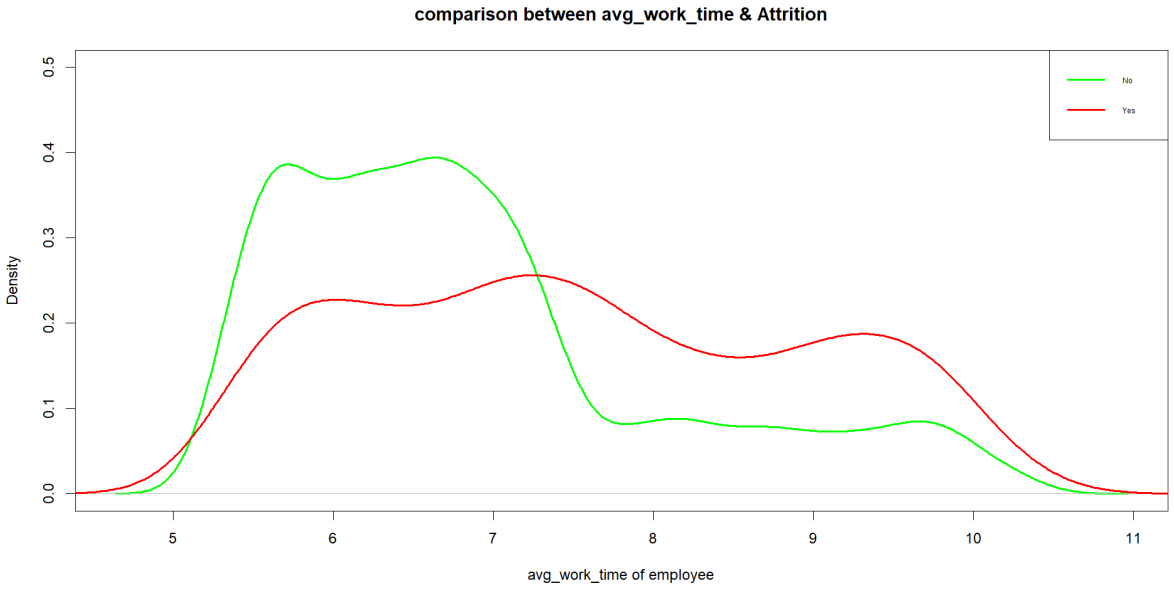
The below curves indicate the values of continuous features of the employees who left the company as well as for those who stayed. The curve for those who left is indicated in red and for those who stayed is indicated in green. The legend indicates the attrition value, that is

1. ‘Yes’ for those who left the company.
2. ‘No’ for those who stayed.



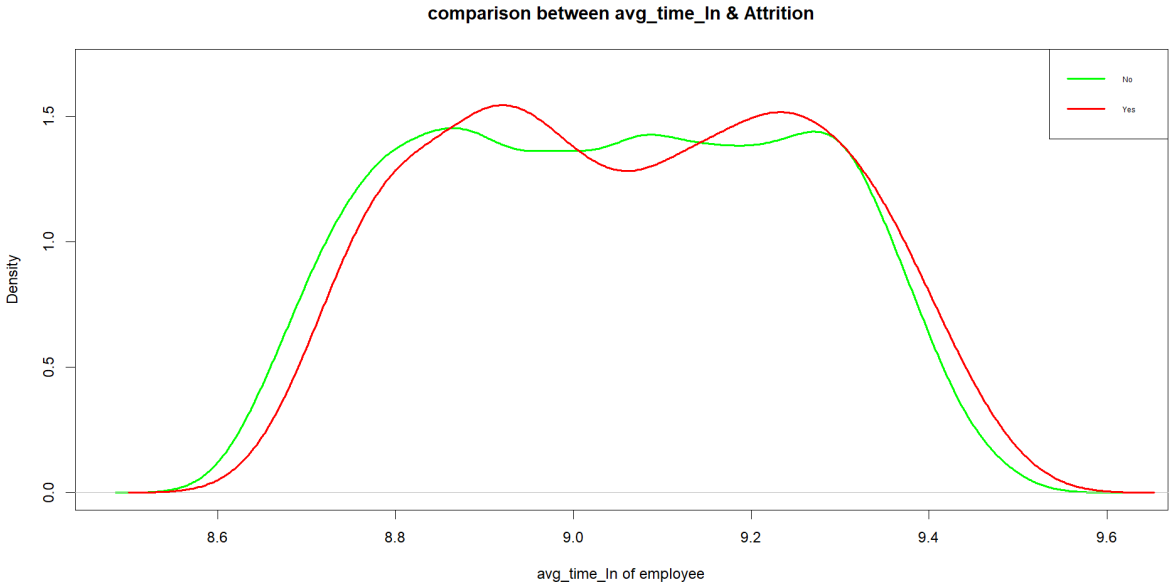
Graph 4.2.1

Graph 4.2.1 indicate the curves for the employees with respect to the employee's age.



Graph 4.2.2

Graph 4.2.2 indicates the curves for the employees with respect to the employee's average daily work time.



Graph 4.2.3

Graph 4.2.3 indicates the curves for the employees with respect to the employee's average daily log-in time.

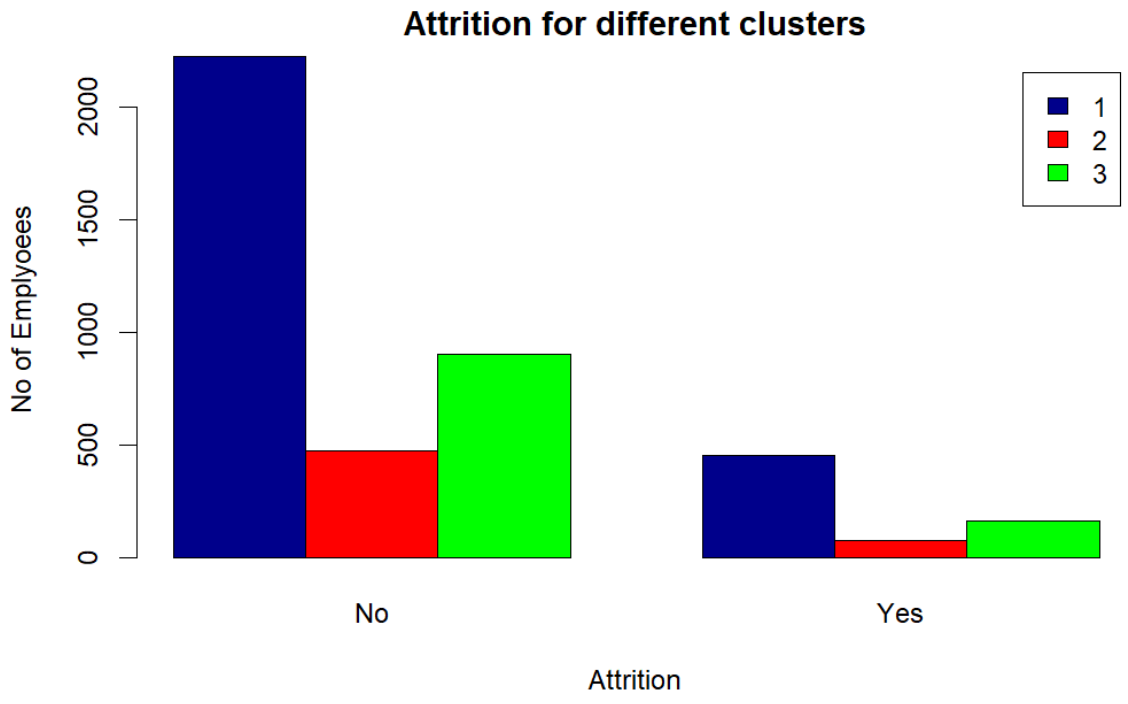
Result from Clustering

Three cohorts is detected from clustering. Interestingly, not much of a difference is observed in the attrition rate for these three different cohorts.

Table 3.4.1: Attrition rate for different clusters

Cluster Number	Attrition Rate
1	16.95
2	13.74
3	15.43

From table 3.4.1, we can observe that the attrition rate for different clusters is not significantly different from the baseline attrition rate of the company. Hence no conclusion can be drawn.



Graph 4.3.1

From graph 4.3.1, we can observe the count of employees who belong to the different clusters who left the company and who have not.

Final recommendations & conclusion

1. Increasing the number of leaves given to an employee might help in reducing the attrition rate for the company.
2. A counterintuitive suggestion is: delaying promotion can also help in reducing the attrition rate.
3. Another counterintuitive suggestion is: not to give large percentages of salary hikes.
4. Since employees who travel frequently are more likely to leave the job, it is suggested that only those employees should be preferred for business traveling who spent a lot of time in this company. As the employees who have given a lot of years to this company are less likely to leave the job.
5. Workload should be properly distributed among the employees. Employees who spent more than average daily time in the company are more likely to leave the job.
6. A minimum number of college-level educated employees should be hired since they contribute to a higher attrition rate.
7. Also the company should prefer to have a minimum number of HRs to reduce the attrition rate.