HR Analytics Case Study

Akhil Verma

Ashesh Rastogi

Ranjani BS

Shriram Bharadhwaj

Objective:

The aim of analysis is to identify the driving factors of the problem (Attrition).

Approach :

Provided in the Next Slide

Derived Measures Details

Derived Metrics	Interpretation	Unit of Measurement (units)	Formula
Time Difference	Total Woring time of an employee in the Company for the day	Hrs	Out Time - Intime
Number of Leaves	Calculated based on NA Values Values for an Employee Other than Common NA (Holidays) Values for All Employees	Numbers	Total Number of Leaves Taken
Average Time Spent	Average Working time in a day of an Employee in the Company	hrs	Average of Time Difference

Assumptions:

Assuming that Current Company is not counted in the Number of Companies Worked and hence the Total Working years is equal to Years at company replacing it with 0. Else Replacing it with 1; such that there are only 1 and 0 values for Number of companies worked.

Inference:

- The Age of the people who resigns from a company goes in the lower Age bracket
- 50-75 percentile of the person who resigns have to travel more distance from home
- Percentage Salary hike not having much impact on the Attrition
- If Total years of experience is less, Higher the attrition Rate
- If number of years in the company is less, Higher the attrition Rate
- Employee tend to Resign after the Promotion with in 2 years
- More Attrition when an Employee Working with Current Manager is less than 5 years
- Number of Leaves Taken not having much impact
- Attrition is High it Average Time Spent by the employee is More

Approach

Objective: to model the **probability of attrition** using a logistic regression.

Data Understanding

Kick OFF and EDA

- Understand the Data with Data Dictionary and Data Provided
- Import the Data Files to R and Merge the Data
- Data Cleanup Check for Missing Values / Not Applicable
- Identify and Derive new Metrics
- Plot the Graphs and Do Outliers Treatment
- Check for Correlation with Dependent Variables
- Do all Analysis and Export the clean data in to another Data frame for Model Building

Logistic Regression

Model Building

- Separate the data in to Train and Test
- Use STEP AIC and Proceed with Modelling

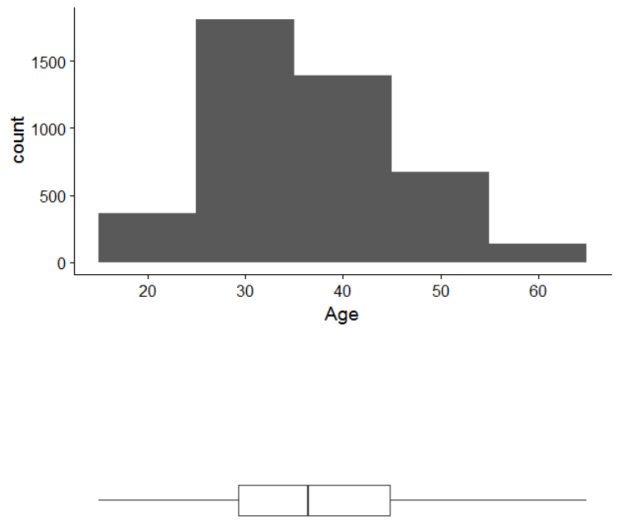
Model Evaluation

- Check the final Model With Test Data
- Confusion Matrix
- Check the Accuracy Sensitivity and Specificity between Train and Test

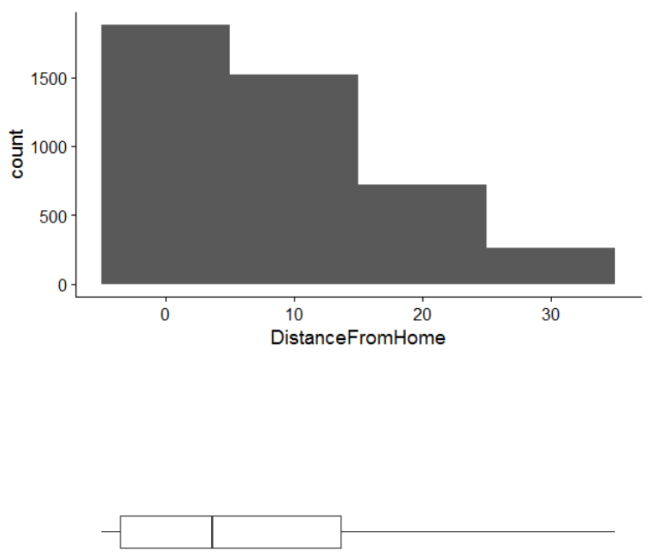
Conclusion

- Conclusion

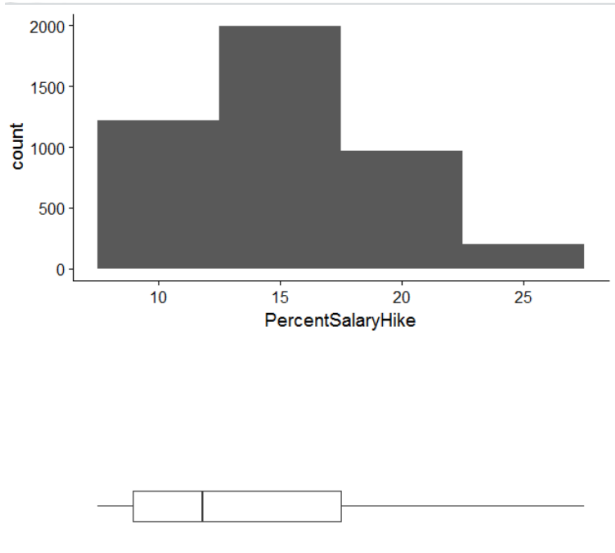
Bar Graph with Box Plot for Continuous Variables



Age

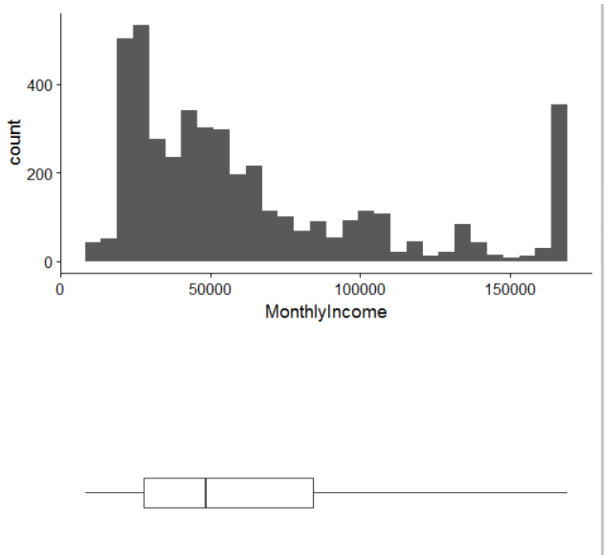


Distance from home

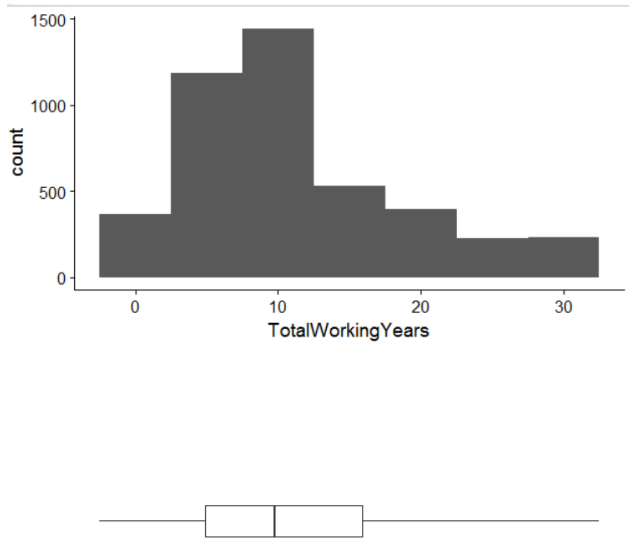


Percentage Salary Hike

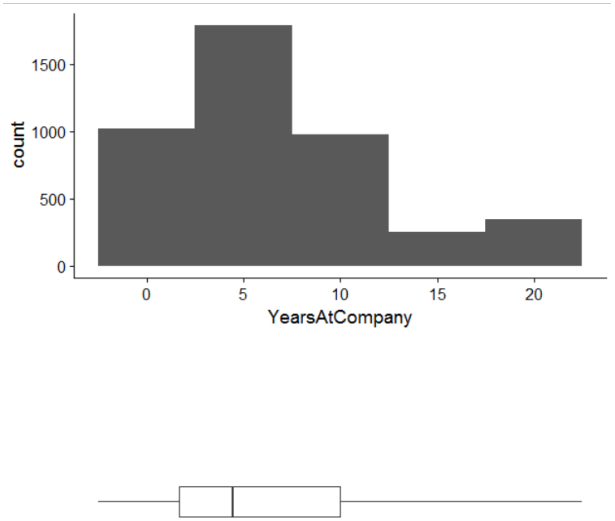
Bar Graph with Box Plot for Continuous Variables



Monthly Income

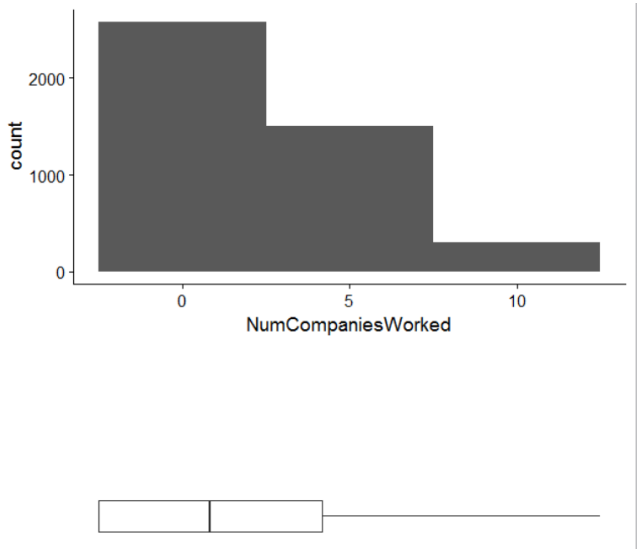


Total Working Years

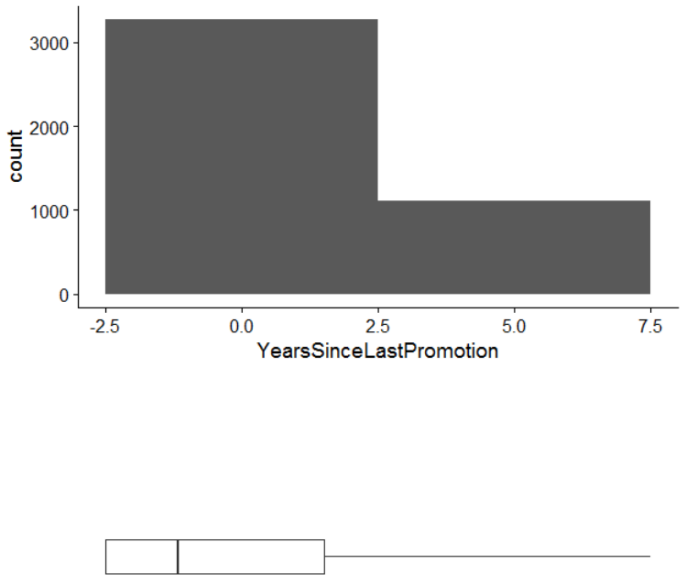


Years at company

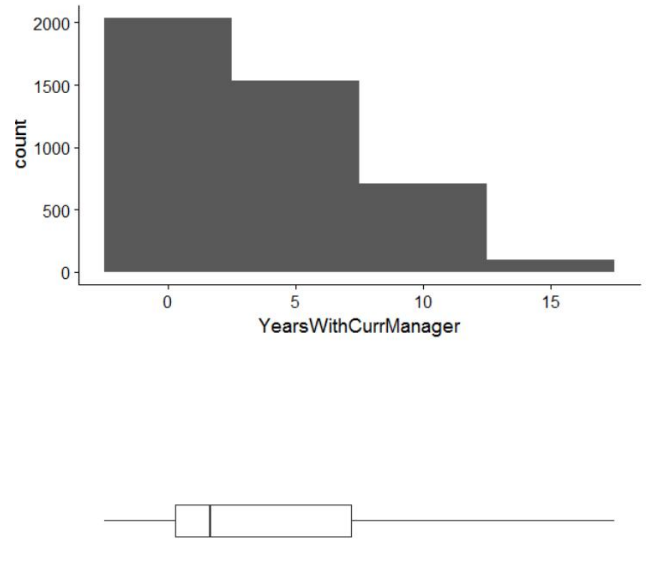
Bar Graph with Box Plot for Continuous Variables



Number of Companies Worked

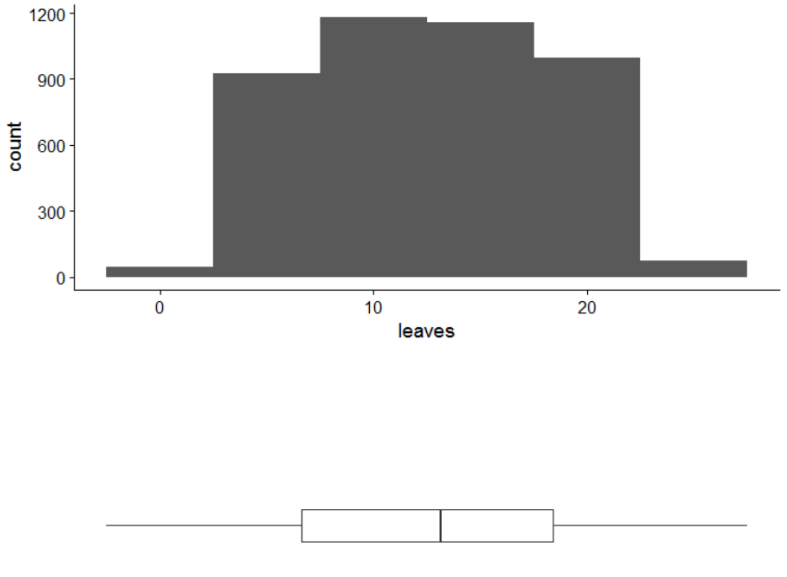


Years Since Last Promotion

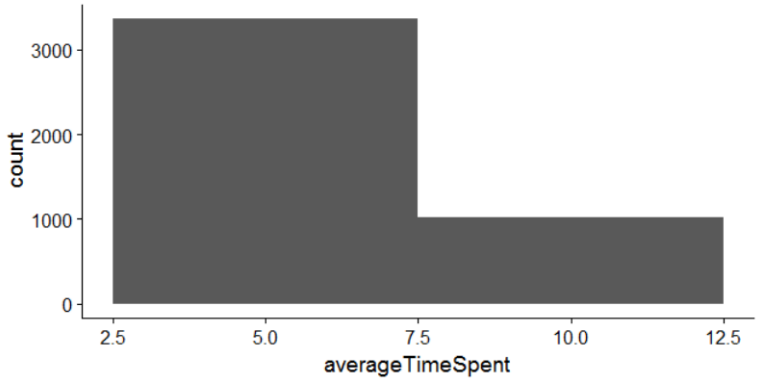


Years With Current Manager

Bar Graph with Box Plot for Derived Variables

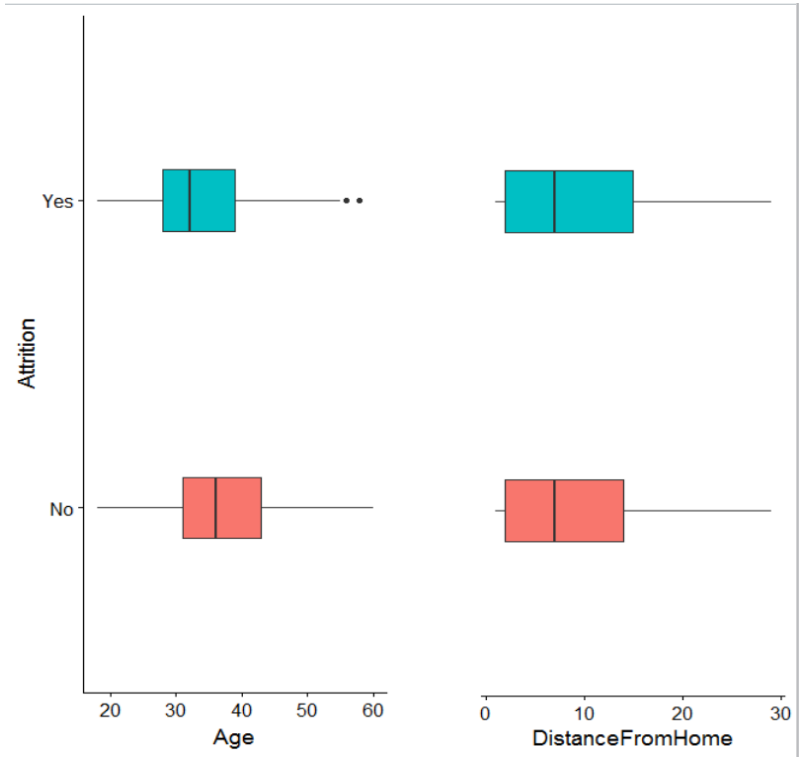


Number of Leaves



Average Time Spent

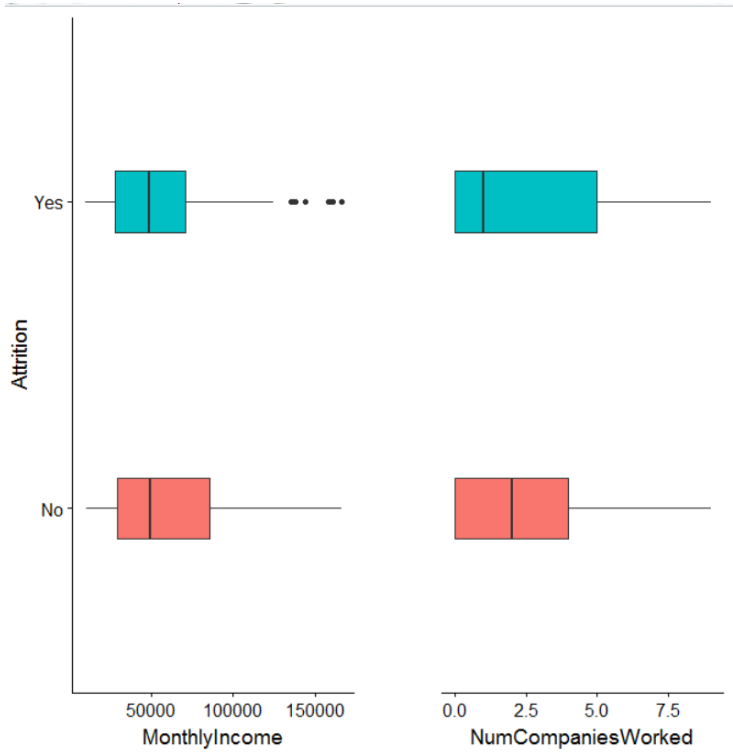
Dependent Variable Vs Continuous Independent Variables



Attrition Vs Age and Distance from Home

Inference-the Age of the people who resigns from a company goes in the lower Age bracket

Inference-The 50-75 percentile of the person who resigns have to travel more distance from home

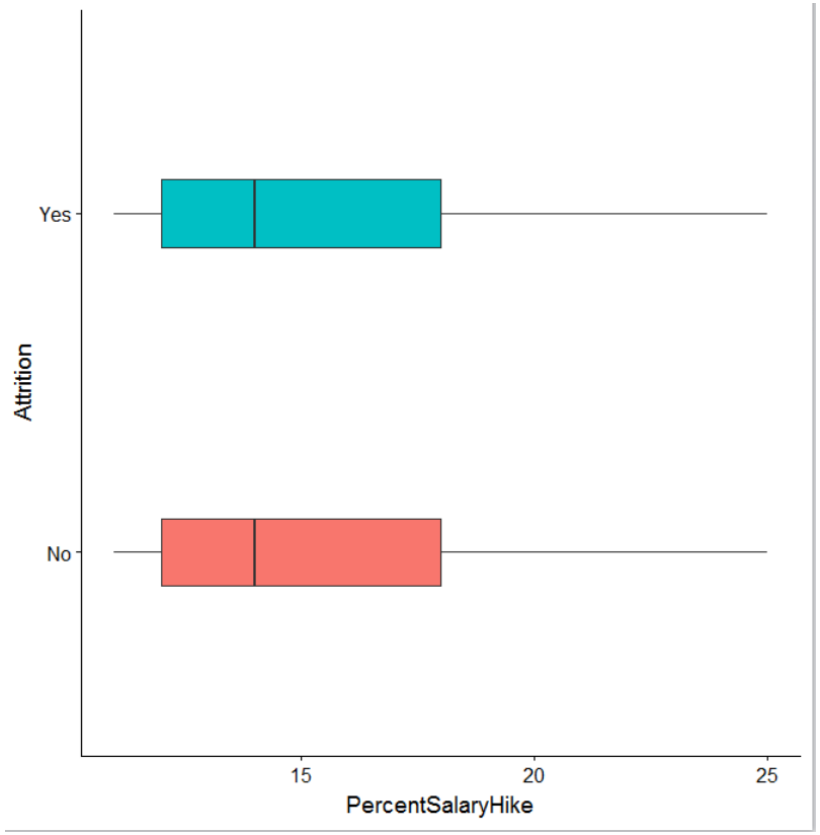


Attrition vs Monthly Income and Number of Companies Worked

Inference- The monthly income is less than the employee who stays with the company

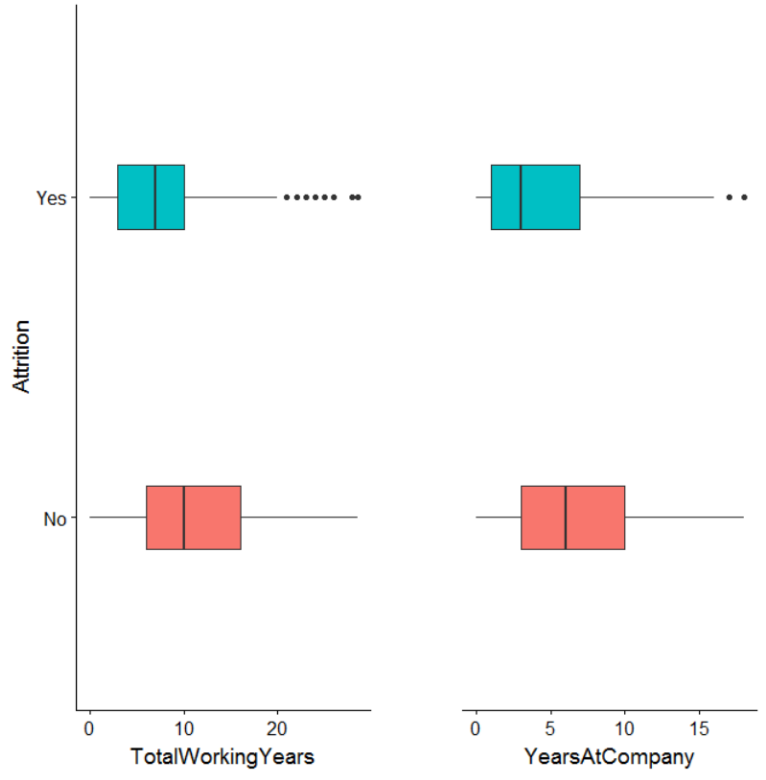
Inference - The Number of company worked is more for the person who resigns

Dependent Variable Vs Continuous Independent Variables



Attrition Vs Percentage Salary Hike

Inference- Percentage Salary hike not having much impact on the Attrition

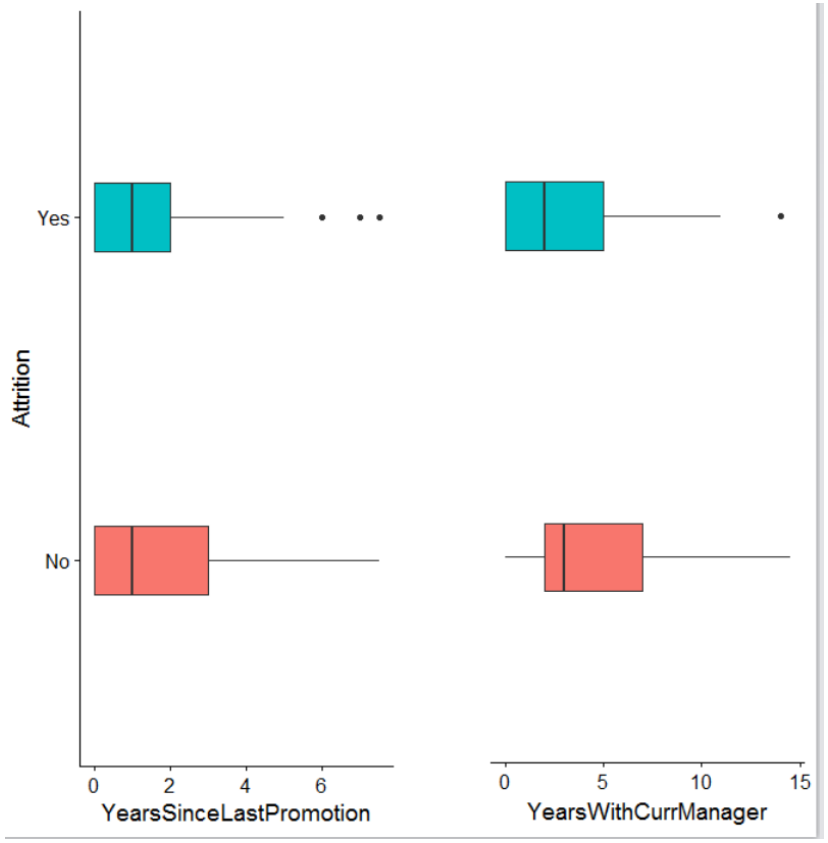


Attrition vs Total Working Years and years at the company

Inference- If Total years of experience is less, Higher the attrition Rate

Inference -If number of years in the company is less, Higher the attrition Rate

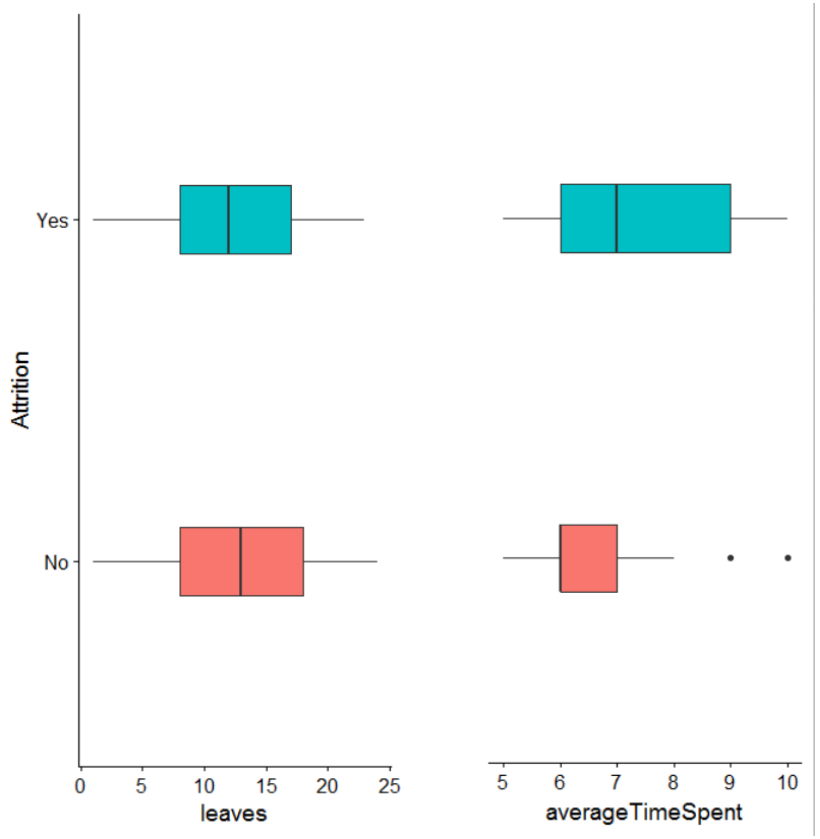
Dependent Variable Vs Continuous Independent Variables



Attrition Vs Years since last Promotion and Years with Current Manager

Inference- Employee tend to Resign after the Promotion with in 2 years

Inference –More Attrition when Employee Working with Current Manager is less than 5 years



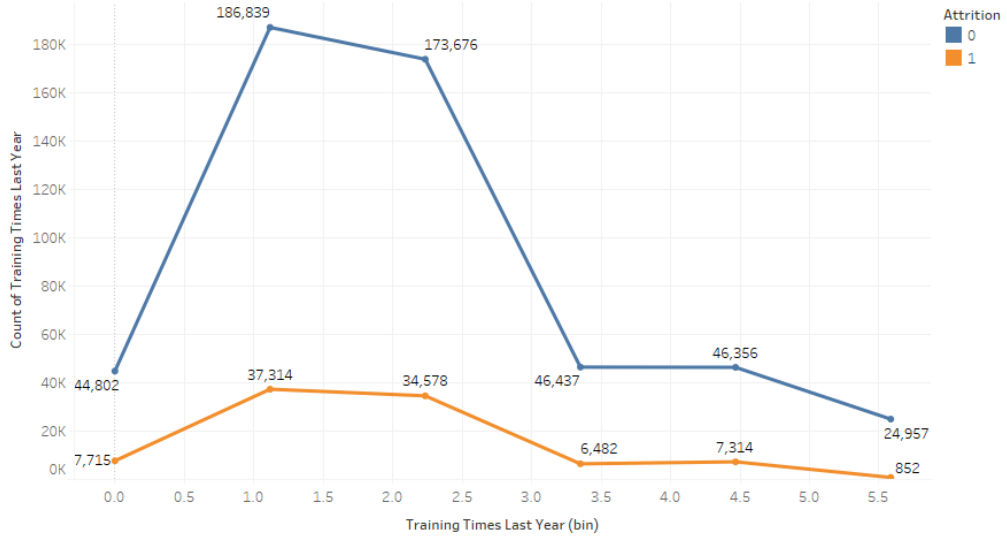
Attrition vs Leaves and Average Time Spent(Average Working Hours)

Inference-Number of Leaves Taken not having much impact

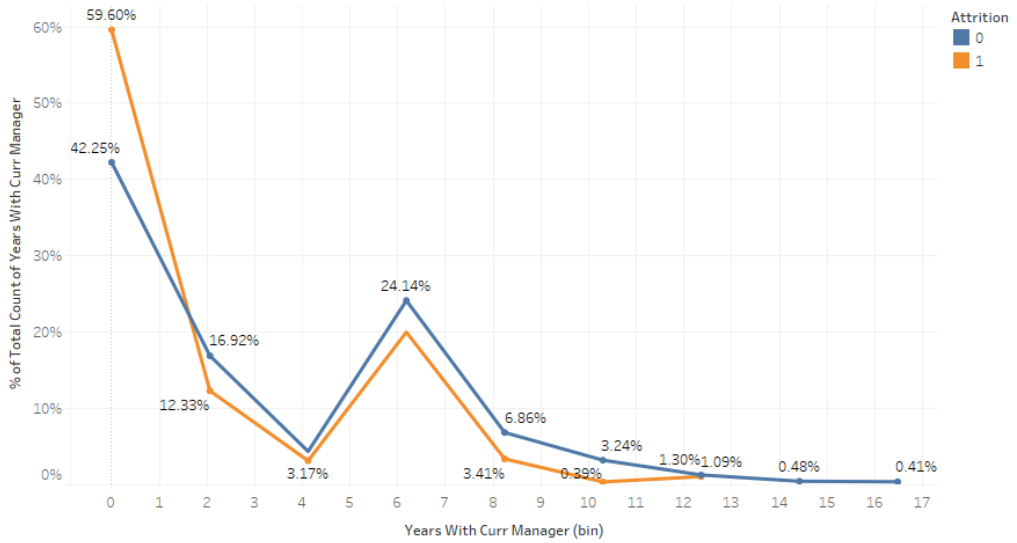
Inference – Attrition is High it Average Time Spent by the employee is More

TABLEAU PLOTS Attrition Vs Independent variables

Attrition Vs Training Times per year

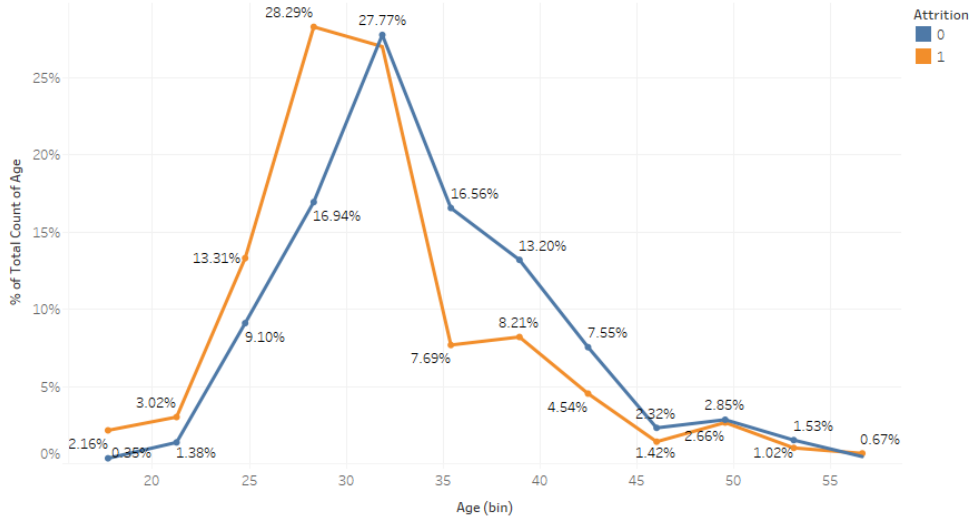


Attrition Vs Years with current manager



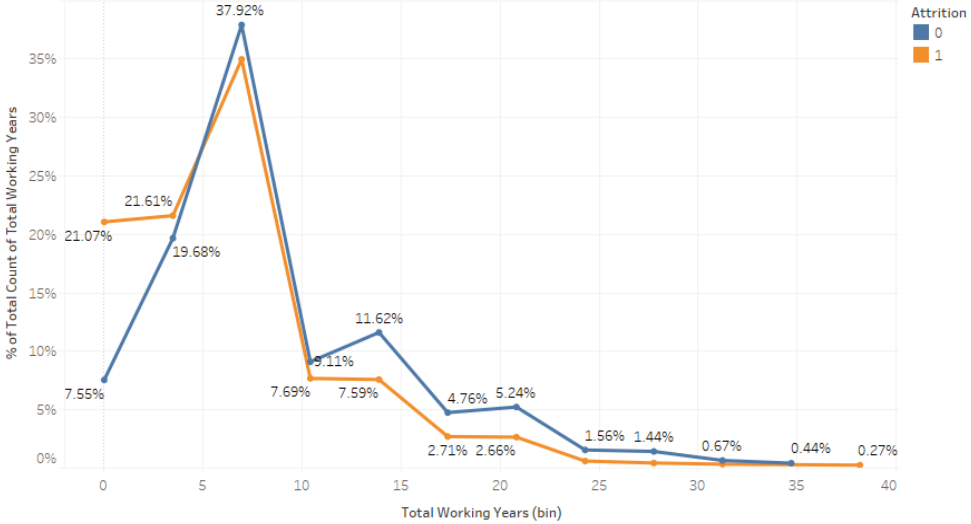
The trend of % of Total Count of Years With Curr Manager for Years With Curr Manager (bin). Color shows details about Attrition.

Attrition vs Age : The plot shows the percentage attrition in the respective age bins.



The trend of % of Total Count of Age for Age (bin). Color shows details about Attrition. The view is filtered on % of Total Count of Age, which keeps all values.

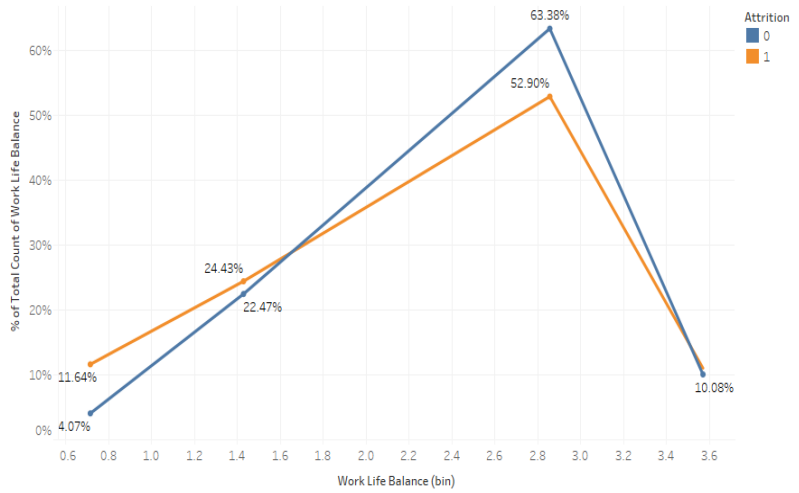
Attrition Vs Total working years



The trend of % of Total Count of Total Working Years for Total Working Years (bin). Color shows details about Attrition.

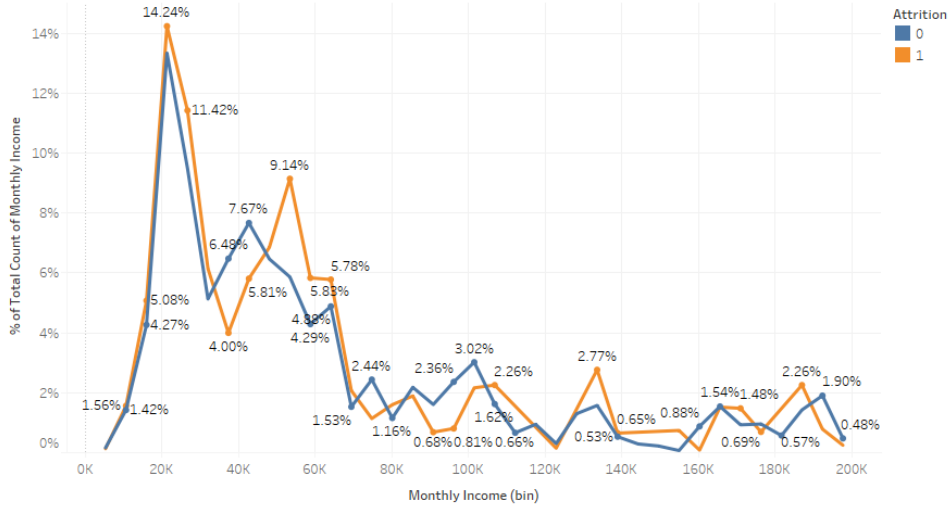
Attrition Vs Independent variables

Attrition Vs work life balance



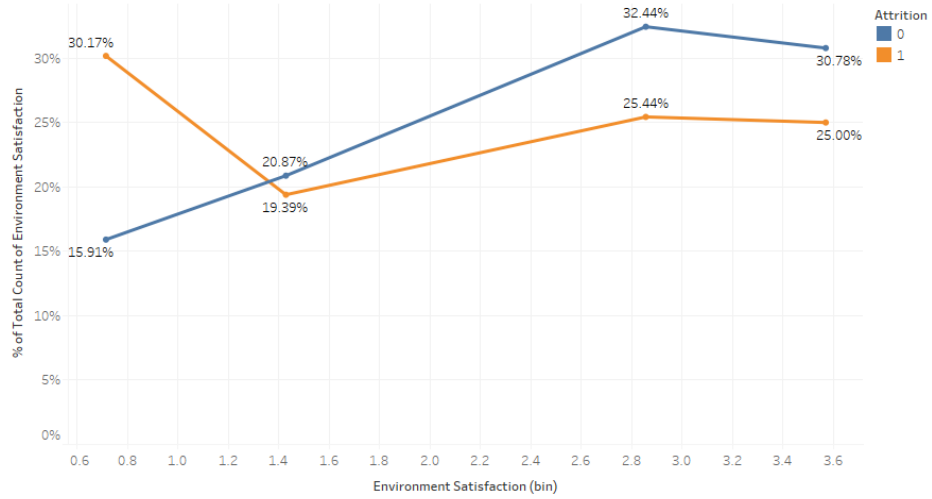
The trend of % of Total Count of Work Life Balance for Work Life Balance (bin). Color shows details about Attrition.

Attrition Vs Monthly income



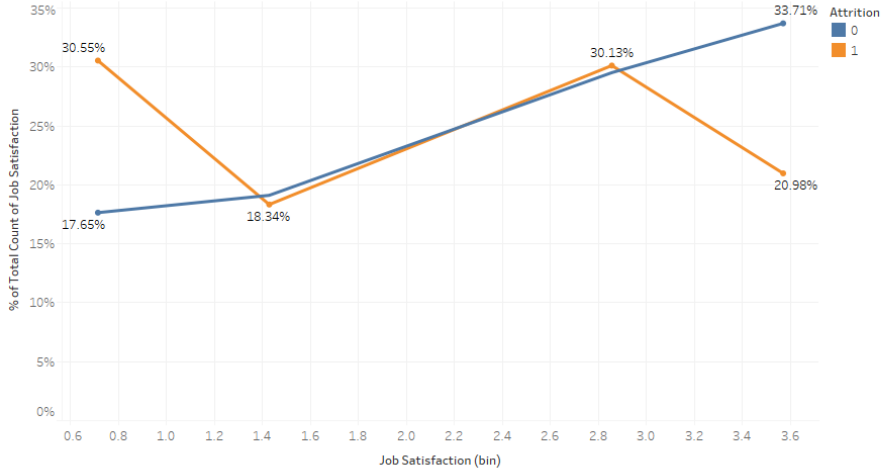
The trend of % of Total Count of Monthly Income for Monthly Income (bin). Color shows details about Attrition.

Attrition Vs Environment Satisfaction



The trend of % of Total Count of Environment Satisfaction for Environment Satisfaction (bin). Color shows details about Attrition.

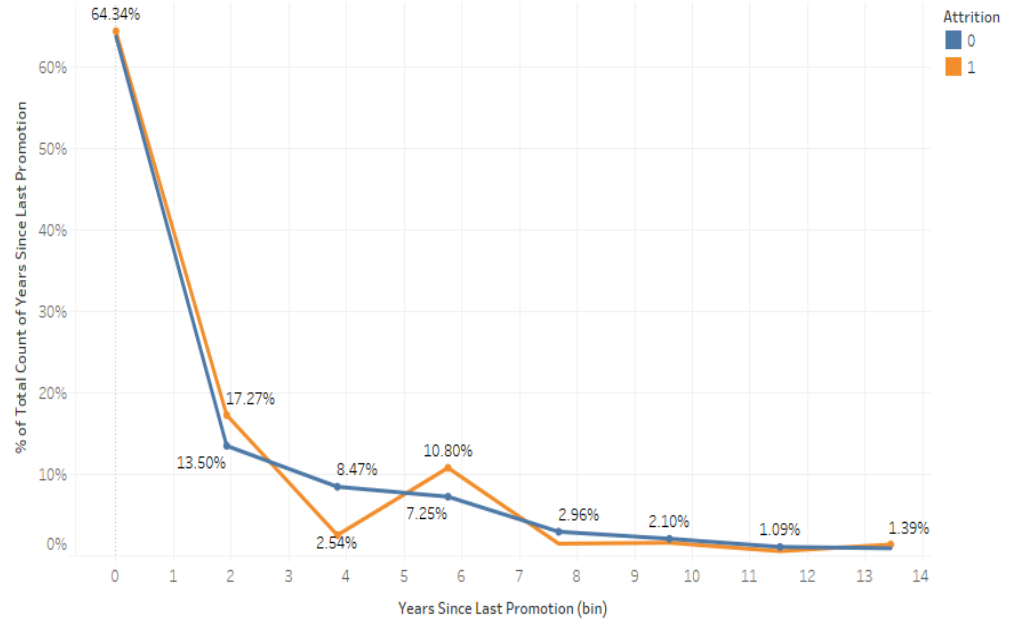
Attrition VS job satisfaction



The trend of % of Total Count of Job Satisfaction for Job Satisfaction (bin). Color shows details about Attrition.

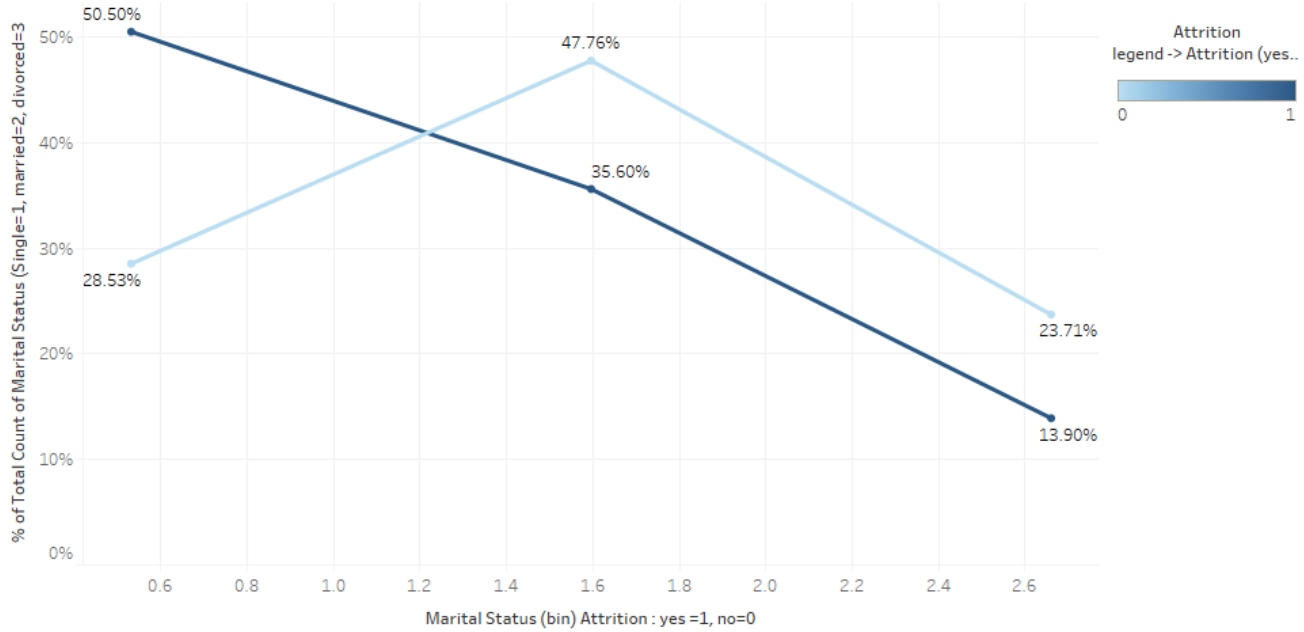
Attrition Vs Independent variables

Attrition Vs years since last promotion



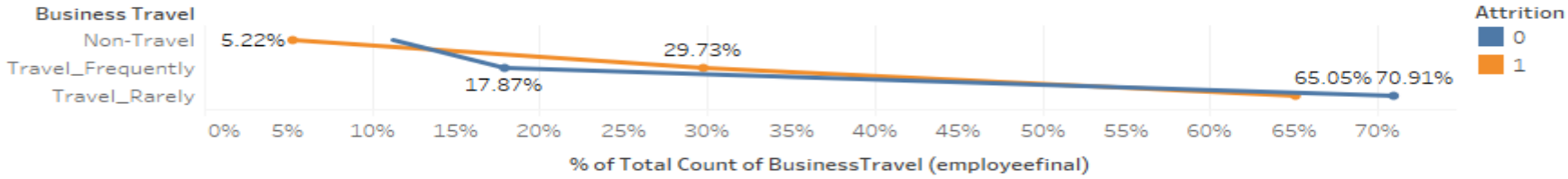
The trend of % of Total Count of Years Since Last Promotion for Years Since Last Promotion (bin). Color shows details about Attrition.

Attrition Vs Marital status



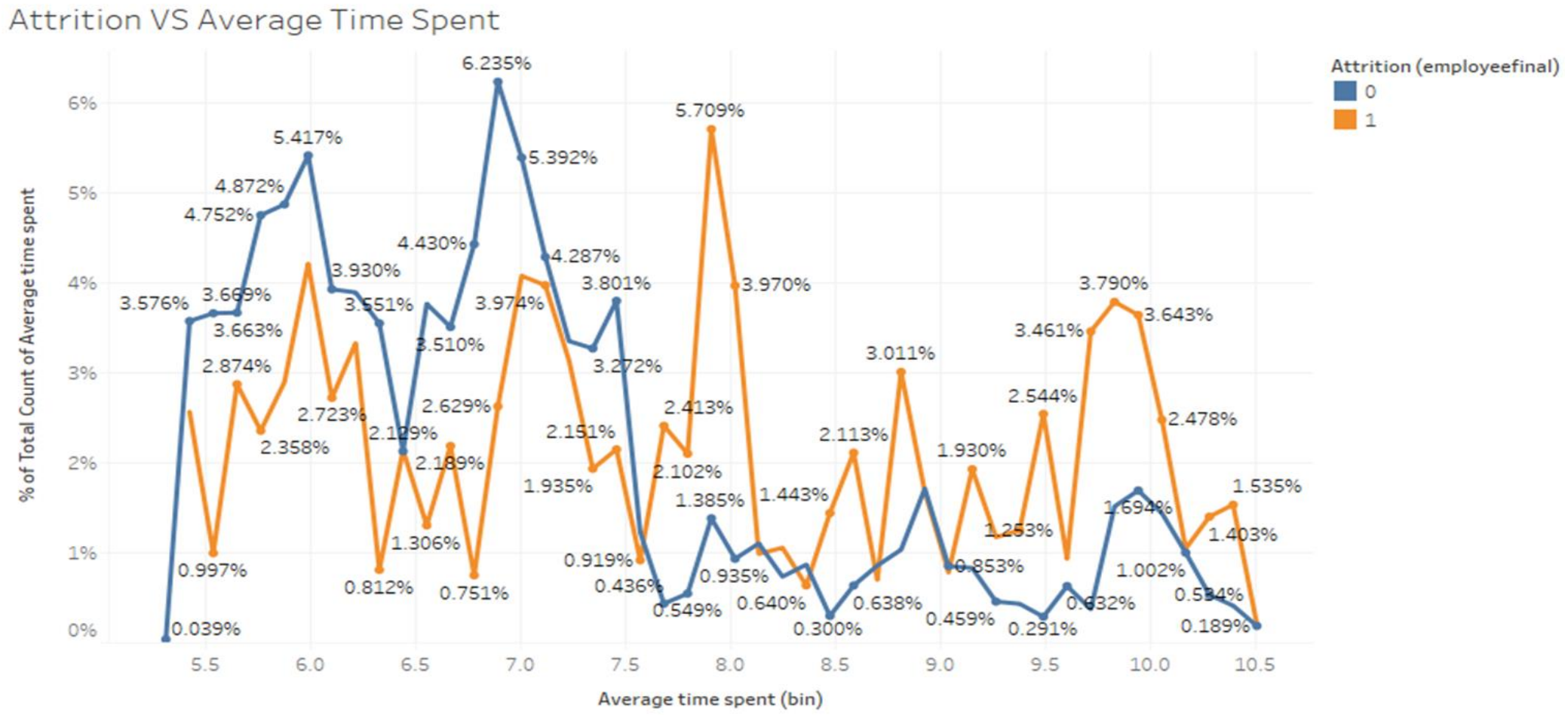
The trend of % of Total Count of Marital Status for Marital Status (bin). Color shows details about Attrition.

Attrition Vs business travel



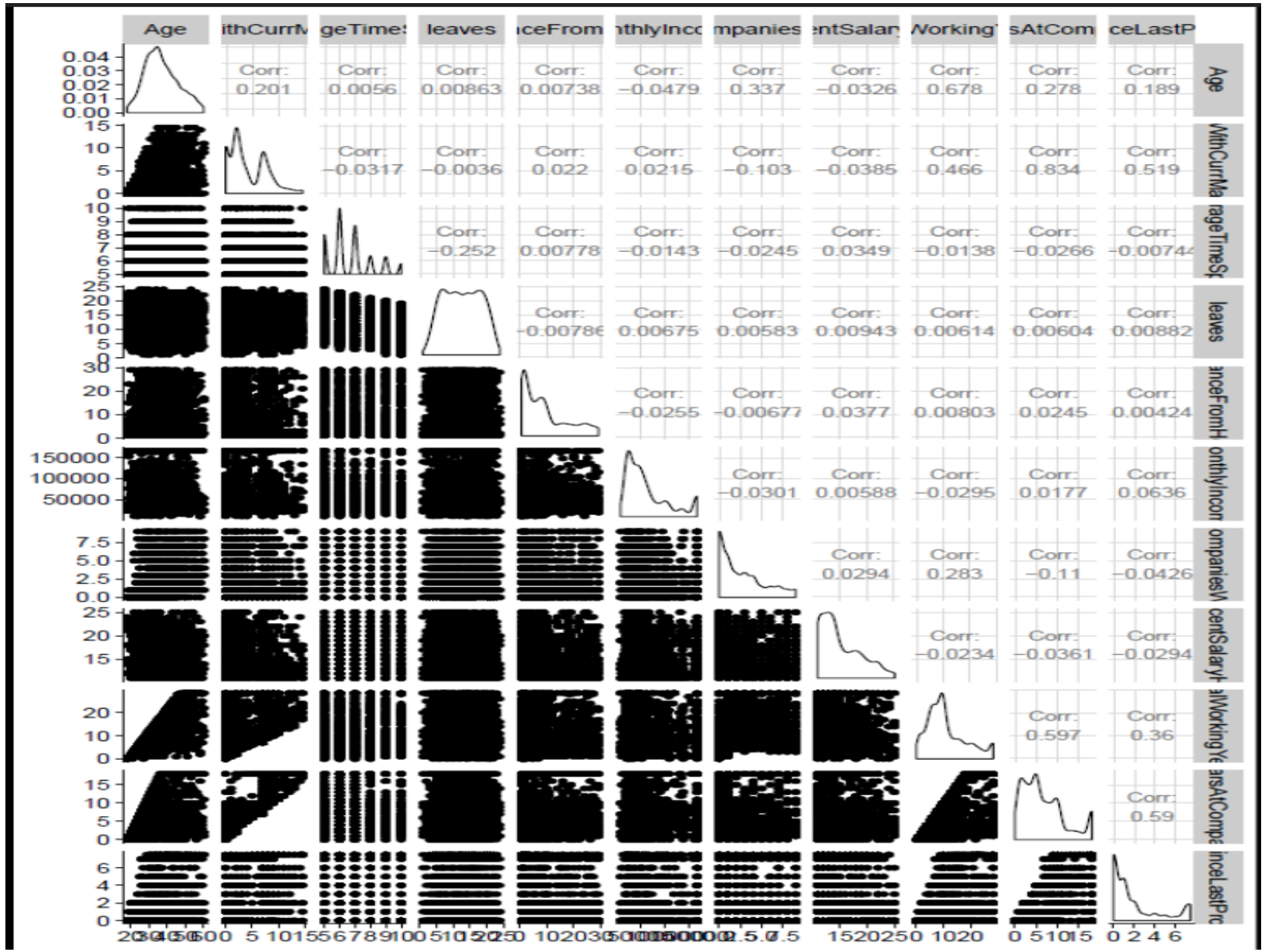
The trend of % of Total Count of BusinessTravel (employeefinal) for Business Travel. Color shows details about Attrition.

Attrition Vs Independent variables



The trend of % of Total Count of Average time spent for Average time spent (bin). Color shows details about Attrition (employeefinal).

Correlation Dependent Vs Independent Variables



Model Evaluation

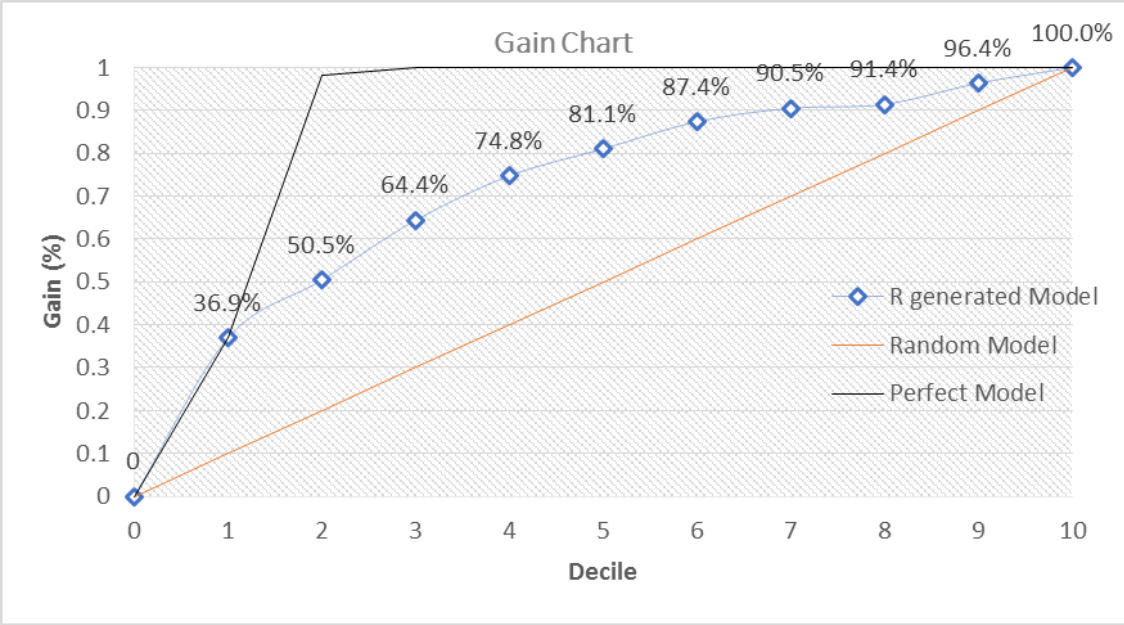
- Confusion Matrix

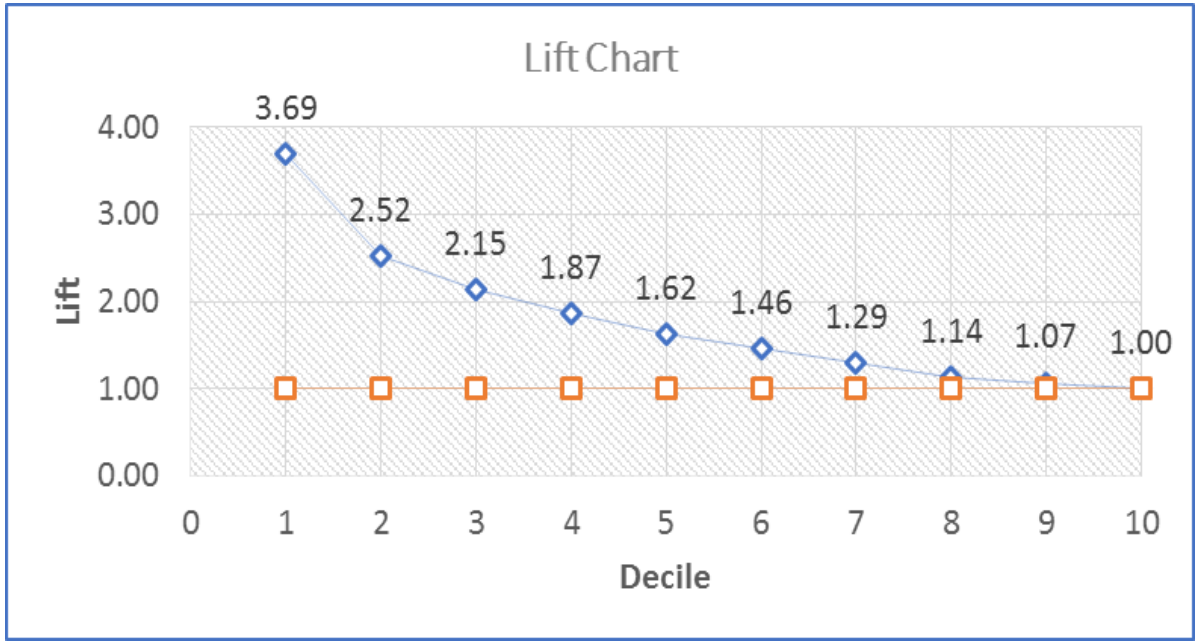
Prediction	No	yes
No	776	66
Yes	317	156

- Accuracy , Specificity, Sensitivity

Accuracy	0.708745
Sensitivity	0.702703
Specificity	0.709973

Gain and Lift Charts





KS Statistic	41.26%