HR Analytics Case Study

SUBMISSION

BY

Krishnaraj Barvathaya

Abstract

A large company named XYZ with employee strength of around 4000 employees has attrition rate of around 15%. Because of this

- Projects gets delayed and difficult to meet deadlines, resulting in a loss of reputation among consumers and partners.

- A sizeable department has to be maintained, for the purposes of recruiting new talent.

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

As part of this analysis need to identify factors they should focus on in order to curb attrition. Analysis should provide guidance to bring changes to their workplace, in order to get most of their employees to stay. Also, need to identify most important attribute that need to be addressed right away.

Problem solving methodology

- Step 1: Data Understanding
  - Five data sets provided
    - General Data about the employee
    - Employee's in and out time with date in two separate data sets
    - Employee review and Manager review data in separate data set
  - Employee id is the unique key to combine these data sets

- Step 2: Data Preparation
  - Missing values will be imputed based on analysis
  - Outliers will be treated by setting their values to whiskers.
  - Standardizing scales of continues variables.
  - Dummy variables will be created for categorical variables

- Step 3: Design the model and test the model
 - Design the model and train 70% of data and test it on 30% of the data

EDA
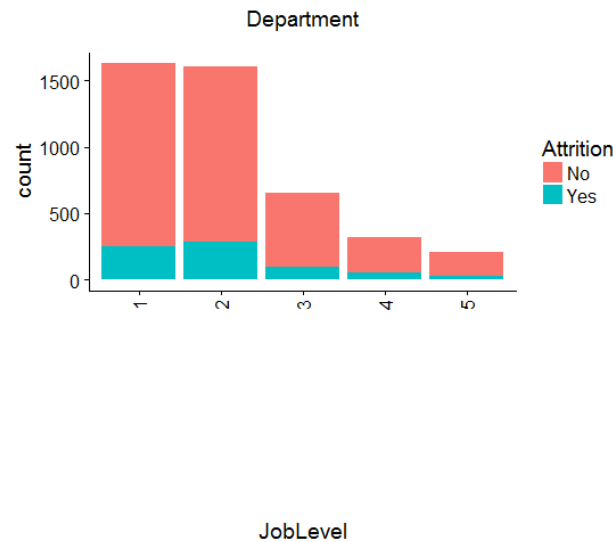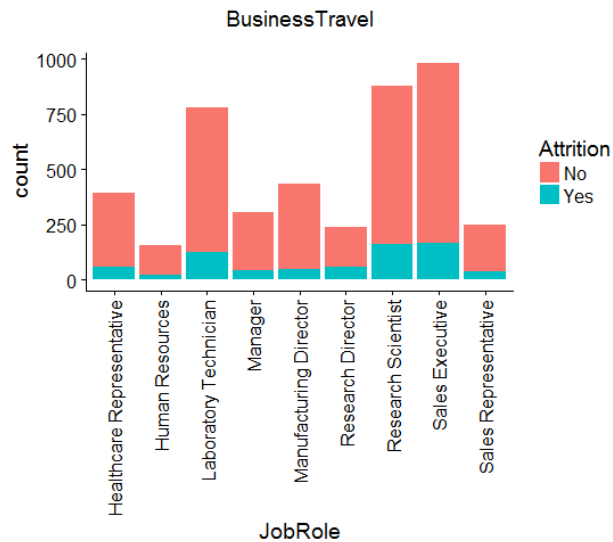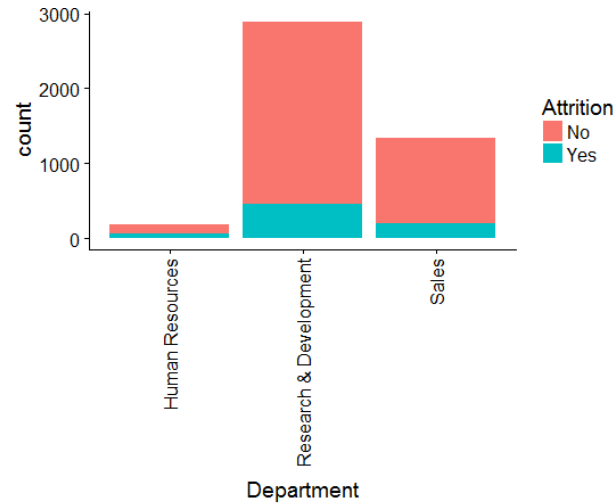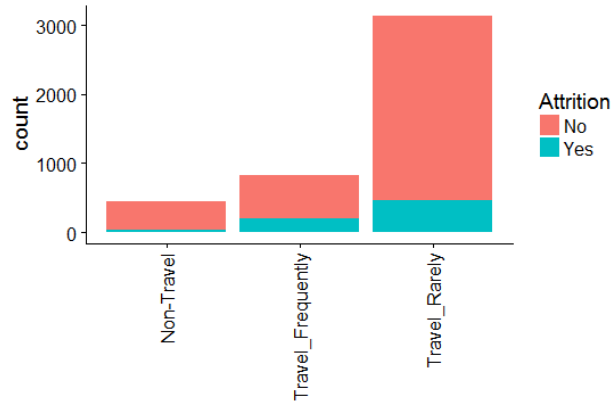
- Observe that an employee who travels rarely are the employee's for which the attrition rate of company will increase, also when we look department wise data it seems that R&D has the highest attrition rate in the company.

- We can observe that job role for an employee is playing vital role where employees attrition rate is high for all types of job roles. So we can say that job satisfaction is directly dependent on the role played by an employee. Hence job role plays a primary key for predicting the results of an employee staying in a company.

- Even we can see the job level factors the attrition rate where job level with 1 and 2 detach from the company more.

EDA
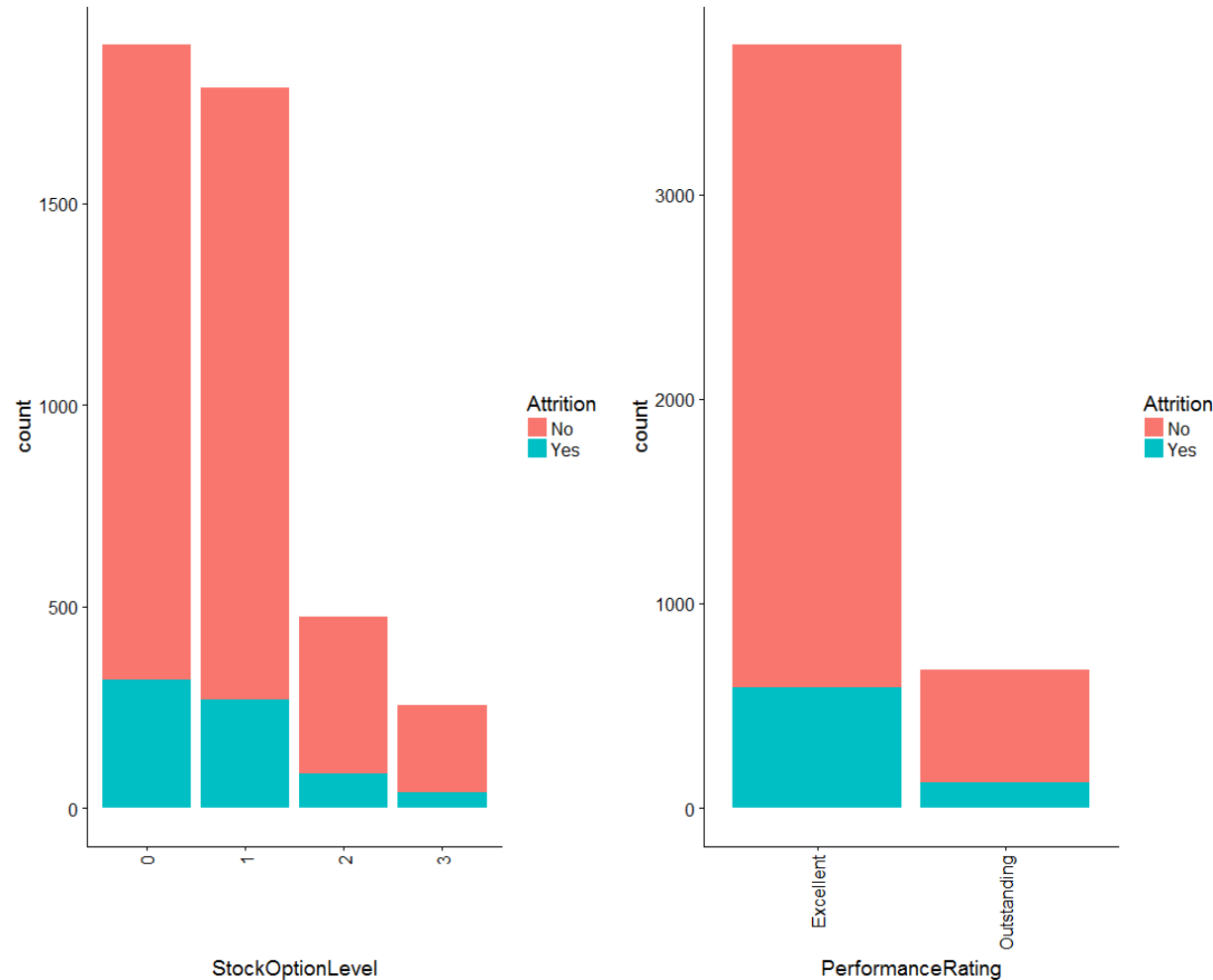
- Providing stocks to employees is also influencing employees to stay longer.

- Also, we can see employees with less rating has high attrition rate and employee's with high rating do not leave company as its common less rating leads to less job satisfaction.

EDA
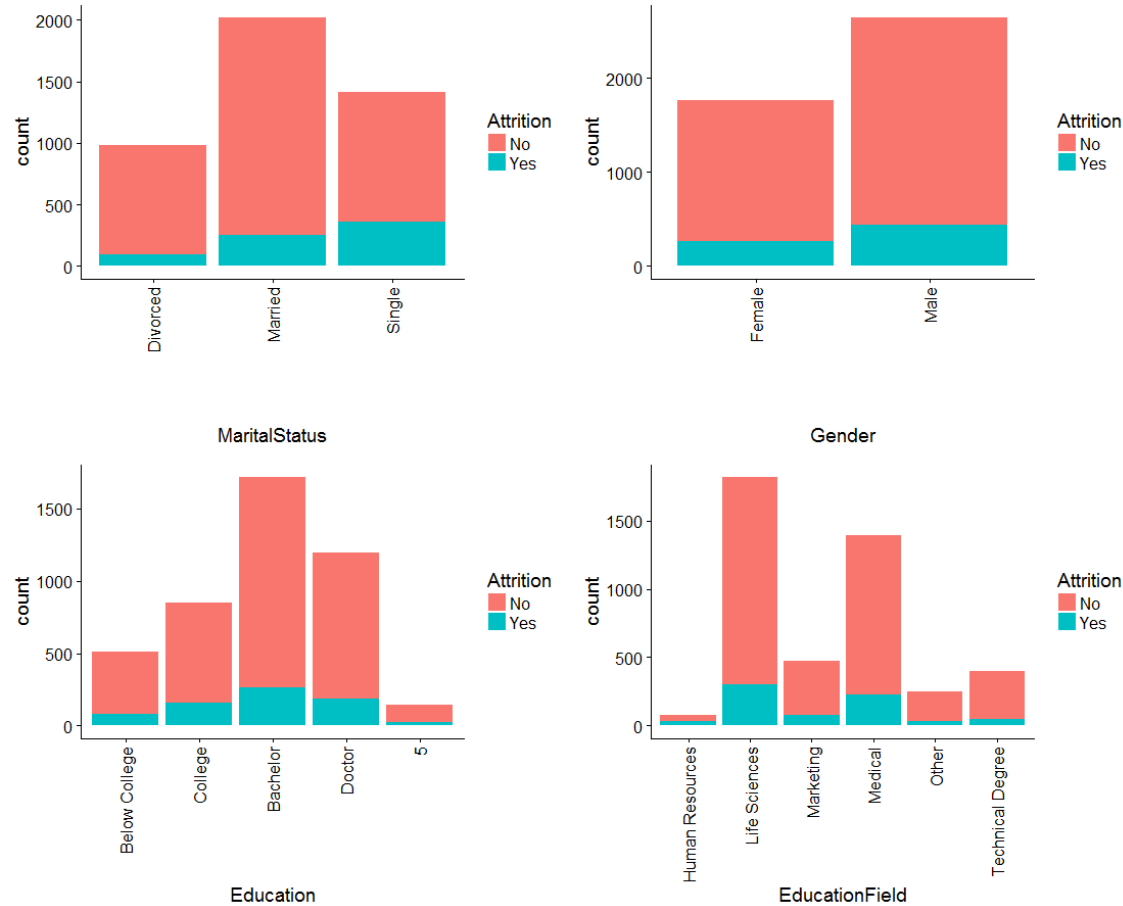
- We can see that most of the single(un married) employees have separated from the company, it quite common as there will be less responsibilities on the employee. Also we can see Job involvement and Stock option do not have too much relation on the attrition rate.

- Comparatively, when we observe Education background of an employee also plays a vital role in predicting whether an employee will quit the company. Also we can see 2 major streams Life Science and Medical streams employees have high attrition rate.

- When we see the gender parameter we can observe that both male and female have nearly equal ratio of quitting the company.

EDA
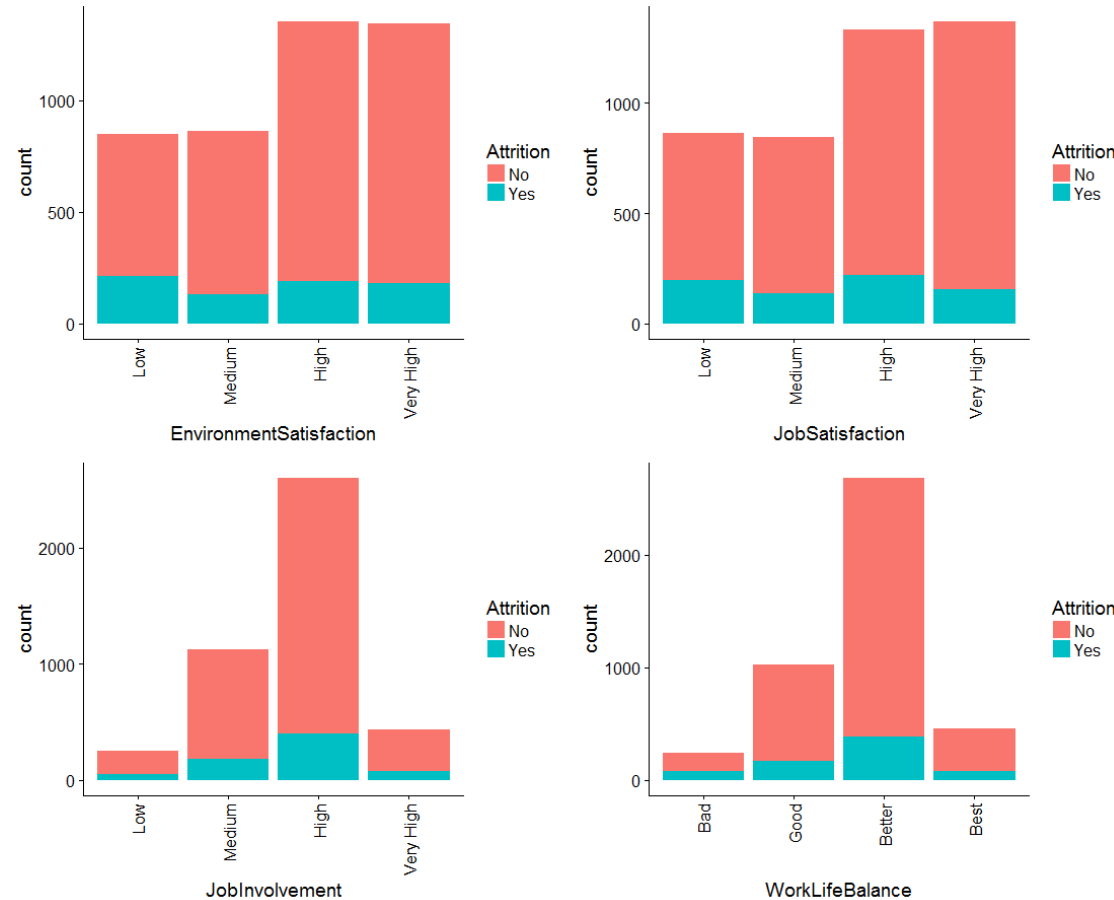
- Considering the attributes like Environment, job satisfaction employee's having less satisfaction quit the company more. We can see from graph most of the employees who have less job and environment quit company.

- If the management looks into the above attributes more then the company could minimize the attrition rate.

Final Model Summary

Deviance Residuals:
  Min   1Q   Median   3Q   Max
-1.6448 -0.5868 -0.3827 -0.2124   3.5528

Coefficients:
  Estimate Std. Error z value Pr(>|z|)
(Intercept)   -1.51313   0.12285 -12.317 < 2e-16 ***
  NumCompaniesWorked   0.32457   0.05513   5.887 3.93e-09 ***
  TotalWorkingYears   -0.67418   0.08261 -8.161 3.33e-16 ***
  YearsSinceLastPromotion   0.40300   0.07178   5.615 1.97e-08 ***
  YearsWithCurrManager   -0.41957   0.08416 -4.985 6.18e-07 ***
  AvgWorkingHour   0.53835   0.05212 10.328 < 2e-16 ***
  BusinessTravel.xTravel_Frequently 0.74103   0.12622   5.871 4.34e-09 ***
  MaritalStatus.xSingle   1.03462   0.11023   9.386 < 2e-16 ***
  EnvironmentSatisfaction.x2   -0.94931   0.16532 -5.742 9.35e-09 ***
  EnvironmentSatisfaction.x3   -0.93577   0.14716 -6.359 2.03e-10 ***
  EnvironmentSatisfaction.x4   -1.20205   0.15216 -7.900 2.80e-15 ***
  JobSatisfaction.x4   -0.76953   0.12886 -5.972 2.35e-09 ***

Model Evaluation
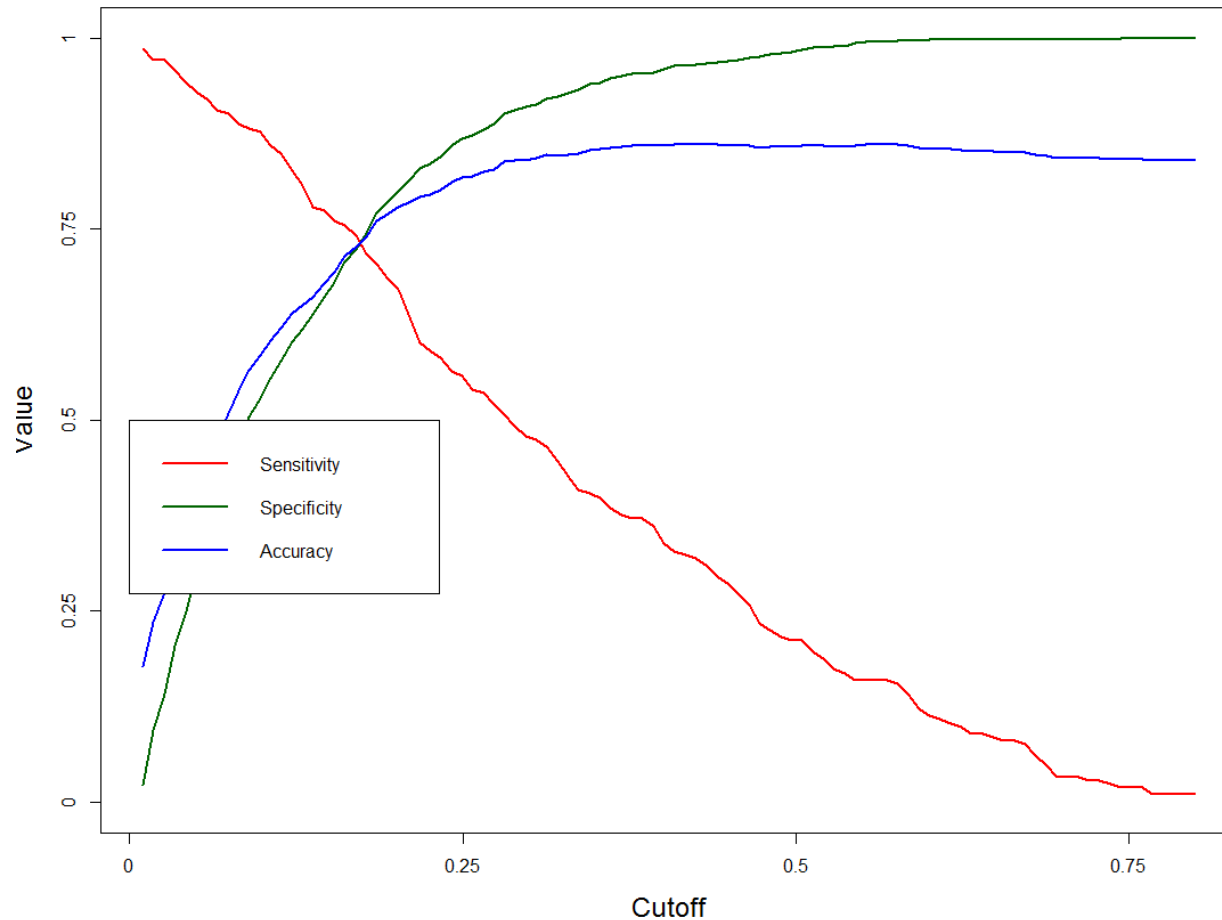
- The reliability of the fitted logistic regression model was assessed using standard goodness of fit tests suitable for logistic regression analysis

- As per calculation fitted model has

  Accuracy of : 85,

  Sensitivity of : 33

  Specificity of : 95

- Model is highly reliable in efficiently classify observations.

Conclusion

Key factors influencing attrition based on model output are
- Number of Companies Worked
- Total Working Years
- Years Since Last Promotion
- Years With Current Manager
- Average working Hours
- Business Travel Frequency
- Marital Status - Single
- Employees having Work Environment Satisfaction Level as Medium, High and Very High
- Employees having very high Job Satisfaction Level

Model also depict "Total Working Years" is the most significant attribute in influencing the attrition.

Organization need to focus on all the above attributes to control the attrition of their employees.