

HR ANALYTICS CASE STUDY

Group Members:

1. Surabhi Varshney
2. Sekhar Sahu
3. Prasasthy K.B
4. Anargha Biswas

Abstract

The HR Analytics case study is being carried for the company who wants to act upon the cause of attrition in the company. Attrition in any company, be it because of the employees leaving voluntarily or being fired harms the companies reputation and the issues must be addressed in a right way.

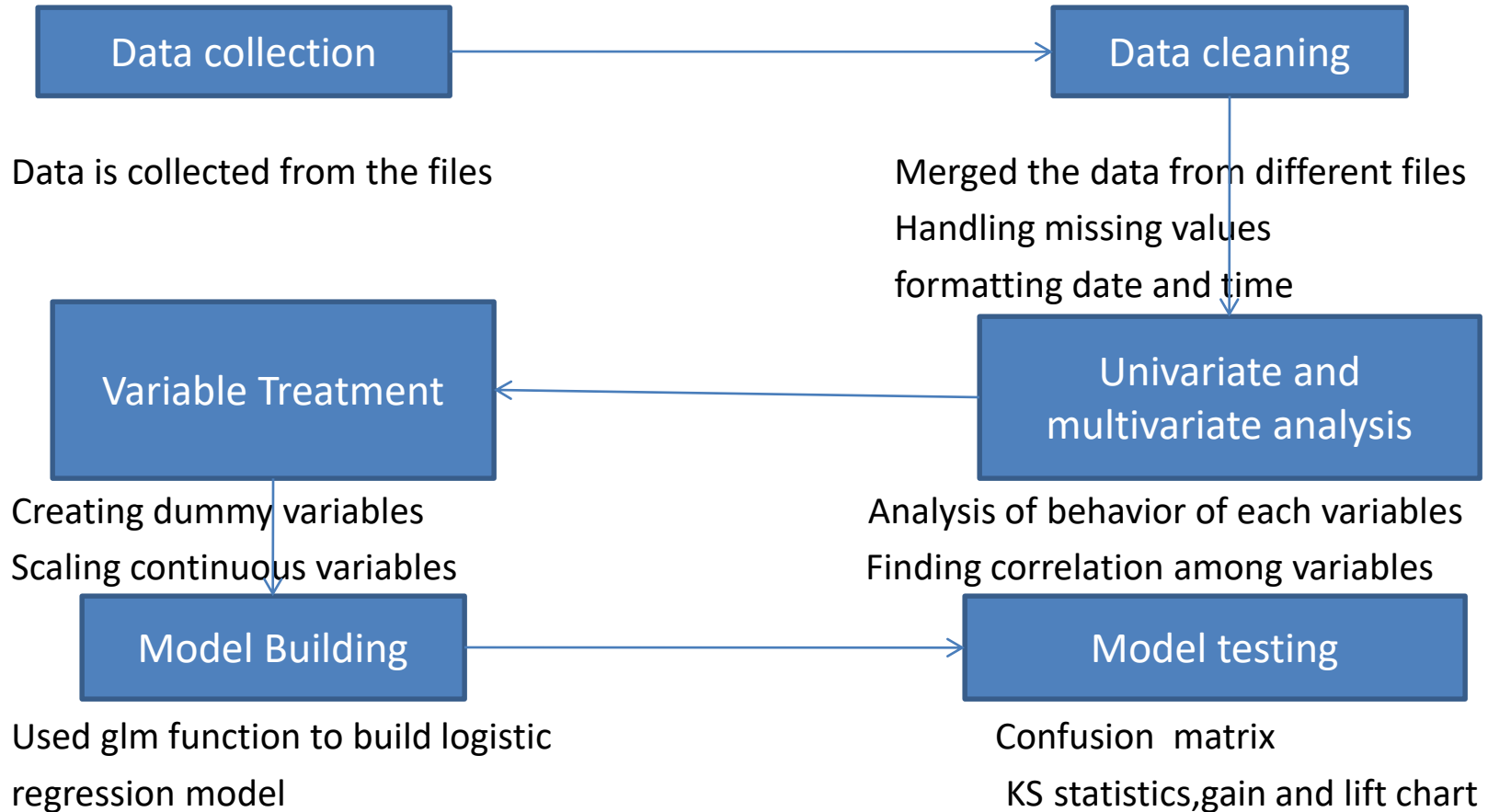
Business Objective :

- Find the factors affecting the attrition rate
- Perform univariate and multivariate analysis of variables of interest
- Express the findings in terms of neat visualizations
- Suggest methods of minimizing the cause of attrition of being a defaulter thereby taking necessary actions.

A glimpse at the data..

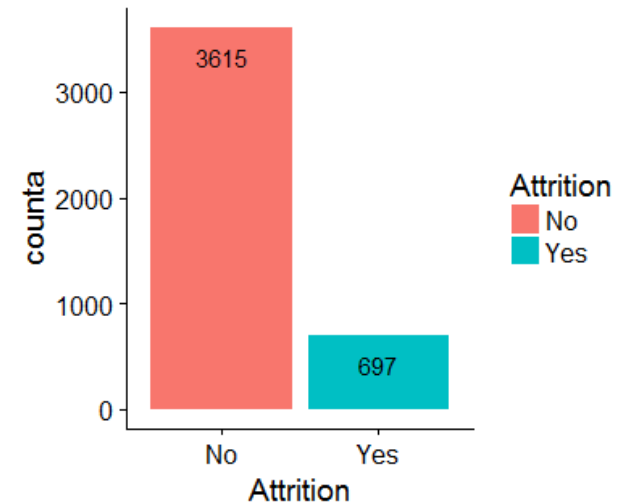
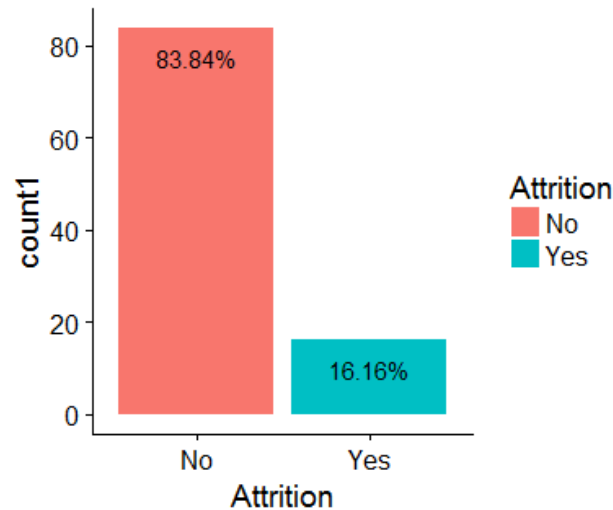
- **Data has been provided in 5 categories -**
- Employee survey data - which provide information on employees emotional well being
- Manager survey data – which tells about the performance of the employee under a manager
- In and out time – which gives the time recorded for employee reporting at office and time at which the employee leaves the office
- General data – This is the master data which contains other details of he employee

Problem solving Methodology



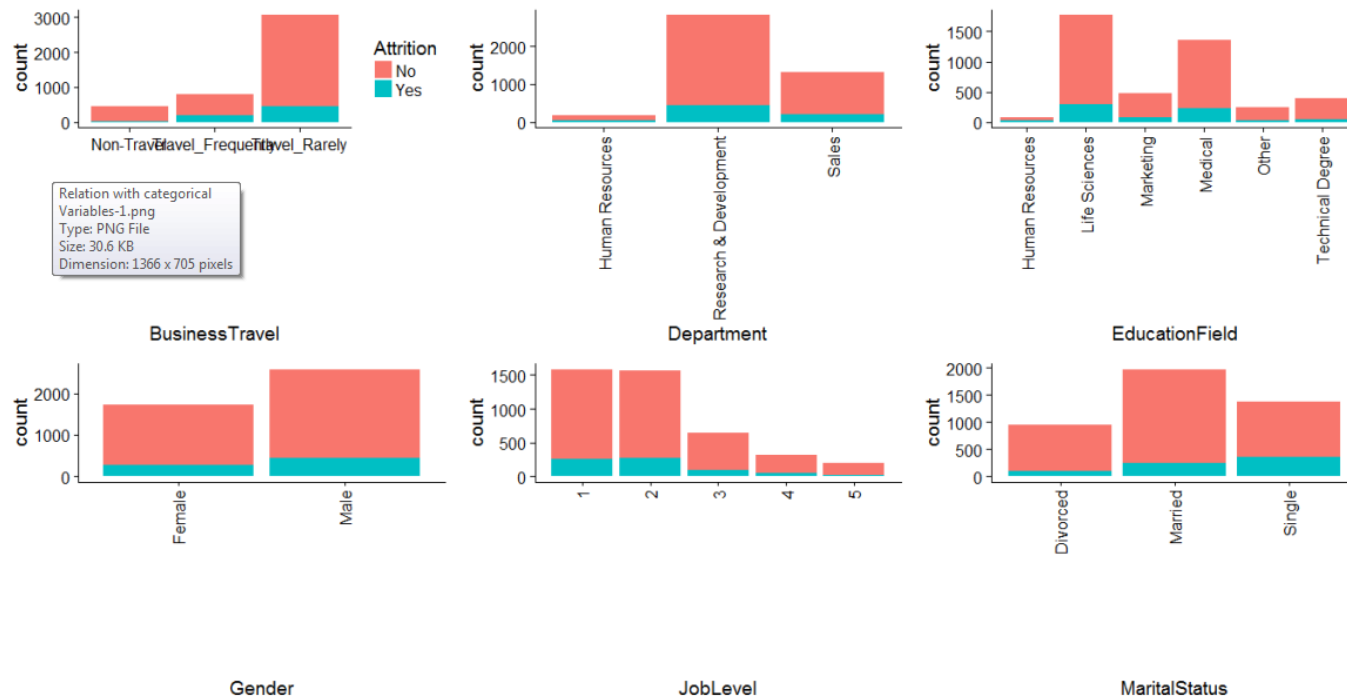
Initial Understanding of data

By analyzing the data set provided we found below stats of attrition
About 16.16% of employees left the company in the year 2015

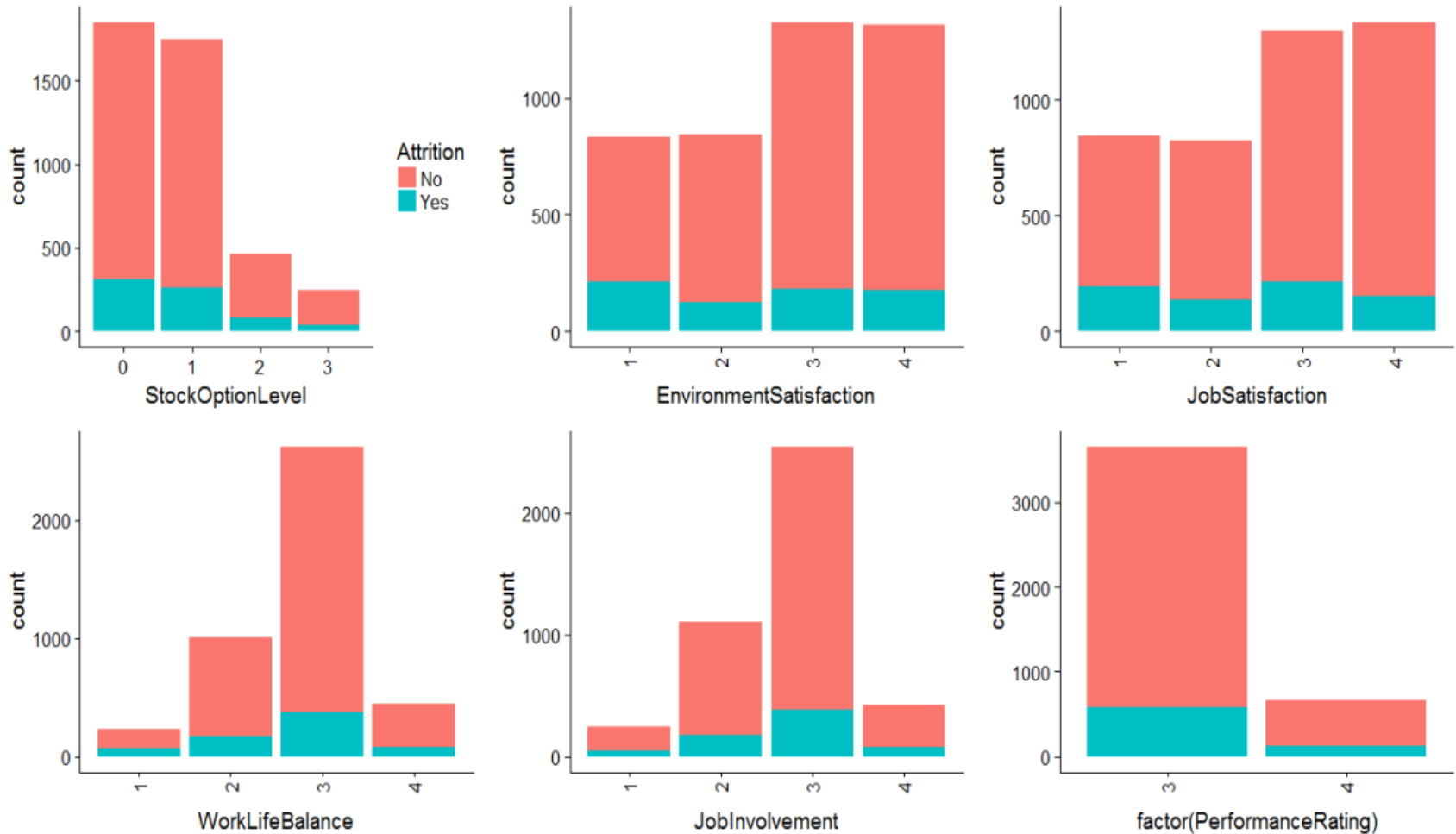


Univariate Data Analysis

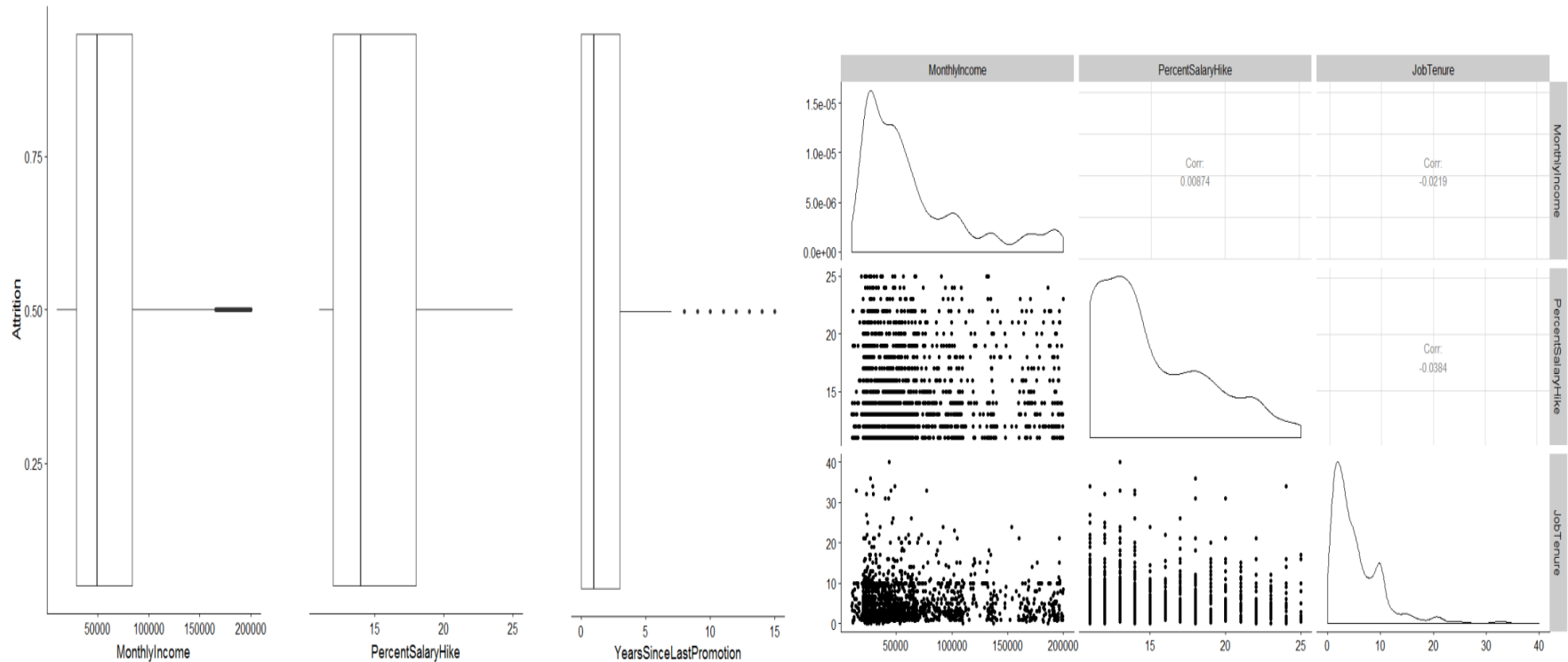
We found the proportion of employees who have left the company against those who are still in the company for each category. Here is the visualization of all categorical variables against attrition.



Univariate Data Analysis(contd..)



Analyzing Continuous variable

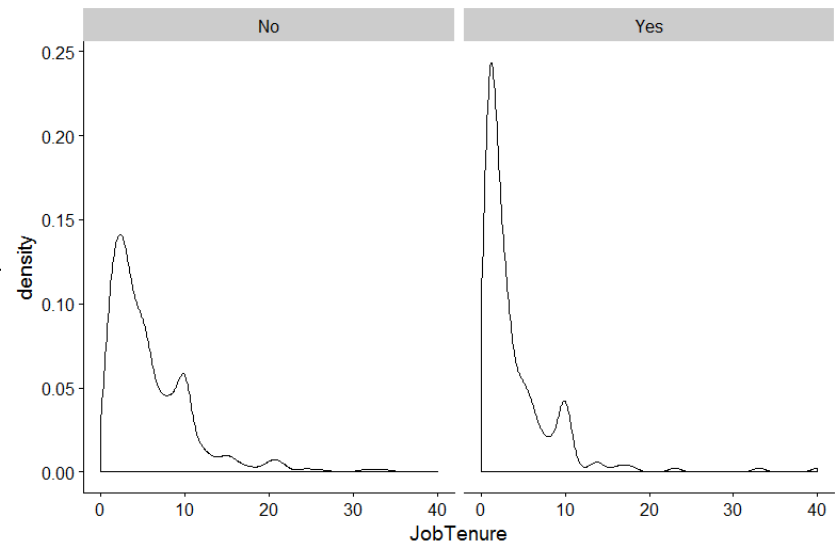
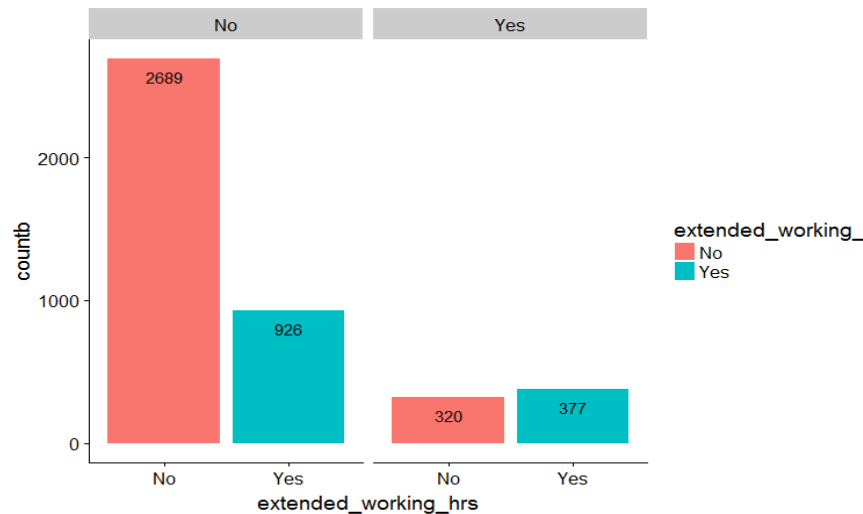


Plot showing the correlation coefficients of different continuous variables with respect to attrition factor.

Analyzing Continuous variable (contd..)

There is a relatively higher number of people working in extended working hours, amongst those who left the company.

Also it clearly shows the trend that as the work experience of employees increases the attrition rate decreases. i.e employees with lesser work experience switch more frequently.



Model Building

Pre-Modeling Stage

- In this stage, we prepared the input variables which were to be given to the model
- We normalized the continuous variables. We scaled them using scale function in R
- We found the nature of the target variable to decide the model to be used –
- Here our target variable is Attrition which is categorical in nature and hence we used logistic regression model
- Attrition rate found in our dataset is 16% only
- We created dummies for categorical variable as part of variable reduction technique
- We split the data into train and test data in ratio 7:3

Modeling Stage

- We used generalized linear model to build our model to predict the attrition
- We used step AIC to get the starting model. Here we are predicting how Attrition behaves with respect to other variables. Since step AIC suggests the optimal model based on Akaike Information Criteria (model having lowest AIC value is preferred), there are some insignificant variables included in the model
- We used vif (Variable Inflation Factor) to check the correlation between the variables
- Finally we used p-values to remove insignificant variables.

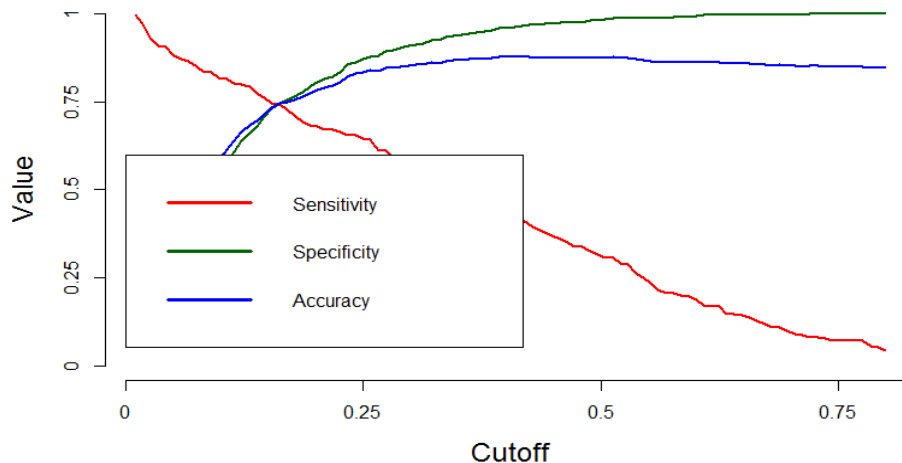
Final model

coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-2.20950	0.18999	-11.629	< 2e-16	***
NumCompaniesWorked	0.15122	0.02454	6.163	7.16e-10	***
YearsSinceLastPromotion	0.17325	0.02378	7.285	3.21e-13	***
YearswithCurrManager	-0.13280	0.02400	-5.534	3.13e-08	***
Age	-0.36046	0.07977	-4.519	6.22e-06	***
TotalWorkingYears	-0.52870	0.10655	-4.962	6.98e-07	***
BusinessTravel.xTF	0.66311	0.13290	4.989	6.06e-07	***
JobRole.xMDir	-0.75620	0.21820	-3.466	0.000529	***
MaritalStatus.xSingle	0.89456	0.11466	7.802	6.10e-15	***
Environmentsatisfaction.x2	-0.67864	0.16483	-4.117	3.84e-05	***
Environmentsatisfaction.x3	-0.91374	0.15371	-5.945	2.77e-09	***
Environmentsatisfaction.x4	-1.25831	0.16152	-7.791	6.67e-15	***
Jobsatisfaction.x4	-0.87025	0.13478	-6.457	1.07e-10	***
extended_working_hrs	1.56824	0.11518	13.616	< 2e-16	***

Model Evaluation

- We ran the model against the test dataset
 - We performed the following model validations - 1. Finding Accuracy, Specificity and Sensitivity through Confusion Matrix
 - In order to find a suitable probability cut-off, we checked the Accuracy, sensitivity and specificity for 1% to 80% probability values
 - The optimum cut-off probability is the one where the value of specificity and sensitivity are close to each other. Here we have taken a safe range of 0.01. Cut-off probability was found to be ~ 0.18
- At Cut-off 0.18



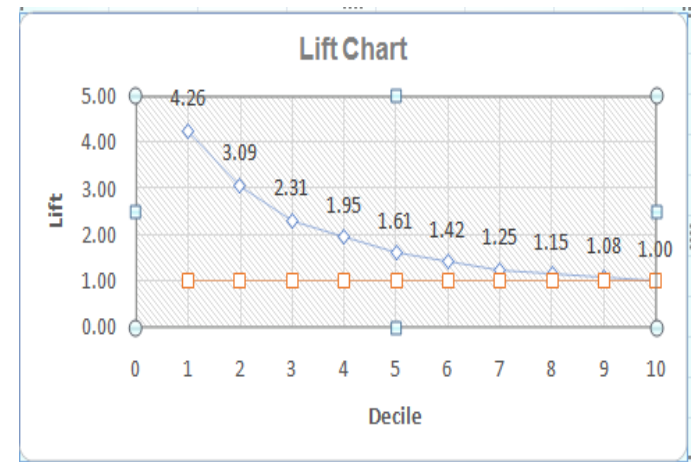
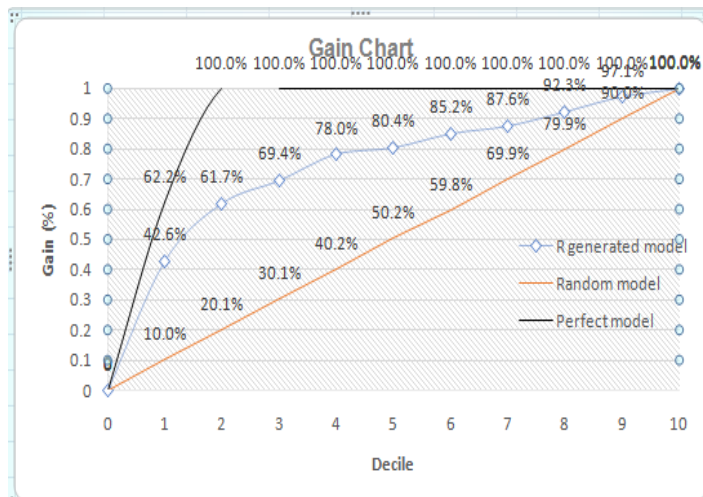
Predictor	No	Yes
No	809	54
Yes	276	155

Model Evaluation(contd..)

Calculating KS Statistics

- KS statistics measures the degree of separation between positive and negative distribution
- The optimal value of KS statistics for a good model should lie between 40-60 and should be within first 3 deciles
- KS-Statistics for our model is 0.513

Gain and Lift Charts



Conclusion

Factors	Inferences	Suggestion
NumCompaniesWorked	The Number of companies worked by a person is more they tend to quit more	Company should focus more on the individual's aspirations and check if they are inline with the companies aspirations before hiring such employees. They should also ask for a reason for switch from these employees to prevent attrition of them in the current company
YearsSinceLastPromotion	The more years gap between two promotions the more tendency of employee to quit	Company should have some policy which will motivate employee even though they don't get promotions
YearsWithCurrManager	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.	Company should ensure a healthy relation between the employees and the manager. The manager should easily be approachable and should be considerate on the employees goals. This will help the company retain the employees
Age	As the age increases the people are less tend to quit, which shows more comfortable with environment	Company should provide suitable opportunities to these employees and reward the critical ones to retain them for long
TotalWorkingYears	People with more experience as they are less likely to leave the company.	Company should look to utilize the experience of such professionals
BusinessTravel.xTF	Those who travel frequently are likely to quit the company. This may be because they do not like to travel and they are forced to or they have found better opportunities in other places	Company should keep a continuous check on employee's aspirations here and see if they are happy with travelling. They should also keep a tap on what other opportunities employee have in other places and keep pace with the changing workforce dynamics
MaritalStatus.xSing	Those who are single are more likely to leave the company	

Conclusion

Factors	Inferences	Suggestion
EnvironmentSatisfaction	Environment Satisfaction, Job Satisfaction are some of the main features that need to be taken for retaining Employees	Company should ensure conducive atmosphere to work where the employee are motivated. Employees should be given rewards in terms of money or coupons from time to time. Company should also ensure that the workspace is clean and safe for the employees to enjoy working
JobSatisfaction	Those who are having a medium, high and very high job satisfaction are likely to stay with company	Company should have a mechanism in place where the employees can share the feedback of the work they are doing. Based on this they should work on the negative ones and promote the positive ones
extended_working_hrs	The more an employee works extended work hours on an average the more are the chances that he/she will leave the company.	Company should ensure that work in the team are divided uniformly and there should not be pressure of a single employee. Knowledge sharing should be increased within the team so that there is no pressure on team members during absence of others

Thank You