

HR ANALYTICS CASE STUDY

SUBMISSION

Group Member Names:

1. Pradeep Meriga
2. Parthasarathy Vijayan
3. Feta Joseph
4. Rama Kumar Jayanti

Business Understanding

Business Abstract:

XYZ Company , with strength of 4000 employees at any point , is facing a problem with annual attrition of 15%. The management believes that this level of attrition (employees leaving, either their own or because they got fired) is bad for the company.

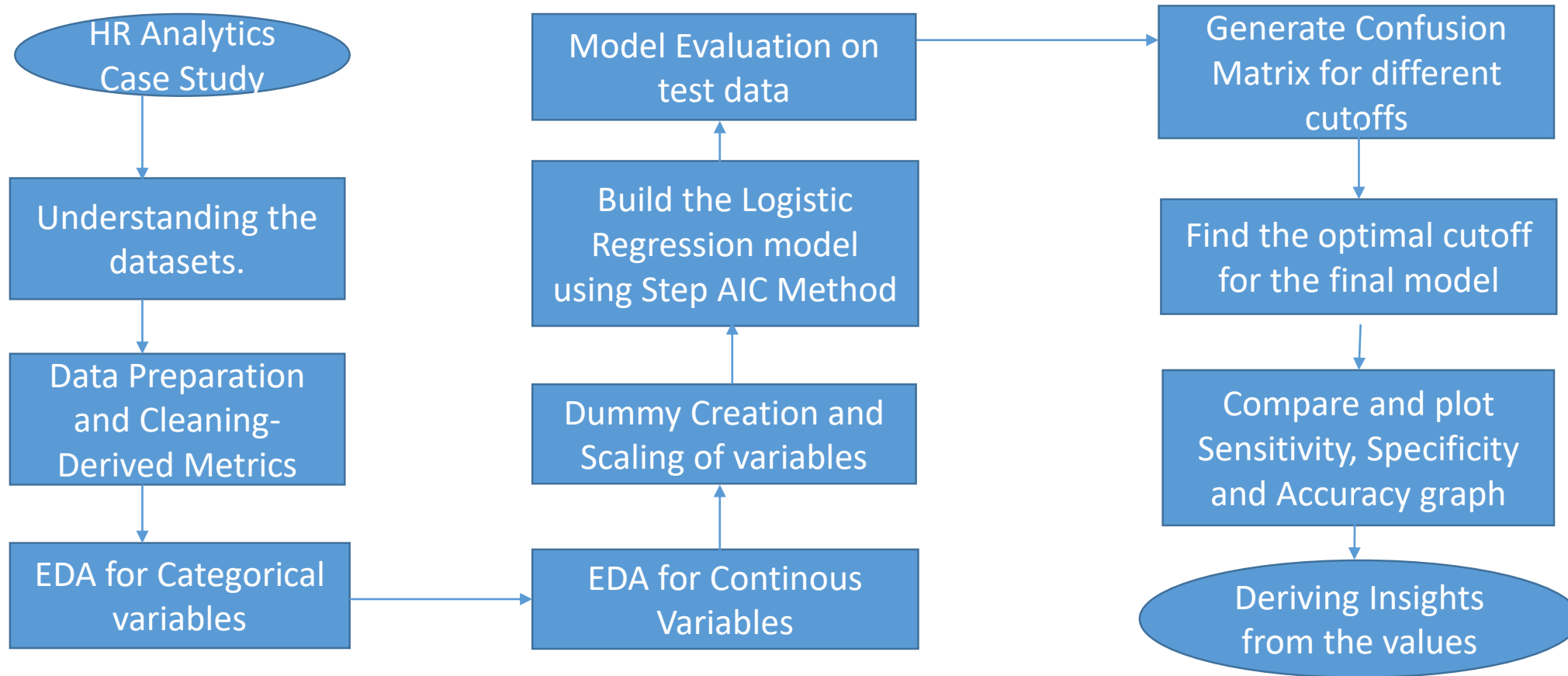
Hence, the management has contracted an HR Analytics firm to understand what factors they should focus on, in order to curb attrition.

Business Objective:

To find the factors which influence the attrition rate of employees using a Logistic Regression and present them to the management so as to take necessary steps to retain the employees.

To Suggest the methods of minimizing the cause of attrition

Problem Solving Methodology



Data Preparation and Cleaning

Data Preparation

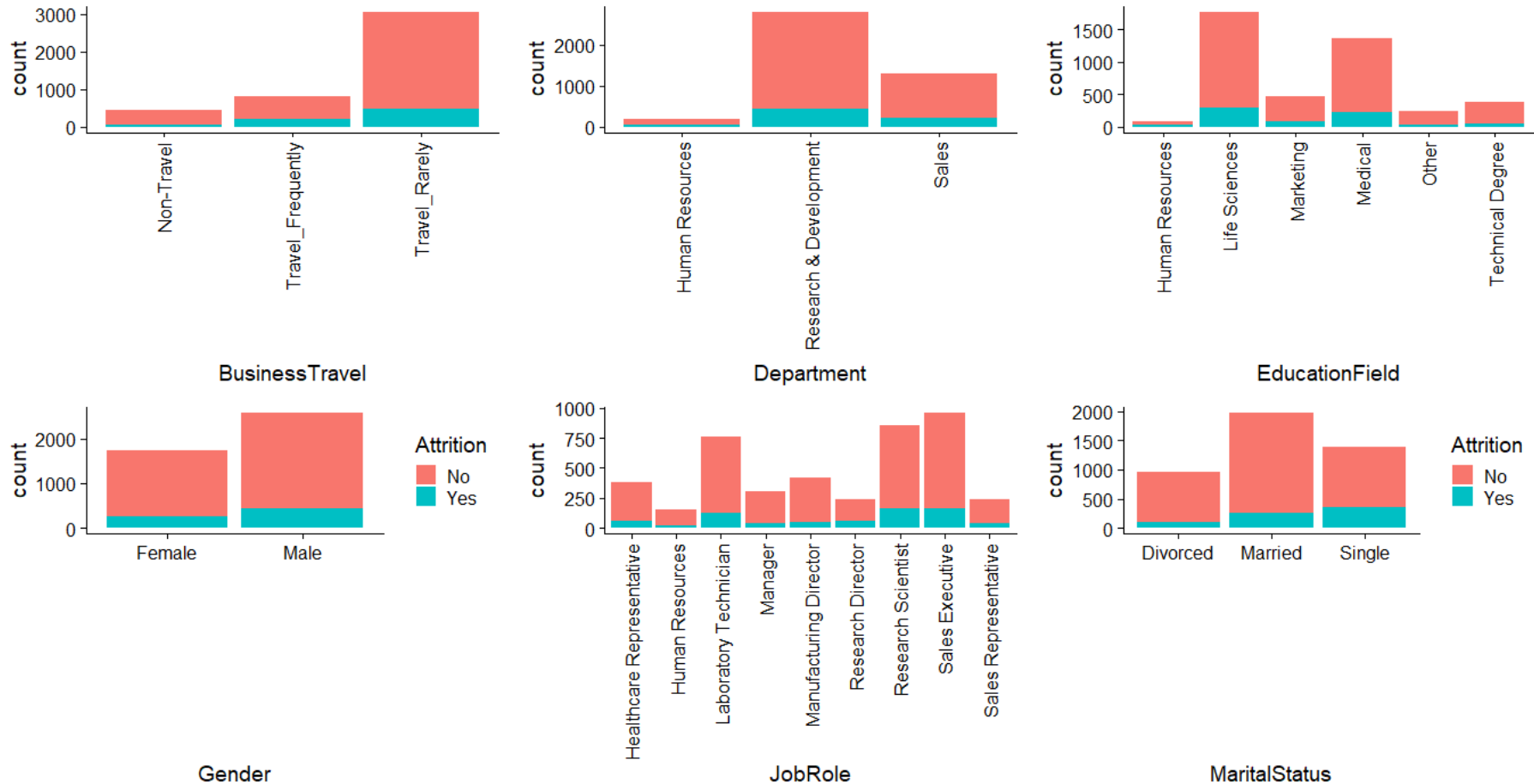
- Merged the files `general_data`, `employee_survey_data`, `manager_survey_data` into the master file `hr_data`.
- Converted `time_in` & `time_out` data to date format and calculated the `work_hours` for each employee and identified the company holidays.
- We Derived the following Metrics : `avg_workhours`, `leave_days`, `no_9hoursMore`, `no_6hoursLess`.

Data Cleaning

- We removed the columns `EmployeeCount`, `Over18`, `StandardHours` as they have the same values for all the rows.
- Imputation of the missing values
- Duplicates are checked
- Removed NA Values from the columns `NumCompaniesWorked` , `TotalworkingYears`, `EnvironmentSatisfaction`, `JobSatisfaction` and `WorkLifeBalance`.
- Treated the Outliers in the data.
- We have Standardised and Derived the following variables – `OptimalDist`, `HighSalary`, `avgYearInComp`, `Attrition`, `Gender`.

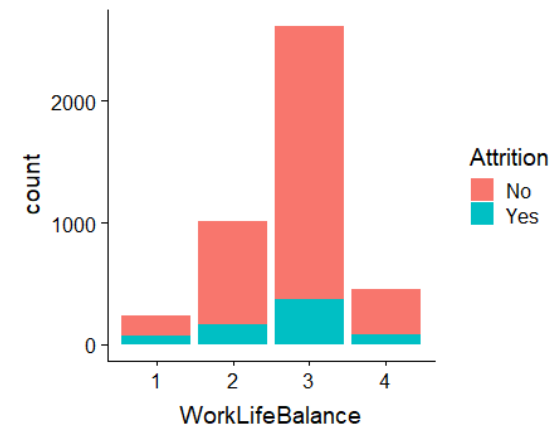
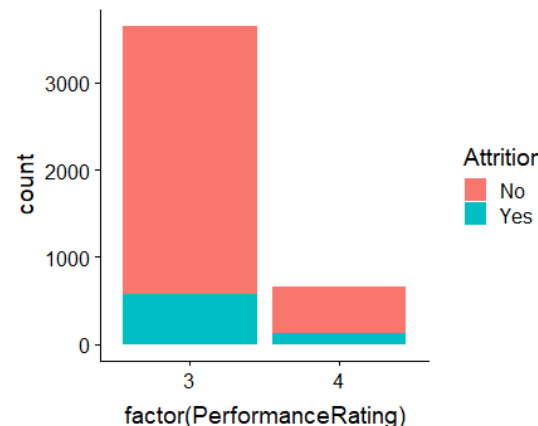
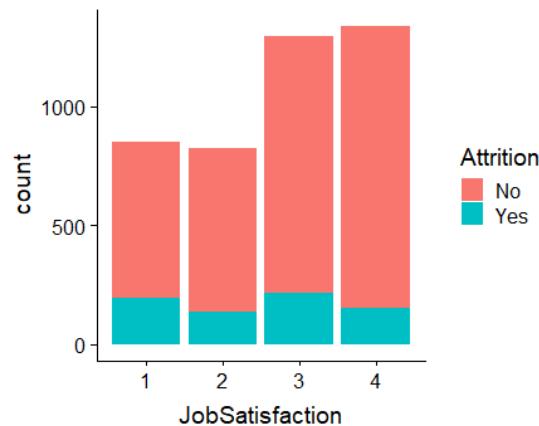
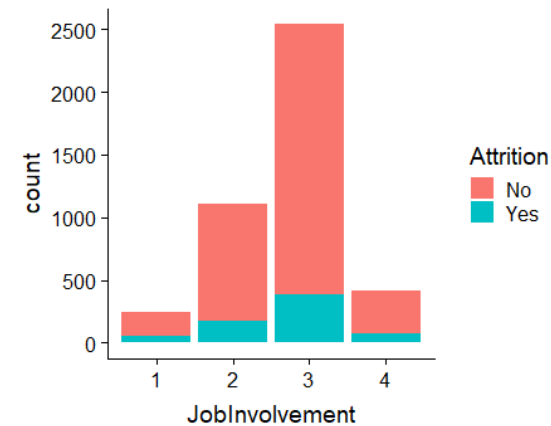
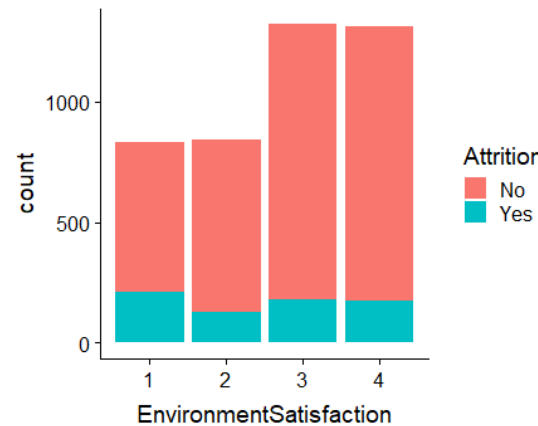
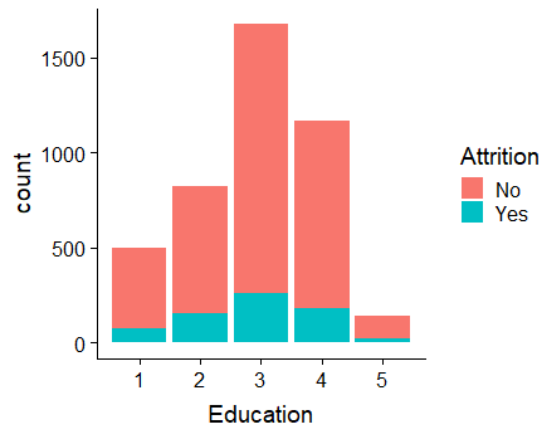
EDA for Categorical variables(nominal)

- Below plot reveals the low contrast of the attrition rate with Business Travel , Department, Education Field, Gender, Job Role and Marital Status.



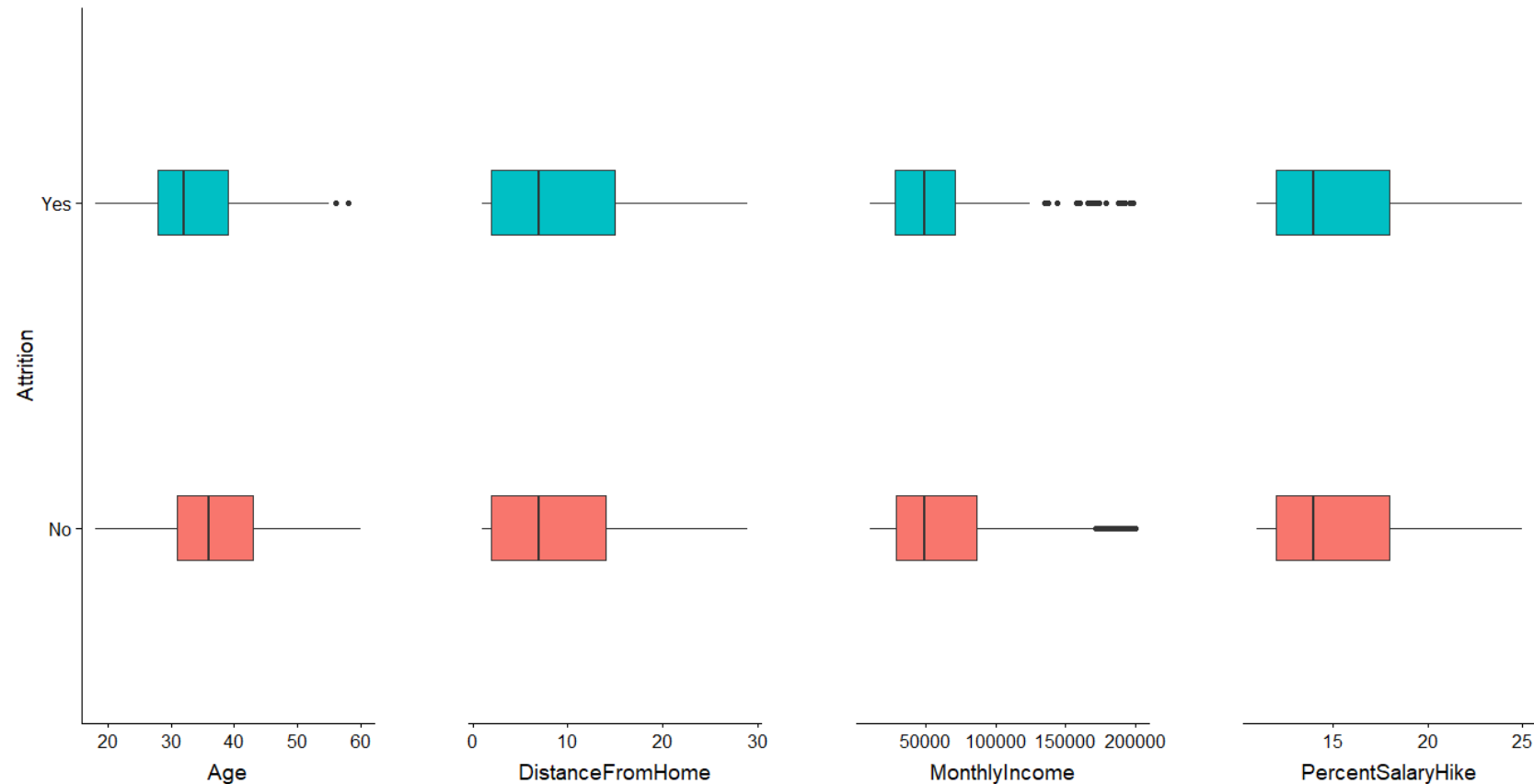
EDA for Categorical Variables(Ordinal)

- Below plot shows the strong contrast of the attrition rate with Education, EnvironmentSatisfaction, JobInvolvement, Job Satisfaction, PerformanceRating and WorkLifeBalance.



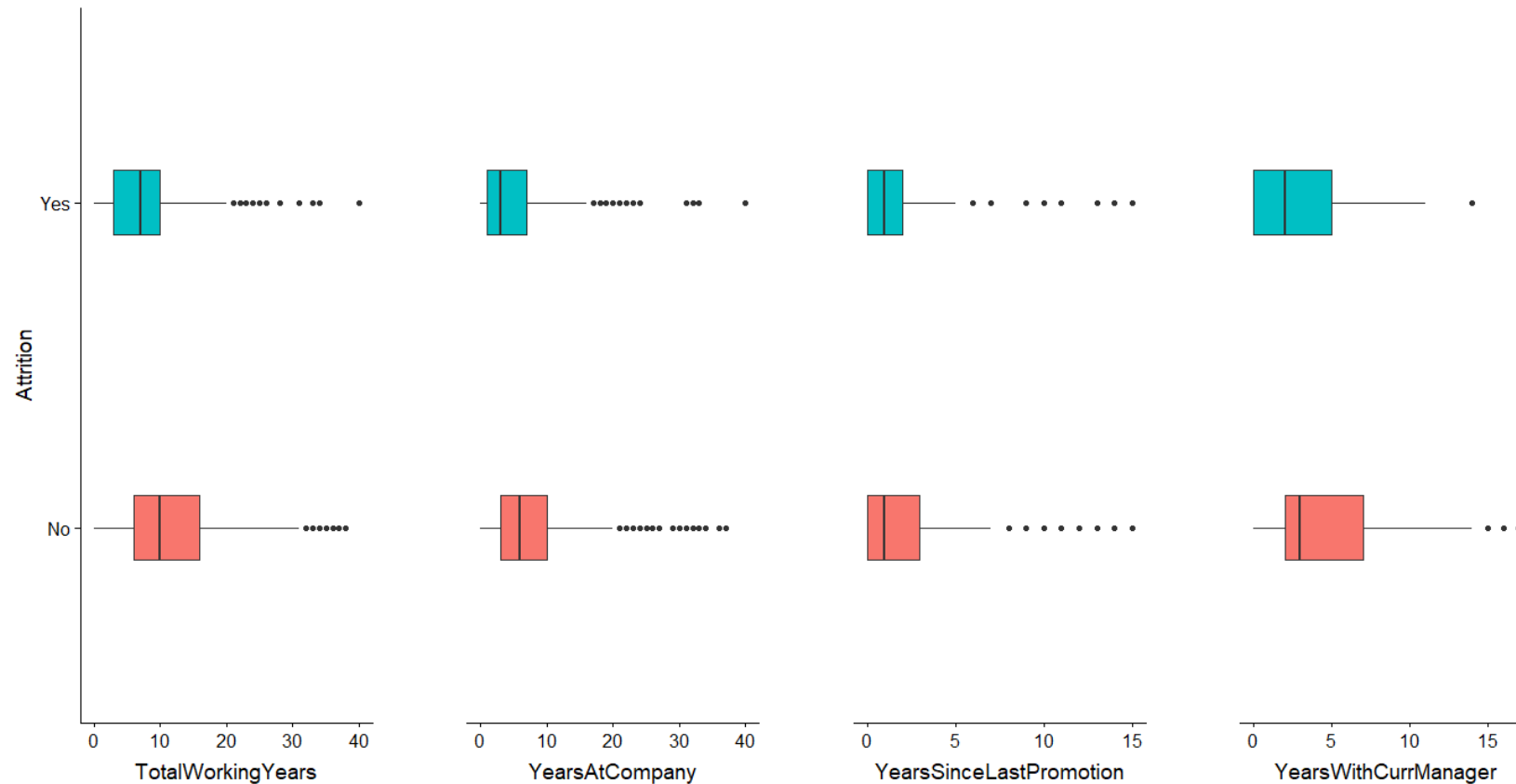
EDA for Continuous Variables

- Employees resigns more from the lower age bracket.
- Employees with lower Age, more DistancefromHome, Less MonthlyIncome and less PercentSalaryHike tends to leave the company.



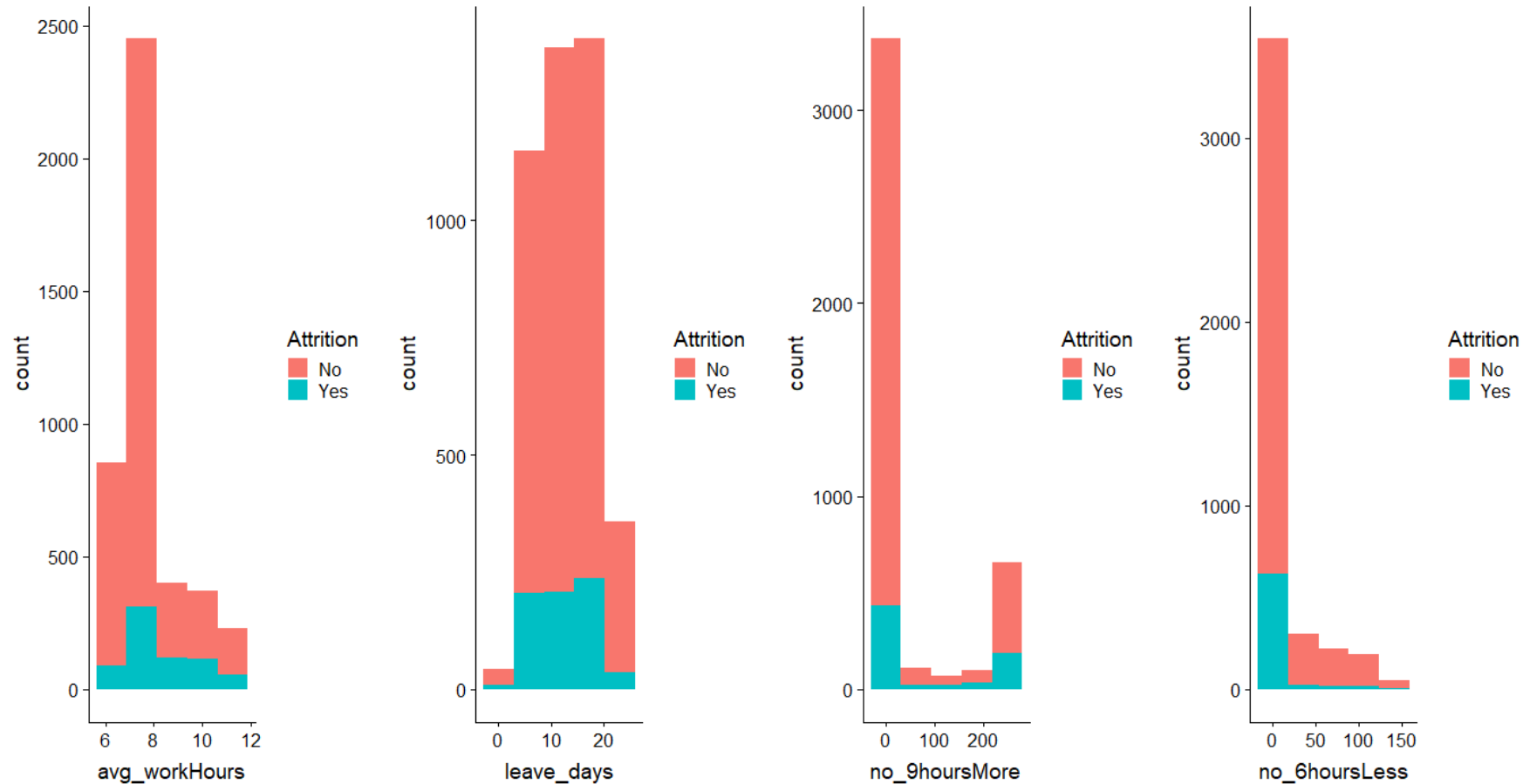
EDA for Continuous Variables

- Employees with less YearsAtCompany and less experienced people resigned more.
- People with less YearsSinceLastPromotion time leaves the company is less.



EDA for Continuous Variables

- Below plot shows the contrast of derived metrics with the attrition.



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 like Business Travel, Department, EducationField, JobRole and Marital Status into dummies. 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
- 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

Modelling Stage

- We used generalized linearmodel to build our model to predict the attrition
- We used stepAIC to get the starting model. Here we are predicting how Attrition behaves with respect to other variables.
- We used VIF(Variable Inflation Factor) to check the correlation between the variables
- Finally we used p-values to remove insignificant variables.

Significant Variables with their coefficient values

Overall , there are 14 variables which influences the employee attrition in the company.

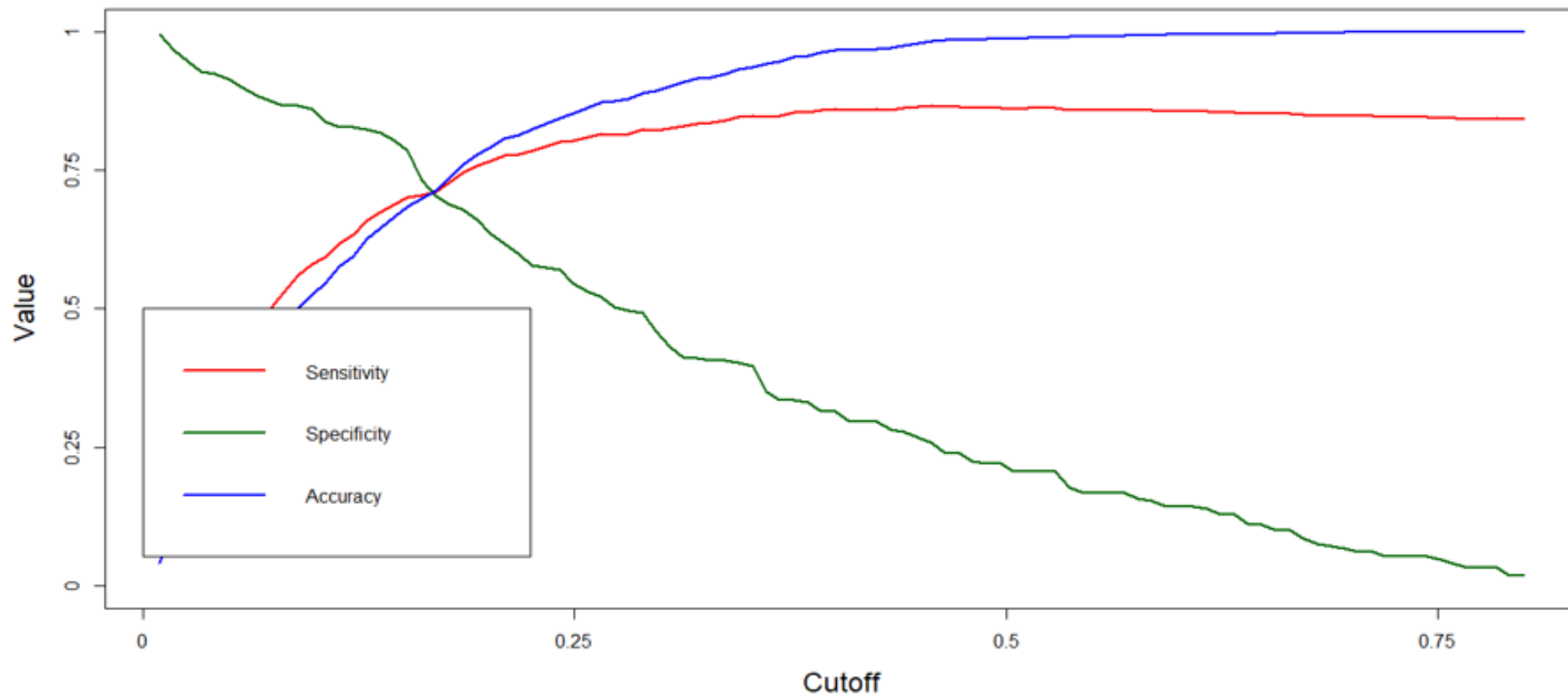
BusinessTravel.xTravel_Rarely	0.004867		JobRole.xManufacturing Director	0.000143
MartialStatus.xSingle	<2E-16		Age	0.000124
NumCompanies Worked	7.04E-10		TotalWorkingYears	5.47E-07
TrainingTimesLastYear	0.000185		YearsSinceLast Promotion	2.23E-10
YearswithCurrManager	2.99E-09		Environment Satisfaction	1.85E-15
JobSatisfaction	5.51E-10		WorkLifeBalance	0.000126
Avg_WorkHours	<2E-16		no_6hoursLess	0.001418

Model Evaluation

- Generating the confusion matrix with different cutoff levels and then comparing the sensitivity , specificity and accuracy of the model.
- Finding the optimal cutoff value and creating the graph showing the optimal cutoff.

Cutoff Probability	Sensitivity	Specificity	Accuracy
50%	0.2392	0.9815	0.8614
45%	0.2679	0.9796	0.8646
Optimal Cutoff (16.96%)	0.7081	0.7285	0.7252

Accuracy , Specificity and Sensitivity



CONCLUSION

From the model, we can conclude that the following factors affects the attrition rate of the employee:

- Employees having 10+ years total experience and earning less than 20000.
- Employees reporting to the same manager for more than 4 years and not getting promoted for 4 or more years duration.
- Employees with total years of experience is more than 20years and having the monthly income less than 80000 and getting percentSalaryHike less than 15%.
- Factors such as EnvironmentSatisfaction , JobSatisfaction and WorkLifeBalance having 2 or less values tends to leave the company more.

Suggestions

With the results thus obtained from our analysis, we recommend the below points to the management to understand what changes they should make to their workplace, in order to get most of their employees to stay.

- Company should provide suitable opportunities to the employees and reward the critical ones to retain them for longer period of time. Company should look to utilize the experience of such professionals.
- Management should keep a continuous check on employees aspirations and see if they are happy with work and should keep tap on what other opportunities employees are looking for with in the company.
- Company should ensure conducive atmosphere to work where the employees are motivated.
- Management should ensure that work in the team is equally distributed among the team members and see that no one is over burden by working extra hours there by reducing work pressure.

Thank You