Midterm Data Analytics Project

**Satyam Singh**

**G01389368**

# **Introduction**

Data collected by any industry is always useful and can help companies a long way. Here, we see data collected by a US company who has data about its employees. This data includes information important to find out the company’s employee attrition rate. To find out certain patterns and trends in attrition you need specific attributes which can help the executives or analysts identify the main reasons behind employees leaving the company.

We have been provided a dataset of 1100 rows and 13 columns. Out of these 13 columns, 7 are categorical and 6 are numeric. The attributes involved are salary, miles from home, job satisfaction, weekly hours, job level, years at company and many more.

We will try to identify the main factors or the key attributes which show some concrete facts about the employees leaving the company, Based on these insights, a company can then change its policy.

## **Data and Model**

We are going to analyse a few variables by finding their relationship with the dependent variable, that is, ‘left company’. Left company is a factor with 2 outcomes, ‘yes’ and ‘no’.

### **Is there a relation between the company’s departments and people leaving.?**

**Summary**

Graphical user interface, application

Description automatically generated

**Plot**

Chart, bar chart

Description automatically generated

**Analysis**

We can see that sales department has the greatest number of employees leaving the company. After Sales , we have product development . The company executives can take a look inside the working of the department as approximately 30% employees who have left are from sales department.

### **Is there any relation between job level and employees leaving the company.?**

**Summary**

Graphical user interface, text, application

Description automatically generated

**Plots**

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

**Analysis**

We can conclude that, executives must investigate the sales department and for the job position, we can spot that managers and associates are the employees who leave the most.

### **How does weekly hours of an employee affect attrition rate?**

**Summary**

Graphical user interface, text, application

Description automatically generated

**Plot**

Chart, scatter chart

Description automatically generated

Chart, histogram

Description automatically generated

**Analysis**

As you can clearly see, weekly hours going above approximately 50 or 55 is a big red flag for employees. This amount of workload is not taken positively. Also, on the Y-Axis, I have plotted the salary, and we can observe that employees below the salary of $100000 leave the company if their workload is greater than 55. In the histogram we can clearly see the employees who leave have a workload over 50.

### **How does company loyalty affect the attrition?**

**Summary**

A screenshot of a computer

Description automatically generated

**Plot**

Chart, histogram

Description automatically generated

**Analysis**

This plot is quite interesting as it gives us many insights. Employees leaving the company after completing 1 year is higher than in the first year. This may be because employees do not fully understand the work culture etc. So, there is a higher chance for employees to leave the company after completing their first year. Also, as years increase, there is a decrease in the trend for employees to leave. Towards the end it is almost negligible.

### **Does number of previous companies help in determining if an employee will leave?**

**Summary**

Graphical user interface, application

Description automatically generated

**Plots**

Chart, histogram

Description automatically generatedChart, box and whisker chart

Description automatically generated

Chart, scatter chart

Description automatically generated

**Analysis**

Previous companies is also an insightful attribute which tells us about the employees behaviour. The histogram shows us how the number of employees who leave the company increases as number of previous companies. Also, in the boxplot, it clearly shows a difference in group. Among all employees who have left the company, 75% of them have worked for more than 4 companies. This means that employees who have worked for more than 4 companies have a higher chance to leave. In the third graph I have compared salary and previous companies in a scatter plot and coloured them based on job. We can observe that the bottom right of the graph which is low salary and more no of companies have higher blue dots signifying job left. This means that employees with low salary and more no of previous companies tend to leave.

### **Is miles from home also an important factor?**

**Summary**

Graphical user interface, application

Description automatically generated

**Plots**

Chart, bar chart, histogram

Description automatically generated

**Analysis**

This plot and summary explains that miles from home is not a strong identifier as there is not much difference in the values. The difference between the average miles from home for people leaving and not leaving the company is approximately 1. Also, in the distribution we can see, more employees have left the company who are closer to their workplace.

### **How does Marital status contribute to Employee attrition?**

Summary

Graphical user interface, application

Description automatically generated

Plot

Chart, bar chart

Description automatically generated

Analysis

Marital status is also a factor which can help in determining if an employee is going to leave or not. From the summary you can see only 62 employees, that is 14% of the employees who are married leave. That means 86% of the employees who are married do not leave the company. This makes sense as married people look for stability and a job would provide that.

## **Conclusion**

After deep analysis on all the attributes of the company’s employees, we can confirm that there are a few factors that can help the executives understand employee attrition in their company. The Factors which seem to be the most promising according to me are salary, department, job level, weekly hours, years at company, previous companies, and marital status. All of these attributes explain the employees behaviour and can be helpful in prediction as well. Also, with analysis, the executives can pinpoint reasons why employees are leaving. For example, salary and weekly hours when plot together can explain that salary less than $100000 and weekly hours more than 50 hours can increase the chances of an employee to leave.