HR Analytics Case Study

Employee Attrition Model

Team Details

- ❑ Vijay Choudhary
- ❑ Aman Kumar
- ❑ Abhinaw Kumar
- ❑ Sanjay Sahoo

Strategy and Business Objectives

Problem Statement

o To understand the employee attrition pattern and identify the key variables on which the organisation should focus to curb the Attrition.

Business Objective

o To build a model to predict the **probability of attrition** using a logistic regression. The results thus obtained will be used by the management to understand what changes they should make to their workplace, in order to get most of their employees to stay.

Business Impact

o To maintain organisation **reputation** among consumers and partners by avoiding unnecessary delay in meeting timelines which usually occurred due to attrition.
o Cost benefit in terms of not maintaining separate **recruitment department** and **training**.

# Data - 4410 unique employee records (Year 2015)

Data

o 4410 unique employee data consist of basic information, Performance review, survey results and login behaviours for a year

o Around 15% attrition rate every year

o There are 5 different files including below no of variables

| Dataset	Records	Columns/Variables
general_data.csv	4410	24
employee_survey_data.csv	4410	4
manager_survey_data.csv	4410	3
In_time.csv	4410	262
Out_time.csv	4410	262

Data Preparation - Variable Correlations

High Correlation variables observed :

- Total Working Years & Age

- YearsAtComany & Total Working Years

- YearsAtComany & YearsWithCurrManager

- YearsSinceLastPromotion & YearsAtComany
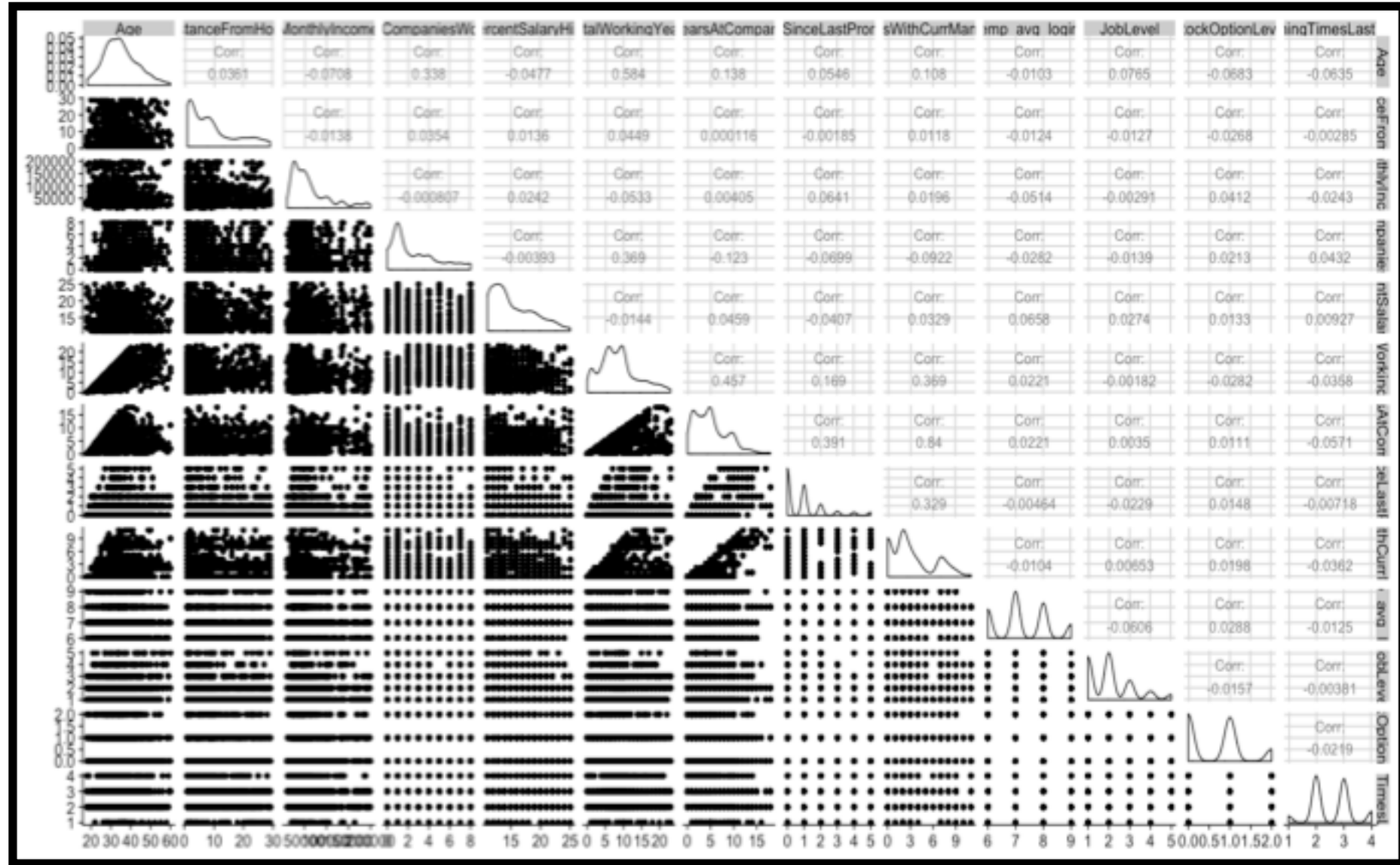
Exploratory Data Analysis - Variable Influence on target variable Attrition
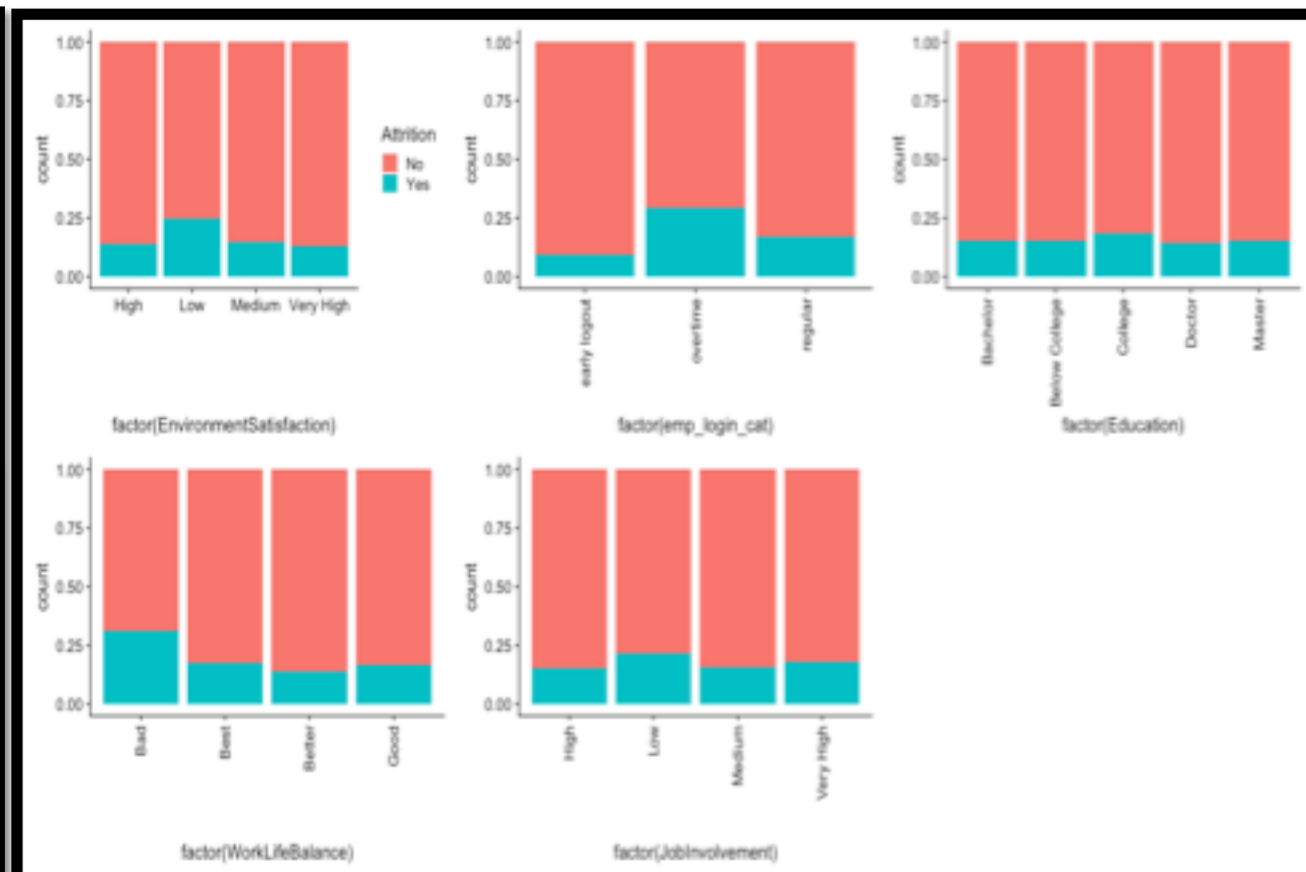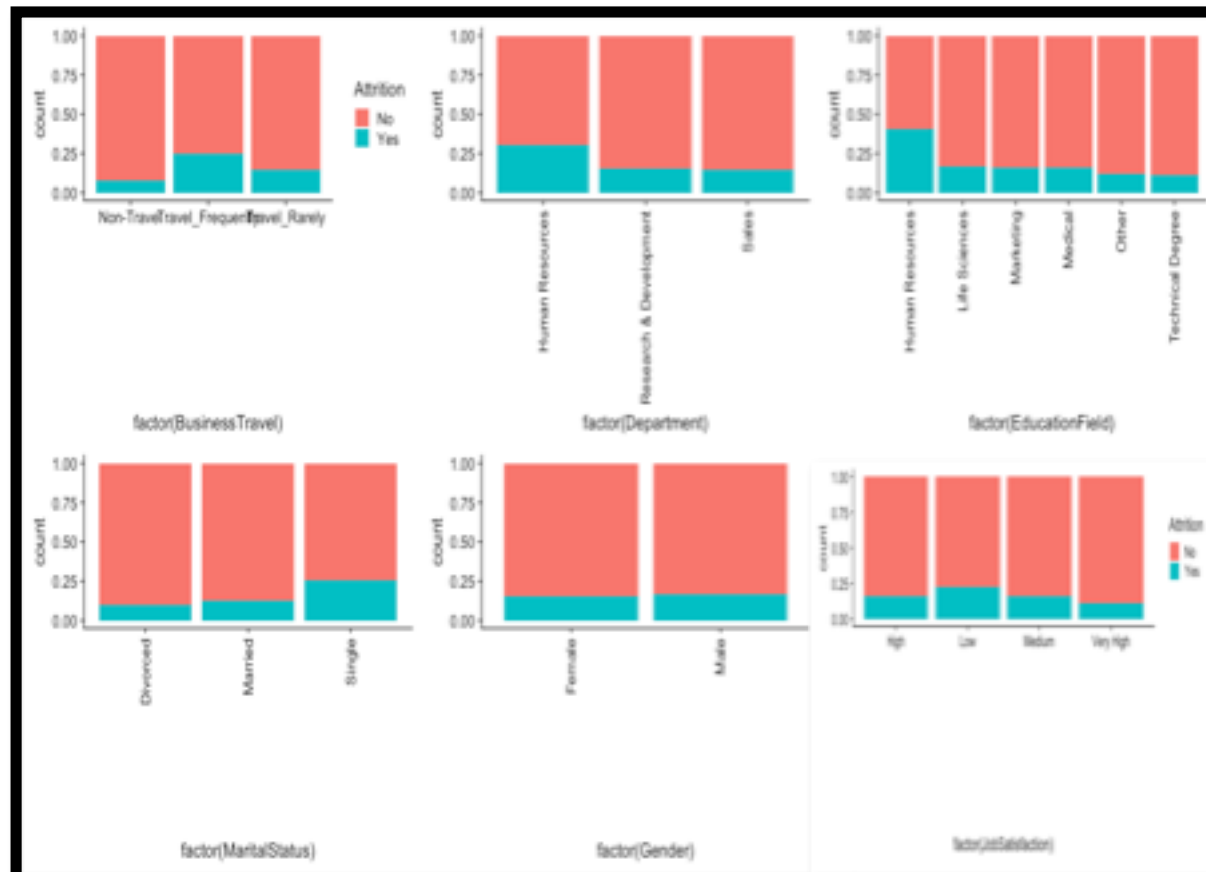
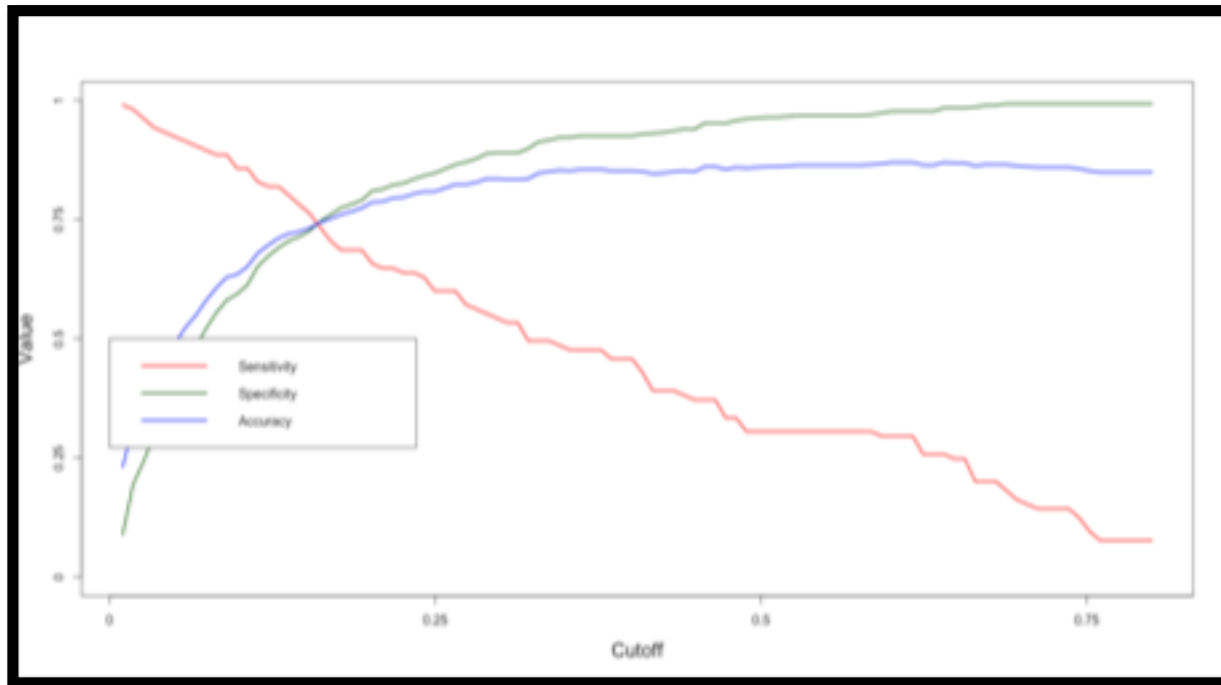**Observation from the graph**

- Frequently Traveling employee have more Attrition.
- HR department has more Attrition.
- Single Employee have more Attrition.
- Employee with low Job Satisfaction have more chances to leave.

- Employee with low Environment Satisfaction have higher chance to leave.
- Employee who do overtime (more than 8 Hours) have high Attrition rate.
- Employee with bad WorkLifeBalance have higher chance to leave.
- Employee with low Job Envolvement have higher chance to leave.

Model Building & Evaluation

We have successfully built a logistic regression model to predict the **probability of attrition** with below characteristics

o   Final model was achieved in 25 iterations at cutoff value is **0.1636**
o   There are total **11 Key variables** as below

- ❑ Accuracy: **75%**
- ❑ Sensitivity(True Positive Rate): **72%**
- ❑ Specificity (True Negative Rate):**75%**
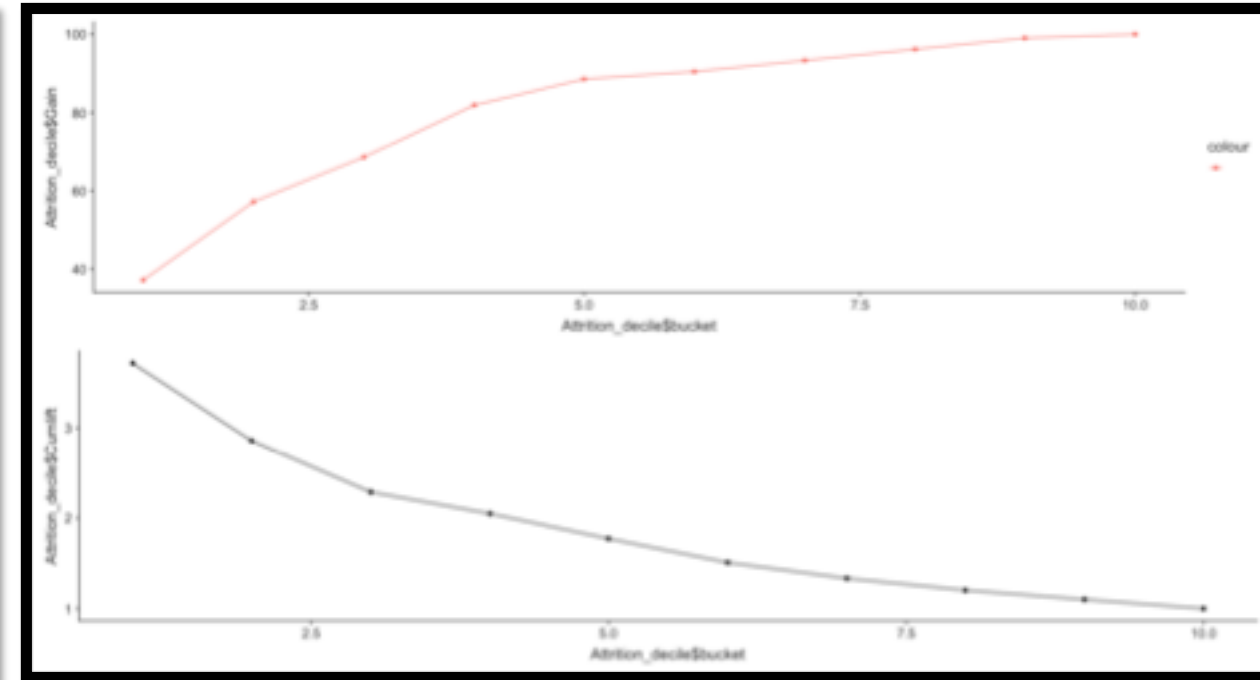
- ❑ KS Statistics: **47.69%**
- ❑ Gain at 4th Decile: **81.9%**
- ❑ Lift at 4th Decile: **2.05%**

Final Model

Observation from Final Model :

- **More Chance of Attrition if**
  - Number of Companies worked for is more.
  - The Employee travels Frequently.
  - Environmental satisfaction is low.
  - Job Satisfaction is low.
  - Employee is doing Overtime.

- **Less Chance of Attrition if**
  - Total Working years is more.
  - Years with Current manager is more.
  - Marital Status is Divorced or Married.
  - Employee does a early Logout

```
Deviance Residuals:
    Min       1Q    Median       3Q       Max
-1.8879   -0.5440   -0.3322   -0.1612    3.2594

Coefficients:
                                Estimate Std. Error z value Pr(>|z|)
(Intercept)                     -1.82003    0.18564  -9.804  < 2e-16 ***
NumCompaniesWorked               0.40561    0.08629   4.700 2.60e-06 ***
TotalWorkingYears               -0.70335    0.11025  -6.380 1.78e-10 ***
YearsWithCurrManager            -0.62518    0.11902  -5.253 1.50e-07 ***
BusinessTravel.xTravel_Frequently 1.12707   0.18396   6.127 8.97e-10 ***
EnvironmentSatisfaction.xLow     0.87718    0.18333   4.785 1.71e-06 ***
MaritalStatus.xDivorced         -1.38050    0.23857  -5.787 7.18e-09 ***
MaritalStatus.xMarried          -1.01287    0.17791  -5.693 1.25e-08 ***
JobSatisfaction.xLow             1.03899    0.18735   5.546 2.93e-08 ***
emp_login_cat.xearly.logout     -0.74046    0.18117  -4.087 4.37e-05 ***
emp_login_cat.xovertime          1.04238    0.24306   4.289 1.80e-05 ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1356.4  on 1559  degrees of freedom
Residual deviance: 1031.2  on 1549  degrees of freedom
AIC: 1053.2

Number of Fisher Scoring iterations: 6
```

Solution :

- **Job And Environmental Satisfaction:** Employee engagement activities should be conducted at regular interval and hold regular meetings to provide management development coaching and feedback.

- **Hiring Strategy:** Candidates who on an average have worked more than 3 years in a company should be preferred, As model shows that employee who have worked in more companies have higher Attrition rate.

- **Update Business Travel Policy:** A service contract should be made with employees who travels for business purpose , This will reduce the Attrition rate.

- **Attention on Early in Careers and Bachelors:** Proper mentorship and career guidance should be provided for early in careers.

- Have flexible work schedules and maintain adequate staff to minimize workload on certain people also schedule meeting with people who are working more to know the main reason for the same.