

# HR ANALYTICS CASE STUDY

---

**SUBMITTED BY:**

ABHIRAG

ADITYA

RATHAN

VAISAKH

# BUSINESS PROBLEM

---

- XYZ company is a large company with 4000 employees at any point of time, However, every year, around 15% of its employees leave the company and need to be replaced 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
- Hence, the management has contracted an HR analytics firm to understand what factors they should focus on, in order to curb attrition. In other words, they want to know what changes they should make to their workplace, in order to get most of their employees to stay. Also, they want to know which of these variables is most important and needs to be addressed right away.
- Goal of this case study is to find the factors affecting the attrition rate of the employee using Logistic Regression and present them in business form so that the company can take appropriate steps to retain the employees

# REMOVING OUTLIERS

Below are the plots after removing outliers

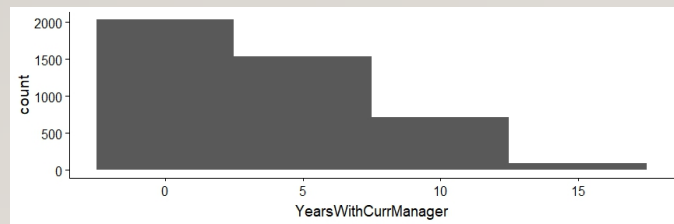
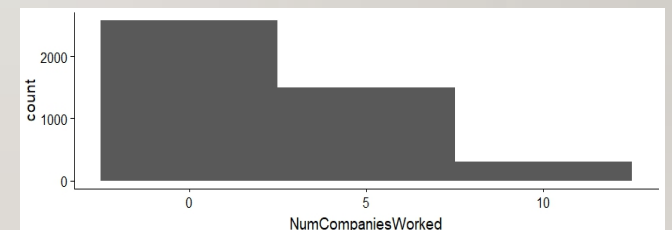
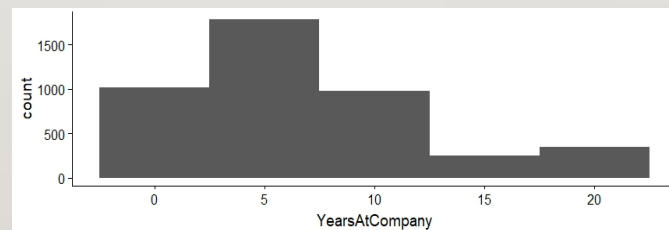
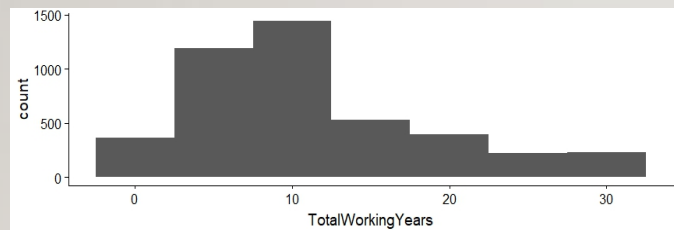
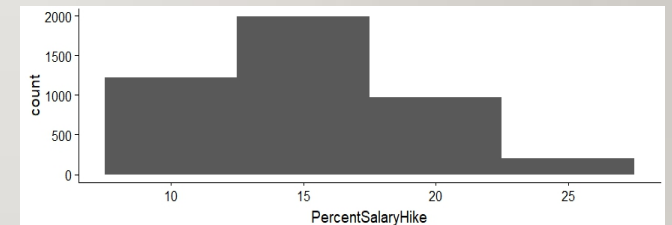
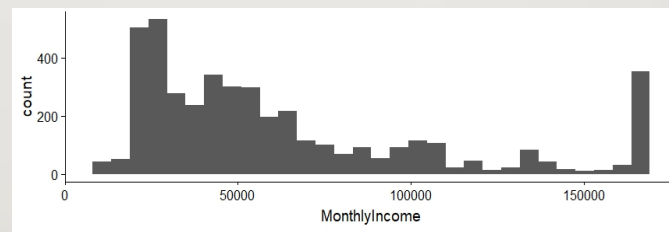
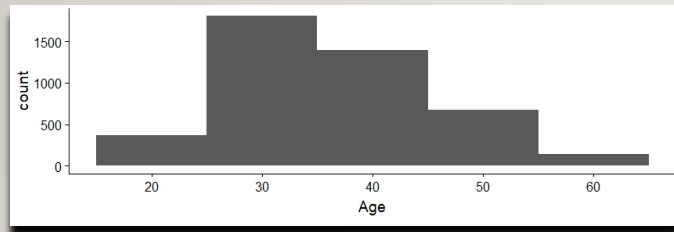
Outliers found in the columns monthly income, total working years, years at company, years since last promotion ,years with current manager.

It is seen that there is a large volume if people working extended hours who tend to leave the company.

People who have a tendency of changing jobs frequently appear to leave the company within short periods of time.This can be found out from Job tenure field.

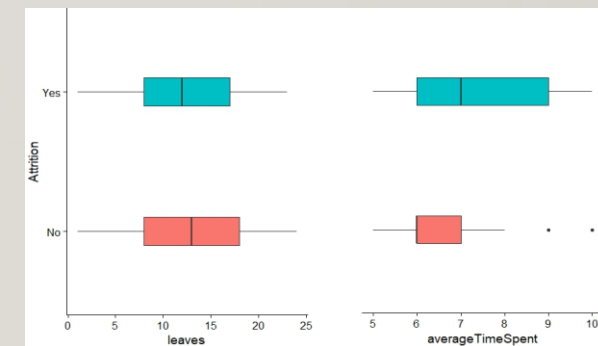
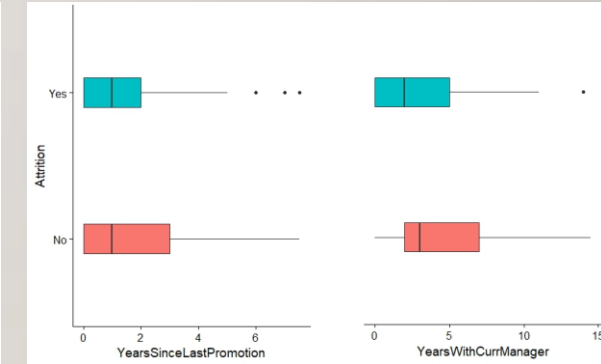
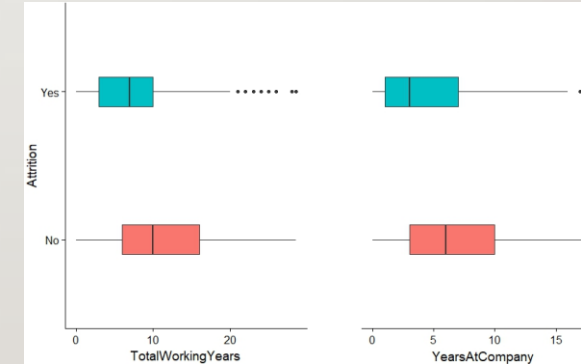
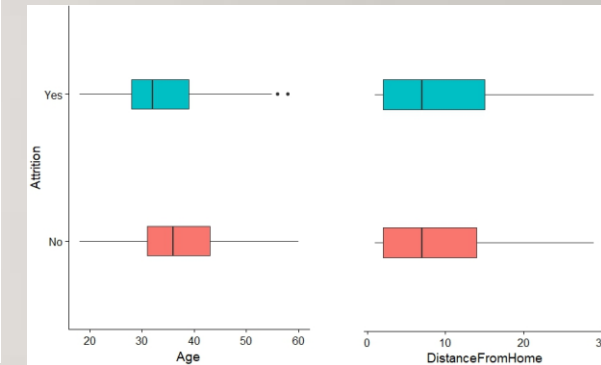
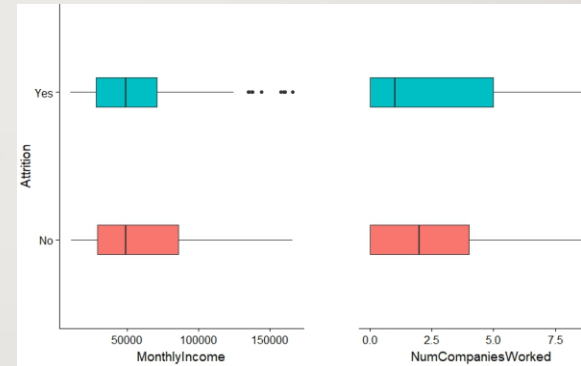
This clearly shows the trend as the number of years per job is less for those who quit

---



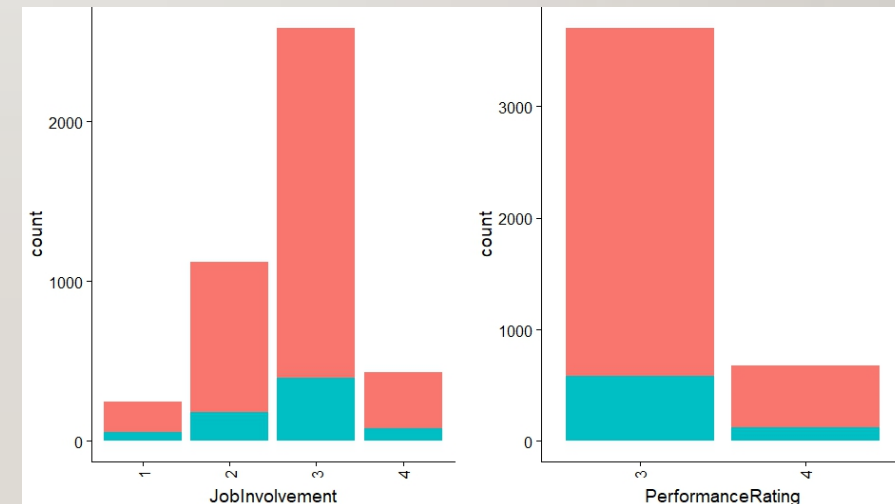
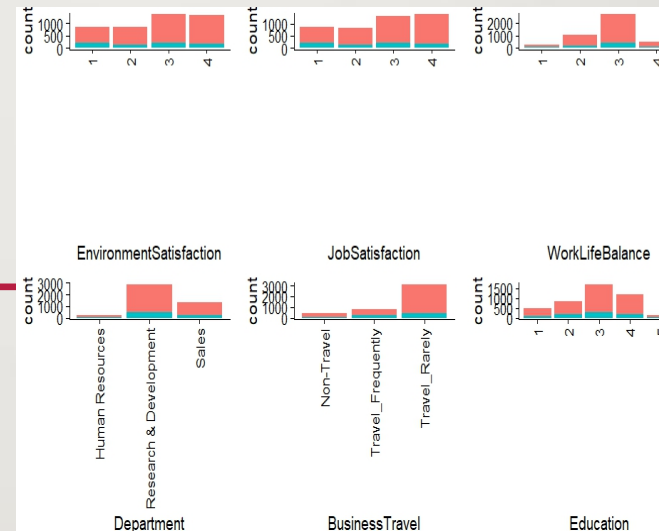
# OBSERVATIONS FOUND USING BOX PLOTS

- People who resigns from a company comes from the lower Age bracket.
- The 50-75 percentile of the person who resigns have to travel more distance from home.
- Monthly income for the persons who resigns is in less bracket than who stays with the company and the number of company worked bracket is more for the person who resigned.
- Less experienced people resign more and person who have spent less years with the company resigned more.
- Promotion bracket for the person who leaves the company is less than person who stays and as expected the years with current manager is also for the person who leaves the company.
- Interestingly that the average time spent by the persons who leaves the company is more than the person who stays with the company



# EDA FOR CATEGORICAL VARIABLES

- First plot will be explaining moderate contrast of attrition rate with Environment satisfaction and the Business Travel and Work Life Balance.
- Below plot reveals strong contrast of attrition rate with Education Field Job Level Job Role and Marital Status
- Below plot reveals strong contrast of attrition rate with Job Involvement





# MODELLING

---

- Below steps are followed for Modelling

Scaling of all the continuous variables.

Converting categorical variables into dummies so that they can be converted into numeric variables.

Total Attrition Rate comes out to be 16.08 percent of total.

Splitting data into training and test data set.

Prepare Model using Logistic Regression and removing the variables having high VIF and low significance.

# SIGNIFICANT VARIABLES AND THEIR RELATION

---

Overall there are 13 significant variables with positive and negative coefficients, positive coefficient means they will be positively more the value of the variable more is the chance of attrition and negative coefficient means they will be affecting attrition rate negatively.

1. NumCompaniesWorked	0.03360575
2. YearsSinceLastPromotion	-0.03068946
3. YearsWithCurrManager	-0.1563666
4. Age	-0.15621
5. TotalWorkingYears	-0.1693355
6. BusinessTravel.xTF	0.1110897
7. JobRole.xMDir	-0.04374527
8. MaritalStatus.xSingle	0.1720183
9. EnvironmentSatisfaction.x2	-0.014466
10. EnvironmentSatisfaction.x3	-0.04393481
11. EnvironmentSatisfaction.x4	-0.04844704
12. JobSatisfaction.x4	-0.08616976
13. extended_working_hrs	0.2295856





# MODEL EVALUATION

- Generate the confusion matrix with different cut-off levels and then compare the sensitivity specificity and accuracy of the model

Cut-Off Probability	Sensitivity	Specificity	Accuracy
40%	0.44976	0.95853	0.8764
30%	0.55024	0.90968	0.8516
25%	0.6316	0.8719	0.8331
Optimal (15.36364%)	0.7272727	0.7336406	0.7326121

# CONCLUSION

---

- When the Number of companies worked by a person is more they tend to quit more.
- If an employee works with the same manager for a longer period of time the lesser are the chances that employee will leave the company.
- As the age increases the people are less tend to quit, which shows they more comfortable with the work environment.
- People with more years of experience are less likely to leave the company.
- Environment Satisfaction, Job Satisfaction are some of the main features that need to be taken into consideration for retaining employees.
- The more an employee works extended work hours on an average the more are the chances that he/she will leave the company.
- Employees who are unmarried are prone to leaving the company.