DDS\_Project2-JR

JR

7/12/2021

### Introduction: DDSAnalytics is an analytics company that has decades of experience and is a leading provide of talent management services for many Fortune 100 companies. At the heart of success in companies is people and at the heart of DDSAnalytics is helping our clients succeed by not only recruiting and hiring the best talent available in the marketplace but also retaining and helping such employees flourish which in turn becomes a win-win for all. As information technology capabilities continue to grow, DDSAnalytics is planning to leverage data science in order to improve their solutions for talent management. The executive leadership has identified predicting employee turnover as its first application of data science for talent management. Before implemnting these tools accross the firm, management has tasked your data science team to conduct an analysis of existing employee data. As such, the following information details this research project.

### Link to YouTube presentation:

### <https://www.youtube.com/watch?v=Y3VLaIGCA50>

### Read in files

Case2 = read.csv("C:/Users/Team Reed/OneDrive/JEFF/SMU/Doing Data Science/Project 2/CaseStudy2-data-JR.csv",header = TRUE)  
head(Case2)

## ID Age Attrition BusinessTravel DailyRate Department  
## 1 254 31 Yes Travel\_Rarely 359 Human Resources  
## 2 398 34 Yes Travel\_Rarely 1107 Human Resources  
## 3 704 26 Yes Travel\_Rarely 920 Human Resources  
## 4 716 24 Yes Travel\_Rarely 240 Human Resources  
## 5 733 27 Yes Travel\_Frequently 1337 Human Resources  
## 6 747 29 Yes Travel\_Rarely 350 Human Resources  
## DistanceFromHome Education EducationField EmployeeCount EmployeeNumber  
## 1 18 5 Human Resources 1 1842  
## 2 9 4 Technical Degree 1 1467  
## 3 20 2 Medical 1 1818  
## 4 22 1 Human Resources 1 1714  
## 5 22 3 Human Resources 1 1944  
## 6 13 3 Human Resources 1 1844  
## EnvironmentSatisfaction Gender HourlyRate JobInvolvement JobLevel  
## 1 4 Male 89 4 1  
## 2 1 Female 52 3 1  
## 3 4 Female 69 3 1  
## 4 4 Male 58 1 1  
## 5 1 Female 58 2 1  
## 6 1 Male 56 2 1  
## JobRole JobSatisfaction MaritalStatus MonthlyIncome MonthlyRate  
## 1 Human Resources 1 Married 2956 21495  
## 2 Human Resources 3 Married 2742 3072  
## 3 Human Resources 2 Married 2148 6889  
## 4 Human Resources 3 Married 1555 11585  
## 5 Human Resources 2 Married 2863 19555  
## 6 Human Resources 1 Divorced 2335 3157  
## NumCompaniesWorked Over18 OverTime PercentSalaryHike PerformanceRating  
## 1 0 Y No 17 3  
## 2 1 Y No 15 3  
## 3 0 Y Yes 11 3  
## 4 1 Y No 11 3  
## 5 1 Y No 12 3  
## 6 4 Y Yes 15 3  
## RelationshipSatisfaction StandardHours StockOptionLevel TotalWorkingYears  
## 1 3 80 0 2  
## 2 4 80 0 2  
## 3 3 80 0 6  
## 4 3 80 1 1  
## 5 1 80 0 1  
## 6 4 80 3 4  
## TrainingTimesLastYear WorkLifeBalance YearsAtCompany YearsInCurrentRole  
## 1 4 3 1 0  
## 2 0 3 2 2  
## 3 3 3 5 1  
## 4 2 3 1 0  
## 5 2 3 1 0  
## 6 3 3 2 2  
## YearsSinceLastPromotion YearsWithCurrManager AttritionCat JobRoleCat  
## 1 0 0 1 1  
## 2 2 2 1 1  
## 3 1 4 1 1  
## 4 0 0 1 1  
## 5 0 0 1 1  
## 6 2 0 1 1  
## JobInvolvement. JobLevel. JobSatisfaction. WorkLifeBalance. JI\_JL JI\_JS  
## 1 0.907 0.001 0.001 0.276 0.4540 0.4540  
## 2 0.316 0.001 0.397 0.277 0.1585 0.3565  
## 3 0.316 0.001 0.206 0.277 0.1585 0.2610  
## 4 0.001 0.001 0.397 0.277 0.0010 0.1990  
## 5 0.055 0.001 0.206 0.277 0.0280 0.1305  
## 6 0.055 0.001 0.001 0.278 0.0280 0.0280  
## JL\_JS JI\_WLB JL\_WLB JI\_JL\_JS JI\_JL\_JS\_WLB Business\_Travel\_Num GenderNum  
## 1 0.0010 0.5915 0.1385 0.30300000 0.29625 1 1  
## 2 0.1990 0.2965 0.1390 0.23800000 0.24775 1 0  
## 3 0.1035 0.2965 0.1390 0.17433333 0.20000 1 0  
## 4 0.1990 0.1390 0.1390 0.13300000 0.16900 1 1  
## 5 0.1035 0.1660 0.1390 0.08733333 0.13475 2 0  
## 6 0.0010 0.1665 0.1395 0.01900000 0.08375 1 1  
## JI1 JS1 JL1 WLB1 JI\_JL\_JS\_2 JI\_JL\_WLB\_2 JIJSJLWLB\_1 LogTotalWorkYrs  
## 1 0 1 1 0 1 0 1 0.3010300  
## 2 0 0 1 0 0 0 0 0.3010300  
## 3 0 0 1 0 0 0 0 0.7781512  
## 4 1 0 1 0 1 1 1 0.0000000  
## 5 0 0 1 0 0 0 0 0.0000000  
## 6 0 1 1 0 1 0 1 0.6020600

### Exploratory Data Analysis

### Create graphs in order to assess relationship of Attrition with variables before researching further

### Start with bar charts and use fill in order to see breakdown of Attrition within categories

### Notable results worth researching further: JobRole, JobInvolvement, WorkLifeBalance, Job Satisfaction

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.0.5

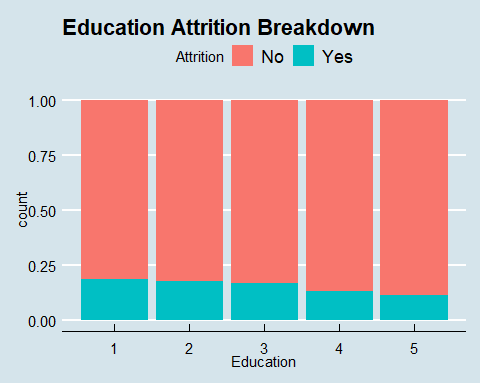
library(magrittr)

## Warning: package 'magrittr' was built under R version 4.0.5

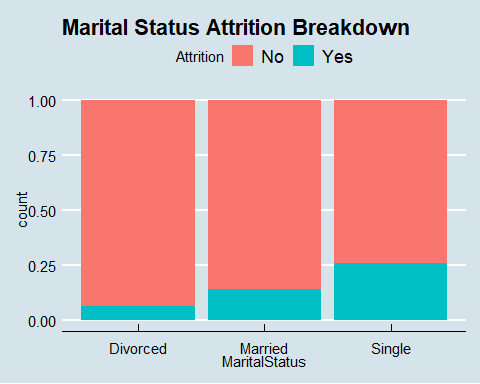
library(ggthemes)

## Warning: package 'ggthemes' was built under R version 4.0.5

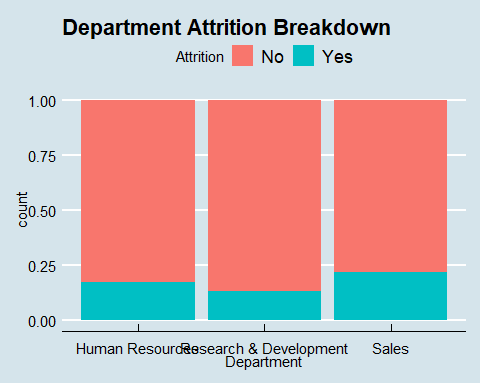
ggplot(Case2, aes(x = Education, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Education Attrition Breakdown") + theme\_economist()



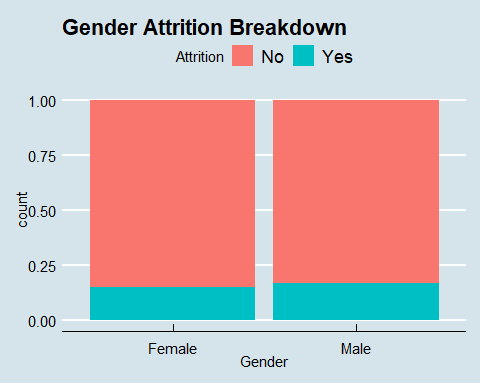
ggplot(Case2, aes(x = MaritalStatus, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Marital Status Attrition Breakdown") + theme\_economist()



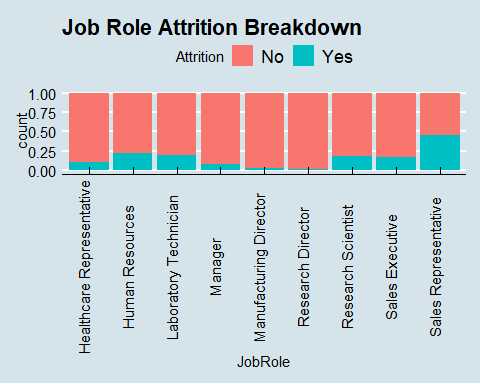
ggplot(Case2, aes(x = Department, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Department Attrition Breakdown") + theme\_economist()



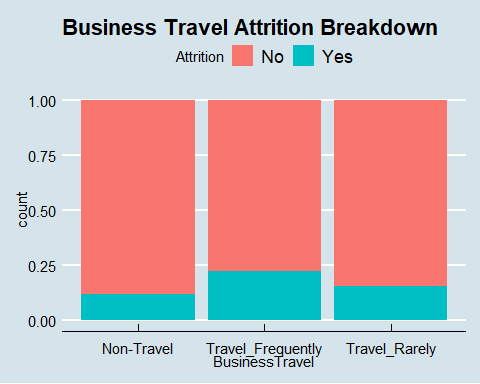
ggplot(Case2, aes(x = Gender, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Gender Attrition Breakdown") + theme\_economist()



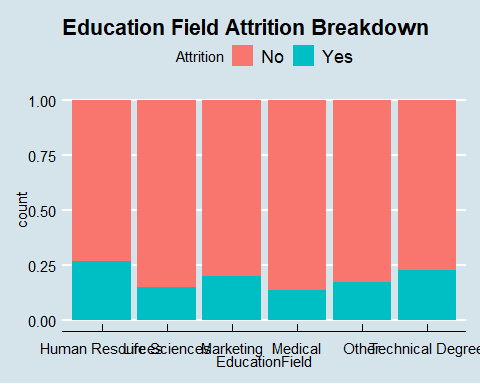
ggplot(Case2, aes(x = JobRole, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Job Role Attrition Breakdown") + theme\_economist() + theme(axis.text.x = element\_text(angle = 90))



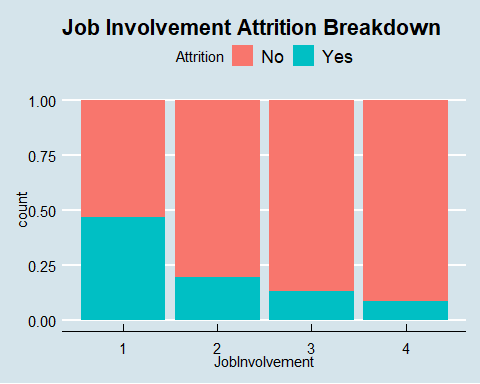
ggplot(Case2, aes(x = BusinessTravel, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Business Travel Attrition Breakdown") + theme\_economist()



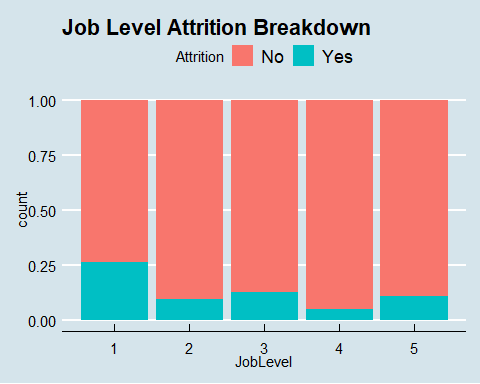
ggplot(Case2, aes(x = EducationField, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Education Field Attrition Breakdown") + theme\_economist()



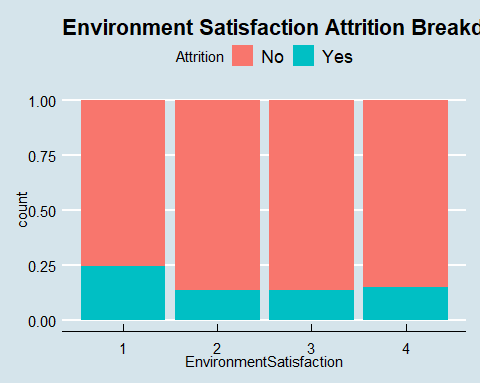
ggplot(Case2, aes(x = JobInvolvement, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Job Involvement Attrition Breakdown") + theme\_economist()



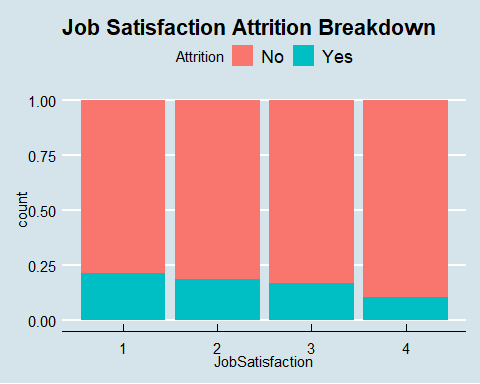
ggplot(Case2, aes(x = JobLevel, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Job Level Attrition Breakdown") + theme\_economist()



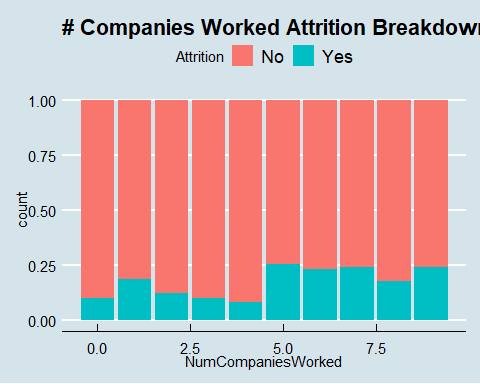
ggplot(Case2, aes(x = EnvironmentSatisfaction, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Environment Satisfaction Attrition Breakdown") + theme\_economist()



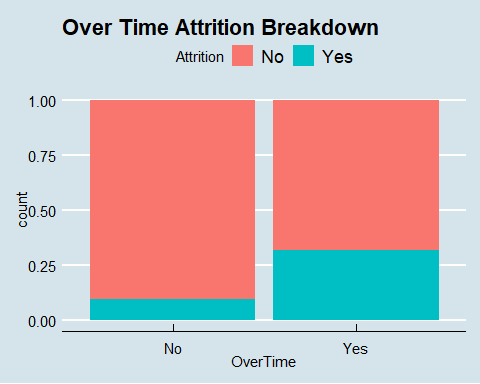
ggplot(Case2, aes(x = JobSatisfaction, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Job Satisfaction Attrition Breakdown") + theme\_economist()



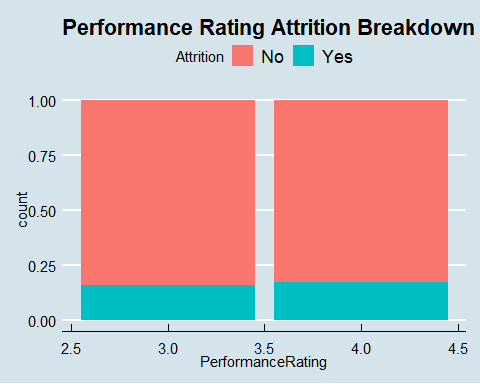
ggplot(Case2, aes(x = NumCompaniesWorked, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("# Companies Worked Attrition Breakdown") + theme\_economist()



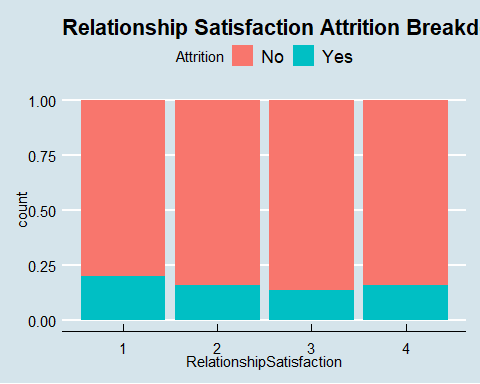
ggplot(Case2, aes(x = OverTime, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Over Time Attrition Breakdown") + theme\_economist()



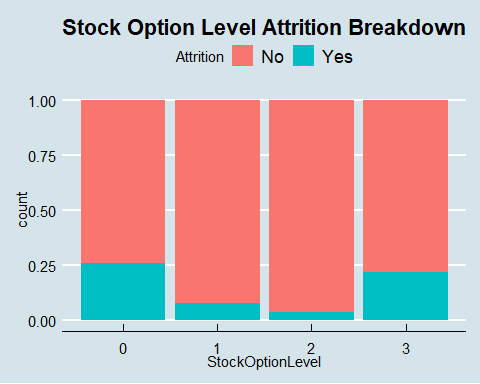
ggplot(Case2, aes(x = PerformanceRating, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Performance Rating Attrition Breakdown") + theme\_economist()



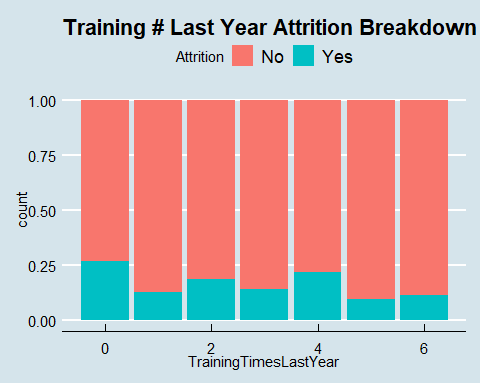
ggplot(Case2, aes(x = RelationshipSatisfaction, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Relationship Satisfaction Attrition Breakdown") + theme\_economist()



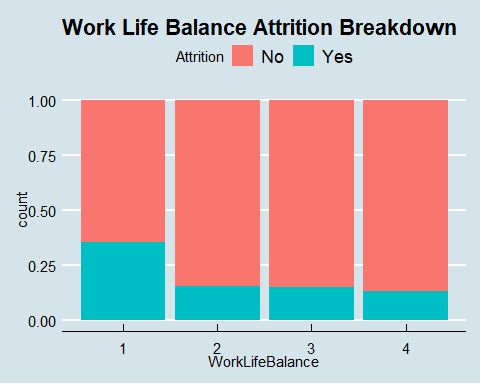
ggplot(Case2, aes(x = StockOptionLevel, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Stock Option Level Attrition Breakdown") + theme\_economist()



ggplot(Case2, aes(x = TrainingTimesLastYear, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Training # Last Year Attrition Breakdown") + theme\_economist()



ggplot(Case2, aes(x = WorkLifeBalance, fill = Attrition)) + geom\_bar(position = "fill") + ggtitle("Work Life Balance Attrition Breakdown") + theme\_economist()

 ### Calculate the percentages of Yes / No Attrition for each of the noted categorical variables above: JobRole, JobInvolvement, WorkLifeBalance, JobSatisfaction ### High attrition in JobRole: Sales Reps (45%), Human Resources (29%), Laboratory Technicians (20%) ### Job Involvement = 1 has had 47% attrition ### Work Life Balance = 1 has had 35% attrition ### Job Satisfaction = 1 has had 21% attrition

library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.0.5

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v tibble 3.1.0 v dplyr 1.0.5  
## v tidyr 1.1.3 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.1  
## v purrr 0.3.4

## Warning: package 'tibble' was built under R version 4.0.5

## Warning: package 'tidyr' was built under R version 4.0.5

## Warning: package 'readr' was built under R version 4.0.5

## Warning: package 'purrr' was built under R version 4.0.5

## Warning: package 'dplyr' was built under R version 4.0.5

## Warning: package 'stringr' was built under R version 4.0.5

## Warning: package 'forcats' was built under R version 4.0.5

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x tidyr::extract() masks magrittr::extract()  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()  
## x purrr::set\_names() masks magrittr::set\_names()

df1 = Case2 %>% select(Attrition, JobRole)   
df2 = as.data.frame(apply(table(df1), 2, function(x) x/sum(x)))  
df2$Attrition = rownames(df2)  
print(df2)

## Healthcare Representative Human Resources Laboratory Technician Manager  
## No 0.8947368 0.7777778 0.8039216 0.92156863  
## Yes 0.1052632 0.2222222 0.1960784 0.07843137  
## Manufacturing Director Research Director Research Scientist Sales Executive  
## No 0.97701149 0.98039216 0.8139535 0.835  
## Yes 0.02298851 0.01960784 0.1860465 0.165  
## Sales Representative Attrition  
## No 0.5471698 No  
## Yes 0.4528302 Yes

df1 = Case2 %>% select(Attrition, JobInvolvement)   
df2 = as.data.frame(apply(table(df1), 2, function(x) x/sum(x)))  
df2$Attrition = rownames(df2)  
print(df2)

## 1 2 3 4 Attrition  
## No 0.5319149 0.8070175 0.8696498 0.91358025 No  
## Yes 0.4680851 0.1929825 0.1303502 0.08641975 Yes

df1 = Case2 %>% select(Attrition, WorkLifeBalance)   
df2 = as.data.frame(apply(table(df1), 2, function(x) x/sum(x)))  
df2$Attrition = rownames(df2)  
print(df2)

## 1 2 3 4 Attrition  
## No 0.6458333 0.84375 0.8496241 0.8673469 No  
## Yes 0.3541667 0.15625 0.1503759 0.1326531 Yes

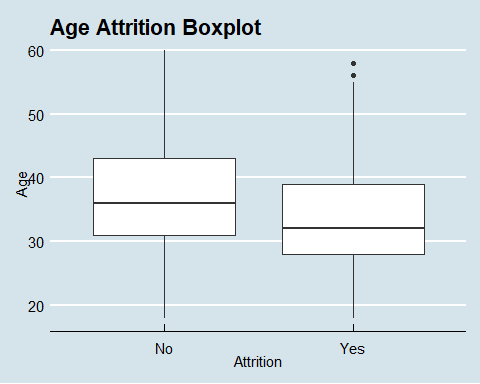
df1 = Case2 %>% select(Attrition, JobSatisfaction)   
df2 = as.data.frame(apply(table(df1), 2, function(x) x/sum(x)))  
df2$Attrition = rownames(df2)  
print(df2)

## 1 2 3 4 Attrition  
## No 0.7877095 0.813253 0.8307087 0.896679 No  
## Yes 0.2122905 0.186747 0.1692913 0.103321 Yes

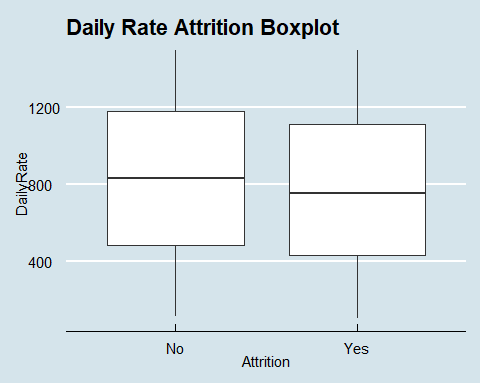
### Next look at boxplots in order to assess distribution of continuous variables relative to Attrition

### Notable results worth researching further: MonthlyIncome, TotalWorkingYears, YearsInCurrentRole, JobSatisfaction, JobLevel

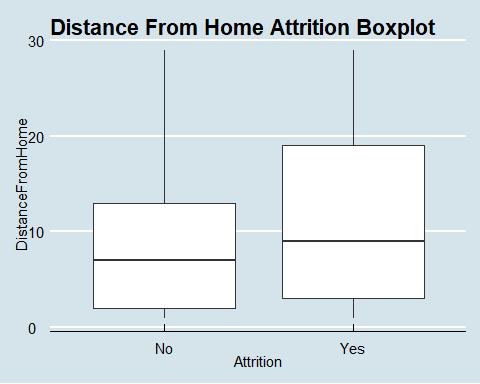
ggplot(data=Case2, mapping = aes(x = Attrition, y = Age)) + geom\_boxplot() + ggtitle("Age Attrition Boxplot") + theme\_economist()



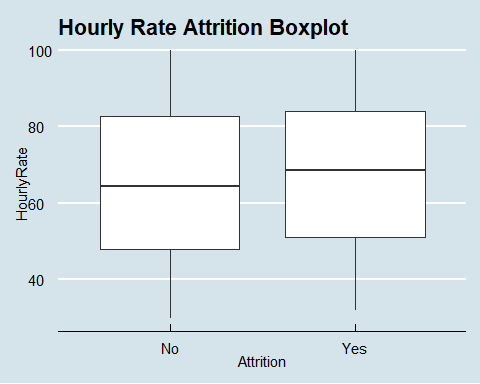
ggplot(data=Case2, mapping = aes(x = Attrition, y = DailyRate)) + geom\_boxplot() + ggtitle("Daily Rate Attrition Boxplot") + theme\_economist()



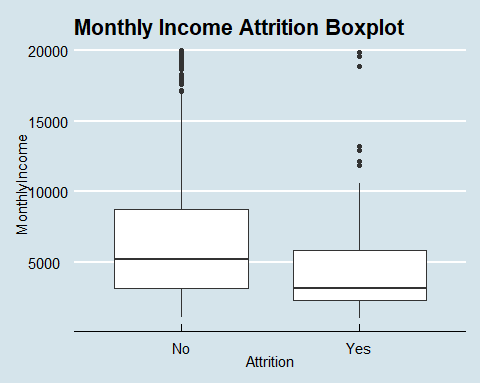
ggplot(data=Case2, mapping = aes(x = Attrition, y = DistanceFromHome)) + geom\_boxplot() + ggtitle("Distance From Home Attrition Boxplot") + theme\_economist()



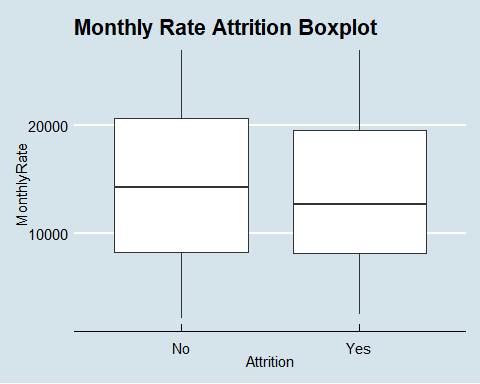
ggplot(data=Case2, mapping = aes(x = Attrition, y = HourlyRate)) + geom\_boxplot() + ggtitle("Hourly Rate Attrition Boxplot") + theme\_economist()



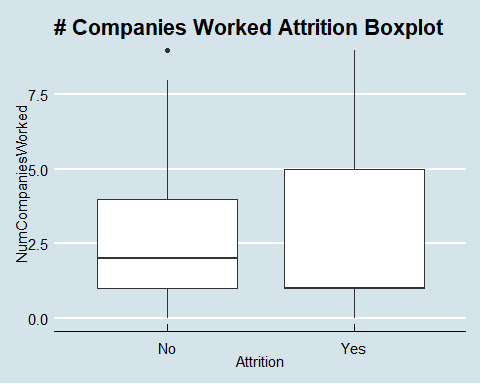
ggplot(data=Case2, mapping = aes(x = Attrition, y = MonthlyIncome)) + geom\_boxplot() + ggtitle("Monthly Income Attrition Boxplot") + theme\_economist()



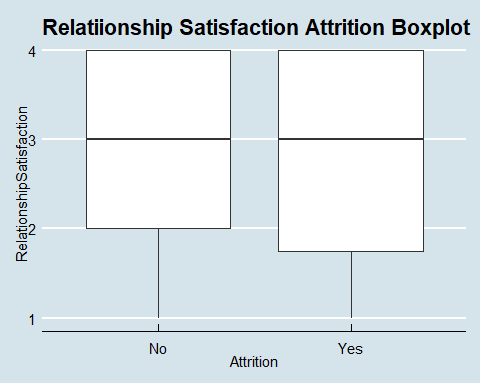
ggplot(data=Case2, mapping = aes(x = Attrition, y = MonthlyRate)) + geom\_boxplot() + ggtitle("Monthly Rate Attrition Boxplot") + theme\_economist()



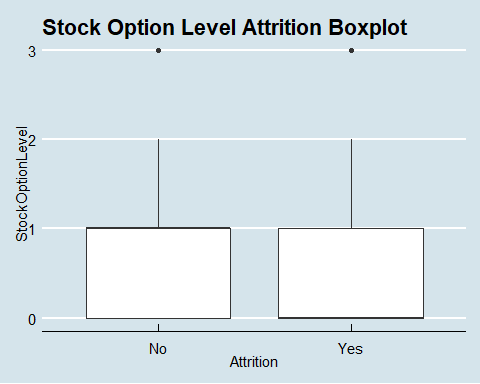
ggplot(data=Case2, mapping = aes(x = Attrition, y = NumCompaniesWorked)) + geom\_boxplot() + ggtitle("# Companies Worked Attrition Boxplot") + theme\_economist()



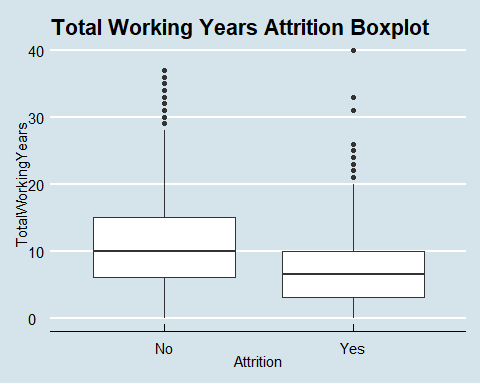
ggplot(data=Case2, mapping = aes(x = Attrition, y = RelationshipSatisfaction)) + geom\_boxplot() + ggtitle("Relatiionship Satisfaction Attrition Boxplot") + theme\_economist()



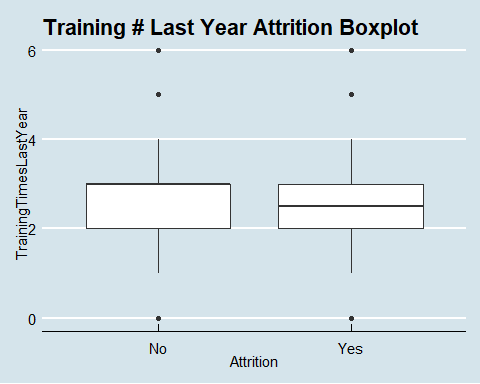
ggplot(data=Case2, mapping = aes(x = Attrition, y = StockOptionLevel)) + geom\_boxplot() + ggtitle("Stock Option Level Attrition Boxplot") + theme\_economist()



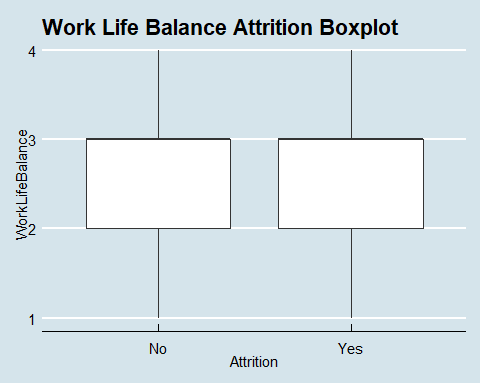
ggplot(data=Case2, mapping = aes(x = Attrition, y = TotalWorkingYears)) + geom\_boxplot()+ ggtitle("Total Working Years Attrition Boxplot") + theme\_economist()



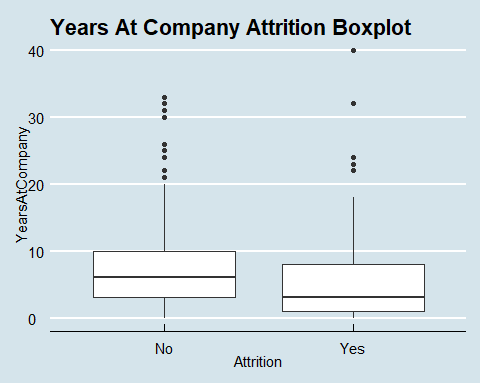
ggplot(data=Case2, mapping = aes(x = Attrition, y = TrainingTimesLastYear)) + geom\_boxplot() + ggtitle("Training # Last Year Attrition Boxplot") + theme\_economist()



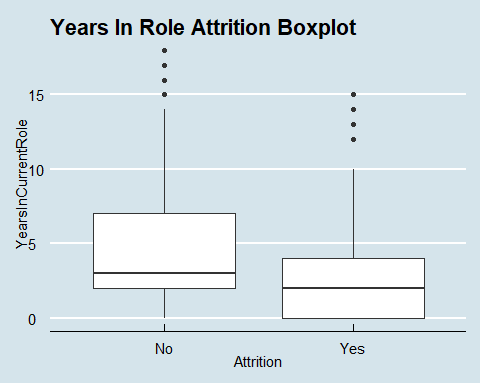
ggplot(data=Case2, mapping = aes(x = Attrition, y = WorkLifeBalance)) + geom\_boxplot() + ggtitle("Work Life Balance Attrition Boxplot") + theme\_economist()



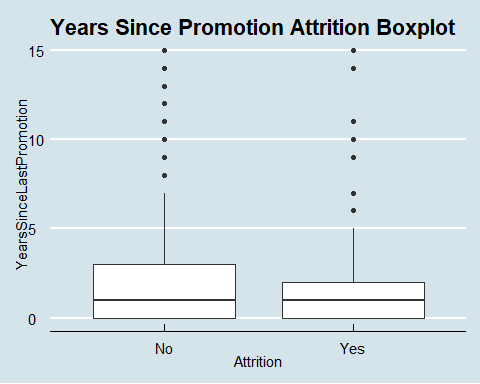
ggplot(data=Case2, mapping = aes(x = Attrition, y = YearsAtCompany)) + geom\_boxplot() + ggtitle("Years At Company Attrition Boxplot") + theme\_economist()



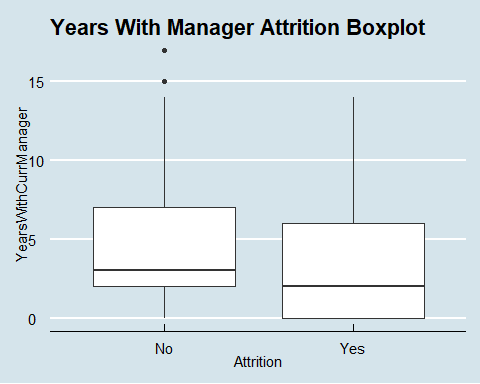
ggplot(data=Case2, mapping = aes(x = Attrition, y = YearsInCurrentRole)) + geom\_boxplot() + ggtitle("Years In Role Attrition Boxplot") + theme\_economist()



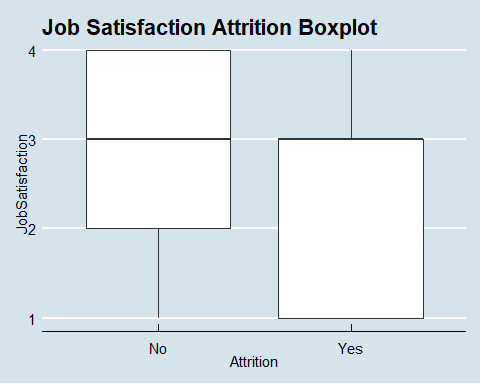
ggplot(data=Case2, mapping = aes(x = Attrition, y = YearsSinceLastPromotion )) + geom\_boxplot() + ggtitle("Years Since Promotion Attrition Boxplot") + theme\_economist()



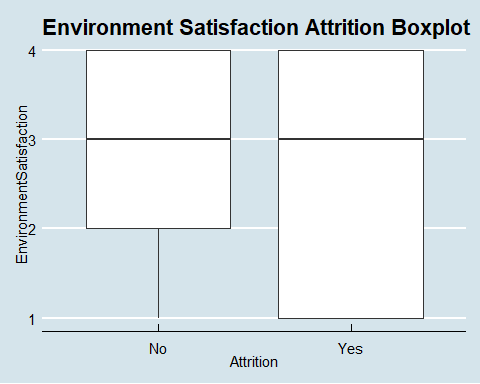
ggplot(data=Case2, mapping = aes(x = Attrition, y = YearsWithCurrManager)) + geom\_boxplot() + ggtitle("Years With Manager Attrition Boxplot") + theme\_economist()



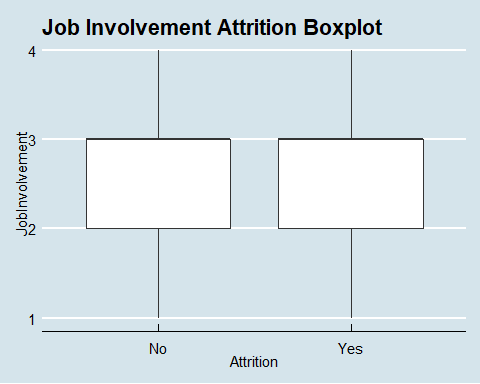
ggplot(data=Case2, mapping = aes(x = Attrition, y = JobSatisfaction)) + geom\_boxplot() + ggtitle("Job Satisfaction Attrition Boxplot") + theme\_economist()



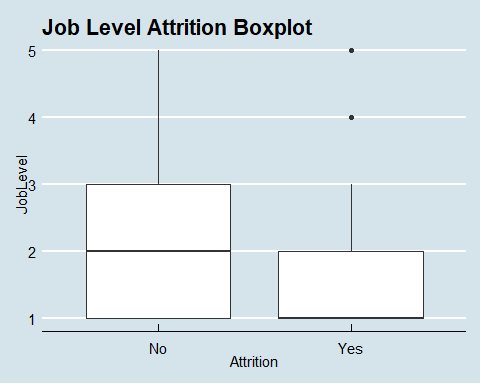
ggplot(data=Case2, mapping = aes(x = Attrition, y = EnvironmentSatisfaction)) + geom\_boxplot() + ggtitle("Environment Satisfaction Attrition Boxplot") + theme\_economist()



ggplot(data=Case2, mapping = aes(x = Attrition, y = JobInvolvement)) + geom\_boxplot() + ggtitle("Job Involvement Attrition Boxplot") + theme\_economist()

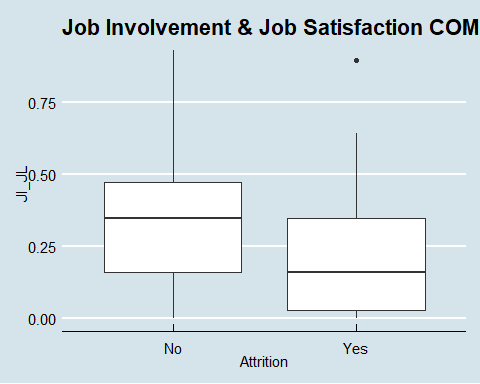


ggplot(data=Case2, mapping = aes(x = Attrition, y = JobLevel)) + geom\_boxplot() + ggtitle("Job Level Attrition Boxplot") + theme\_economist()

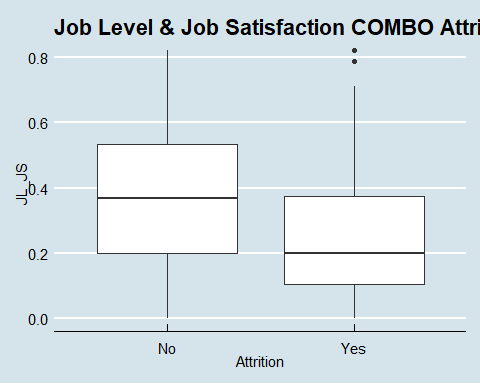


### Create combo variables in excel then graph additional boxplots

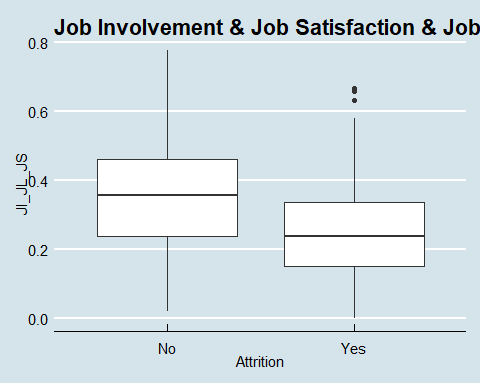
ggplot(data=Case2, mapping = aes(x = Attrition, y = JI\_JL)) + geom\_boxplot() + ggtitle("Job Involvement & Job Satisfaction COMBO Attrition Boxplot") + theme\_economist()



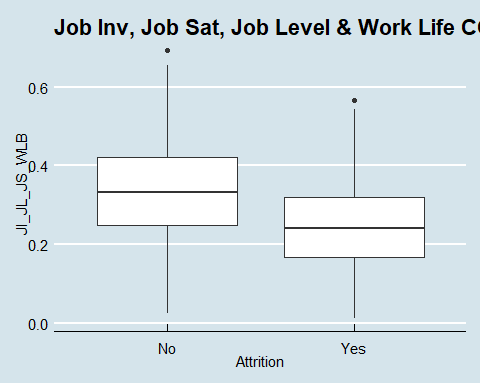
ggplot(data=Case2, mapping = aes(x = Attrition, y = JL\_JS)) + geom\_boxplot() + ggtitle("Job Level & Job Satisfaction COMBO Attrition Boxplot") + theme\_economist()



ggplot(data=Case2, mapping = aes(x = Attrition, y = JI\_JL\_JS)) + geom\_boxplot() + ggtitle("Job Involvement & Job Satisfaction & Job Level COMBO Attrition Boxplot") + theme\_economist()



ggplot(data=Case2, mapping = aes(x = Attrition, y = JI\_JL\_JS\_WLB)) + geom\_boxplot() + ggtitle("Job Inv, Job Sat, Job Level & Work Life COMBO Boxplot") + theme\_economist()



### Initial results from EDA:

### Notable variables to analyze further: JobRole, JobInvolvement, WorkLifeBalance, MonthlyIncome, TotalWorkingYears, YearsInCurrentRole, JobSatisfaction, JobLevel

### Calculate correlation between categorical variables with Attrition with chi squared test: JobRole, JobInvolvement, WorkLifeBalance, JobSatisfaction

### All p-values are statistically significant though the lowest p-value / highest correlation with Attrition is the following from high to low: JobInvolvement, JobRole, WorkLifeBalance, JobSatisfaction

dfJR = Case2[,c(3,17)]  
tblJR = table(dfJR$Attrition, dfJR$JobRole)  
tblJR

##   
## Healthcare Representative Human Resources Laboratory Technician Manager  
## No 68 21 123 47  
## Yes 8 6 30 4  
##   
## Manufacturing Director Research Director Research Scientist  
## No 85 50 140  
## Yes 2 1 32  
##   
## Sales Executive Sales Representative  
## No 167 29  
## Yes 33 24

chisqJR = chisq.test(tblJR)

## Warning in chisq.test(tblJR): Chi-squared approximation may be incorrect

chisqJR

##   
## Pearson's Chi-squared test  
##   
## data: tblJR  
## X-squared = 60.543, df = 8, p-value = 3.647e-10

dfJI = Case2[,c(3,15)]  
tblJI = table(dfJI$Attrition, dfJI$JobInvolvement)  
tblJI

##   
## 1 2 3 4  
## No 25 184 447 74  
## Yes 22 44 67 7

chisqJI = chisq.test(tblJI)  
chisqJI

##   
## Pearson's Chi-squared test  
##   
## data: tblJI  
## X-squared = 41.465, df = 3, p-value = 5.211e-09

dfWL = Case2[,c(3,32)]  
tblWL = table(dfWL$Attrition, dfWL$WorkLifeBalance)  
tblWL

##   
## 1 2 3 4  
## No 31 162 452 85  
## Yes 17 30 80 13

chisqWL = chisq.test(tblWL)  
chisqWL

##   
## Pearson's Chi-squared test  
##   
## data: tblWL  
## X-squared = 14.325, df = 3, p-value = 0.002495

dfJS = Case2[,c(3,18)]  
tblJS = table(dfJS$Attrition, dfJS$JobSatisfaction)  
tblJS

##   
## 1 2 3 4  
## No 141 135 211 243  
## Yes 38 31 43 28

chisqJS = chisq.test(tblJS)  
chisqJS

##   
## Pearson's Chi-squared test  
##   
## data: tblJS  
## X-squared = 11.109, df = 3, p-value = 0.01115

### Calculate correlation between continous variables with Attrition with pearson correlation

### Highest correlation with Attrition (highest to lowest) = TotalWorkingYears, YearsInCurrentRole, MonthlyIncome, DistanceFromHome

library(ltm)

## Warning: package 'ltm' was built under R version 4.0.5

## Loading required package: MASS

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

## Loading required package: msm

## Warning: package 'msm' was built under R version 4.0.5

## Loading required package: polycor

## Warning: package 'polycor' was built under R version 4.0.5

biserial.cor(Case2$DistanceFromHome, Case2$AttritionCat)

## [1] -0.08713629

biserial.cor(Case2$MonthlyIncome, Case2$AttritionCat)

## [1] 0.154915

biserial.cor(Case2$TotalWorkingYears, Case2$AttritionCat)

## [1] 0.1672061

biserial.cor(Case2$YearsInCurrentRole, Case2$AttritionCat)

## [1] 0.1562157

### Calculate correlation between notable continuous variables to test for lack of independence

### High correlation between the following: TotalWorkingYears & MonthlyIncome

### Therefore when testing model be mindful of such interaction

df4 = Case2[,c(7,20,30,34)]  
summary(df4)

## DistanceFromHome MonthlyIncome TotalWorkingYears YearsInCurrentRole  
## Min. : 1.000 Min. : 1081 Min. : 0.00 Min. : 0.000   
## 1st Qu.: 2.000 1st Qu.: 2840 1st Qu.: 6.00 1st Qu.: 2.000   
## Median : 7.000 Median : 4946 Median :10.00 Median : 3.000   
## Mean : 9.339 Mean : 6390 Mean :11.05 Mean : 4.205   
## 3rd Qu.:14.000 3rd Qu.: 8182 3rd Qu.:15.00 3rd Qu.: 7.000   
## Max. :29.000 Max. :19999 Max. :40.00 Max. :18.000

corr = cor(df4)  
round(corr, 2)

## DistanceFromHome MonthlyIncome TotalWorkingYears  
## DistanceFromHome 1.00 -0.01 0.00  
## MonthlyIncome -0.01 1.00 0.78  
## TotalWorkingYears 0.00 0.78 1.00  
## YearsInCurrentRole -0.01 0.36 0.49  
## YearsInCurrentRole  
## DistanceFromHome -0.01  
## MonthlyIncome 0.36  
## TotalWorkingYears 0.49  
## YearsInCurrentRole 1.00

### Calculate correlation between categorical variables to test for lack of independence with chi squared test: JobRole, JobInvolvement, JobSatisfaction, WorkLifeBalance

### All p-values are greater than alpha of 0.05 so we conclude that there is independence between these variables

dfJRJI = Case2[,c(15,17)]  
tblJRJI = table(dfJRJI$JobInvolvement, dfJR$JobRole)  
tblJRJI

##   
## Healthcare Representative Human Resources Laboratory Technician Manager  
## 1 2 1 10 4  
## 2 19 7 44 12  
## 3 44 16 89 32  
## 4 11 3 10 3  
##   
## Manufacturing Director Research Director Research Scientist Sales Executive  
## 1 3 1 8 14  
## 2 30 11 42 48  
## 3 49 32 99 121  
## 4 5 7 23 17  
##   
## Sales Representative  
## 1 4  
## 2 15  
## 3 32  
## 4 2

chisqJRJI = chisq.test(tblJRJI)

## Warning in chisq.test(tblJRJI): Chi-squared approximation may be incorrect

chisqJRJI

##   
## Pearson's Chi-squared test  
##   
## data: tblJRJI  
## X-squared = 21.286, df = 24, p-value = 0.6218

dfJIWL = Case2[,c(15,32)]  
tblJIWL = table(dfJIWL$JobInvolvement, dfJIWL$WorkLifeBalance)  
tblJIWL

##   
## 1 2 3 4  
## 1 3 10 32 2  
## 2 13 47 142 26  
## 3 28 119 302 65  
## 4 4 16 56 5

chisqJIWL = chisq.test(tblJIWL)

## Warning in chisq.test(tblJIWL): Chi-squared approximation may be incorrect

chisqJIWL

##   
## Pearson's Chi-squared test  
##   
## data: tblJIWL  
## X-squared = 7.3667, df = 9, p-value = 0.599

dfJRWL = Case2[,c(17,32)]  
tblJRWL = table(dfJRWL$JobRole, dfJRWL$WorkLifeBalance)  
tblJRWL

##   
## 1 2 3 4  
## Healthcare Representative 7 19 42 8  
## Human Resources 1 3 19 4  
## Laboratory Technician 11 28 100 14  
## Manager 3 11 32 5  
## Manufacturing Director 2 20 54 11  
## Research Director 4 7 32 8  
## Research Scientist 12 48 93 19  
## Sales Executive 8 43 126 23  
## Sales Representative 0 13 34 6

chisqJRWL = chisq.test(tblJRWL)

## Warning in chisq.test(tblJRWL): Chi-squared approximation may be incorrect

chisqJRWL

##   
## Pearson's Chi-squared test  
##   
## data: tblJRWL  
## X-squared = 21.608, df = 24, p-value = 0.6027

dfJIJS = Case2[,c(15,18)]  
tblJIJS = table(dfJIJS$JobInvolvement, dfJIJS$JobSatisfaction)  
tblJIJS

##   
## 1 2 3 4  
## 1 8 8 15 16  
## 2 43 44 68 73  
## 3 106 97 151 160  
## 4 22 17 20 22

chisqJIJS = chisq.test(tblJIJS)  
chisqJIJS

##   
## Pearson's Chi-squared test  
##   
## data: tblJIJS  
## X-squared = 3.9127, df = 9, p-value = 0.9171

dfJRJS = Case2[,c(17,18)]  
tblJRJS = table(dfJRJS$JobRole, dfJRJS$JobSatisfaction)  
tblJRJS

##   
## 1 2 3 4  
## Healthcare Representative 16 9 23 28  
## Human Resources 5 8 8 6  
## Laboratory Technician 32 31 43 47  
## Manager 12 14 12 13  
## Manufacturing Director 12 23 29 23  
## Research Director 13 11 16 11  
## Research Scientist 32 31 48 61  
## Sales Executive 48 25 61 66  
## Sales Representative 9 14 14 16

chisqJRJS = chisq.test(tblJRJS)  
chisqJRJS

##   
## Pearson's Chi-squared test  
##   
## data: tblJRJS  
## X-squared = 26.048, df = 24, p-value = 0.3507

dfWLJS = Case2[,c(32,18)]  
tblWLJS = table(dfWLJS$WorkLifeBalance, dfWLJS$JobSatisfaction)  
tblWLJS

##   
## 1 2 3 4  
## 1 10 13 12 13  
## 2 30 34 57 71  
## 3 121 100 153 158  
## 4 18 19 32 29

chisqWLJS = chisq.test(tblWLJS)  
chisqWLJS

##   
## Pearson's Chi-squared test  
##   
## data: tblWLJS  
## X-squared = 9.0913, df = 9, p-value = 0.4289

### Analyze data further to answer following questions to better understand breakdown of:

### Which Job Roles have highest Job Involvement

### Which Job Roles have highest Job Satisfaction

### Summary notes before building prediction models:

### Notable variables to analyze further: JobRole, JobInvolvement, WorkLifeBalance, MonthlyIncome, TotalWorkingYears, YearsInCurrentRole, JobSatisfaction, JobLevel

### High correlation between the following: JobLevel & MonthlyIncome, JobLevel & TotalWorkingYears, TotalWorkingYears & MonthlyIncome

### Therefore when testing model don’t use the aforementioned pairs together

### Highest continuous correlation with Attrition (highest to lowest) = TotalWorkingYears, JobLevel, YearsInCurrentRole, MonthlyIncome, JobSatisfaction, DistanceFromHome

### Highest categorical correlation with Attrition is the following from high to low: JobInvolvement, JobRole, WorkLifeBalance, JobSatisfaction

### KNN Prediction Model

### Had difficulty generating adequate Sensitivity and Specificity so tested multiple different combinations and also created new variables in order to further test for such improvements

### New variables created =

### JI1 = Seperate Job Involvment = 1 from others (1 or 0)

### JS1, JL1, WLB1 = JobSatisfaction, JobLevel and WorkLifeBalance based on same criteria

### Then created several variables based on the number of iterations of the aforementioned variables that had one of them equal to 1. Trying to consolidate analysis on these variables that appear to have correlation with Attrition.

library(class)

## Warning: package 'class' was built under R version 4.0.5

library(caret)

## Warning: package 'caret' was built under R version 4.0.5

## Loading required package: lattice

##   
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':  
##   
## lift

library(e1071)

## Warning: package 'e1071' was built under R version 4.0.5

set.seed(8)  
splitPerc = .7  
trainIndices = sample(1:dim(Case2)[1],round(splitPerc \* dim(Case2)[1]))  
train = Case2[trainIndices,]  
test = Case2[-trainIndices,]  
  
# test iterations = MonthlyIncome, JI\_JL\_JS  
classifications = knn(train[,c(30,55)],test[,c(30,55)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8199234

# test iterations = MonthlyIncome, JI\_JL\_JS  
classifications = knn(train[,c(30,48)],test[,c(30,48)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8199234

# test iterations = JobRole, MonthlyIncome, JobLevel  
classifications = knn(train[,c(38,40,15)],test[,c(38,40,15)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8199234

# test iterations = JobRole, MonthlyIncome, JobInvolvement  
classifications = knn(train[,c(38,20,15)],test[,c(38,20,15)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8199234

# test iterations = JobRole, TotalWorkingYears, JobInvolvement  
classifications = knn(train[,c(38,30,15)],test[,c(38,30,15)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8199234

# test iterations = JobRole, WorkLifeBalance, JobInvolvement  
classifications = knn(train[,c(38,32,15)],test[,c(38,32,15)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8199234

# test iterations = JobRole, MonthlyIncome  
classifications = knn(train[,c(38,20)],test[,c(38,20)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8199234

# test iterations = JobInvolvement, JobSatisfaction, JobLevel  
classifications = knn(train[,c(15,18,16)],test[,c(15,18,16)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 46  
## Yes 0 1

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 46  
## Yes 0 1  
##   
## Accuracy : 0.8238   
## 95% CI : (0.772, 0.868)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.4747   
##   
## Kappa : 0.0344   
##   
## Mcnemar's Test P-Value : 3.247e-11   
##   
## Sensitivity : 1.00000   
## Specificity : 0.02128   
## Pos Pred Value : 0.82308   
## Neg Pred Value : 1.00000   
## Prevalence : 0.81992   
## Detection Rate : 0.81992   
## Detection Prevalence : 0.99617   
## Balanced Accuracy : 0.51064   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8237548

# test iterations = JobLevel, Job Satisfaction  
classifications = knn(train[,c(16,18)],test[,c(16,18)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.8199234

# test iterations = MonthlyIncome, Job Satisfaction  
classifications = knn(train[,c(20,45)],test[,c(20,45)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 214 47  
## Yes 0 0

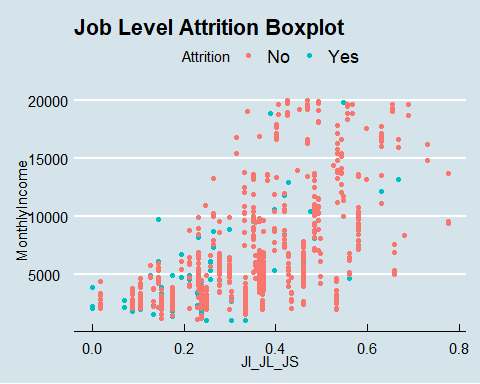
confusionMatrix(table(classifications,test$Attrition))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 214 47  
## Yes 0 0  
##   
## Accuracy : 0.8199   
## 95% CI : (0.7678, 0.8646)  
## No Information Rate : 0.8199   
## P-Value [Acc > NIR] : 0.5389   
##   
## Kappa : 0   
##   
## Mcnemar's Test P-Value : 1.949e-11   
##   
## Sensitivity : 1.0000   
## Specificity : 0.0000   
## Pos Pred Value : 0.8199   
## Neg Pred Value : NaN   
## Prevalence : 0.8199   
## Detection Rate : 0.8199   
## Detection Prevalence : 1.0000   
## Balanced Accuracy : 0.5000   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,test$Attrition))  
CM$overall[1]

## Accuracy   
## 0.8199234

ggplot(data=Case2, mapping = aes(x = JI\_JL\_JS, y = MonthlyIncome, color = Attrition)) + geom\_point() + ggtitle("Job Level Attrition Boxplot") + theme\_economist()



### KNN attempt number 2 after narrowing down the sample set based on select variables of which one of them equals 1

library(class)  
library(caret)  
library(e1071)  
  
Case2\_adj = Case2 %>% filter(JobInvolvement == 1 | JobLevel == 1 | JobSatisfaction == 1 | WorkLifeBalance == 1)  
  
  
set.seed(8)  
splitPerc = .7  
trainIndices = sample(1:dim(Case2\_adj)[1],round(splitPerc \* dim(Case2\_adj)[1]))  
train = Case2\_adj[trainIndices,]  
test = Case2\_adj[-trainIndices,]  
  
# test iterations = MonthlyIncome, JobRole  
classifications = knn(train[,c(20,38)],test[,c(20,38)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 107 35  
## Yes 4 1

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 107 35  
## Yes 4 1  
##   
## Accuracy : 0.7347   
## 95% CI : (0.6556, 0.804)  
## No Information Rate : 0.7551   
## P-Value [Acc > NIR] : 0.7518   
##   
## Kappa : -0.0116   
##   
## Mcnemar's Test P-Value : 1.556e-06   
##   
## Sensitivity : 0.96396   
## Specificity : 0.02778   
## Pos Pred Value : 0.75352   
## Neg Pred Value : 0.20000   
## Prevalence : 0.75510   
## Detection Rate : 0.72789   
## Detection Prevalence : 0.96599   
## Balanced Accuracy : 0.49587   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.7346939

# test iterations = MonthlyIncome, TotalWorkingYears  
classifications = knn(train[,c(20,30)],test[,c(20,30)],train$Attrition, prob = TRUE, k = 15)  
t = table(classifications,test$Attrition)  
t

##   
## classifications No Yes  
## No 107 35  
## Yes 4 1

confusionMatrix(table(classifications,as.factor(test$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 107 35  
## Yes 4 1  
##   
## Accuracy : 0.7347   
## 95% CI : (0.6556, 0.804)  
## No Information Rate : 0.7551   
## P-Value [Acc > NIR] : 0.7518   
##   
## Kappa : -0.0116   
##   
## Mcnemar's Test P-Value : 1.556e-06   
##   
## Sensitivity : 0.96396   
## Specificity : 0.02778   
## Pos Pred Value : 0.75352   
## Neg Pred Value : 0.20000   
## Prevalence : 0.75510   
## Detection Rate : 0.72789   
## Detection Prevalence : 0.96599   
## Balanced Accuracy : 0.49587   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(test$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.7346939

# internal cross validation test iterations = MonthlyIncome, TotalWorkingYears  
classifications = knn.cv(Case2\_adj[,c(20,30)],Case2\_adj$Attrition, k = 15)  
t = table(classifications,Case2\_adj$Attrition)  
t

##   
## classifications No Yes  
## No 364 113  
## Yes 12 0

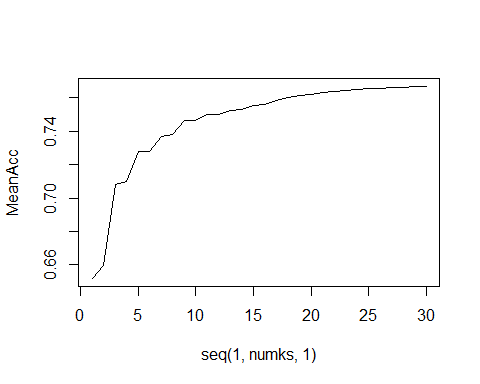
confusionMatrix(table(classifications,as.factor(Case2\_adj$Attrition)))

## Confusion Matrix and Statistics  
##   
##   
## classifications No Yes  
## No 364 113  
## Yes 12 0  
##   
## Accuracy : 0.7444   
## 95% CI : (0.7033, 0.7825)  
## No Information Rate : 0.7689   
## P-Value [Acc > NIR] : 0.9089   
##   
## Kappa : -0.0464   
##   
## Mcnemar's Test P-Value : <2e-16   
##   
## Sensitivity : 0.9681   
## Specificity : 0.0000   
## Pos Pred Value : 0.7631   
## Neg Pred Value : 0.0000   
## Prevalence : 0.7689   
## Detection Rate : 0.7444   
## Detection Prevalence : 0.9755   
## Balanced Accuracy : 0.4840   
##   
## 'Positive' Class : No   
##

CM = confusionMatrix(table(classifications,as.factor(Case2\_adj$Attrition)))  
CM$overall[1]

## Accuracy   
## 0.7443763

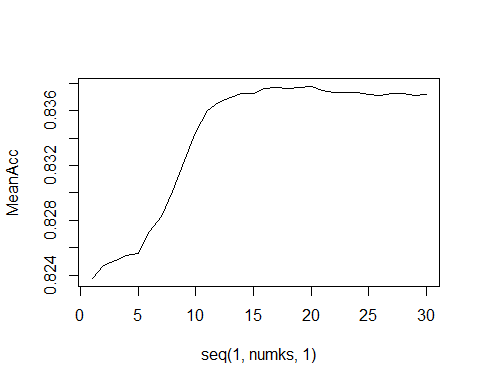
library(class)  
library(caret)  
library(e1071)  
  
Case2\_adj = Case2 %>% filter(JobInvolvement == 1 | JobLevel == 1 | JobSatisfaction == 1 | WorkLifeBalance == 1)  
  
iterations = 500  
numks = 30  
masterAcc = matrix(nrow = iterations, ncol = numks)  
   
for(j in 1:iterations)  
{  
accs = data.frame(accuracy = numeric(30), k = numeric(30))  
trainIndices = sample(1:dim(Case2\_adj)[1],round(splitPerc \* dim(Case2\_adj)[1]))  
train = Case2\_adj[trainIndices,]  
test = Case2\_adj[-trainIndices,]  
for(i in 1:numks)  
{  
 classifications = knn(train[,c(20,30)],test[,c(20,30)],train$Attrition, prob = TRUE, k = i)  
 table(classifications,test$Attrition)  
 CM = confusionMatrix(table(classifications,test$Attrition))  
 masterAcc[j,i] = CM$overall[1]  
}  
}  
MeanAcc = colMeans(masterAcc)  
plot(seq(1,numks,1),MeanAcc, type = "l")



### Repeat these two models with more iterations to determine best k amount to use

### Based on the graph, k = 15 looks ideal as the accuracy is near the peak yet is mindful of cost of each k for client

set.seed(8)  
splitPerc = .7  
trainIndices = sample(1:dim(Case2)[1],round(splitPerc \* dim(Case2)[1]))  
train = Case2[trainIndices,]  
test = Case2[-trainIndices,]  
  
iterations = 500  
numks = 30  
masterAcc = matrix(nrow = iterations, ncol = numks)  
   
for(j in 1:iterations)  
{  
accs = data.frame(accuracy = numeric(30), k = numeric(30))  
trainIndices = sample(1:dim(Case2)[1],round(splitPerc \* dim(Case2)[1]))  
train = Case2[trainIndices,]  
test = Case2[-trainIndices,]  
for(i in 1:numks)  
{  
 classifications = knn(train[,c(30,58)],test[,c(30,58)],train$Attrition, prob = TRUE, k = i)  
 table(classifications,test$Attrition)  
 CM = confusionMatrix(table(classifications,test$Attrition))  
 masterAcc[j,i] = CM$overall[1]  
}  
}  
MeanAcc = colMeans(masterAcc)  
plot(seq(1,numks,1),MeanAcc, type = "l")



### Unable to model data with sufficient efficacy using KNN so moved on to NaiveBayes

model = naiveBayes(Case2[,c(17,55)],Case2$Attrition,laplace = 1)  
table(predict(model,Case2[,c(17,55)]),Case2$Attrition)

##   
## No Yes  
## No 699 123  
## Yes 31 17

CM = confusionMatrix(table(predict(model,Case2[,c(17,55)]),Case2$Attrition))  
CM

## Confusion Matrix and Statistics  
##   
##   
## No Yes  
## No 699 123  
## Yes 31 17  
##   
## Accuracy : 0.823   
## 95% CI : (0.796, 0.8478)  
## No Information Rate : 0.8391   
## P-Value [Acc > NIR] : 0.9083   
##   
## Kappa : 0.1075   
##   
## Mcnemar's Test P-Value : 2.251e-13   
##   
## Sensitivity : 0.9575   
## Specificity : 0.1214   
## Pos Pred Value : 0.8504   
## Neg Pred Value : 0.3542   
## Prevalence : 0.8391   
## Detection Rate : 0.8034   
## Detection Prevalence : 0.9448   
## Balanced Accuracy : 0.5395   
##   
## 'Positive' Class : No   
##

model = naiveBayes(Case2[,c(17,56)],Case2$Attrition,laplace = 1)  
table(predict(model,Case2[,c(17,56)]),Case2$Attrition)

##   
## No Yes  
## No 680 108  
## Yes 50 32

CM = confusionMatrix(table(predict(model,Case2[,c(17,56)]),Case2$Attrition))  
CM

## Confusion Matrix and Statistics  
##   
##   
## No Yes  
## No 680 108  
## Yes 50 32  
##   
## Accuracy : 0.8184   
## 95% CI : (0.7911, 0.8435)  
## No Information Rate : 0.8391   
## P-Value [Acc > NIR] : 0.9543   
##   
## Kappa : 0.1923   
##   
## Mcnemar's Test P-Value : 5.769e-06   
##   
## Sensitivity : 0.9315   
## Specificity : 0.2286   
## Pos Pred Value : 0.8629   
## Neg Pred Value : 0.3902   
## Prevalence : 0.8391   
## Detection Rate : 0.7816   
## Detection Prevalence : 0.9057   
## Balanced Accuracy : 0.5800   
##   
## 'Positive' Class : No   
##

model = naiveBayes(Case2[,c(30,55)],Case2$Attrition,laplace = 1)  
table(predict(model,Case2[,c(30,55)]),Case2$Attrition)

##   
## No Yes  
## No 699 123  
## Yes 31 17

CM = confusionMatrix(table(predict(model,Case2[,c(30,55)]),Case2$Attrition))  
CM

## Confusion Matrix and Statistics  
##   
##   
## No Yes  
## No 699 123  
## Yes 31 17  
##   
## Accuracy : 0.823   
## 95% CI : (0.796, 0.8478)  
## No Information Rate : 0.8391   
## P-Value [Acc > NIR] : 0.9083   
##   
## Kappa : 0.1075   
##   
## Mcnemar's Test P-Value : 2.251e-13   
##   
## Sensitivity : 0.9575   
## Specificity : 0.1214   
## Pos Pred Value : 0.8504   
## Neg Pred Value : 0.3542   
## Prevalence : 0.8391   
## Detection Rate : 0.8034   
## Detection Prevalence : 0.9448   
## Balanced Accuracy : 0.5395   
##   
## 'Positive' Class : No   
##

### These last 2 were the best NB models I generated thus far

model = naiveBayes(Case2[,c(30,58)],Case2$Attrition,laplace = 1)  
table(predict(model,Case2[,c(30,58)]),Case2$Attrition)

##   
## No Yes  
## No 667 98  
## Yes 63 42

CM = confusionMatrix(table(predict(model,Case2[,c(30,58)]),Case2$Attrition))  
CM

## Confusion Matrix and Statistics  
##   
##   
## No Yes  
## No 667 98  
## Yes 63 42  
##   
## Accuracy : 0.8149   
## 95% CI : (0.7875, 0.8402)  
## No Information Rate : 0.8391   
## P-Value [Acc > NIR] : 0.974729   
##   
## Kappa : 0.2377   
##   
## Mcnemar's Test P-Value : 0.007372   
##   
## Sensitivity : 0.9137   
## Specificity : 0.3000   
## Pos Pred Value : 0.8719   
## Neg Pred Value : 0.4000   
## Prevalence : 0.8391   
## Detection Rate : 0.7667   
## Detection Prevalence : 0.8793   
## Balanced Accuracy : 0.6068   
##   
## 'Positive' Class : No   
##

model = naiveBayes(Case2[,c(17,58)],Case2$Attrition,laplace = 1)  
table(predict(model,Case2[,c(17,58)]),Case2$Attrition)

##   
## No Yes  
## No 667 98  
## Yes 63 42

CM = confusionMatrix(table(predict(model,Case2[,c(17,58)]),Case2$Attrition))  
CM

## Confusion Matrix and Statistics  
##   
##   
## No Yes  
## No 667 98  
## Yes 63 42  
##   
## Accuracy : 0.8149   
## 95% CI : (0.7875, 0.8402)  
## No Information Rate : 0.8391   
## P-Value [Acc > NIR] : 0.974729   
##   
## Kappa : 0.2377   
##   
## Mcnemar's Test P-Value : 0.007372   
##   
## Sensitivity : 0.9137   
## Specificity : 0.3000   
## Pos Pred Value : 0.8719   
## Neg Pred Value : 0.4000   
## Prevalence : 0.8391   
## Detection Rate : 0.7667   
## Detection Prevalence : 0.8793   
## Balanced Accuracy : 0.6068   
##   
## 'Positive' Class : No   
##

### Repeat NB model with aforementioned adjustments above

Case2\_adj = Case2 %>% filter(JobInvolvement == 1 | JobLevel == 1 | JobSatisfaction == 1 | WorkLifeBalance == 1)  
  
set.seed(8)  
splitPerc = .7  
trainIndices = sample(1:dim(Case2\_adj)[1],round(splitPerc \* dim(Case2\_adj)[1]))  
train = Case2\_adj[trainIndices,]  
test = Case2\_adj[-trainIndices,]  
  
model = naiveBayes(Case2\_adj[,c(30,58)],Case2\_adj$Attrition,laplace = 1)  
table(predict(model,Case2\_adj[,c(30,58)]),Case2\_adj$Attrition)

##   
## No Yes  
## No 323 75  
## Yes 53 38

CM = confusionMatrix(table(predict(model,Case2\_adj[,c(30,58)]),Case2\_adj$Attrition))  
CM

## Confusion Matrix and Statistics  
##   
##   
## No Yes  
## No 323 75  
## Yes 53 38  
##   
## Accuracy : 0.7382   
## 95% CI : (0.6969, 0.7767)  
## No Information Rate : 0.7689   
## P-Value [Acc > NIR] : 0.95024   
##   
## Kappa : 0.2096   
##   
## Mcnemar's Test P-Value : 0.06343   
##   
## Sensitivity : 0.8590   
## Specificity : 0.3363   
## Pos Pred Value : 0.8116   
## Neg Pred Value : 0.4176   
## Prevalence : 0.7689   
## Detection Rate : 0.6605   
## Detection Prevalence : 0.8139   
## Balanced Accuracy : 0.5977   
##   
## 'Positive' Class : No   
##

model

##   
## Naive Bayes Classifier for Discrete Predictors  
##   
## Call:  
## naiveBayes.default(x = Case2\_adj[, c(30, 58)], y = Case2\_adj$Attrition,   
## laplace = 1)  
##   
## A-priori probabilities:  
## Case2\_adj$Attrition  
## No Yes   
## 0.7689162 0.2310838   
##   
## Conditional probabilities:  
## TotalWorkingYears  
## Case2\_adj$Attrition [,1] [,2]  
## No 9.101064 6.509513  
## Yes 7.000000 6.587325  
##   
## JIJSJLWLB\_1  
## Case2\_adj$Attrition [,1] [,2]  
## No 0.1675532 0.3739667  
## Yes 0.3716814 0.4854065

### This was the highest I was able to get Specificity with my models so am going with this one  
### Build model then load select columns and write to csv  
model1 = naiveBayes(Case2\_adj[,c(17,58)],Case2\_adj$Attrition,laplace = 1)  
summary(model1)

## Length Class Mode   
## apriori 2 table numeric   
## tables 2 -none- list   
## levels 2 -none- character  
## isnumeric 2 -none- logical   
## call 4 -none- call

NBmodel = table(predict(model1,Case2\_adj[,c(17,58)]),Case2\_adj$Attrition)  
NBmodelPred = predict(model1, newdata = test)  
NBmodelPred

## [1] Yes No Yes No No No No Yes Yes No Yes Yes No No No No No No   
## [19] Yes Yes No No No Yes Yes No No No Yes No Yes No No No No No   
## [37] No Yes No No No No No Yes No No No No No Yes Yes No Yes No   
## [55] No No Yes Yes Yes No No No No No No No Yes Yes No No No No   
## [73] No No No No No No No Yes No No No No No Yes No No No Yes  
## [91] No No No No No No No No No No No Yes No Yes No No Yes No   
## [109] No No No No No No No No Yes Yes No Yes No Yes No No No No   
## [127] No No No No Yes No No No Yes No No Yes No Yes No No No No   
## [145] No No No   
## Levels: No Yes

NBM = as.data.frame(NBmodelPred)  
NBM

## NBmodelPred  
## 1 Yes  
## 2 No  
## 3 Yes  
## 4 No  
## 5 No  
## 6 No  
## 7 No  
## 8 Yes  
## 9 Yes  
## 10 No  
## 11 Yes  
## 12 Yes  
## 13 No  
## 14 No  
## 15 No  
## 16 No  
## 17 No  
## 18 No  
## 19 Yes  
## 20 Yes  
## 21 No  
## 22 No  
## 23 No  
## 24 Yes  
## 25 Yes  
## 26 No  
## 27 No  
## 28 No  
## 29 Yes  
## 30 No  
## 31 Yes  
## 32 No  
## 33 No  
## 34 No  
## 35 No  
## 36 No  
## 37 No  
## 38 Yes  
## 39 No  
## 40 No  
## 41 No  
## 42 No  
## 43 No  
## 44 Yes  
## 45 No  
## 46 No  
## 47 No  
## 48 No  
## 49 No  
## 50 Yes  
## 51 Yes  
## 52 No  
## 53 Yes  
## 54 No  
## 55 No  
## 56 No  
## 57 Yes  
## 58 Yes  
## 59 Yes  
## 60 No  
## 61 No  
## 62 No  
## 63 No  
## 64 No  
## 65 No  
## 66 No  
## 67 Yes  
## 68 Yes  
## 69 No  
## 70 No  
## 71 No  
## 72 No  
## 73 No  
## 74 No  
## 75 No  
## 76 No  
## 77 No  
## 78 No  
## 79 No  
## 80 Yes  
## 81 No  
## 82 No  
## 83 No  
## 84 No  
## 85 No  
## 86 Yes  
## 87 No  
## 88 No  
## 89 No  
## 90 Yes  
## 91 No  
## 92 No  
## 93 No  
## 94 No  
## 95 No  
## 96 No  
## 97 No  
## 98 No  
## 99 No  
## 100 No  
## 101 No  
## 102 Yes  
## 103 No  
## 104 Yes  
## 105 No  
## 106 No  
## 107 Yes  
## 108 No  
## 109 No  
## 110 No  
## 111 No  
## 112 No  
## 113 No  
## 114 No  
## 115 No  
## 116 No  
## 117 Yes  
## 118 Yes  
## 119 No  
## 120 Yes  
## 121 No  
## 122 Yes  
## 123 No  
## 124 No  
## 125 No  
## 126 No  
## 127 No  
## 128 No  
## 129 No  
## 130 No  
## 131 Yes  
## 132 No  
## 133 No  
## 134 No  
## 135 Yes  
## 136 No  
## 137 No  
## 138 Yes  
## 139 No  
## 140 Yes  
## 141 No  
## 142 No  
## 143 No  
## 144 No  
## 145 No  
## 146 No  
## 147 No

test

## ID Age Attrition BusinessTravel DailyRate Department  
## 1 254 31 Yes Travel\_Rarely 359 Human Resources  
## 2 398 34 Yes Travel\_Rarely 1107 Human Resources  
## 4 716 24 Yes Travel\_Rarely 240 Human Resources  
## 5 733 27 Yes Travel\_Frequently 1337 Human Resources  
## 9 81 37 Yes Travel\_Rarely 1141 Research & Development  
## 15 162 28 Yes Non-Travel 1366 Research & Development  
## 16 177 55 Yes Travel\_Rarely 725 Research & Development  
## 17 178 26 Yes Travel\_Frequently 342 Research & Development  
## 18 204 28 Yes Travel\_Rarely 654 Research & Development  
## 23 287 22 Yes Travel\_Rarely 391 Research & Development  
## 25 299 31 Yes Non-Travel 335 Research & Development  
## 28 313 31 Yes Travel\_Frequently 561 Research & Development  
## 29 333 44 Yes Travel\_Frequently 429 Research & Development  
## 32 380 30 Yes Travel\_Rarely 138 Research & Development  
## 33 385 51 Yes Travel\_Rarely 1323 Research & Development  
## 35 433 49 Yes Travel\_Frequently 1475 Research & Development  
## 42 520 56 Yes Travel\_Rarely 1162 Research & Development  
## 44 530 56 Yes Travel\_Rarely 441 Research & Development  
## 48 613 33 Yes Travel\_Rarely 465 Research & Development  
## 50 625 28 Yes Travel\_Frequently 289 Research & Development  
## 55 655 22 Yes Travel\_Frequently 1368 Research & Development  
## 70 28 33 Yes Travel\_Rarely 603 Sales  
## 73 38 36 Yes Travel\_Rarely 1456 Sales  
## 84 244 25 Yes Travel\_Rarely 383 Sales  
## 87 362 29 Yes Travel\_Frequently 746 Sales  
## 88 363 52 Yes Travel\_Rarely 266 Sales  
## 89 390 21 Yes Travel\_Frequently 756 Sales  
## 92 439 50 Yes Travel\_Rarely 869 Sales  
## 93 469 26 Yes Non-Travel 265 Sales  
## 99 539 24 Yes Travel\_Rarely 1448 Sales  
## 103 605 29 Yes Travel\_Rarely 428 Sales  
## 104 614 19 Yes Travel\_Rarely 419 Sales  
## 106 699 18 Yes Travel\_Frequently 544 Sales  
## 107 751 20 Yes Travel\_Frequently 769 Sales  
## 109 760 45 Yes Travel\_Rarely 1449 Sales  
## 113 851 31 Yes Travel\_Rarely 542 Sales  
## 114 12 28 No Non-Travel 280 Human Resources  
## 115 15 46 No Travel\_Rarely 991 Human Resources  
## 116 239 25 No Travel\_Rarely 309 Human Resources  
## 123 550 45 No Travel\_Rarely 788 Human Resources  
## 125 690 38 No Non-Travel 1336 Human Resources  
## 126 701 29 No Travel\_Rarely 352 Human Resources  
## 127 705 59 No Travel\_Rarely 818 Human Resources  
## 129 765 34 No Travel\_Rarely 829 Human Resources  
## 130 777 30 No Travel\_Rarely 330 Human Resources  
## 134 6 27 No Travel\_Frequently 294 Research & Development  
## 135 7 41 No Travel\_Rarely 1283 Research & Development  
## 137 19 34 No Travel\_Rarely 181 Research & Development  
## 141 34 35 No Travel\_Rarely 982 Research & Development  
## 142 36 32 No Travel\_Frequently 1311 Research & Development  
## 143 43 49 No Travel\_Rarely 464 Research & Development  
## 145 46 42 No Travel\_Frequently 748 Research & Development  
## 147 51 38 No Travel\_Rarely 362 Research & Development  
## 148 52 58 No Travel\_Rarely 1055 Research & Development  
## 149 53 44 No Travel\_Rarely 1117 Research & Development  
## 153 66 32 No Travel\_Rarely 334 Research & Development  
## 156 69 35 No Travel\_Frequently 664 Research & Development  
## 165 101 27 No Travel\_Rarely 1377 Research & Development  
## 166 104 23 No Travel\_Rarely 885 Research & Development  
## 169 111 28 No Travel\_Frequently 773 Research & Development  
## 173 129 43 No Travel\_Rarely 930 Research & Development  
## 176 149 24 No Travel\_Rarely 477 Research & Development  
## 182 164 42 No Travel\_Rarely 916 Research & Development  
## 184 170 29 No Travel\_Rarely 1107 Research & Development  
## 186 183 43 No Travel\_Frequently 957 Research & Development  
## 187 185 39 No Travel\_Frequently 443 Research & Development  
## 193 206 44 No Travel\_Rarely 661 Research & Development  
## 194 209 30 No Travel\_Rarely 201 Research & Development  
## 195 213 37 No Travel\_Rarely 1225 Research & Development  
## 198 219 30 No Travel\_Rarely 438 Research & Development  
## 208 246 29 No Travel\_Rarely 1086 Research & Development  
## 212 260 25 No Travel\_Rarely 977 Research & Development  
## 214 266 35 No Travel\_Rarely 809 Research & Development  
## 215 267 34 No Travel\_Frequently 829 Research & Development  
## 216 270 30 No Travel\_Rarely 1176 Research & Development  
## 227 306 24 No Non-Travel 1269 Research & Development  
## 233 323 30 No Travel\_Frequently 1012 Research & Development  
## 234 324 46 No Travel\_Rarely 1003 Research & Development  
## 237 334 27 No Travel\_Rarely 608 Research & Development  
## 238 340 27 No Travel\_Rarely 1240 Research & Development  
## 240 342 28 No Travel\_Frequently 193 Research & Development  
## 245 353 40 No Travel\_Frequently 902 Research & Development  
## 246 357 26 No Travel\_Rarely 841 Research & Development  
## 250 371 31 No Travel\_Frequently 853 Research & Development  
## 253 376 50 No Travel\_Frequently 333 Research & Development  
## 258 396 31 No Travel\_Rarely 192 Research & Development  
## 270 427 30 No Non-Travel 990 Research & Development  
## 271 428 21 No Travel\_Rarely 996 Research & Development  
## 277 446 47 No Travel\_Rarely 465 Research & Development  
## 280 455 48 No Travel\_Rarely 1236 Research & Development  
## 283 462 36 No Travel\_Rarely 311 Research & Development  
## 284 466 38 No Travel\_Rarely 371 Research & Development  
## 286 470 53 No Travel\_Rarely 1084 Research & Development  
## 288 473 45 No Travel\_Rarely 193 Research & Development  
## 292 481 29 No Travel\_Rarely 136 Research & Development  
## 293 482 31 No Travel\_Rarely 1222 Research & Development  
## 295 488 31 No Travel\_Rarely 670 Research & Development  
## 297 492 50 No Travel\_Rarely 1322 Research & Development  
## 305 521 34 No Travel\_Frequently 702 Research & Development  
## 307 525 59 No Travel\_Rarely 142 Research & Development  
## 311 536 43 No Non-Travel 1344 Research & Development  
## 312 542 33 No Travel\_Rarely 867 Research & Development  
## 320 563 27 No Travel\_Rarely 1134 Research & Development  
## 330 590 35 No Travel\_Rarely 1142 Research & Development  
## 331 591 22 No Non-Travel 457 Research & Development  
## 335 604 29 No Travel\_Rarely 1401 Research & Development  
## 336 609 26 No Travel\_Rarely 652 Research & Development  
## 337 617 33 No Travel\_Rarely 1099 Research & Development  
## 339 620 35 No Non-Travel 727 Research & Development  
## 342 631 55 No Travel\_Rarely 836 Research & Development  
## 345 636 35 No Travel\_Frequently 636 Research & Development  
## 349 642 40 No Travel\_Rarely 750 Research & Development  
## 353 656 26 No Travel\_Frequently 496 Research & Development  
## 359 684 57 No Travel\_Rarely 334 Research & Development  
## 362 709 38 No Travel\_Rarely 1153 Research & Development  
## 365 713 48 No Travel\_Rarely 969 Research & Development  
## 366 723 34 No Travel\_Frequently 560 Research & Development  
## 372 746 24 No Travel\_Rarely 350 Research & Development  
## 375 750 33 No Travel\_Rarely 134 Research & Development  
## 377 756 44 No Non-Travel 381 Research & Development  
## 379 759 38 No Travel\_Rarely 1495 Research & Development  
## 380 766 45 No Travel\_Rarely 1015 Research & Development  
## 386 790 25 No Travel\_Rarely 266 Research & Development  
## 389 800 36 No Travel\_Rarely 1040 Research & Development  
## 393 809 30 No Travel\_Rarely 793 Research & Development  
## 398 819 30 No Travel\_Rarely 1427 Research & Development  
## 400 829 38 No Travel\_Frequently 594 Research & Development  
## 404 837 37 No Travel\_Rarely 1439 Research & Development  
## 405 846 28 No Travel\_Rarely 1300 Research & Development  
## 407 850 26 No Travel\_Frequently 1096 Research & Development  
## 410 859 41 No Travel\_Rarely 933 Research & Development  
## 414 14 30 No Travel\_Rarely 202 Sales  
## 420 63 31 No Travel\_Rarely 1154 Sales  
## 424 90 46 No Travel\_Rarely 705 Sales  
## 429 148 33 No Travel\_Rarely 1242 Sales  
## 432 205 32 No Travel\_Rarely 1401 Sales  
## 444 317 34 No Travel\_Frequently 303 Sales  
## 447 356 34 No Travel\_Rarely 1326 Sales  
## 452 383 48 No Travel\_Rarely 1221 Sales  
## 453 384 23 No Travel\_Rarely 541 Sales  
## 454 391 34 No Travel\_Rarely 1111 Sales  
## 468 608 38 No Travel\_Rarely 322 Sales  
## 473 673 28 No Travel\_Frequently 467 Sales  
## 477 696 47 No Travel\_Rarely 1454 Sales  
## 480 763 41 No Non-Travel 256 Sales  
## 484 785 47 No Non-Travel 543 Sales  
## 487 828 34 No Travel\_Rarely 131 Sales  
## DistanceFromHome Education EducationField EmployeeCount EmployeeNumber  
## 1 18 5 Human Resources 1 1842  
## 2 9 4 Technical Degree 1 1467  
## 4 22 1 Human Resources 1 1714  
## 5 22 3 Human Resources 1 1944  
## 9 11 2 Medical 1 1033  
## 15 24 2 Technical Degree 1 1082  
## 16 2 3 Medical 1 787  
## 17 2 3 Life Sciences 1 1053  
## 18 1 2 Life Sciences 1 741  
## 23 7 1 Life Sciences 1 1878  
## 25 9 2 Medical 1 991  
## 28 3 3 Life Sciences 1 1537  
## 29 1 2 Medical 1 1792  
## 32 22 3 Life Sciences 1 1004  
## 33 4 4 Life Sciences 1 1081  
## 35 28 2 Life Sciences 1 1420  
## 42 24 2 Life Sciences 1 1907  
## 44 14 4 Life Sciences 1 161  
## 48 2 2 Life Sciences 1 328  
## 50 2 2 Medical 1 1504  
## 55 4 1 Technical Degree 1 593  
## 70 9 4 Marketing 1 1157  
## 73 13 5 Marketing 1 1733  
## 84 9 2 Life Sciences 1 1439  
## 87 24 3 Technical Degree 1 1928  
## 88 2 1 Marketing 1 1038  
## 89 1 1 Technical Degree 1 478  
## 92 3 2 Marketing 1 47  
## 93 29 2 Medical 1 1037  
## 99 1 1 Technical Degree 1 554  
## 103 9 3 Marketing 1 1752  
## 104 21 3 Other 1 959  
## 106 3 2 Medical 1 1624  
## 107 9 3 Marketing 1 1077  
## 109 2 3 Marketing 1 1277  
## 113 20 3 Life Sciences 1 175  
## 114 1 2 Life Sciences 1 1858  
## 115 1 2 Life Sciences 1 1314  
## 116 2 3 Human Resources 1 1987  
## 123 24 4 Medical 1 1049  
## 125 2 3 Human Resources 1 1805  
## 126 6 1 Medical 1 1865  
## 127 6 2 Medical 1 321  
## 129 3 2 Human Resources 1 847  
## 130 1 3 Life Sciences 1 1499  
## 134 10 2 Life Sciences 1 733  
## 135 5 5 Medical 1 1448  
## 137 2 4 Medical 1 1755  
## 141 1 4 Medical 1 1172  
## 142 7 3 Life Sciences 1 359  
## 143 16 3 Medical 1 1674  
## 145 9 2 Medical 1 1480  
## 147 1 1 Life Sciences 1 662  
## 148 1 3 Medical 1 1423  
## 149 2 1 Life Sciences 1 1246  
## 153 5 2 Life Sciences 1 21  
## 156 1 3 Medical 1 88  
## 165 11 1 Life Sciences 1 1434  
## 166 4 3 Medical 1 705  
## 169 6 3 Life Sciences 1 1154  
## 173 6 3 Medical 1 1402  
## 176 24 3 Medical 1 1173  
## 182 17 2 Life Sciences 1 347  
## 184 28 4 Life Sciences 1 1120  
## 186 28 3 Medical 1 171  
## 187 8 1 Life Sciences 1 602  
## 193 9 2 Life Sciences 1 913  
## 194 5 3 Technical Degree 1 197  
## 195 10 2 Life Sciences 1 715  
## 198 18 3 Life Sciences 1 194  
## 208 7 1 Medical 1 912  
## 212 2 1 Other 1 1992  
## 214 16 3 Medical 1 14  
## 215 15 3 Medical 1 1485  
## 216 20 3 Other 1 1084  
## 227 4 1 Life Sciences 1 888  
## 233 5 4 Life Sciences 1 861  
## 234 8 4 Life Sciences 1 1080  
## 237 1 2 Life Sciences 1 725  
## 238 2 4 Life Sciences 1 54  
## 240 2 3 Life Sciences 1 1296  
## 245 26 2 Medical 1 1180  
## 246 6 3 Other 1 164  
## 250 1 1 Life Sciences 1 1011  
## 253 22 5 Medical 1 1539  
## 258 2 4 Life Sciences 1 426  
## 270 7 3 Technical Degree 1 1482  
## 271 3 2 Medical 1 379  
## 277 1 3 Technical Degree 1 1438  
## 280 1 4 Life Sciences 1 664  
## 283 7 3 Life Sciences 1 1659  
## 284 2 3 Life Sciences 1 24  
## 286 13 2 Medical 1 250  
## 288 6 4 Other 1 101  
## 292 1 3 Life Sciences 1 1954  
## 293 11 4 Life Sciences 1 895  
## 295 26 1 Life Sciences 1 16  
## 297 28 3 Life Sciences 1 1317  
## 305 16 4 Life Sciences 1 838  
## 307 3 3 Life Sciences 1 309  
## 311 7 3 Medical 1 262  
## 312 8 4 Life Sciences 1 1798  
## 320 16 4 Technical Degree 1 1001  
## 330 23 4 Medical 1 75  
## 331 26 2 Other 1 1605  
## 335 6 1 Medical 1 1192  
## 336 7 3 Other 1 1417  
## 337 4 4 Medical 1 1502  
## 339 3 3 Life Sciences 1 704  
## 342 2 4 Technical Degree 1 1873  
## 345 4 4 Other 1 1185  
## 349 12 3 Life Sciences 1 1829  
## 353 11 2 Medical 1 390  
## 359 24 2 Life Sciences 1 223  
## 362 6 2 Other 1 1782  
## 365 2 2 Technical Degree 1 1258  
## 366 1 4 Other 1 1431  
## 372 21 2 Technical Degree 1 1551  
## 375 2 3 Life Sciences 1 242  
## 377 24 3 Medical 1 744  
## 379 4 2 Medical 1 1687  
## 380 5 5 Medical 1 1611  
## 386 1 3 Medical 1 1303  
## 389 3 2 Life Sciences 1 1664  
## 393 16 1 Life Sciences 1 1729  
## 398 2 1 Medical 1 198  
## 400 2 2 Medical 1 1760  
## 404 4 1 Life Sciences 1 1394  
## 405 17 2 Medical 1 536  
## 407 6 3 Other 1 1918  
## 410 9 4 Life Sciences 1 200  
## 414 2 1 Technical Degree 1 508  
## 420 2 2 Life Sciences 1 1996  
## 424 2 4 Marketing 1 38  
## 429 8 4 Life Sciences 1 1560  
## 432 4 2 Life Sciences 1 330  
## 444 2 4 Marketing 1 216  
## 447 3 3 Other 1 1478  
## 452 7 3 Marketing 1 1466  
## 453 2 1 Technical Degree 1 113  
## 454 8 2 Life Sciences 1 808  
## 468 7 2 Medical 1 382  
## 473 7 3 Life Sciences 1 1507  
## 477 2 4 Life Sciences 1 925  
## 480 10 2 Medical 1 1329  
## 484 2 4 Marketing 1 1731  
## 487 2 3 Marketing 1 1281  
## EnvironmentSatisfaction Gender HourlyRate JobInvolvement JobLevel  
## 1 4 Male 89 4 1  
## 2 1 Female 52 3 1  
## 4 4 Male 58 1 1  
## 5 1 Female 58 2 1  
## 9 1 Female 61 1 2  
## 15 2 Male 72 2 3  
## 16 4 Male 78 3 5  
## 17 1 Male 57 3 1  
## 18 1 Female 67 1 1  
## 23 4 Male 75 3 1  
## 25 3 Male 46 2 1  
## 28 4 Female 33 3 1  
## 29 3 Male 99 3 1  
## 32 1 Female 48 3 1  
## 33 1 Male 34 3 1  
## 35 1 Male 97 2 2  
## 42 1 Male 97 3 1  
## 44 2 Female 72 3 1  
## 48 1 Female 39 3 1  
## 50 3 Male 38 2 1  
## 55 3 Male 99 2 1  
## 70 1 Female 77 3 2  
## 73 2 Male 96 2 2  
## 84 1 Male 68 2 1  
## 87 3 Male 45 4 1  
## 88 1 Female 57 1 5  
## 89 1 Female 99 2 1  
## 92 1 Male 86 2 1  
## 93 2 Male 79 1 2  
## 99 1 Female 62 3 1  
## 103 2 Female 52 1 1  
## 104 4 Male 37 2 1  
## 106 2 Female 70 3 1  
## 107 4 Female 54 3 1  
## 109 1 Female 94 1 5  
## 113 2 Female 71 1 2  
## 114 3 Male 43 3 1  
## 115 4 Female 44 3 1  
## 116 3 Female 82 3 1  
## 123 2 Male 36 3 1  
## 125 1 Male 100 3 1  
## 126 4 Male 87 2 1  
## 127 2 Male 52 3 1  
## 129 3 Male 88 3 1  
## 130 3 Male 46 3 1  
## 134 4 Male 32 3 3  
## 135 2 Male 90 4 1  
## 137 4 Male 97 4 1  
## 141 4 Male 58 2 1  
## 142 2 Male 100 4 1  
## 143 4 Female 74 3 1  
## 145 1 Female 74 3 1  
## 147 3 Female 43 3 1  
## 148 4 Female 76 3 5  
## 149 1 Female 72 4 1  
## 153 1 Male 80 4 1  
## 156 2 Male 79 3 1  
## 165 2 Male 91 3 1  
## 166 1 Male 58 4 1  
## 169 3 Male 39 2 1  
## 173 1 Female 73 2 2  
## 176 4 Male 49 3 1  
## 182 4 Female 82 4 2  
## 184 3 Female 93 3 1  
## 186 2 Female 72 4 1  
## 187 3 Female 48 3 1  
## 193 2 Male 61 3 1  
## 194 4 Female 84 3 1  
## 195 4 Male 80 4 1  
## 198 1 Female 75 3 1  
## 208 1 Female 62 2 1  
## 212 4 Male 57 3 1  
## 214 1 Male 84 4 1  
## 215 2 Male 71 3 4  
## 216 3 Male 85 3 2  
## 227 1 Male 46 2 1  
## 233 2 Male 75 2 1  
## 234 4 Female 74 2 2  
## 237 3 Female 68 3 3  
## 238 4 Female 33 3 1  
## 240 4 Male 52 2 1  
## 245 3 Female 92 2 2  
## 246 3 Female 46 2 1  
## 250 3 Female 96 3 2  
## 253 3 Male 88 1 4  
## 258 3 Male 32 3 1  
## 270 3 Male 64 3 1  
## 271 4 Male 100 2 1  
## 277 1 Male 74 3 1  
## 280 4 Female 40 2 4  
## 283 1 Male 77 3 1  
## 284 4 Male 45 3 1  
## 286 4 Female 57 4 2  
## 288 4 Male 52 3 3  
## 292 1 Male 89 3 2  
## 293 4 Male 48 3 1  
## 295 1 Male 31 3 1  
## 297 4 Female 43 3 4  
## 305 3 Female 100 2 1  
## 307 3 Male 70 2 1  
## 311 4 Male 37 4 1  
## 312 4 Male 90 4 1  
## 320 3 Female 37 3 1  
## 330 3 Female 30 3 1  
## 331 2 Female 85 2 1  
## 335 2 Female 54 3 1  
## 336 3 Male 100 4 1  
## 337 1 Female 82 2 1  
## 339 3 Male 41 2 1  
## 342 2 Male 98 2 1  
## 345 4 Male 47 2 1  
## 349 2 Female 47 3 2  
## 353 1 Male 60 3 2  
## 359 3 Male 83 4 3  
## 362 4 Female 40 2 1  
## 365 4 Male 76 4 1  
## 366 4 Male 91 3 1  
## 372 3 Male 57 2 1  
## 375 3 Male 90 3 1  
## 377 1 Male 49 1 1  
## 379 4 Female 87 3 1  
## 380 3 Female 50 1 2  
## 386 4 Female 40 3 1  
## 389 4 Male 79 4 2  
## 393 2 Male 33 3 1  
## 398 2 Male 35 2 1  
## 400 3 Female 75 2 1  
## 404 3 Male 54 3 1  
## 405 3 Male 79 3 2  
## 407 3 Male 61 4 1  
## 410 3 Male 94 3 1  
## 414 3 Male 72 3 1  
## 420 1 Male 54 3 1  
## 424 2 Female 83 3 5  
## 429 1 Male 46 3 2  
## 432 3 Female 56 3 1  
## 444 3 Female 75 3 1  
## 447 4 Male 81 1 2  
## 452 3 Male 96 3 2  
## 453 3 Male 62 3 1  
## 454 3 Female 93 3 2  
## 468 1 Female 44 4 2  
## 473 3 Male 55 3 2  
## 477 4 Female 65 2 1  
## 480 3 Male 40 1 2  
## 484 3 Male 87 3 2  
## 487 3 Female 86 3 2  
## JobRole JobSatisfaction MaritalStatus MonthlyIncome  
## 1 Human Resources 1 Married 2956  
## 2 Human Resources 3 Married 2742  
## 4 Human Resources 3 Married 1555  
## 5 Human Resources 2 Married 2863  
## 9 Healthcare Representative 2 Married 4777  
## 15 Healthcare Representative 1 Single 8722  
## 16 Manager 1 Married 19859  
## 17 Research Scientist 1 Married 2042  
## 18 Research Scientist 2 Single 2216  
## 23 Research Scientist 2 Single 2472  
## 25 Research Scientist 1 Single 2321  
## 28 Research Scientist 3 Single 4084  
## 29 Research Scientist 2 Divorced 2342  
## 32 Research Scientist 3 Married 2132  
## 33 Research Scientist 3 Married 2461  
## 35 Laboratory Technician 1 Single 4284  
## 42 Laboratory Technician 4 Single 2587  
## 44 Research Scientist 2 Married 4963  
## 48 Laboratory Technician 1 Married 2707  
## 50 Laboratory Technician 1 Single 2561  
## 55 Laboratory Technician 3 Single 3894  
## 70 Sales Executive 1 Single 8224  
## 73 Sales Executive 1 Divorced 6134  
## 84 Sales Representative 1 Married 4400  
## 87 Sales Representative 1 Single 1091  
## 88 Manager 4 Married 19845  
## 89 Sales Representative 2 Single 2174  
## 92 Sales Representative 3 Married 2683  
## 93 Sales Executive 1 Single 4969  
## 99 Sales Representative 2 Single 3202  
## 103 Sales Representative 2 Single 2760  
## 104 Sales Representative 2 Single 2121  
## 106 Sales Representative 4 Single 1569  
## 107 Sales Representative 4 Single 2323  
## 109 Manager 2 Single 18824  
## 113 Sales Executive 3 Married 4559  
## 114 Human Resources 4 Divorced 2706  
## 115 Human Resources 1 Single 3423  
## 116 Human Resources 2 Married 2187  
## 123 Human Resources 2 Single 2177  
## 125 Human Resources 2 Divorced 2592  
## 126 Human Resources 2 Married 2804  
## 127 Human Resources 3 Married 2267  
## 129 Human Resources 4 Married 3737  
## 130 Human Resources 3 Divorced 2064  
## 134 Manufacturing Director 1 Divorced 8793  
## 135 Research Scientist 3 Married 2127  
## 137 Research Scientist 4 Married 2932  
## 141 Laboratory Technician 3 Married 2258  
## 142 Laboratory Technician 2 Married 2794  
## 143 Laboratory Technician 1 Divorced 2587  
## 145 Laboratory Technician 4 Single 3673  
## 147 Research Scientist 1 Single 2619  
## 148 Research Director 1 Married 19701  
## 149 Research Scientist 4 Married 2011  
## 153 Research Scientist 2 Divorced 3298  
## 156 Research Scientist 1 Married 2194  
## 165 Laboratory Technician 1 Married 2099  
## 166 Research Scientist 1 Married 2819  
## 169 Research Scientist 3 Divorced 2703  
## 173 Research Scientist 3 Single 4081  
## 176 Laboratory Technician 2 Single 3597  
## 182 Research Scientist 1 Single 6545  
## 184 Research Scientist 4 Divorced 2514  
## 186 Research Scientist 3 Single 4739  
## 187 Laboratory Technician 3 Married 3755  
## 193 Research Scientist 1 Married 2559  
## 194 Research Scientist 1 Divorced 3204  
## 195 Research Scientist 4 Single 4680  
## 198 Research Scientist 3 Single 2632  
## 208 Laboratory Technician 4 Divorced 2532  
## 212 Laboratory Technician 3 Divorced 3977  
## 214 Laboratory Technician 2 Married 2426  
## 215 Research Director 1 Divorced 17007  
## 216 Manufacturing Director 1 Married 9957  
## 227 Laboratory Technician 4 Married 3162  
## 233 Research Scientist 4 Divorced 3761  
## 234 Research Scientist 1 Divorced 4615  
## 237 Manufacturing Director 1 Married 7412  
## 238 Laboratory Technician 1 Divorced 2341  
## 240 Laboratory Technician 4 Married 3867  
## 245 Research Scientist 4 Married 4422  
## 246 Research Scientist 2 Married 2368  
## 250 Manufacturing Director 1 Married 4148  
## 253 Research Director 4 Single 14411  
## 258 Research Scientist 4 Divorced 2695  
## 270 Research Scientist 3 Divorced 1274  
## 271 Research Scientist 3 Single 3230  
## 277 Research Scientist 4 Married 3420  
## 280 Manager 1 Married 15402  
## 283 Laboratory Technician 2 Single 2013  
## 284 Research Scientist 4 Single 3944  
## 286 Manufacturing Director 1 Divorced 4450  
## 288 Research Director 1 Married 13245  
## 292 Healthcare Representative 1 Married 5373  
## 293 Research Scientist 4 Married 2356  
## 295 Research Scientist 3 Divorced 2911  
## 297 Research Director 1 Married 16880  
## 305 Research Scientist 4 Single 2553  
## 307 Research Scientist 4 Married 2177  
## 311 Research Scientist 4 Divorced 2089  
## 312 Research Scientist 1 Married 3143  
## 320 Laboratory Technician 2 Married 2811  
## 330 Laboratory Technician 1 Married 4014  
## 331 Research Scientist 3 Married 2814  
## 335 Laboratory Technician 4 Married 3131  
## 336 Laboratory Technician 1 Single 3578  
## 337 Laboratory Technician 2 Married 3838  
## 339 Laboratory Technician 3 Married 1281  
## 342 Research Scientist 4 Married 2662  
## 345 Laboratory Technician 4 Married 2376  
## 349 Healthcare Representative 1 Divorced 4448  
## 353 Healthcare Representative 1 Married 4741  
## 359 Healthcare Representative 4 Divorced 9439  
## 362 Laboratory Technician 3 Married 3702  
## 365 Laboratory Technician 2 Single 2559  
## 366 Research Scientist 1 Divorced 2996  
## 372 Laboratory Technician 1 Divorced 2296  
## 375 Research Scientist 4 Single 2500  
## 377 Laboratory Technician 3 Single 3708  
## 379 Laboratory Technician 3 Married 3306  
## 380 Laboratory Technician 1 Single 5769  
## 386 Research Scientist 2 Single 2096  
## 389 Healthcare Representative 1 Divorced 6842  
## 393 Research Scientist 4 Married 2862  
## 398 Laboratory Technician 4 Single 2720  
## 400 Laboratory Technician 2 Married 2468  
## 404 Research Scientist 3 Married 2996  
## 405 Laboratory Technician 1 Divorced 4558  
## 407 Laboratory Technician 4 Married 2544  
## 410 Laboratory Technician 1 Married 2238  
## 414 Sales Representative 2 Married 2476  
## 420 Sales Representative 3 Married 3067  
## 424 Manager 1 Single 18947  
## 429 Sales Executive 1 Married 6392  
## 432 Sales Representative 2 Married 3931  
## 444 Sales Representative 3 Married 2231  
## 447 Sales Executive 1 Single 4759  
## 452 Sales Executive 1 Divorced 5486  
## 453 Sales Representative 1 Divorced 2322  
## 454 Sales Executive 1 Married 6500  
## 468 Sales Executive 1 Divorced 5605  
## 473 Sales Executive 1 Single 4898  
## 477 Sales Representative 4 Single 3294  
## 480 Sales Executive 2 Single 6151  
## 484 Sales Executive 2 Married 4978  
## 487 Sales Executive 1 Single 4538  
## MonthlyRate NumCompaniesWorked Over18 OverTime PercentSalaryHike  
## 1 21495 0 Y No 17  
## 2 3072 1 Y No 15  
## 4 11585 1 Y No 11  
## 5 19555 1 Y No 12  
## 9 14382 5 Y No 15  
## 15 12355 1 Y No 12  
## 16 21199 5 Y Yes 13  
## 17 15346 6 Y Yes 14  
## 18 3872 7 Y Yes 13  
## 23 26092 1 Y Yes 23  
## 25 10322 0 Y Yes 22  
## 28 4156 1 Y No 12  
## 29 11092 1 Y Yes 12  
## 32 11539 4 Y Yes 11  
## 33 10332 9 Y Yes 12  
## 35 22710 3 Y No 20  
## 42 10261 1 Y No 16  
## 44 4510 9 Y Yes 18  
## 48 21509 7 Y No 20  
## 50 5355 7 Y No 11  
## 55 9129 5 Y No 16  
## 70 18385 0 Y Yes 17  
## 73 8658 5 Y Yes 13  
## 84 15182 3 Y No 12  
## 87 10642 1 Y No 17  
## 88 25846 1 Y No 15  
## 89 9150 1 Y Yes 11  
## 92 3810 1 Y Yes 14  
## 93 21813 8 Y No 18  
## 99 21972 1 Y Yes 16  
## 103 14630 1 Y No 13  
## 104 9947 1 Y Yes 13  
## 106 18420 1 Y Yes 12  
## 107 17205 1 Y Yes 14  
## 109 2493 2 Y Yes 16  
## 113 24788 3 Y Yes 11  
## 114 10494 1 Y No 15  
## 115 22957 6 Y No 12  
## 116 19655 4 Y No 14  
## 123 8318 1 Y No 16  
## 125 7129 5 Y No 13  
## 126 15434 1 Y No 11  
## 127 25657 8 Y No 17  
## 129 2243 0 Y No 19  
## 130 15428 0 Y No 21  
## 134 4809 1 Y No 21  
## 135 5561 2 Y Yes 12  
## 137 5586 0 Y Yes 14  
## 141 16340 6 Y No 12  
## 142 26062 1 Y No 20  
## 143 24941 4 Y Yes 16  
## 145 16458 1 Y No 13  
## 147 14561 3 Y No 17  
## 148 22456 3 Y Yes 21  
## 149 19982 1 Y No 13  
## 153 15053 0 Y Yes 12  
## 156 5868 4 Y No 13  
## 165 7679 0 Y No 14  
## 166 8544 2 Y No 16  
## 169 22088 1 Y Yes 14  
## 173 20003 1 Y Yes 14  
## 176 6409 8 Y No 22  
## 182 23016 3 Y Yes 13  
## 184 26968 4 Y No 22  
## 186 16090 4 Y No 12  
## 187 17872 1 Y No 11  
## 193 7508 1 Y Yes 13  
## 194 10415 5 Y No 14  
## 195 15232 3 Y No 17  
## 198 23910 1 Y No 14  
## 208 6054 6 Y No 14  
## 212 7298 6 Y Yes 19  
## 214 16479 0 Y No 13  
## 215 11929 7 Y No 14  
## 216 9096 0 Y No 15  
## 227 10778 0 Y No 17  
## 233 2373 9 Y No 12  
## 234 21029 8 Y Yes 23  
## 237 6009 1 Y No 11  
## 238 19715 1 Y No 13  
## 240 14222 1 Y Yes 12  
## 245 21203 3 Y Yes 13  
## 246 23300 1 Y No 19  
## 250 11275 1 Y No 12  
## 253 24450 1 Y Yes 13  
## 258 7747 0 Y Yes 18  
## 270 7152 1 Y No 13  
## 271 10531 1 Y No 17  
## 277 10205 7 Y No 12  
## 280 17997 7 Y No 11  
## 283 10950 2 Y No 11  
## 284 4306 5 Y Yes 11  
## 286 26250 1 Y No 11  
## 288 15067 4 Y Yes 14  
## 292 6225 0 Y No 12  
## 293 14871 3 Y Yes 19  
## 295 15170 1 Y No 17  
## 297 22422 4 Y Yes 11  
## 305 8306 1 Y No 16  
## 307 8456 3 Y No 17  
## 311 5228 4 Y No 14  
## 312 6076 6 Y No 19  
## 320 12086 9 Y No 14  
## 330 16002 3 Y Yes 15  
## 331 10293 1 Y Yes 14  
## 335 26342 1 Y No 13  
## 336 23577 0 Y No 12  
## 337 8192 8 Y No 11  
## 339 16900 1 Y No 18  
## 342 7975 8 Y No 20  
## 345 26537 1 Y No 13  
## 349 10748 2 Y No 12  
## 353 22722 1 Y Yes 13  
## 359 23402 3 Y Yes 16  
## 362 16376 1 Y No 11  
## 365 16620 5 Y No 11  
## 366 20284 5 Y No 14  
## 372 10036 0 Y No 14  
## 375 10515 0 Y No 14  
## 377 2104 2 Y No 14  
## 379 26176 7 Y No 19  
## 380 23447 1 Y Yes 14  
## 386 18830 1 Y No 18  
## 389 26308 6 Y No 20  
## 393 3811 1 Y No 12  
## 398 11162 0 Y No 13  
## 400 15963 4 Y No 14  
## 404 5182 7 Y Yes 15  
## 405 13535 1 Y No 12  
## 407 7102 0 Y No 18  
## 410 6961 2 Y No 21  
## 414 17434 1 Y No 18  
## 420 6393 0 Y No 19  
## 424 22822 3 Y No 12  
## 429 10589 2 Y No 13  
## 432 20990 2 Y No 11  
## 444 11314 6 Y No 18  
## 447 15891 3 Y No 18  
## 452 24795 4 Y No 11  
## 453 9518 3 Y No 13  
## 454 13305 5 Y No 17  
## 468 19191 1 Y Yes 24  
## 473 11827 0 Y No 14  
## 477 13137 1 Y Yes 18  
## 480 22074 1 Y No 13  
## 484 3536 7 Y No 11  
## 487 6039 0 Y Yes 12  
## PerformanceRating RelationshipSatisfaction StandardHours StockOptionLevel  
## 1 3 3 80 0  
## 2 3 4 80 0  
## 4 3 3 80 1  
## 5 3 1 80 0  
## 9 3 1 80 0  
## 15 3 1 80 0  
## 16 3 4 80 1  
## 17 3 2 80 1  
## 18 3 4 80 0  
## 23 4 1 80 0  
## 25 4 1 80 0  
## 28 3 1 80 0  
## 29 3 3 80 3  
## 32 3 2 80 0  
## 33 3 3 80 3  
## 35 4 1 80 0  
## 42 3 4 80 0  
## 44 3 1 80 3  
## 48 4 1 80 0  
## 50 3 3 80 0  
## 55 3 3 80 0  
## 70 3 1 80 0  
## 73 3 2 80 3  
## 84 3 1 80 0  
## 87 3 4 80 0  
## 88 3 4 80 1  
## 89 3 3 80 0  
## 92 3 3 80 0  
## 93 3 4 80 0  
## 99 3 2 80 0  
## 103 3 3 80 0  
## 104 3 2 80 0  
## 106 3 3 80 0  
## 107 3 2 80 0  
## 109 3 1 80 0  
## 113 3 3 80 1  
## 114 3 2 80 1  
## 115 3 3 80 0  
## 116 3 3 80 0  
## 123 3 1 80 0  
## 125 3 4 80 3  
## 126 3 4 80 0  
## 127 3 4 80 0  
## 129 3 3 80 1  
## 130 4 1 80 1  
## 134 4 3 80 2  
## 135 3 1 80 0  
## 137 3 1 80 3  
## 141 3 2 80 1  
## 142 4 3 80 0  
## 143 3 2 80 1  
## 145 3 3 80 0  
## 147 3 4 80 0  
## 148 4 3 80 1  
## 149 3 4 80 1  
## 153 3 4 80 2  
## 156 3 4 80 1  
## 165 3 2 80 0  
## 166 3 1 80 1  
## 169 3 4 80 1  
## 173 3 1 80 0  
## 176 4 4 80 0  
## 182 3 3 80 0  
## 184 4 1 80 1  
## 186 3 4 80 0  
## 187 3 1 80 1  
## 193 3 4 80 0  
## 194 3 4 80 1  
## 195 3 1 80 0  
## 198 3 3 80 0  
## 208 3 3 80 3  
## 212 3 3 80 1  
## 214 3 3 80 1  
## 215 3 4 80 2  
## 216 3 3 80 1  
## 227 3 4 80 0  
## 233 3 2 80 1  
## 234 4 1 80 3  
## 237 3 4 80 0  
## 238 3 4 80 1  
## 240 3 2 80 1  
## 245 3 4 80 1  
## 246 3 3 80 0  
## 250 3 3 80 1  
## 253 3 4 80 0  
## 258 3 2 80 1  
## 270 3 2 80 2  
## 271 3 1 80 0  
## 277 3 3 80 1  
## 280 3 1 80 1  
## 283 3 3 80 0  
## 284 3 3 80 0  
## 286 3 3 80 2  
## 288 3 2 80 0  
## 292 3 1 80 1  
## 293 3 2 80 1  
## 295 3 4 80 1  
## 297 3 2 80 0  
## 305 3 3 80 0  
## 307 3 1 80 1  
## 311 3 4 80 3  
## 312 3 2 80 1  
## 320 3 2 80 1  
## 330 3 3 80 1  
## 331 3 2 80 0  
## 335 3 1 80 1  
## 336 3 4 80 0  
## 337 3 4 80 0  
## 339 3 3 80 2  
## 342 4 2 80 1  
## 345 3 2 80 1  
## 349 3 2 80 1  
## 353 3 3 80 1  
## 359 3 2 80 1  
## 362 3 2 80 1  
## 365 3 3 80 0  
## 366 3 3 80 2  
## 372 3 2 80 3  
## 375 3 1 80 0  
## 377 3 3 80 0  
## 379 3 4 80 1  
## 380 3 1 80 0  
## 386 3 4 80 0  
## 389 4 1 80 1  
## 393 3 2 80 1  
## 398 3 4 80 0  
## 400 3 2 80 1  
## 404 3 4 80 0  
## 405 3 4 80 1  
## 407 3 1 80 1  
## 410 4 4 80 1  
## 414 3 1 80 1  
## 420 3 3 80 1  
## 424 3 4 80 0  
## 429 3 4 80 1  
## 432 3 1 80 1  
## 444 3 4 80 1  
## 447 3 4 80 0  
## 452 3 1 80 3  
## 453 3 3 80 1  
## 454 3 2 80 1  
## 468 4 3 80 1  
## 473 3 4 80 0  
## 477 3 1 80 0  
## 480 3 1 80 0  
## 484 3 4 80 1  
## 487 3 4 80 0  
## TotalWorkingYears TrainingTimesLastYear WorkLifeBalance YearsAtCompany  
## 1 2 4 3 1  
## 2 2 0 3 2  
## 4 1 2 3 1  
## 5 1 2 3 1  
## 9 15 2 1 1  
## 15 10 2 2 10  
## 16 24 2 3 5  
## 17 6 2 3 3  
## 18 10 4 3 7  
## 23 1 2 3 1  
## 25 4 0 3 3  
## 28 7 2 1 7  
## 29 6 2 2 5  
## 32 7 2 3 5  
## 33 18 2 4 10  
## 35 20 2 3 4  
## 42 5 3 3 4  
## 44 7 2 3 5  
## 48 13 3 4 9  
## 50 8 2 2 0  
## 55 4 3 3 2  
## 70 6 3 3 5  
## 73 16 3 3 2  
## 84 6 2 3 3  
## 87 1 3 3 1  
## 88 33 3 3 32  
## 89 3 3 3 3  
## 92 3 2 3 3  
## 93 7 6 3 2  
## 99 6 4 3 5  
## 103 2 3 3 2  
## 104 1 3 4 1  
## 106 0 2 4 0  
## 107 2 3 3 2  
## 109 26 2 3 24  
## 113 4 2 3 2  
## 114 3 2 3 3  
## 115 10 3 4 7  
## 116 6 3 3 2  
## 123 6 3 3 6  
## 125 13 3 3 11  
## 126 1 3 3 1  
## 127 7 2 2 2  
## 129 4 1 1 3  
## 130 6 3 4 5  
## 134 9 4 2 9  
## 135 7 5 2 4  
## 137 6 3 3 5  
## 141 10 2 3 8  
## 142 5 3 1 5  
## 143 17 2 2 2  
## 145 12 3 3 12  
## 147 8 3 2 0  
## 148 32 3 3 9  
## 149 10 5 3 10  
## 153 7 5 2 6  
## 156 5 2 2 3  
## 165 6 3 4 5  
## 166 5 3 4 3  
## 169 3 2 3 3  
## 173 20 3 1 20  
## 176 6 2 3 4  
## 182 10 1 3 3  
## 184 11 1 3 7  
## 186 18 2 3 3  
## 187 8 3 3 8  
## 193 8 0 3 8  
## 194 8 3 3 3  
## 195 4 2 3 1  
## 198 5 4 2 5  
## 208 8 5 3 4  
## 212 7 2 2 2  
## 214 6 5 3 5  
## 215 16 3 2 14  
## 216 7 1 2 6  
## 227 6 2 2 5  
## 233 10 3 2 5  
## 234 19 2 3 16  
## 237 9 3 3 9  
## 238 1 6 3 1  
## 240 2 2 3 2  
## 245 16 3 1 1  
## 246 5 3 2 5  
## 250 4 1 3 4  
## 253 32 2 3 32  
## 258 3 2 1 2  
## 270 1 2 2 1  
## 271 3 4 4 3  
## 277 17 2 2 6  
## 280 21 3 1 3  
## 283 15 4 3 4  
## 284 6 3 3 3  
## 286 5 3 3 4  
## 288 17 3 4 0  
## 292 6 5 2 5  
## 293 8 2 3 6  
## 295 5 1 2 5  
## 297 25 2 3 3  
## 305 6 3 3 5  
## 307 7 6 3 1  
## 311 7 3 4 5  
## 312 14 1 3 10  
## 320 4 2 3 2  
## 330 4 3 3 2  
## 331 4 2 2 4  
## 335 10 5 3 10  
## 336 8 2 3 7  
## 337 8 5 3 5  
## 339 1 3 3 1  
## 342 19 2 4 5  
## 345 2 2 4 2  
## 349 15 3 3 7  
## 353 5 3 3 5  
## 359 12 2 1 5  
## 362 5 3 3 5  
## 365 7 4 2 1  
## 366 10 2 3 4  
## 372 2 3 3 1  
## 375 4 2 4 3  
## 377 9 5 3 5  
## 379 7 5 2 0  
## 380 10 3 3 10  
## 386 2 3 2 2  
## 389 13 3 3 5  
## 393 10 2 2 10  
## 398 6 3 3 5  
## 400 9 4 2 6  
## 404 8 2 3 6  
## 405 10 2 3 10  
## 407 8 3 3 7  
## 410 7 2 3 5  
## 414 1 3 3 1  
## 420 3 1 3 2  
## 424 22 2 2 2  
## 429 8 6 1 2  
## 432 6 5 3 4  
## 444 6 3 3 4  
## 447 15 2 3 13  
## 452 15 3 3 2  
## 453 3 3 3 0  
## 454 6 1 3 3  
## 468 8 3 3 8  
## 473 5 5 3 4  
## 477 3 3 2 3  
## 480 19 4 3 19  
## 484 4 3 1 1  
## 487 4 3 3 3  
## YearsInCurrentRole YearsSinceLastPromotion YearsWithCurrManager  
## 1 0 0 0  
## 2 2 2 2  
## 4 0 0 0  
## 5 0 0 0  
## 9 0 0 0  
## 15 7 1 9  
## 16 2 1 4  
## 17 2 1 2  
## 18 7 3 7  
## 23 0 0 0  
## 25 2 1 2  
## 28 2 7 7  
## 29 3 2 3  
## 32 2 0 1  
## 33 0 2 7  
## 35 3 1 3  
## 42 2 1 0  
## 44 4 4 3  
## 48 7 1 7  
## 50 0 0 0  
## 55 2 1 2  
## 70 2 0 3  
## 73 2 2 2  
## 84 2 2 2  
## 87 0 0 0  
## 88 14 6 9  
## 89 2 1 2  
## 92 2 0 2  
## 93 2 2 2  
## 99 3 1 4  
## 103 2 2 2  
## 104 0 0 0  
## 106 0 0 0  
## 107 2 0 2  
## 109 10 1 11  
## 113 2 2 2  
## 114 2 2 2  
## 115 6 5 7  
## 116 0 1 2  
## 123 3 0 4  
## 125 10 3 8  
## 126 0 0 0  
## 127 2 2 2  
## 129 2 0 2  
## 130 3 1 3  
## 134 7 1 7  
## 135 2 0 3  
## 137 0 1 2  
## 141 0 1 7  
## 142 1 0 3  
## 143 2 2 2  
## 145 9 5 8  
## 147 0 0 0  
## 148 8 1 5  
## 149 5 7 7  
## 153 2 0 5  
## 156 2 1 2  
## 165 0 1 4  
## 166 2 0 2  
## 169 1 0 2  
## 173 7 1 8  
## 176 3 1 2  
## 182 2 0 2  
## 184 5 1 7  
## 186 2 1 2  
## 187 3 0 7  
## 193 7 7 1  
## 194 2 2 2  
## 195 0 0 0  
## 198 4 0 4  
## 208 3 0 3  
## 212 2 0 2  
## 214 4 0 3  
## 215 8 6 9  
## 216 2 0 2  
## 227 2 3 4  
## 233 4 0 3  
## 234 13 1 7  
## 237 7 0 7  
## 238 0 0 0  
## 240 2 2 2  
## 245 1 0 0  
## 246 4 4 3  
## 250 3 0 3  
## 253 6 13 9  
## 258 2 2 2  
## 270 0 0 0  
## 271 2 1 0  
## 277 5 1 2  
## 280 2 0 2  
## 283 3 1 3  
## 284 2 1 2  
## 286 2 1 3  
## 288 0 0 0  
## 292 3 0 2  
## 293 4 0 2  
## 295 2 4 3  
## 297 2 1 2  
## 305 2 1 3  
## 307 0 0 0  
## 311 4 2 2  
## 312 8 7 6  
## 320 2 2 2  
## 330 2 2 2  
## 331 2 1 3  
## 335 8 0 8  
## 336 7 0 7  
## 337 4 0 2  
## 339 0 0 0  
## 342 2 0 4  
## 345 2 2 2  
## 349 4 7 7  
## 353 3 3 3  
## 359 3 1 4  
## 362 4 0 4  
## 365 0 0 0  
## 366 3 1 3  
## 372 1 0 0  
## 375 1 0 2  
## 377 2 1 4  
## 379 0 0 0  
## 380 7 1 4  
## 386 2 2 1  
## 389 4 0 4  
## 393 0 0 8  
## 398 3 1 2  
## 400 1 0 5  
## 404 4 1 3  
## 405 0 1 8  
## 407 7 7 7  
## 410 0 1 4  
## 414 0 0 0  
## 420 2 1 2  
## 424 2 2 1  
## 429 2 2 2  
## 432 3 1 2  
## 444 3 1 2  
## 447 9 3 12  
## 452 2 2 2  
## 453 0 0 0  
## 454 2 1 2  
## 468 0 7 7  
## 473 2 1 3  
## 477 2 1 2  
## 480 2 11 9  
## 484 0 0 0  
## 487 2 0 2  
## AttritionCat JobRoleCat JobInvolvement. JobLevel. JobSatisfaction.  
## 1 1 1 0.907 0.001 0.001  
## 2 1 1 0.316 0.001 0.397  
## 4 1 1 0.001 0.001 0.397  
## 5 1 1 0.055 0.001 0.206  
## 9 1 5 0.001 0.378 0.206  
## 15 1 5 0.055 0.737 0.001  
## 16 1 2 0.316 0.957 0.001  
## 17 1 3 0.316 0.001 0.001  
## 18 1 3 0.001 0.001 0.206  
## 23 1 3 0.316 0.001 0.206  
## 25 1 3 0.055 0.001 0.001  
## 28 1 3 0.316 0.001 0.397  
## 29 1 3 0.316 0.001 0.206  
## 32 1 3 0.316 0.001 0.397  
## 33 1 3 0.316 0.001 0.397  
## 35 1 7 0.055 0.378 0.001  
## 42 1 7 0.316 0.001 0.688  
## 44 1 3 0.316 0.001 0.206  
## 48 1 7 0.316 0.001 0.001  
## 50 1 7 0.055 0.001 0.001  
## 55 1 7 0.055 0.001 0.397  
## 70 1 4 0.316 0.378 0.001  
## 73 1 4 0.055 0.378 0.001  
## 84 1 6 0.055 0.001 0.001  
## 87 1 6 0.907 0.001 0.001  
## 88 1 2 0.001 0.957 0.688  
## 89 1 6 0.055 0.001 0.206  
## 92 1 6 0.055 0.001 0.397  
## 93 1 4 0.001 0.378 0.001  
## 99 1 6 0.316 0.001 0.206  
## 103 1 6 0.001 0.001 0.206  
## 104 1 6 0.055 0.001 0.206  
## 106 1 6 0.316 0.001 0.688  
## 107 1 6 0.316 0.001 0.688  
## 109 1 2 0.001 0.957 0.206  
## 113 1 4 0.001 0.378 0.397  
## 114 0 1 0.316 0.001 0.688  
## 115 0 1 0.316 0.001 0.001  
## 116 0 1 0.316 0.001 0.206  
## 123 0 1 0.316 0.001 0.206  
## 125 0 1 0.316 0.001 0.206  
## 126 0 1 0.055 0.001 0.206  
## 127 0 1 0.316 0.001 0.397  
## 129 0 1 0.316 0.001 0.688  
## 130 0 1 0.316 0.001 0.397  
## 134 0 8 0.316 0.737 0.001  
## 135 0 3 0.907 0.001 0.397  
## 137 0 3 0.907 0.001 0.688  
## 141 0 7 0.055 0.001 0.397  
## 142 0 7 0.907 0.001 0.206  
## 143 0 7 0.316 0.001 0.001  
## 145 0 7 0.316 0.001 0.688  
## 147 0 3 0.316 0.001 0.001  
## 148 0 9 0.316 0.957 0.001  
## 149 0 3 0.907 0.001 0.688  
## 153 0 3 0.907 0.001 0.206  
## 156 0 3 0.316 0.001 0.001  
## 165 0 7 0.316 0.001 0.001  
## 166 0 3 0.907 0.001 0.001  
## 169 0 3 0.055 0.001 0.397  
## 173 0 3 0.055 0.378 0.397  
## 176 0 7 0.316 0.001 0.206  
## 182 0 3 0.907 0.378 0.001  
## 184 0 3 0.316 0.001 0.688  
## 186 0 3 0.907 0.001 0.397  
## 187 0 7 0.316 0.001 0.397  
## 193 0 3 0.316 0.001 0.001  
## 194 0 3 0.316 0.001 0.001  
## 195 0 3 0.907 0.001 0.688  
## 198 0 3 0.316 0.001 0.397  
## 208 0 7 0.055 0.001 0.688  
## 212 0 7 0.316 0.001 0.397  
## 214 0 7 0.907 0.001 0.206  
## 215 0 9 0.316 0.888 0.001  
## 216 0 8 0.316 0.378 0.001  
## 227 0 7 0.055 0.001 0.688  
## 233 0 3 0.055 0.001 0.688  
## 234 0 3 0.055 0.378 0.001  
## 237 0 8 0.316 0.737 0.001  
## 238 0 7 0.316 0.001 0.001  
## 240 0 7 0.055 0.001 0.688  
## 245 0 3 0.055 0.378 0.688  
## 246 0 3 0.055 0.001 0.206  
## 250 0 8 0.316 0.378 0.001  
## 253 0 9 0.001 0.888 0.688  
## 258 0 3 0.316 0.001 0.688  
## 270 0 3 0.316 0.001 0.397  
## 271 0 3 0.055 0.001 0.397  
## 277 0 3 0.316 0.001 0.688  
## 280 0 2 0.055 0.888 0.001  
## 283 0 7 0.316 0.001 0.206  
## 284 0 3 0.316 0.001 0.688  
## 286 0 8 0.907 0.378 0.001  
## 288 0 9 0.316 0.737 0.001  
## 292 0 5 0.316 0.378 0.001  
## 293 0 3 0.316 0.001 0.688  
## 295 0 3 0.316 0.001 0.397  
## 297 0 9 0.316 0.888 0.001  
## 305 0 3 0.055 0.001 0.688  
## 307 0 3 0.055 0.001 0.688  
## 311 0 3 0.907 0.001 0.688  
## 312 0 3 0.907 0.001 0.001  
## 320 0 7 0.316 0.001 0.206  
## 330 0 7 0.316 0.001 0.001  
## 331 0 3 0.055 0.001 0.397  
## 335 0 7 0.316 0.001 0.688  
## 336 0 7 0.907 0.001 0.001  
## 337 0 7 0.055 0.001 0.206  
## 339 0 7 0.055 0.001 0.397  
## 342 0 3 0.055 0.001 0.688  
## 345 0 7 0.055 0.001 0.688  
## 349 0 5 0.316 0.378 0.001  
## 353 0 5 0.316 0.378 0.001  
## 359 0 5 0.907 0.737 0.688  
## 362 0 7 0.055 0.001 0.397  
## 365 0 7 0.907 0.001 0.206  
## 366 0 3 0.316 0.001 0.001  
## 372 0 7 0.055 0.001 0.001  
## 375 0 3 0.316 0.001 0.688  
## 377 0 7 0.001 0.001 0.397  
## 379 0 7 0.316 0.001 0.397  
## 380 0 7 0.001 0.378 0.001  
## 386 0 3 0.316 0.001 0.206  
## 389 0 5 0.907 0.378 0.001  
## 393 0 3 0.316 0.001 0.688  
## 398 0 7 0.055 0.001 0.688  
## 400 0 7 0.055 0.001 0.206  
## 404 0 3 0.316 0.001 0.397  
## 405 0 7 0.316 0.378 0.001  
## 407 0 7 0.907 0.001 0.688  
## 410 0 7 0.316 0.001 0.001  
## 414 0 6 0.316 0.001 0.206  
## 420 0 6 0.316 0.001 0.397  
## 424 0 2 0.316 0.957 0.001  
## 429 0 4 0.316 0.378 0.001  
## 432 0 6 0.316 0.001 0.206  
## 444 0 6 0.316 0.001 0.397  
## 447 0 4 0.001 0.378 0.001  
## 452 0 4 0.316 0.378 0.001  
## 453 0 6 0.316 0.001 0.001  
## 454 0 4 0.316 0.378 0.001  
## 468 0 4 0.907 0.378 0.001  
## 473 0 4 0.316 0.378 0.001  
## 477 0 6 0.055 0.001 0.688  
## 480 0 4 0.001 0.378 0.206  
## 484 0 4 0.316 0.378 0.206  
## 487 0 4 0.316 0.378 0.001  
## WorkLifeBalance. JI\_JL JI\_JS JL\_JS JI\_WLB JL\_WLB JI\_JL\_JS JI\_JL\_JS\_WLB  
## 1 0.276 0.4540 0.4540 0.0010 0.5915 0.1385 0.30300000 0.29625  
## 2 0.277 0.1585 0.3565 0.1990 0.2965 0.1390 0.23800000 0.24775  
## 4 0.277 0.0010 0.1990 0.1990 0.1390 0.1390 0.13300000 0.16900  
## 5 0.277 0.0280 0.1305 0.1035 0.1660 0.1390 0.08733333 0.13475  
## 9 0.001 0.1895 0.1035 0.2920 0.0010 0.1895 0.19500000 0.14650  
## 15 0.052 0.3960 0.0280 0.3690 0.0535 0.3945 0.26433333 0.21125  
## 16 0.276 0.6365 0.1585 0.4790 0.2960 0.6165 0.42466667 0.38750  
## 17 0.276 0.1585 0.1585 0.0010 0.2960 0.1385 0.10600000 0.14850  
## 18 0.277 0.0010 0.1035 0.1035 0.1390 0.1390 0.06933333 0.12125  
## 23 0.276 0.1585 0.2610 0.1035 0.2960 0.1385 0.17433333 0.19975  
## 25 0.277 0.0280 0.0280 0.0010 0.1660 0.1390 0.01900000 0.08350  
## 28 0.001 0.1585 0.3565 0.1990 0.1585 0.0010 0.23800000 0.17875  
## 29 0.050 0.1585 0.2610 0.1035 0.1830 0.0255 0.17433333 0.14325  
## 32 0.276 0.1585 0.3565 0.1990 0.2960 0.1385 0.23800000 0.24750  
## 33 0.885 0.1585 0.3565 0.1990 0.6005 0.4430 0.23800000 0.39975  
## 35 0.276 0.2165 0.0280 0.1895 0.1655 0.3270 0.14466667 0.17750  
## 42 0.273 0.1585 0.5020 0.3445 0.2945 0.1370 0.33500000 0.31950  
## 44 0.274 0.1585 0.2610 0.1035 0.2950 0.1375 0.17433333 0.19925  
## 48 0.883 0.1585 0.1585 0.0010 0.5995 0.4420 0.10600000 0.30025  
## 50 0.049 0.0280 0.0280 0.0010 0.0520 0.0250 0.01900000 0.02650  
## 55 0.270 0.0280 0.2260 0.1990 0.1625 0.1355 0.15100000 0.18075  
## 70 0.268 0.3470 0.1585 0.1895 0.2920 0.3230 0.23166667 0.24075  
## 73 0.268 0.2165 0.0280 0.1895 0.1615 0.3230 0.14466667 0.17550  
## 84 0.262 0.0280 0.0280 0.0010 0.1585 0.1315 0.01900000 0.07975  
## 87 0.263 0.4540 0.4540 0.0010 0.5850 0.1320 0.30300000 0.29300  
## 88 0.263 0.4790 0.3445 0.8225 0.1320 0.6100 0.54866667 0.47725  
## 89 0.263 0.0280 0.1305 0.1035 0.1590 0.1320 0.08733333 0.13125  
## 92 0.262 0.0280 0.2260 0.1990 0.1585 0.1315 0.15100000 0.17875  
## 93 0.262 0.1895 0.0010 0.1895 0.1315 0.3200 0.12666667 0.16050  
## 99 0.262 0.1585 0.2610 0.1035 0.2890 0.1315 0.17433333 0.19625  
## 103 0.264 0.0010 0.1035 0.1035 0.1325 0.1325 0.06933333 0.11800  
## 104 0.881 0.0280 0.1305 0.1035 0.4680 0.4410 0.08733333 0.28575  
## 106 0.882 0.1585 0.5020 0.3445 0.5990 0.4415 0.33500000 0.47175  
## 107 0.265 0.1585 0.5020 0.3445 0.2905 0.1330 0.33500000 0.31750  
## 109 0.266 0.4790 0.1035 0.5815 0.1335 0.6115 0.38800000 0.35750  
## 113 0.266 0.1895 0.1990 0.3875 0.1335 0.3220 0.25866667 0.26050  
## 114 0.265 0.1585 0.5020 0.3445 0.2905 0.1330 0.33500000 0.31750  
## 115 0.883 0.1585 0.1585 0.0010 0.5995 0.4420 0.10600000 0.30025  
## 116 0.266 0.1585 0.2610 0.1035 0.2910 0.1335 0.17433333 0.19725  
## 123 0.270 0.1585 0.2610 0.1035 0.2930 0.1355 0.17433333 0.19825  
## 125 0.270 0.1585 0.2610 0.1035 0.2930 0.1355 0.17433333 0.19825  
## 126 0.271 0.0280 0.1305 0.1035 0.1630 0.1360 0.08733333 0.13325  
## 127 0.045 0.1585 0.3565 0.1990 0.1805 0.0230 0.23800000 0.18975  
## 129 0.001 0.1585 0.5020 0.3445 0.1585 0.0010 0.33500000 0.25150  
## 130 0.882 0.1585 0.3565 0.1990 0.5990 0.4415 0.23800000 0.39900  
## 134 0.044 0.5265 0.1585 0.3690 0.1800 0.3905 0.35133333 0.27450  
## 135 0.044 0.4540 0.6520 0.1990 0.4755 0.0225 0.43500000 0.33725  
## 137 0.270 0.4540 0.7975 0.3445 0.5885 0.1355 0.53200000 0.46650  
## 141 0.272 0.0280 0.2260 0.1990 0.1635 0.1365 0.15100000 0.18125  
## 142 0.001 0.4540 0.5565 0.1035 0.4540 0.0010 0.37133333 0.27875  
## 143 0.043 0.1585 0.1585 0.0010 0.1795 0.0220 0.10600000 0.09025  
## 145 0.272 0.1585 0.5020 0.3445 0.2940 0.1365 0.33500000 0.31925  
## 147 0.042 0.1585 0.1585 0.0010 0.1790 0.0215 0.10600000 0.09000  
## 148 0.270 0.6365 0.1585 0.4790 0.2930 0.6135 0.42466667 0.38600  
## 149 0.271 0.4540 0.7975 0.3445 0.5890 0.1360 0.53200000 0.46675  
## 153 0.041 0.4540 0.5565 0.1035 0.4740 0.0210 0.37133333 0.28875  
## 156 0.042 0.1585 0.1585 0.0010 0.1790 0.0215 0.10600000 0.09000  
## 165 0.884 0.1585 0.1585 0.0010 0.6000 0.4425 0.10600000 0.30050  
## 166 0.885 0.4540 0.4540 0.0010 0.8960 0.4430 0.30300000 0.44850  
## 169 0.276 0.0280 0.2260 0.1990 0.1655 0.1385 0.15100000 0.18225  
## 173 0.001 0.2165 0.2260 0.3875 0.0280 0.1895 0.27666667 0.20775  
## 176 0.280 0.1585 0.2610 0.1035 0.2980 0.1405 0.17433333 0.20075  
## 182 0.278 0.6425 0.4540 0.1895 0.5925 0.3280 0.42866667 0.39100  
## 184 0.280 0.1585 0.5020 0.3445 0.2980 0.1405 0.33500000 0.32125  
## 186 0.280 0.4540 0.6520 0.1990 0.5935 0.1405 0.43500000 0.39625  
## 187 0.281 0.1585 0.3565 0.1990 0.2985 0.1410 0.23800000 0.24875  
## 193 0.281 0.1585 0.1585 0.0010 0.2985 0.1410 0.10600000 0.14975  
## 194 0.281 0.1585 0.1585 0.0010 0.2985 0.1410 0.10600000 0.14975  
## 195 0.282 0.4540 0.7975 0.3445 0.5945 0.1415 0.53200000 0.46950  
## 198 0.044 0.1585 0.3565 0.1990 0.1800 0.0225 0.23800000 0.18950  
## 208 0.277 0.0280 0.3715 0.3445 0.1660 0.1390 0.24800000 0.25525  
## 212 0.044 0.1585 0.3565 0.1990 0.1800 0.0225 0.23800000 0.18950  
## 214 0.273 0.4540 0.5565 0.1035 0.5900 0.1370 0.37133333 0.34675  
## 215 0.045 0.6020 0.1585 0.4445 0.1805 0.4665 0.40166667 0.31250  
## 216 0.045 0.3470 0.1585 0.1895 0.1805 0.2115 0.23166667 0.18500  
## 227 0.044 0.0280 0.3715 0.3445 0.0495 0.0225 0.24800000 0.19700  
## 233 0.045 0.0280 0.3715 0.3445 0.0500 0.0230 0.24800000 0.19725  
## 234 0.272 0.2165 0.0280 0.1895 0.1635 0.3250 0.14466667 0.17650  
## 237 0.271 0.5265 0.1585 0.3690 0.2935 0.5040 0.35133333 0.33125  
## 238 0.270 0.1585 0.1585 0.0010 0.2930 0.1355 0.10600000 0.14700  
## 240 0.272 0.0280 0.3715 0.3445 0.1635 0.1365 0.24800000 0.25400  
## 245 0.001 0.2165 0.3715 0.5330 0.0280 0.1895 0.37366667 0.28050  
## 246 0.043 0.0280 0.1305 0.1035 0.0490 0.0220 0.08733333 0.07625  
## 250 0.270 0.3470 0.1585 0.1895 0.2930 0.3240 0.23166667 0.24125  
## 253 0.269 0.4445 0.3445 0.7880 0.1350 0.5785 0.52566667 0.46150  
## 258 0.002 0.1585 0.5020 0.3445 0.1590 0.0015 0.33500000 0.25175  
## 270 0.041 0.1585 0.3565 0.1990 0.1785 0.0210 0.23800000 0.18875  
## 271 0.873 0.0280 0.2260 0.1990 0.4640 0.4370 0.15100000 0.33150  
## 277 0.042 0.1585 0.5020 0.3445 0.1790 0.0215 0.33500000 0.26175  
## 280 0.002 0.4715 0.0280 0.4445 0.0285 0.4450 0.31466667 0.23650  
## 283 0.260 0.1585 0.2610 0.1035 0.2880 0.1305 0.17433333 0.19575  
## 284 0.259 0.1585 0.5020 0.3445 0.2875 0.1300 0.33500000 0.31600  
## 286 0.260 0.6425 0.4540 0.1895 0.5835 0.3190 0.42866667 0.38650  
## 288 0.873 0.5265 0.1585 0.3690 0.5945 0.8050 0.35133333 0.48175  
## 292 0.039 0.3470 0.1585 0.1895 0.1775 0.2085 0.23166667 0.18350  
## 293 0.254 0.1585 0.5020 0.3445 0.2850 0.1275 0.33500000 0.31475  
## 295 0.039 0.1585 0.3565 0.1990 0.1775 0.0200 0.23800000 0.18825  
## 297 0.255 0.6020 0.1585 0.4445 0.2855 0.5715 0.40166667 0.36500  
## 305 0.249 0.0280 0.3715 0.3445 0.1520 0.1250 0.24800000 0.24825  
## 307 0.251 0.0280 0.3715 0.3445 0.1530 0.1260 0.24800000 0.24875  
## 311 0.867 0.4540 0.7975 0.3445 0.8870 0.4340 0.53200000 0.61575  
## 312 0.248 0.4540 0.4540 0.0010 0.5775 0.1245 0.30300000 0.28925  
## 320 0.242 0.1585 0.2610 0.1035 0.2790 0.1215 0.17433333 0.19125  
## 330 0.248 0.1585 0.1585 0.0010 0.2820 0.1245 0.10600000 0.14150  
## 331 0.038 0.0280 0.2260 0.1990 0.0465 0.0195 0.15100000 0.12275  
## 335 0.243 0.1585 0.5020 0.3445 0.2795 0.1220 0.33500000 0.31200  
## 336 0.245 0.4540 0.4540 0.0010 0.5760 0.1230 0.30300000 0.28850  
## 337 0.247 0.0280 0.1305 0.1035 0.1510 0.1240 0.08733333 0.12725  
## 339 0.249 0.0280 0.2260 0.1990 0.1520 0.1250 0.15100000 0.17550  
## 342 0.868 0.0280 0.3715 0.3445 0.4615 0.4345 0.24800000 0.40300  
## 345 0.870 0.0280 0.3715 0.3445 0.4625 0.4355 0.24800000 0.40350  
## 349 0.256 0.3470 0.1585 0.1895 0.2860 0.3170 0.23166667 0.23775  
## 353 0.257 0.3470 0.1585 0.1895 0.2865 0.3175 0.23166667 0.23800  
## 359 0.003 0.8220 0.7975 0.7125 0.4550 0.3700 0.77733333 0.58375  
## 362 0.259 0.0280 0.2260 0.1990 0.1570 0.1300 0.15100000 0.17800  
## 365 0.036 0.4540 0.5565 0.1035 0.4715 0.0185 0.37133333 0.28750  
## 366 0.256 0.1585 0.1585 0.0010 0.2860 0.1285 0.10600000 0.14350  
## 372 0.256 0.0280 0.0280 0.0010 0.1555 0.1285 0.01900000 0.07825  
## 375 0.886 0.1585 0.5020 0.3445 0.6010 0.4435 0.33500000 0.47275  
## 377 0.258 0.0010 0.1990 0.1990 0.1295 0.1295 0.13300000 0.16425  
## 379 0.025 0.1585 0.3565 0.1990 0.1705 0.0130 0.23800000 0.18475  
## 380 0.248 0.1895 0.0010 0.1895 0.1245 0.3130 0.12666667 0.15700  
## 386 0.022 0.1585 0.2610 0.1035 0.1690 0.0115 0.17433333 0.13625  
## 389 0.253 0.6425 0.4540 0.1895 0.5800 0.3155 0.42866667 0.38475  
## 393 0.023 0.1585 0.5020 0.3445 0.1695 0.0120 0.33500000 0.25700  
## 398 0.250 0.0280 0.3715 0.3445 0.1525 0.1255 0.24800000 0.24850  
## 400 0.025 0.0280 0.1305 0.1035 0.0400 0.0130 0.08733333 0.07175  
## 404 0.255 0.1585 0.3565 0.1990 0.2855 0.1280 0.23800000 0.24225  
## 405 0.256 0.3470 0.1585 0.1895 0.2860 0.3170 0.23166667 0.23775  
## 407 0.255 0.4540 0.7975 0.3445 0.5810 0.1280 0.53200000 0.46275  
## 410 0.262 0.1585 0.1585 0.0010 0.2890 0.1315 0.10600000 0.14500  
## 414 0.260 0.1585 0.2610 0.1035 0.2880 0.1305 0.17433333 0.19575  
## 420 0.261 0.1585 0.3565 0.1990 0.2885 0.1310 0.23800000 0.24375  
## 424 0.026 0.6365 0.1585 0.4790 0.1710 0.4915 0.42466667 0.32500  
## 429 0.005 0.3470 0.1585 0.1895 0.1605 0.1915 0.23166667 0.17500  
## 432 0.270 0.1585 0.2610 0.1035 0.2930 0.1355 0.17433333 0.19825  
## 444 0.291 0.1585 0.3565 0.1990 0.3035 0.1460 0.23800000 0.25125  
## 447 0.300 0.1895 0.0010 0.1895 0.1505 0.3390 0.12666667 0.17000  
## 452 0.307 0.3470 0.1585 0.1895 0.3115 0.3425 0.23166667 0.25050  
## 453 0.309 0.1585 0.1585 0.0010 0.3125 0.1550 0.10600000 0.15675  
## 454 0.309 0.3470 0.1585 0.1895 0.3125 0.3435 0.23166667 0.25100  
## 468 0.365 0.6425 0.4540 0.1895 0.6360 0.3715 0.42866667 0.41275  
## 473 0.340 0.3470 0.1585 0.1895 0.3280 0.3590 0.23166667 0.25875  
## 477 0.047 0.0280 0.3715 0.3445 0.0510 0.0240 0.24800000 0.19775  
## 480 0.269 0.1895 0.1035 0.2920 0.1350 0.3235 0.19500000 0.21350  
## 484 0.055 0.3470 0.2610 0.2920 0.1855 0.2165 0.30000000 0.23875  
## 487 0.300 0.3470 0.1585 0.1895 0.3080 0.3390 0.23166667 0.24875  
## Business\_Travel\_Num GenderNum JI1 JS1 JL1 WLB1 JI\_JL\_JS\_2 JI\_JL\_WLB\_2  
## 1 1 1 0 1 1 0 1 0  
## 2 1 0 0 0 1 0 0 0  
## 4 1 1 1 0 1 0 1 1  
## 5 2 0 0 0 1 0 0 0  
## 9 1 0 1 0 0 1 0 1  
## 15 3 1 0 1 0 0 0 0  
## 16 1 1 0 1 0 0 0 0  
## 17 2 1 0 1 1 0 1 0  
## 18 1 0 1 0 1 0 1 1  
## 23 1 1 0 0 1 0 0 0  
## 25 3 1 0 1 1 0 1 0  
## 28 2 0 0 0 1 1 0 1  
## 29 2 1 0 0 1 0 0 0  
## 32 1 0 0 0 1 0 0 0  
## 33 1 1 0 0 1 0 0 0  
## 35 2 1 0 1 0 0 0 0  
## 42 1 1 0 0 1 0 0 0  
## 44 1 0 0 0 1 0 0 0  
## 48 1 0 0 1 1 0 1 0  
## 50 2 1 0 1 1 0 1 0  
## 55 2 1 0 0 1 0 0 0  
## 70 1 0 0 1 0 0 0 0  
## 73 1 1 0 1 0 0 0 0  
## 84 1 1 0 1 1 0 1 0  
## 87 2 1 0 1 1 0 1 0  
## 88 1 0 1 0 0 0 0 0  
## 89 2 0 0 0 1 0 0 0  
## 92 1 1 0 0 1 0 0 0  
## 93 3 1 1 1 0 0 1 0  
## 99 1 0 0 0 1 0 0 0  
## 103 1 0 1 0 1 0 1 1  
## 104 1 1 0 0 1 0 0 0  
## 106 2 0 0 0 1 0 0 0  
## 107 2 0 0 0 1 0 0 0  
## 109 1 0 1 0 0 0 0 0  
## 113 1 0 1 0 0 0 0 0  
## 114 3 1 0 0 1 0 0 0  
## 115 1 0 0 1 1 0 1 0  
## 116 1 0 0 0 1 0 0 0  
## 123 1 1 0 0 1 0 0 0  
## 125 3 1 0 0 1 0 0 0  
## 126 1 1 0 0 1 0 0 0  
## 127 1 1 0 0 1 0 0 0  
## 129 1 1 0 0 1 1 0 1  
## 130 1 1 0 0 1 0 0 0  
## 134 2 1 0 1 0 0 0 0  
## 135 1 1 0 0 1 0 0 0  
## 137 1 1 0 0 1 0 0 0  
## 141 1 1 0 0 1 0 0 0  
## 142 2 1 0 0 1 1 0 1  
## 143 1 0 0 1 1 0 1 0  
## 145 2 0 0 0 1 0 0 0  
## 147 1 0 0 1 1 0 1 0  
## 148 1 0 0 1 0 0 0 0  
## 149 1 0 0 0 1 0 0 0  
## 153 1 1 0 0 1 0 0 0  
## 156 2 1 0 1 1 0 1 0  
## 165 1 1 0 1 1 0 1 0  
## 166 1 1 0 1 1 0 1 0  
## 169 2 1 0 0 1 0 0 0  
## 173 1 0 0 0 0 1 0 0  
## 176 1 1 0 0 1 0 0 0  
## 182 1 0 0 1 0 0 0 0  
## 184 1 0 0 0 1 0 0 0  
## 186 2 0 0 0 1 0 0 0  
## 187 2 0 0 0 1 0 0 0  
## 193 1 1 0 1 1 0 1 0  
## 194 1 0 0 1 1 0 1 0  
## 195 1 1 0 0 1 0 0 0  
## 198 1 0 0 0 1 0 0 0  
## 208 1 0 0 0 1 0 0 0  
## 212 1 1 0 0 1 0 0 0  
## 214 1 1 0 0 1 0 0 0  
## 215 2 1 0 1 0 0 0 0  
## 216 1 1 0 1 0 0 0 0  
## 227 3 1 0 0 1 0 0 0  
## 233 2 1 0 0 1 0 0 0  
## 234 1 0 0 1 0 0 0 0  
## 237 1 0 0 1 0 0 0 0  
## 238 1 0 0 1 1 0 1 0  
## 240 2 1 0 0 1 0 0 0  
## 245 2 0 0 0 0 1 0 0  
## 246 1 0 0 0 1 0 0 0  
## 250 2 0 0 1 0 0 0 0  
## 253 2 1 1 0 0 0 0 0  
## 258 1 1 0 0 1 1 0 1  
## 270 3 1 0 0 1 0 0 0  
## 271 1 1 0 0 1 0 0 0  
## 277 1 1 0 0 1 0 0 0  
## 280 1 0 0 1 0 1 0 0  
## 283 1 1 0 0 1 0 0 0  
## 284 1 1 0 0 1 0 0 0  
## 286 1 0 0 1 0 0 0 0  
## 288 1 1 0 1 0 0 0 0  
## 292 1 1 0 1 0 0 0 0  
## 293 1 1 0 0 1 0 0 0  
## 295 1 1 0 0 1 0 0 0  
## 297 1 0 0 1 0 0 0 0  
## 305 2 0 0 0 1 0 0 0  
## 307 1 1 0 0 1 0 0 0  
## 311 3 1 0 0 1 0 0 0  
## 312 1 1 0 1 1 0 1 0  
## 320 1 0 0 0 1 0 0 0  
## 330 1 0 0 1 1 0 1 0  
## 331 3 0 0 0 1 0 0 0  
## 335 1 0 0 0 1 0 0 0  
## 336 1 1 0 1 1 0 1 0  
## 337 1 0 0 0 1 0 0 0  
## 339 3 1 0 0 1 0 0 0  
## 342 1 1 0 0 1 0 0 0  
## 345 2 1 0 0 1 0 0 0  
## 349 1 0 0 1 0 0 0 0  
## 353 2 1 0 1 0 0 0 0  
## 359 1 1 0 0 0 1 0 0  
## 362 1 0 0 0 1 0 0 0  
## 365 1 1 0 0 1 0 0 0  
## 366 2 1 0 1 1 0 1 0  
## 372 1 1 0 1 1 0 1 0  
## 375 1 1 0 0 1 0 0 0  
## 377 3 1 1 0 1 0 1 1  
## 379 1 0 0 0 1 0 0 0  
## 380 1 0 1 1 0 0 1 0  
## 386 1 0 0 0 1 0 0 0  
## 389 1 1 0 1 0 0 0 0  
## 393 1 1 0 0 1 0 0 0  
## 398 1 1 0 0 1 0 0 0  
## 400 2 0 0 0 1 0 0 0  
## 404 1 1 0 0 1 0 0 0  
## 405 1 1 0 1 0 0 0 0  
## 407 2 1 0 0 1 0 0 0  
## 410 1 1 0 1 1 0 1 0  
## 414 1 1 0 0 1 0 0 0  
## 420 1 1 0 0 1 0 0 0  
## 424 1 0 0 1 0 0 0 0  
## 429 1 1 0 1 0 1 0 0  
## 432 1 0 0 0 1 0 0 0  
## 444 2 0 0 0 1 0 0 0  
## 447 1 1 1 1 0 0 1 0  
## 452 1 1 0 1 0 0 0 0  
## 453 1 1 0 1 1 0 1 0  
## 454 1 0 0 1 0 0 0 0  
## 468 1 0 0 1 0 0 0 0  
## 473 2 1 0 1 0 0 0 0  
## 477 1 0 0 0 1 0 0 0  
## 480 3 1 1 0 0 0 0 0  
## 484 3 1 0 0 0 1 0 0  
## 487 1 0 0 1 0 0 0 0  
## JIJSJLWLB\_1 LogTotalWorkYrs  
## 1 1 0.3010300  
## 2 0 0.3010300  
## 4 1 0.0000000  
## 5 0 0.0000000  
## 9 1 1.1760913  
## 15 0 1.0000000  
## 16 0 1.3802112  
## 17 1 0.7781512  
## 18 1 1.0000000  
## 23 0 0.0000000  
## 25 1 0.6020600  
## 28 1 0.8450980  
## 29 0 0.7781512  
## 32 0 0.8450980  
## 33 0 1.2552725  
## 35 0 1.3010300  
## 42 0 0.6989700  
## 44 0 0.8450980  
## 48 1 1.1139434  
## 50 1 0.9030900  
## 55 0 0.6020600  
## 70 0 0.7781512  
## 73 0 1.2041200  
## 84 1 0.7781512  
## 87 1 0.0000000  
## 88 0 1.5185139  
## 89 0 0.4771213  
## 92 0 0.4771213  
## 93 1 0.8450980  
## 99 0 0.7781512  
## 103 1 0.3010300  
## 104 0 0.0000000  
## 106 0 0.0000000  
## 107 0 0.3010300  
## 109 0 1.4149733  
## 113 0 0.6020600  
## 114 0 0.4771213  
## 115 1 1.0000000  
## 116 0 0.7781512  
## 123 0 0.7781512  
## 125 0 1.1139434  
## 126 0 0.0000000  
## 127 0 0.8450980  
## 129 1 0.6020600  
## 130 0 0.7781512  
## 134 0 0.9542425  
## 135 0 0.8450980  
## 137 0 0.7781512  
## 141 0 1.0000000  
## 142 1 0.6989700  
## 143 1 1.2304489  
## 145 0 1.0791812  
## 147 1 0.9030900  
## 148 0 1.5051500  
## 149 0 1.0000000  
## 153 0 0.8450980  
## 156 1 0.6989700  
## 165 1 0.7781512  
## 166 1 0.6989700  
## 169 0 0.4771213  
## 173 0 1.3010300  
## 176 0 0.7781512  
## 182 0 1.0000000  
## 184 0 1.0413927  
## 186 0 1.2552725  
## 187 0 0.9030900  
## 193 1 0.9030900  
## 194 1 0.9030900  
## 195 0 0.6020600  
## 198 0 0.6989700  
## 208 0 0.9030900  
## 212 0 0.8450980  
## 214 0 0.7781512  
## 215 0 1.2041200  
## 216 0 0.8450980  
## 227 0 0.7781512  
## 233 0 1.0000000  
## 234 0 1.2787536  
## 237 0 0.9542425  
## 238 1 0.0000000  
## 240 0 0.3010300  
## 245 0 1.2041200  
## 246 0 0.6989700  
## 250 0 0.6020600  
## 253 0 1.5051500  
## 258 1 0.4771213  
## 270 0 0.0000000  
## 271 0 0.4771213  
## 277 0 1.2304489  
## 280 1 1.3222193  
## 283 0 1.1760913  
## 284 0 0.7781512  
## 286 0 0.6989700  
## 288 0 1.2304489  
## 292 0 0.7781512  
## 293 0 0.9030900  
## 295 0 0.6989700  
## 297 0 1.3979400  
## 305 0 0.7781512  
## 307 0 0.8450980  
## 311 0 0.8450980  
## 312 1 1.1461280  
## 320 0 0.6020600  
## 330 1 0.6020600  
## 331 0 0.6020600  
## 335 0 1.0000000  
## 336 1 0.9030900  
## 337 0 0.9030900  
## 339 0 0.0000000  
## 342 0 1.2787536  
## 345 0 0.3010300  
## 349 0 1.1760913  
## 353 0 0.6989700  
## 359 0 1.0791812  
## 362 0 0.6989700  
## 365 0 0.8450980  
## 366 1 1.0000000  
## 372 1 0.3010300  
## 375 0 0.6020600  
## 377 1 0.9542425  
## 379 0 0.8450980  
## 380 1 1.0000000  
## 386 0 0.3010300  
## 389 0 1.1139434  
## 393 0 1.0000000  
## 398 0 0.7781512  
## 400 0 0.9542425  
## 404 0 0.9030900  
## 405 0 1.0000000  
## 407 0 0.9030900  
## 410 1 0.8450980  
## 414 0 0.0000000  
## 420 0 0.4771213  
## 424 0 1.3424227  
## 429 1 0.9030900  
## 432 0 0.7781512  
## 444 0 0.7781512  
## 447 1 1.1760913  
## 452 0 1.1760913  
## 453 1 0.4771213  
## 454 0 0.7781512  
## 468 0 0.9030900  
## 473 0 0.6989700  
## 477 0 0.4771213  
## 480 0 1.2787536  
## 484 0 0.6020600  
## 487 0 0.6020600

CM = confusionMatrix(table(predict(model1,Case2\_adj[,c(17,58)]),Case2\_adj$Attrition))  
CM

## Confusion Matrix and Statistics  
##   
##   
## No Yes  
## No 314 72  
## Yes 62 41  
##   
## Accuracy : 0.726   
## 95% CI : (0.6841, 0.7651)  
## No Information Rate : 0.7689   
## P-Value [Acc > NIR] : 0.9884   
##   
## Kappa : 0.2043   
##   
## Mcnemar's Test P-Value : 0.4369   
##   
## Sensitivity : 0.8351   
## Specificity : 0.3628   
## Pos Pred Value : 0.8135   
## Neg Pred Value : 0.3981   
## Prevalence : 0.7689   
## Detection Rate : 0.6421   
## Detection Prevalence : 0.7894   
## Balanced Accuracy : 0.5990   
##   
## 'Positive' Class : No   
##

IDdf = as.data.frame(test[,c(1)])  
IDdf

## test[, c(1)]  
## 1 254  
## 2 398  
## 3 716  
## 4 733  
## 5 81  
## 6 162  
## 7 177  
## 8 178  
## 9 204  
## 10 287  
## 11 299  
## 12 313  
## 13 333  
## 14 380  
## 15 385  
## 16 433  
## 17 520  
## 18 530  
## 19 613  
## 20 625  
## 21 655  
## 22 28  
## 23 38  
## 24 244  
## 25 362  
## 26 363  
## 27 390  
## 28 439  
## 29 469  
## 30 539  
## 31 605  
## 32 614  
## 33 699  
## 34 751  
## 35 760  
## 36 851  
## 37 12  
## 38 15  
## 39 239  
## 40 550  
## 41 690  
## 42 701  
## 43 705  
## 44 765  
## 45 777  
## 46 6  
## 47 7  
## 48 19  
## 49 34  
## 50 36  
## 51 43  
## 52 46  
## 53 51  
## 54 52  
## 55 53  
## 56 66  
## 57 69  
## 58 101  
## 59 104  
## 60 111  
## 61 129  
## 62 149  
## 63 164  
## 64 170  
## 65 183  
## 66 185  
## 67 206  
## 68 209  
## 69 213  
## 70 219  
## 71 246  
## 72 260  
## 73 266  
## 74 267  
## 75 270  
## 76 306  
## 77 323  
## 78 324  
## 79 334  
## 80 340  
## 81 342  
## 82 353  
## 83 357  
## 84 371  
## 85 376  
## 86 396  
## 87 427  
## 88 428  
## 89 446  
## 90 455  
## 91 462  
## 92 466  
## 93 470  
## 94 473  
## 95 481  
## 96 482  
## 97 488  
## 98 492  
## 99 521  
## 100 525  
## 101 536  
## 102 542  
## 103 563  
## 104 590  
## 105 591  
## 106 604  
## 107 609  
## 108 617  
## 109 620  
## 110 631  
## 111 636  
## 112 642  
## 113 656  
## 114 684  
## 115 709  
## 116 713  
## 117 723  
## 118 746  
## 119 750  
## 120 756  
## 121 759  
## 122 766  
## 123 790  
## 124 800  
## 125 809  
## 126 819  
## 127 829  
## 128 837  
## 129 846  
## 130 850  
## 131 859  
## 132 14  
## 133 63  
## 134 90  
## 135 148  
## 136 205  
## 137 317  
## 138 356  
## 139 383  
## 140 384  
## 141 391  
## 142 608  
## 143 673  
## 144 696  
## 145 763  
## 146 785  
## 147 828

PredNB = cbind(IDdf, NBM)  
PredNB

## test[, c(1)] NBmodelPred  
## 1 254 Yes  
## 2 398 No  
## 3 716 Yes  
## 4 733 No  
## 5 81 No  
## 6 162 No  
## 7 177 No  
## 8 178 Yes  
## 9 204 Yes  
## 10 287 No  
## 11 299 Yes  
## 12 313 Yes  
## 13 333 No  
## 14 380 No  
## 15 385 No  
## 16 433 No  
## 17 520 No  
## 18 530 No  
## 19 613 Yes  
## 20 625 Yes  
## 21 655 No  
## 22 28 No  
## 23 38 No  
## 24 244 Yes  
## 25 362 Yes  
## 26 363 No  
## 27 390 No  
## 28 439 No  
## 29 469 Yes  
## 30 539 No  
## 31 605 Yes  
## 32 614 No  
## 33 699 No  
## 34 751 No  
## 35 760 No  
## 36 851 No  
## 37 12 No  
## 38 15 Yes  
## 39 239 No  
## 40 550 No  
## 41 690 No  
## 42 701 No  
## 43 705 No  
## 44 765 Yes  
## 45 777 No  
## 46 6 No  
## 47 7 No  
## 48 19 No  
## 49 34 No  
## 50 36 Yes  
## 51 43 Yes  
## 52 46 No  
## 53 51 Yes  
## 54 52 No  
## 55 53 No  
## 56 66 No  
## 57 69 Yes  
## 58 101 Yes  
## 59 104 Yes  
## 60 111 No  
## 61 129 No  
## 62 149 No  
## 63 164 No  
## 64 170 No  
## 65 183 No  
## 66 185 No  
## 67 206 Yes  
## 68 209 Yes  
## 69 213 No  
## 70 219 No  
## 71 246 No  
## 72 260 No  
## 73 266 No  
## 74 267 No  
## 75 270 No  
## 76 306 No  
## 77 323 No  
## 78 324 No  
## 79 334 No  
## 80 340 Yes  
## 81 342 No  
## 82 353 No  
## 83 357 No  
## 84 371 No  
## 85 376 No  
## 86 396 Yes  
## 87 427 No  
## 88 428 No  
## 89 446 No  
## 90 455 Yes  
## 91 462 No  
## 92 466 No  
## 93 470 No  
## 94 473 No  
## 95 481 No  
## 96 482 No  
## 97 488 No  
## 98 492 No  
## 99 521 No  
## 100 525 No  
## 101 536 No  
## 102 542 Yes  
## 103 563 No  
## 104 590 Yes  
## 105 591 No  
## 106 604 No  
## 107 609 Yes  
## 108 617 No  
## 109 620 No  
## 110 631 No  
## 111 636 No  
## 112 642 No  
## 113 656 No  
## 114 684 No  
## 115 709 No  
## 116 713 No  
## 117 723 Yes  
## 118 746 Yes  
## 119 750 No  
## 120 756 Yes  
## 121 759 No  
## 122 766 Yes  
## 123 790 No  
## 124 800 No  
## 125 809 No  
## 126 819 No  
## 127 829 No  
## 128 837 No  
## 129 846 No  
## 130 850 No  
## 131 859 Yes  
## 132 14 No  
## 133 63 No  
## 134 90 No  
## 135 148 Yes  
## 136 205 No  
## 137 317 No  
## 138 356 Yes  
## 139 383 No  
## 140 384 Yes  
## 141 391 No  
## 142 608 No  
## 143 673 No  
## 144 696 No  
## 145 763 No  
## 146 785 No  
## 147 828 No

colnames(PredNB) = c("ID", "Attrition")  
PredNB

## ID Attrition  
## 1 254 Yes  
## 2 398 No  
## 3 716 Yes  
## 4 733 No  
## 5 81 No  
## 6 162 No  
## 7 177 No  
## 8 178 Yes  
## 9 204 Yes  
## 10 287 No  
## 11 299 Yes  
## 12 313 Yes  
## 13 333 No  
## 14 380 No  
## 15 385 No  
## 16 433 No  
## 17 520 No  
## 18 530 No  
## 19 613 Yes  
## 20 625 Yes  
## 21 655 No  
## 22 28 No  
## 23 38 No  
## 24 244 Yes  
## 25 362 Yes  
## 26 363 No  
## 27 390 No  
## 28 439 No  
## 29 469 Yes  
## 30 539 No  
## 31 605 Yes  
## 32 614 No  
## 33 699 No  
## 34 751 No  
## 35 760 No  
## 36 851 No  
## 37 12 No  
## 38 15 Yes  
## 39 239 No  
## 40 550 No  
## 41 690 No  
## 42 701 No  
## 43 705 No  
## 44 765 Yes  
## 45 777 No  
## 46 6 No  
## 47 7 No  
## 48 19 No  
## 49 34 No  
## 50 36 Yes  
## 51 43 Yes  
## 52 46 No  
## 53 51 Yes  
## 54 52 No  
## 55 53 No  
## 56 66 No  
## 57 69 Yes  
## 58 101 Yes  
## 59 104 Yes  
## 60 111 No  
## 61 129 No  
## 62 149 No  
## 63 164 No  
## 64 170 No  
## 65 183 No  
## 66 185 No  
## 67 206 Yes  
## 68 209 Yes  
## 69 213 No  
## 70 219 No  
## 71 246 No  
## 72 260 No  
## 73 266 No  
## 74 267 No  
## 75 270 No  
## 76 306 No  
## 77 323 No  
## 78 324 No  
## 79 334 No  
## 80 340 Yes  
## 81 342 No  
## 82 353 No  
## 83 357 No  
## 84 371 No  
## 85 376 No  
## 86 396 Yes  
## 87 427 No  
## 88 428 No  
## 89 446 No  
## 90 455 Yes  
## 91 462 No  
## 92 466 No  
## 93 470 No  
## 94 473 No  
## 95 481 No  
## 96 482 No  
## 97 488 No  
## 98 492 No  
## 99 521 No  
## 100 525 No  
## 101 536 No  
## 102 542 Yes  
## 103 563 No  
## 104 590 Yes  
## 105 591 No  
## 106 604 No  
## 107 609 Yes  
## 108 617 No  
## 109 620 No  
## 110 631 No  
## 111 636 No  
## 112 642 No  
## 113 656 No  
## 114 684 No  
## 115 709 No  
## 116 713 No  
## 117 723 Yes  
## 118 746 Yes  
## 119 750 No  
## 120 756 Yes  
## 121 759 No  
## 122 766 Yes  
## 123 790 No  
## 124 800 No  
## 125 809 No  
## 126 819 No  
## 127 829 No  
## 128 837 No  
## 129 846 No  
## 130 850 No  
## 131 859 Yes  
## 132 14 No  
## 133 63 No  
## 134 90 No  
## 135 148 Yes  
## 136 205 No  
## 137 317 No  
## 138 356 Yes  
## 139 383 No  
## 140 384 Yes  
## 141 391 No  
## 142 608 No  
## 143 673 No  
## 144 696 No  
## 145 763 No  
## 146 785 No  
## 147 828 No

write.csv(PredNB,"C:/Users/Team Reed/OneDrive/JEFF/SMU/Doing Data Science/Project 2/Case2PredictionsReedAttrition.csv")

### Salary / MonthlyIncome Linear Regression Model

### Check all correlations then focus on correlation between the most notable MonthlyIncome and JobLevel, TotalWorkingYears, YearsAtCompany and Age

### Key variables based on correlation = JobLevel,TotalWorkingYears

df5 = Case2[,c(20,16,30,33,2)]  
 summary(df5)

## MonthlyIncome JobLevel TotalWorkingYears YearsAtCompany   
## Min. : 1081 Min. :1.000 Min. : 0.00 Min. : 0.000   
## 1st Qu.: 2840 1st Qu.:1.000 1st Qu.: 6.00 1st Qu.: 3.000   
## Median : 4946 Median :2.000 Median :10.00 Median : 5.000   
## Mean : 6390 Mean :2.039 Mean :11.05 Mean : 6.962   
## 3rd Qu.: 8182 3rd Qu.:3.000 3rd Qu.:15.00 3rd Qu.:10.000   
## Max. :19999 Max. :5.000 Max. :40.00 Max. :40.000   
## Age   
## Min. :18.00   
## 1st Qu.:30.00   
## Median :35.00   
## Mean :36.83   
## 3rd Qu.:43.00   
## Max. :60.00

corr = cor(df5)  
 round(corr, 2)

## MonthlyIncome JobLevel TotalWorkingYears YearsAtCompany Age  
## MonthlyIncome 1.00 0.95 0.78 0.49 0.48  
## JobLevel 0.95 1.00 0.78 0.52 0.48  
## TotalWorkingYears 0.78 0.78 1.00 0.64 0.65  
## YearsAtCompany 0.49 0.52 0.64 1.00 0.29  
## Age 0.48 0.48 0.65 0.29 1.00

### Create linear regression model to forecast salary / MonthlyIncome

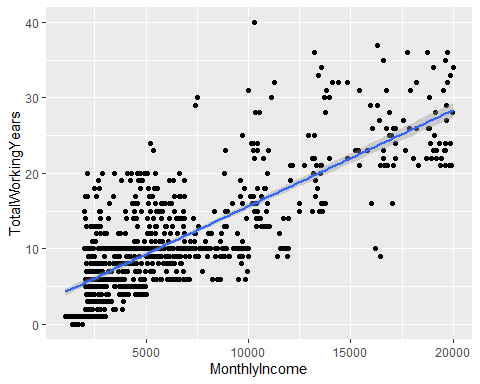
### Start with plotting the models

set.seed(10)  
splitPerc = .7  
trainIndices = sample(1:dim(Case2)[1],round(splitPerc \* dim(Case2)[1]))  
train = Case2[trainIndices,]  
test = Case2[-trainIndices,]  
  
### Simple Linear Regression between MonthlyIncome and TotalWorking Years  
fit1 <- lm(MonthlyIncome ~ TotalWorkingYears, data = train)  
summary(fit1)

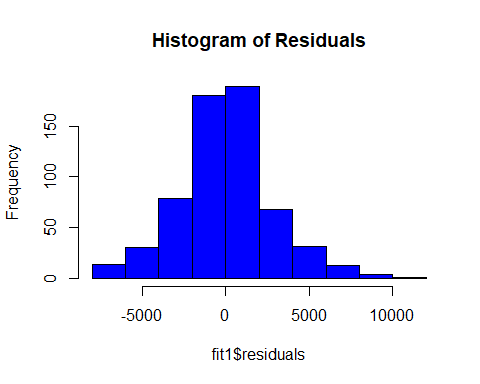
##   
## Call:  
## lm(formula = MonthlyIncome ~ TotalWorkingYears, data = train)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7813.7 -1745.5 6.7 1422.9 10995.6   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1134.7 209.0 5.43 8.16e-08 \*\*\*  
## TotalWorkingYears 476.9 15.8 30.17 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2864 on 607 degrees of freedom  
## Multiple R-squared: 0.6, Adjusted R-squared: 0.5993   
## F-statistic: 910.5 on 1 and 607 DF, p-value: < 2.2e-16

Case2 %>% ggplot(aes(x = MonthlyIncome, y = TotalWorkingYears)) + geom\_point() + geom\_smooth(method = "lm")

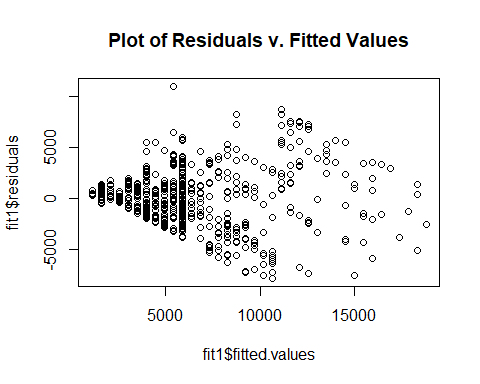
## `geom\_smooth()` using formula 'y ~ x'



hist(fit1$residuals, col = "blue", main = "Histogram of Residuals")



plot(fit1$fitted.values,fit1$residuals, main = "Plot of Residuals v. Fitted Values")



Model1\_Preds = predict(fit1, newdata = test)  
as.data.frame(Model1\_Preds)

## Model1\_Preds  
## 14 5903.226  
## 16 4949.523  
## 21 5903.226  
## 23 6380.078  
## 24 4949.523  
## 28 3042.115  
## 29 3518.967  
## 30 4472.671  
## 34 3995.819  
## 37 9718.041  
## 38 3042.115  
## 44 20208.782  
## 47 3995.819  
## 49 3518.967  
## 52 4472.671  
## 53 1611.560  
## 57 4949.523  
## 61 5903.226  
## 65 5903.226  
## 66 5426.375  
## 76 1611.560  
## 82 3995.819  
## 83 8287.486  
## 85 8764.338  
## 87 1134.708  
## 99 3995.819  
## 100 2565.264  
## 104 1611.560  
## 108 6856.930  
## 113 1611.560  
## 124 3518.967  
## 130 2565.264  
## 131 7810.634  
## 135 3995.819  
## 138 4949.523  
## 144 17824.523  
## 150 4949.523  
## 153 3042.115  
## 156 3995.819  
## 161 1611.560  
## 162 4472.671  
## 163 4949.523  
## 168 16393.967  
## 170 11148.597  
## 172 3995.819  
## 174 4472.671  
## 175 4949.523  
## 180 3995.819  
## 183 5903.226  
## 187 5903.226  
## 190 9241.189  
## 193 6856.930  
## 197 5903.226  
## 201 5903.226  
## 203 5903.226  
## 208 3518.967  
## 210 6856.930  
## 211 3995.819  
## 216 7333.782  
## 222 9241.189  
## 226 3518.967  
## 232 13532.856  
## 236 7333.782  
## 237 7810.634  
## 240 10671.745  
## 242 4472.671  
## 243 5426.375  
## 246 16870.819  
## 247 2088.412  
## 248 7333.782  
## 250 4949.523  
## 251 3995.819  
## 253 6380.078  
## 255 9718.041  
## 257 5903.226  
## 258 5903.226  
## 259 14009.708  
## 260 5903.226  
## 262 6380.078  
## 265 2088.412  
## 268 5426.375  
## 269 4949.523  
## 272 14486.560  
## 276 1611.560  
## 281 7810.634  
## 282 4949.523  
## 287 3518.967  
## 293 1611.560  
## 295 6856.930  
## 296 1611.560  
## 297 3042.115  
## 300 12579.152  
## 303 5903.226  
## 306 5903.226  
## 310 1611.560  
## 312 8764.338  
## 313 4472.671  
## 315 3995.819  
## 316 3995.819  
## 318 5426.375  
## 329 2565.264  
## 331 6380.078  
## 332 4472.671  
## 336 4949.523  
## 343 11625.449  
## 345 3042.115  
## 352 1611.560  
## 353 12102.301  
## 355 10671.745  
## 357 5426.375  
## 358 4472.671  
## 359 1611.560  
## 360 3995.819  
## 363 1611.560  
## 364 3518.967  
## 366 3995.819  
## 374 17347.671  
## 375 11625.449  
## 383 16393.967  
## 386 8287.486  
## 387 4949.523  
## 393 8764.338  
## 395 6380.078  
## 397 5426.375  
## 401 11625.449  
## 404 6856.930  
## 405 3042.115  
## 406 7810.634  
## 407 5903.226  
## 411 1611.560  
## 413 7333.782  
## 415 1611.560  
## 419 5426.375  
## 421 9241.189  
## 422 10194.893  
## 424 5903.226  
## 427 11148.597  
## 428 3518.967  
## 433 3995.819  
## 434 3518.967  
## 441 8764.338  
## 442 3995.819  
## 445 10194.893  
## 446 3518.967  
## 447 5903.226  
## 448 11625.449  
## 449 13056.004  
## 450 6380.078  
## 451 4949.523  
## 459 8287.486  
## 473 5903.226  
## 476 7810.634  
## 480 5903.226  
## 481 10194.893  
## 492 3995.819  
## 495 4472.671  
## 496 4472.671  
## 497 5903.226  
## 502 5426.375  
## 505 4949.523  
## 514 5426.375  
## 517 4949.523  
## 518 5903.226  
## 519 5903.226  
## 522 5426.375  
## 528 12102.301  
## 529 5426.375  
## 535 3995.819  
## 545 4949.523  
## 546 3995.819  
## 549 2565.264  
## 550 5903.226  
## 551 18301.375  
## 552 8764.338  
## 553 3518.967  
## 563 7810.634  
## 566 13056.004  
## 567 7810.634  
## 579 3518.967  
## 582 3042.115  
## 587 2088.412  
## 588 7810.634  
## 589 3995.819  
## 590 3042.115  
## 591 16393.967  
## 595 10671.745  
## 597 13532.856  
## 598 5903.226  
## 608 17347.671  
## 609 2088.412  
## 615 11148.597  
## 616 7333.782  
## 617 11148.597  
## 620 2088.412  
## 623 6856.930  
## 624 4472.671  
## 629 3995.819  
## 632 5426.375  
## 633 5426.375  
## 635 1611.560  
## 639 14963.412  
## 642 5903.226  
## 643 1611.560  
## 648 5426.375  
## 651 4472.671  
## 653 5903.226  
## 655 4472.671  
## 661 1611.560  
## 674 3995.819  
## 677 12102.301  
## 678 10671.745  
## 682 12102.301  
## 689 9718.041  
## 693 13056.004  
## 697 4472.671  
## 707 5903.226  
## 710 6856.930  
## 713 3518.967  
## 720 3995.819  
## 722 3042.115  
## 733 5903.226  
## 735 6380.078  
## 736 5903.226  
## 738 5903.226  
## 743 5903.226  
## 744 5426.375  
## 750 5903.226  
## 755 5903.226  
## 756 2088.412  
## 758 8287.486  
## 759 2565.264  
## 762 3995.819  
## 766 15917.115  
## 778 6380.078  
## 785 5903.226  
## 788 7810.634  
## 790 5903.226  
## 791 2565.264  
## 792 15440.264  
## 793 5426.375  
## 805 4949.523  
## 812 5426.375  
## 818 8764.338  
## 819 6856.930  
## 821 7333.782  
## 826 13532.856  
## 827 7333.782  
## 832 8764.338  
## 833 8764.338  
## 834 16393.967  
## 836 5903.226  
## 841 7333.782  
## 845 3995.819  
## 846 10194.893  
## 847 10194.893  
## 848 10671.745  
## 859 2565.264  
## 862 3042.115  
## 864 5903.226  
## 865 3995.819  
## 867 12102.301

# Model 1  
MSPE = data.frame(Observed = test$MonthlyIncome, Predicted = Model1\_Preds)  
MSPE$Residual = MSPE$Observed - MSPE$Predicted  
MSPE$SquaredResidual = MSPE$Residual^2  
MSPE

## Observed Predicted Residual SquaredResidual  
## 14 3919 5903.226 -1984.226478 3.937155e+06  
## 16 2404 4949.523 -2545.522771 6.479686e+06  
## 21 2216 5903.226 -3687.226478 1.359564e+07  
## 23 2362 6380.078 -4018.078331 1.614495e+07  
## 24 2436 4949.523 -2513.522771 6.317797e+06  
## 28 2321 3042.115 -721.115359 5.200074e+05  
## 29 3464 3518.967 -54.967212 3.021394e+03  
## 30 2479 4472.671 -1993.670918 3.974724e+06  
## 34 2546 3995.819 -1449.819065 2.101975e+06  
## 37 2461 9718.041 -7257.041304 5.266465e+07  
## 38 3730 3042.115 687.884641 4.731853e+05  
## 44 10312 20208.782 -9896.782075 9.794630e+07  
## 47 2028 3995.819 -1967.819065 3.872312e+06  
## 49 2587 3518.967 -931.967212 8.685629e+05  
## 52 2119 4472.671 -2353.670918 5.539767e+06  
## 53 2693 1611.560 1081.440201 1.169513e+06  
## 57 2561 4949.523 -2388.522771 5.705041e+06  
## 61 2166 5903.226 -3737.226478 1.396686e+07  
## 65 2022 5903.226 -3881.226478 1.506392e+07  
## 66 7553 5426.375 2126.625375 4.522535e+06  
## 76 2398 1611.560 786.440201 6.184882e+05  
## 82 8224 3995.819 4228.180935 1.787751e+07  
## 83 10448 8287.486 2160.514256 4.667822e+06  
## 85 6134 8764.338 -2630.337597 6.918676e+06  
## 87 1878 1134.708 743.292054 5.524831e+05  
## 99 4400 3995.819 404.180935 1.633622e+05  
## 100 2785 2565.264 219.736495 4.828413e+04  
## 104 1091 1611.560 -520.559799 2.709825e+05  
## 108 6172 6856.930 -684.930184 4.691294e+05  
## 113 1081 1611.560 -530.559799 2.814937e+05  
## 124 3041 3518.967 -477.967212 2.284527e+05  
## 130 2302 2565.264 -263.263505 6.930767e+04  
## 131 7314 7810.634 -496.633891 2.466452e+05  
## 135 9854 3995.819 5858.180935 3.431828e+07  
## 138 6380 4949.523 1430.477229 2.046265e+06  
## 144 19636 17824.523 1811.477191 3.281450e+06  
## 150 2942 4949.523 -2007.522771 4.030148e+06  
## 153 2696 3042.115 -346.115359 1.197958e+05  
## 156 4936 3995.819 940.180935 8.839402e+05  
## 161 2804 1611.560 1192.440201 1.421914e+06  
## 162 2267 4472.671 -2205.670918 4.864984e+06  
## 163 2109 4949.523 -2840.522771 8.068570e+06  
## 168 18200 16393.967 1806.032751 3.261754e+06  
## 170 19626 11148.597 8477.403136 7.186636e+07  
## 172 3760 3995.819 -235.819065 5.561063e+04  
## 174 2127 4472.671 -2345.670918 5.502172e+06  
## 175 5063 4949.523 113.477229 1.287708e+04  
## 180 2932 3995.819 -1063.819065 1.131711e+06  
## 183 5679 5903.226 -224.226478 5.027751e+04  
## 187 2258 5903.226 -3645.226478 1.328768e+07  
## 190 2587 9241.189 -6654.189451 4.427824e+07  
## 193 3673 6856.930 -3183.930184 1.013741e+07  
## 197 2011 5903.226 -3892.226478 1.514943e+07  
## 201 4723 5903.226 -1180.226478 1.392935e+06  
## 203 7094 5903.226 1190.773522 1.417942e+06  
## 208 2194 3518.967 -1324.967212 1.755538e+06  
## 210 2782 6856.930 -4074.930184 1.660506e+07  
## 211 4766 3995.819 770.180935 5.931787e+05  
## 216 5974 7333.782 -1359.782038 1.849007e+06  
## 222 5163 9241.189 -4078.189451 1.663163e+07  
## 226 2819 3518.967 -699.967212 4.899541e+05  
## 232 17099 13532.856 3566.143870 1.271738e+07  
## 236 2451 7333.782 -4882.782038 2.384156e+07  
## 237 6687 7810.634 -1123.633891 1.262553e+06  
## 240 4081 10671.745 -6590.745010 4.343792e+07  
## 242 5295 4472.671 822.329082 6.762251e+05  
## 243 5257 5426.375 -169.374625 2.868776e+04  
## 246 13402 16870.819 -3468.819102 1.203271e+07  
## 247 2372 2088.412 283.588348 8.042235e+04  
## 248 10845 7333.782 3511.217962 1.232865e+07  
## 250 2756 4949.523 -2193.522771 4.811542e+06  
## 251 3597 3995.819 -398.819065 1.590566e+05  
## 253 6870 6380.078 489.921669 2.400232e+05  
## 255 10820 9718.041 1101.958696 1.214313e+06  
## 257 5126 5903.226 -777.226478 6.040810e+05  
## 258 6545 5903.226 641.773522 4.118733e+05  
## 259 16413 14009.708 2403.292017 5.775813e+06  
## 260 16184 5903.226 10280.773522 1.056943e+08  
## 262 2514 6380.078 -3866.078331 1.494656e+07  
## 265 2070 2088.412 -18.411652 3.389889e+02  
## 268 8376 5426.375 2949.625375 8.700290e+06  
## 269 3755 4949.523 -1194.522771 1.426885e+06  
## 272 19144 14486.560 4657.440164 2.169175e+07  
## 276 2723 1611.560 1111.440201 1.235299e+06  
## 281 6162 7810.634 -1648.633891 2.717994e+06  
## 282 2559 4949.523 -2390.522771 5.714599e+06  
## 287 2045 3518.967 -1473.967212 2.172579e+06  
## 293 1563 1611.560 -48.559799 2.358054e+03  
## 295 3748 6856.930 -3108.930184 9.665447e+06  
## 296 1951 1611.560 339.440201 1.152197e+05  
## 297 2328 3042.115 -714.115359 5.099607e+05  
## 300 18880 12579.152 6300.847577 3.970068e+07  
## 303 7406 5903.226 1502.773522 2.258328e+06  
## 306 5056 5903.226 -847.226478 7.177927e+05  
## 310 2008 1611.560 396.440201 1.571648e+05  
## 312 13757 8764.338 4992.662403 2.492668e+07  
## 313 2438 4472.671 -2034.670918 4.139886e+06  
## 315 2774 3995.819 -1221.819065 1.492842e+06  
## 316 2426 3995.819 -1569.819065 2.464332e+06  
## 318 5968 5426.375 541.625375 2.933580e+05  
## 329 2661 2565.264 95.736495 9.165476e+03  
## 331 5674 6380.078 -706.078331 4.985466e+05  
## 332 2279 4472.671 -2193.670918 4.812192e+06  
## 336 3537 4949.523 -1412.522771 1.995221e+06  
## 343 13237 11625.449 1611.551283 2.597098e+06  
## 345 4617 3042.115 1574.884641 2.480262e+06  
## 352 2318 1611.560 706.440201 4.990578e+05  
## 353 19049 12102.301 6946.699430 4.825663e+07  
## 355 2133 10671.745 -8538.745010 7.291017e+07  
## 357 7412 5426.375 1985.625375 3.942708e+06  
## 358 7756 4472.671 3283.329082 1.078025e+07  
## 359 2341 1611.560 729.440201 5.320830e+05  
## 360 3419 3995.819 -576.819065 3.327202e+05  
## 363 2083 1611.560 471.440201 2.222559e+05  
## 364 4033 3518.967 514.032788 2.642297e+05  
## 366 2290 3995.819 -1705.819065 2.909819e+06  
## 374 19999 17347.671 2651.329044 7.029546e+06  
## 375 19436 11625.449 7810.551283 6.100471e+07  
## 383 14411 16393.967 -1982.967249 3.932159e+06  
## 386 13582 8287.486 5294.514256 2.803188e+07  
## 387 4262 4949.523 -687.522771 4.726876e+05  
## 393 10209 8764.338 1444.662403 2.087049e+06  
## 395 5265 6380.078 -1115.078331 1.243400e+06  
## 397 8008 5426.375 2581.625375 6.664790e+06  
## 401 17159 11625.449 5533.551283 3.062019e+07  
## 404 2564 6856.930 -4292.930184 1.842925e+07  
## 405 3688 3042.115 645.884641 4.171670e+05  
## 406 6474 7810.634 -1336.633891 1.786590e+06  
## 407 11935 5903.226 6031.773522 3.638229e+07  
## 411 1274 1611.560 -337.559799 1.139466e+05  
## 413 5957 7333.782 -1376.782038 1.895529e+06  
## 415 2109 1611.560 497.440201 2.474468e+05  
## 419 2974 5426.375 -2452.374625 6.014141e+06  
## 421 3420 9241.189 -5821.189451 3.388625e+07  
## 422 4735 10194.893 -5459.893157 2.981043e+07  
## 424 2406 5903.226 -3497.226478 1.223059e+07  
## 427 15402 11148.597 4253.403136 1.809144e+07  
## 428 2066 3518.967 -1452.967212 2.111114e+06  
## 433 3944 3995.819 -51.819065 2.685215e+03  
## 434 3038 3518.967 -480.967212 2.313295e+05  
## 441 13744 8764.338 4979.662403 2.479704e+07  
## 442 5373 3995.819 1377.180935 1.896627e+06  
## 445 4257 10194.893 -5937.893157 3.525858e+07  
## 446 2911 3518.967 -607.967212 3.696241e+05  
## 447 6142 5903.226 238.773522 5.701279e+04  
## 448 18665 11625.449 7039.551283 4.955528e+07  
## 449 16880 13056.004 3823.995723 1.462294e+07  
## 450 8606 6380.078 2225.921669 4.954727e+06  
## 451 2654 4949.523 -2295.522771 5.269425e+06  
## 459 13458 8287.486 5170.514256 2.673422e+07  
## 473 3491 5903.226 -2412.226478 5.818837e+06  
## 476 6553 7810.634 -1257.633891 1.581643e+06  
## 480 5094 5903.226 -809.226478 6.548475e+05  
## 481 12061 10194.893 1866.106843 3.482355e+06  
## 492 2436 3995.819 -1559.819065 2.433036e+06  
## 495 5993 4472.671 1520.329082 2.311401e+06  
## 496 3162 4472.671 -1310.670918 1.717858e+06  
## 497 5042 5903.226 -861.226478 7.417110e+05  
## 502 6465 5426.375 1038.625375 1.078743e+06  
## 505 3629 4949.523 -1320.522771 1.743780e+06  
## 514 11416 5426.375 5989.625375 3.587561e+07  
## 517 3578 4949.523 -1371.522771 1.881075e+06  
## 518 5363 5903.226 -540.226478 2.918446e+05  
## 519 6811 5903.226 907.773522 8.240528e+05  
## 522 9991 5426.375 4564.625375 2.083580e+07  
## 528 10976 12102.301 -1126.300570 1.268553e+06  
## 529 8621 5426.375 3194.625375 1.020563e+07  
## 535 2725 3995.819 -1270.819065 1.614981e+06  
## 545 6032 4949.523 1082.477229 1.171757e+06  
## 546 6499 3995.819 2503.180935 6.265915e+06  
## 549 3917 2565.264 1351.736495 1.827192e+06  
## 550 6294 5903.226 390.773522 1.527039e+05  
## 551 17779 18301.375 -522.374662 2.728753e+05  
## 552 13664 8764.338 4899.662403 2.400669e+07  
## 553 3517 3518.967 -1.967212 3.869922e+00  
## 563 6781 7810.634 -1029.633891 1.060146e+06  
## 566 19081 13056.004 6024.995723 3.630057e+07  
## 567 6854 7810.634 -956.633891 9.151484e+05  
## 579 4898 3518.967 1379.032788 1.901731e+06  
## 582 4765 3042.115 1722.884641 2.968331e+06  
## 587 2296 2088.412 207.588348 4.309292e+04  
## 588 5343 7810.634 -2467.633891 6.089217e+06  
## 589 3907 3995.819 -88.819065 7.888826e+03  
## 590 2500 3042.115 -542.115359 2.938891e+05  
## 591 11245 16393.967 -5148.967249 2.651186e+07  
## 595 6651 10671.745 -4020.745010 1.616639e+07  
## 597 16752 13532.856 3219.143870 1.036289e+07  
## 598 5769 5903.226 -134.226478 1.801675e+04  
## 608 19038 17347.671 1690.329044 2.857212e+06  
## 609 2096 2088.412 7.588348 5.758302e+01  
## 615 13496 11148.597 2347.403136 5.510301e+06  
## 616 6842 7333.782 -491.782038 2.418496e+05  
## 617 17068 11148.597 5919.403136 3.503933e+07  
## 620 2323 2088.412 234.588348 5.503169e+04  
## 623 3072 6856.930 -3784.930184 1.432570e+07  
## 624 3229 4472.671 -1243.670918 1.546717e+06  
## 629 3312 3995.819 -683.819065 4.676085e+05  
## 632 2468 5426.375 -2958.374625 8.751980e+06  
## 633 3633 5426.375 -1793.374625 3.216193e+06  
## 635 2552 1611.560 940.440201 8.844278e+05  
## 639 19665 14963.412 4701.588310 2.210493e+07  
## 642 4558 5903.226 -1345.226478 1.809634e+06  
## 643 1483 1611.560 -128.559799 1.652762e+04  
## 648 5661 5426.375 234.625375 5.504907e+04  
## 651 2238 4472.671 -2234.670918 4.993754e+06  
## 653 9380 5903.226 3476.773522 1.208795e+07  
## 655 4477 4472.671 4.329082 1.874095e+01  
## 661 2476 1611.560 864.440201 7.472569e+05  
## 674 8463 3995.819 4467.180935 1.995571e+07  
## 677 8865 12102.301 -3237.300570 1.048011e+07  
## 678 4960 10671.745 -5711.745010 3.262403e+07  
## 682 16606 12102.301 4503.699430 2.028331e+07  
## 689 10761 9718.041 1042.958696 1.087763e+06  
## 693 13225 13056.004 168.995723 2.855955e+04  
## 697 6502 4472.671 2029.329082 4.118177e+06  
## 707 8380 5903.226 2476.773522 6.134407e+06  
## 710 8686 6856.930 1829.069816 3.345496e+06  
## 713 4221 3518.967 702.032788 4.928500e+05  
## 720 6274 3995.819 2278.180935 5.190108e+06  
## 722 4999 3042.115 1956.884641 3.829397e+06  
## 733 5154 5903.226 -749.226478 5.613403e+05  
## 735 5647 6380.078 -733.078331 5.374038e+05  
## 736 7847 5903.226 1943.773522 3.778256e+06  
## 738 9738 5903.226 3834.773522 1.470549e+07  
## 743 5405 5903.226 -498.226478 2.482296e+05  
## 744 4163 5426.375 -1263.374625 1.596115e+06  
## 750 6929 5903.226 1025.773522 1.052211e+06  
## 755 8412 5903.226 2508.773522 6.293945e+06  
## 756 2728 2088.412 639.588348 4.090733e+05  
## 758 5486 8287.486 -2801.485744 7.848322e+06  
## 759 2322 2565.264 -243.263505 5.917713e+04  
## 762 6500 3995.819 2504.180935 6.270922e+06  
## 766 15427 15917.115 -490.115396 2.402131e+05  
## 778 5441 6380.078 -939.078331 8.818681e+05  
## 785 5673 5903.226 -230.226478 5.300423e+04  
## 788 10596 7810.634 2785.366109 7.758264e+06  
## 790 5343 5903.226 -560.226478 3.138537e+05  
## 791 2610 2565.264 44.736495 2.001354e+03  
## 792 7525 15440.264 -7915.263543 6.265140e+07  
## 793 5006 5426.375 -420.374625 1.767148e+05  
## 805 4779 4949.523 -170.522771 2.907802e+04  
## 812 5454 5426.375 27.625375 7.631614e+02  
## 818 4312 8764.338 -4452.337597 1.982331e+07  
## 819 5220 6856.930 -1636.930184 2.679540e+06  
## 821 10368 7333.782 3034.217962 9.206479e+06  
## 826 18789 13532.856 5256.143870 2.762705e+07  
## 827 2329 7333.782 -5004.782038 2.504784e+07  
## 832 10377 8764.338 1612.662403 2.600680e+06  
## 833 6230 8764.338 -2534.337597 6.422867e+06  
## 834 14118 16393.967 -2275.967249 5.180027e+06  
## 836 7587 5903.226 1683.773522 2.835093e+06  
## 841 5171 7333.782 -2162.782038 4.677626e+06  
## 845 4907 3995.819 911.180935 8.302507e+05  
## 846 6151 10194.893 -4043.893157 1.635307e+07  
## 847 6347 10194.893 -3847.893157 1.480628e+07  
## 848 10932 10671.745 260.254990 6.773266e+04  
## 859 2899 2565.264 333.736495 1.113800e+05  
## 862 4538 3042.115 1495.884641 2.237671e+06  
## 864 5337 5903.226 -566.226478 3.206124e+05  
## 865 6029 3995.819 2033.180935 4.133825e+06  
## 867 10231 12102.301 -1871.300570 3.501766e+06

mean(MSPE$SquaredResidual)

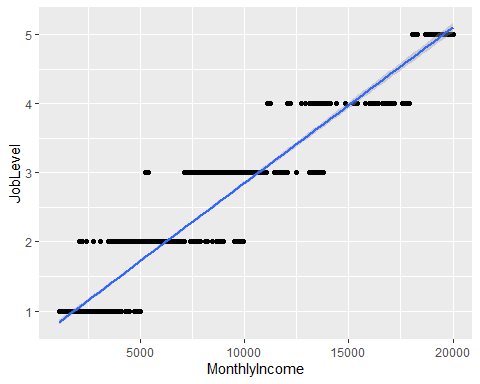
## [1] 8652707

### Simple Linear Regression between MonthlyIncome and JobLevel  
fit2 <- lm(MonthlyIncome ~ JobLevel, data = Case2)  
summary(fit2)

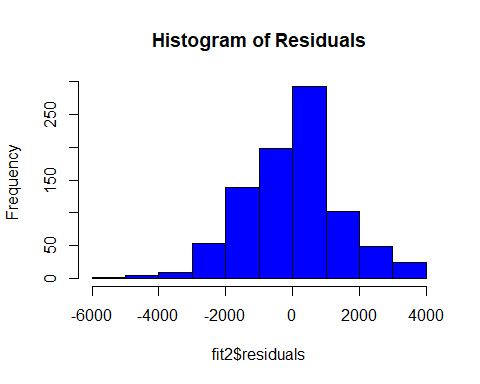
##   
## Call:  
## lm(formula = MonthlyIncome ~ JobLevel, data = Case2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5037.1 -928.2 80.1 697.1 3723.6   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1793.93 101.68 -17.64 <2e-16 \*\*\*  
## JobLevel 4013.67 43.98 91.26 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1413 on 868 degrees of freedom  
## Multiple R-squared: 0.9056, Adjusted R-squared: 0.9055   
## F-statistic: 8329 on 1 and 868 DF, p-value: < 2.2e-16

Case2 %>% ggplot(aes(x = MonthlyIncome, y = JobLevel)) + geom\_point() + geom\_smooth(method = "lm")

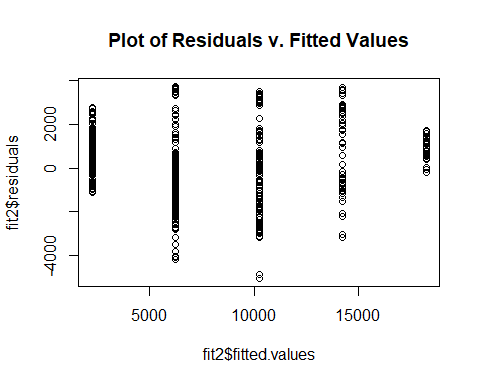
## `geom\_smooth()` using formula 'y ~ x'



hist(fit2$residuals, col = "blue", main = "Histogram of Residuals")



plot(fit2$fitted.values,fit2$residuals, main = "Plot of Residuals v. Fitted Values")



Model2\_Preds = predict(fit2, newdata = test)  
as.data.frame(Model2\_Preds)

## Model2\_Preds  
## 14 2219.737  
## 16 2219.737  
## 21 2219.737  
## 23 2219.737  
## 24 2219.737  
## 28 2219.737  
## 29 2219.737  
## 30 2219.737  
## 34 2219.737  
## 37 2219.737  
## 38 2219.737  
## 44 10247.080  
## 47 2219.737  
## 49 2219.737  
## 52 2219.737  
## 53 2219.737  
## 57 2219.737  
## 61 2219.737  
## 65 2219.737  
## 66 10247.080  
## 76 2219.737  
## 82 6233.408  
## 83 10247.080  
## 85 6233.408  
## 87 2219.737  
## 99 2219.737  
## 100 2219.737  
## 104 2219.737  
## 108 6233.408  
## 113 2219.737  
## 124 2219.737  
## 130 2219.737  
## 131 10247.080  
## 135 6233.408  
## 138 6233.408  
## 144 18274.422  
## 150 2219.737  
## 153 2219.737  
## 156 2219.737  
## 161 2219.737  
## 162 2219.737  
## 163 2219.737  
## 168 18274.422  
## 170 18274.422  
## 172 2219.737  
## 174 2219.737  
## 175 6233.408  
## 180 2219.737  
## 183 6233.408  
## 187 2219.737  
## 190 2219.737  
## 193 2219.737  
## 197 2219.737  
## 201 2219.737  
## 203 10247.080  
## 208 2219.737  
## 210 2219.737  
## 211 2219.737  
## 216 6233.408  
## 222 6233.408  
## 226 2219.737  
## 232 14260.751  
## 236 2219.737  
## 237 6233.408  
## 240 6233.408  
## 242 6233.408  
## 243 6233.408  
## 246 14260.751  
## 247 2219.737  
## 248 10247.080  
## 250 2219.737  
## 251 2219.737  
## 253 6233.408  
## 255 10247.080  
## 257 6233.408  
## 258 6233.408  
## 259 14260.751  
## 260 14260.751  
## 262 2219.737  
## 265 2219.737  
## 268 10247.080  
## 269 2219.737  
## 272 18274.422  
## 276 2219.737  
## 281 6233.408  
## 282 2219.737  
## 287 2219.737  
## 293 2219.737  
## 295 2219.737  
## 296 2219.737  
## 297 2219.737  
## 300 18274.422  
## 303 10247.080  
## 306 6233.408  
## 310 2219.737  
## 312 10247.080  
## 313 2219.737  
## 315 2219.737  
## 316 2219.737  
## 318 6233.408  
## 329 2219.737  
## 331 6233.408  
## 332 2219.737  
## 336 2219.737  
## 343 14260.751  
## 345 6233.408  
## 352 2219.737  
## 353 18274.422  
## 355 6233.408  
## 357 10247.080  
## 358 6233.408  
## 359 2219.737  
## 360 2219.737  
## 363 2219.737  
## 364 6233.408  
## 366 2219.737  
## 374 18274.422  
## 375 18274.422  
## 383 14260.751  
## 386 10247.080  
## 387 6233.408  
## 393 10247.080  
## 395 6233.408  
## 397 10247.080  
## 401 14260.751  
## 404 2219.737  
## 405 2219.737  
## 406 6233.408  
## 407 10247.080  
## 411 2219.737  
## 413 6233.408  
## 415 2219.737  
## 419 2219.737  
## 421 2219.737  
## 422 6233.408  
## 424 6233.408  
## 427 14260.751  
## 428 2219.737  
## 433 2219.737  
## 434 2219.737  
## 441 10247.080  
## 442 6233.408  
## 445 2219.737  
## 446 2219.737  
## 447 6233.408  
## 448 18274.422  
## 449 14260.751  
## 450 10247.080  
## 451 2219.737  
## 459 10247.080  
## 473 6233.408  
## 476 6233.408  
## 480 6233.408  
## 481 10247.080  
## 492 2219.737  
## 495 6233.408  
## 496 2219.737  
## 497 6233.408  
## 502 6233.408  
## 505 2219.737  
## 514 10247.080  
## 517 2219.737  
## 518 6233.408  
## 519 6233.408  
## 522 10247.080  
## 528 10247.080  
## 529 10247.080  
## 535 2219.737  
## 545 6233.408  
## 546 6233.408  
## 549 2219.737  
## 550 6233.408  
## 551 14260.751  
## 552 10247.080  
## 553 2219.737  
## 563 6233.408  
## 566 18274.422  
## 567 6233.408  
## 579 6233.408  
## 582 6233.408  
## 587 2219.737  
## 588 6233.408  
## 589 2219.737  
## 590 2219.737  
## 591 14260.751  
## 595 6233.408  
## 597 14260.751  
## 598 6233.408  
## 608 18274.422  
## 609 2219.737  
## 615 14260.751  
## 616 6233.408  
## 617 14260.751  
## 620 2219.737  
## 623 6233.408  
## 624 2219.737  
## 629 2219.737  
## 632 2219.737  
## 633 6233.408  
## 635 2219.737  
## 639 18274.422  
## 642 6233.408  
## 643 2219.737  
## 648 6233.408  
## 651 2219.737  
## 653 10247.080  
## 655 2219.737  
## 661 2219.737  
## 674 6233.408  
## 677 10247.080  
## 678 6233.408  
## 682 14260.751  
## 689 10247.080  
## 693 14260.751  
## 697 6233.408  
## 707 10247.080  
## 710 6233.408  
## 713 6233.408  
## 720 6233.408  
## 722 6233.408  
## 733 6233.408  
## 735 6233.408  
## 736 6233.408  
## 738 10247.080  
## 743 6233.408  
## 744 6233.408  
## 750 6233.408  
## 755 10247.080  
## 756 2219.737  
## 758 6233.408  
## 759 2219.737  
## 762 6233.408  
## 766 14260.751  
## 778 6233.408  
## 785 6233.408  
## 788 10247.080  
## 790 6233.408  
## 791 2219.737  
## 792 10247.080  
## 793 6233.408  
## 805 6233.408  
## 812 6233.408  
## 818 6233.408  
## 819 6233.408  
## 821 10247.080  
## 826 18274.422  
## 827 2219.737  
## 832 10247.080  
## 833 6233.408  
## 834 14260.751  
## 836 10247.080  
## 841 6233.408  
## 845 6233.408  
## 846 6233.408  
## 847 6233.408  
## 848 10247.080  
## 859 2219.737  
## 862 6233.408  
## 864 6233.408  
## 865 6233.408  
## 867 10247.080

# Model 2  
MSPE = data.frame(Observed = test$MonthlyIncome, Predicted = Model2\_Preds)  
MSPE$Residual = MSPE$Observed - MSPE$Predicted  
MSPE$SquaredResidual = MSPE$Residual^2  
MSPE

## Observed Predicted Residual SquaredResidual  
## 14 3919 2219.737 1699.263047 2.887495e+06  
## 16 2404 2219.737 184.263047 3.395287e+04  
## 21 2216 2219.737 -3.736953 1.396482e+01  
## 23 2362 2219.737 142.263047 2.023877e+04  
## 24 2436 2219.737 216.263047 4.676971e+04  
## 28 2321 2219.737 101.263047 1.025420e+04  
## 29 3464 2219.737 1244.263047 1.548191e+06  
## 30 2479 2219.737 259.263047 6.721733e+04  
## 34 2546 2219.737 326.263047 1.064476e+05  
## 37 2461 2219.737 241.263047 5.820786e+04  
## 38 3730 2219.737 1510.263047 2.280894e+06  
## 44 10312 10247.080 64.920456 4.214666e+03  
## 47 2028 2219.737 -191.736953 3.676306e+04  
## 49 2587 2219.737 367.263047 1.348821e+05  
## 52 2119 2219.737 -100.736953 1.014793e+04  
## 53 2693 2219.737 473.263047 2.239779e+05  
## 57 2561 2219.737 341.263047 1.164605e+05  
## 61 2166 2219.737 -53.736953 2.887660e+03  
## 65 2022 2219.737 -197.736953 3.909990e+04  
## 66 7553 10247.080 -2694.079544 7.258065e+06  
## 76 2398 2219.737 178.263047 3.177771e+04  
## 82 8224 6233.408 1990.591752 3.962456e+06  
## 83 10448 10247.080 200.920456 4.036903e+04  
## 85 6134 6233.408 -99.408248 9.882000e+03  
## 87 1878 2219.737 -341.736953 1.167841e+05  
## 99 4400 2219.737 2180.263047 4.753547e+06  
## 100 2785 2219.737 565.263047 3.195223e+05  
## 104 1091 2219.737 -1128.736953 1.274047e+06  
## 108 6172 6233.408 -61.408248 3.770973e+03  
## 113 1081 2219.737 -1138.736953 1.296722e+06  
## 124 3041 2219.737 821.263047 6.744730e+05  
## 130 2302 2219.737 82.263047 6.767209e+03  
## 131 7314 10247.080 -2933.079544 8.602956e+06  
## 135 9854 6233.408 3620.591752 1.310868e+07  
## 138 6380 6233.408 146.591752 2.148914e+04  
## 144 19636 18274.422 1361.577865 1.853894e+06  
## 150 2942 2219.737 722.263047 5.216639e+05  
## 153 2696 2219.737 476.263047 2.268265e+05  
## 156 4936 2219.737 2716.263047 7.378085e+06  
## 161 2804 2219.737 584.263047 3.413633e+05  
## 162 2267 2219.737 47.263047 2.233796e+03  
## 163 2109 2219.737 -110.736953 1.226267e+04  
## 168 18200 18274.422 -74.422135 5.538654e+03  
## 170 19626 18274.422 1351.577865 1.826763e+06  
## 172 3760 2219.737 1540.263047 2.372410e+06  
## 174 2127 2219.737 -92.736953 8.600142e+03  
## 175 5063 6233.408 -1170.408248 1.369855e+06  
## 180 2932 2219.737 712.263047 5.073186e+05  
## 183 5679 6233.408 -554.408248 3.073685e+05  
## 187 2258 2219.737 38.263047 1.464061e+03  
## 190 2587 2219.737 367.263047 1.348821e+05  
## 193 3673 2219.737 1453.263047 2.111973e+06  
## 197 2011 2219.737 -208.736953 4.357112e+04  
## 201 4723 2219.737 2503.263047 6.266326e+06  
## 203 7094 10247.080 -3153.079544 9.941911e+06  
## 208 2194 2219.737 -25.736953 6.623907e+02  
## 210 2782 2219.737 562.263047 3.161397e+05  
## 211 4766 2219.737 2546.263047 6.483456e+06  
## 216 5974 6233.408 -259.408248 6.729264e+04  
## 222 5163 6233.408 -1070.408248 1.145774e+06  
## 226 2819 2219.737 599.263047 3.591162e+05  
## 232 17099 14260.751 2838.249161 8.055658e+06  
## 236 2451 2219.737 231.263047 5.348260e+04  
## 237 6687 6233.408 453.591752 2.057455e+05  
## 240 4081 6233.408 -2152.408248 4.632861e+06  
## 242 5295 6233.408 -938.408248 8.806100e+05  
## 243 5257 6233.408 -976.408248 9.533731e+05  
## 246 13402 14260.751 -858.750839 7.374530e+05  
## 247 2372 2219.737 152.263047 2.318404e+04  
## 248 10845 10247.080 597.920456 3.575089e+05  
## 250 2756 2219.737 536.263047 2.875781e+05  
## 251 3597 2219.737 1377.263047 1.896854e+06  
## 253 6870 6233.408 636.591752 4.052491e+05  
## 255 10820 10247.080 572.920456 3.282378e+05  
## 257 5126 6233.408 -1107.408248 1.226353e+06  
## 258 6545 6233.408 311.591752 9.708942e+04  
## 259 16413 14260.751 2152.249161 4.632176e+06  
## 260 16184 14260.751 1923.249161 3.698887e+06  
## 262 2514 2219.737 294.263047 8.659074e+04  
## 265 2070 2219.737 -149.736953 2.242116e+04  
## 268 8376 10247.080 -1871.079544 3.500939e+06  
## 269 3755 2219.737 1535.263047 2.357033e+06  
## 272 19144 18274.422 869.577865 7.561657e+05  
## 276 2723 2219.737 503.263047 2.532737e+05  
## 281 6162 6233.408 -71.408248 5.099138e+03  
## 282 2559 2219.737 339.263047 1.150994e+05  
## 287 2045 2219.737 -174.736953 3.053300e+04  
## 293 1563 2219.737 -656.736953 4.313034e+05  
## 295 3748 2219.737 1528.263047 2.335588e+06  
## 296 1951 2219.737 -268.736953 7.221955e+04  
## 297 2328 2219.737 108.263047 1.172089e+04  
## 300 18880 18274.422 605.577865 3.667246e+05  
## 303 7406 10247.080 -2841.079544 8.071733e+06  
## 306 5056 6233.408 -1177.408248 1.386290e+06  
## 310 2008 2219.737 -211.736953 4.483254e+04  
## 312 13757 10247.080 3509.920456 1.231954e+07  
## 313 2438 2219.737 218.263047 4.763876e+04  
## 315 2774 2219.737 554.263047 3.072075e+05  
## 316 2426 2219.737 206.263047 4.254444e+04  
## 318 5968 6233.408 -265.408248 7.044154e+04  
## 329 2661 2219.737 441.263047 1.947131e+05  
## 331 5674 6233.408 -559.408248 3.129376e+05  
## 332 2279 2219.737 59.263047 3.512109e+03  
## 336 3537 2219.737 1317.263047 1.735182e+06  
## 343 13237 14260.751 -1023.750839 1.048066e+06  
## 345 4617 6233.408 -1616.408248 2.612776e+06  
## 352 2318 2219.737 98.263047 9.655626e+03  
## 353 19049 18274.422 774.577865 5.999709e+05  
## 355 2133 6233.408 -4100.408248 1.681335e+07  
## 357 7412 10247.080 -2835.079544 8.037676e+06  
## 358 7756 6233.408 1522.591752 2.318286e+06  
## 359 2341 2219.737 121.263047 1.470473e+04  
## 360 3419 2219.737 1199.263047 1.438232e+06  
## 363 2083 2219.737 -136.736953 1.869699e+04  
## 364 4033 6233.408 -2200.408248 4.841796e+06  
## 366 2290 2219.737 70.263047 4.936896e+03  
## 374 19999 18274.422 1724.577865 2.974169e+06  
## 375 19436 18274.422 1161.577865 1.349263e+06  
## 383 14411 14260.751 150.249161 2.257481e+04  
## 386 13582 10247.080 3334.920456 1.112169e+07  
## 387 4262 6233.408 -1971.408248 3.886450e+06  
## 393 10209 10247.080 -38.079544 1.450052e+03  
## 395 5265 6233.408 -968.408248 9.378145e+05  
## 397 8008 10247.080 -2239.079544 5.013477e+06  
## 401 17159 14260.751 2898.249161 8.399848e+06  
## 404 2564 2219.737 344.263047 1.185170e+05  
## 405 3688 2219.737 1468.263047 2.155796e+06  
## 406 6474 6233.408 240.591752 5.788439e+04  
## 407 11935 10247.080 1687.920456 2.849075e+06  
## 411 1274 2219.737 -945.736953 8.944184e+05  
## 413 5957 6233.408 -276.408248 7.640152e+04  
## 415 2109 2219.737 -110.736953 1.226267e+04  
## 419 2974 2219.737 754.263047 5.689127e+05  
## 421 3420 2219.737 1200.263047 1.440631e+06  
## 422 4735 6233.408 -1498.408248 2.245227e+06  
## 424 2406 6233.408 -3827.408248 1.464905e+07  
## 427 15402 14260.751 1141.249161 1.302450e+06  
## 428 2066 2219.737 -153.736953 2.363505e+04  
## 433 3944 2219.737 1724.263047 2.973083e+06  
## 434 3038 2219.737 818.263047 6.695544e+05  
## 441 13744 10247.080 3496.920456 1.222845e+07  
## 442 5373 6233.408 -860.408248 7.403024e+05  
## 445 4257 2219.737 2037.263047 4.150441e+06  
## 446 2911 2219.737 691.263047 4.778446e+05  
## 447 6142 6233.408 -91.408248 8.355468e+03  
## 448 18665 18274.422 390.577865 1.525511e+05  
## 449 16880 14260.751 2619.249161 6.860466e+06  
## 450 8606 10247.080 -1641.079544 2.693142e+06  
## 451 2654 2219.737 434.263047 1.885844e+05  
## 459 13458 10247.080 3210.920456 1.031001e+07  
## 473 3491 6233.408 -2742.408248 7.520803e+06  
## 476 6553 6233.408 319.591752 1.021389e+05  
## 480 5094 6233.408 -1139.408248 1.298251e+06  
## 481 12061 10247.080 1813.920456 3.290307e+06  
## 492 2436 2219.737 216.263047 4.676971e+04  
## 495 5993 6233.408 -240.408248 5.779613e+04  
## 496 3162 2219.737 942.263047 8.878597e+05  
## 497 5042 6233.408 -1191.408248 1.419454e+06  
## 502 6465 6233.408 231.591752 5.363474e+04  
## 505 3629 2219.737 1409.263047 1.986022e+06  
## 514 11416 10247.080 1168.920456 1.366375e+06  
## 517 3578 2219.737 1358.263047 1.844879e+06  
## 518 5363 6233.408 -870.408248 7.576105e+05  
## 519 6811 6233.408 577.591752 3.336122e+05  
## 522 9991 10247.080 -256.079544 6.557673e+04  
## 528 10976 10247.080 728.920456 5.313250e+05  
## 529 8621 10247.080 -1626.079544 2.644135e+06  
## 535 2725 2219.737 505.263047 2.552907e+05  
## 545 6032 6233.408 -201.408248 4.056528e+04  
## 546 6499 6233.408 265.591752 7.053898e+04  
## 549 3917 2219.737 1697.263047 2.880702e+06  
## 550 6294 6233.408 60.591752 3.671360e+03  
## 551 17779 14260.751 3518.249161 1.237808e+07  
## 552 13664 10247.080 3416.920456 1.167535e+07  
## 553 3517 2219.737 1297.263047 1.682891e+06  
## 563 6781 6233.408 547.591752 2.998567e+05  
## 566 19081 18274.422 806.577865 6.505679e+05  
## 567 6854 6233.408 620.591752 3.851341e+05  
## 579 4898 6233.408 -1335.408248 1.783315e+06  
## 582 4765 6233.408 -1468.408248 2.156223e+06  
## 587 2296 2219.737 76.263047 5.816052e+03  
## 588 5343 6233.408 -890.408248 7.928268e+05  
## 589 3907 2219.737 1687.263047 2.846857e+06  
## 590 2500 2219.737 280.263047 7.854738e+04  
## 591 11245 14260.751 -3015.750839 9.094753e+06  
## 595 6651 6233.408 417.591752 1.743829e+05  
## 597 16752 14260.751 2491.249161 6.206322e+06  
## 598 5769 6233.408 -464.408248 2.156750e+05  
## 608 19038 18274.422 763.577865 5.830512e+05  
## 609 2096 2219.737 -123.736953 1.531083e+04  
## 615 13496 14260.751 -764.750839 5.848438e+05  
## 616 6842 6233.408 608.591752 3.703839e+05  
## 617 17068 14260.751 2807.249161 7.880648e+06  
## 620 2323 2219.737 103.263047 1.066326e+04  
## 623 3072 6233.408 -3161.408248 9.994502e+06  
## 624 3229 2219.737 1009.263047 1.018612e+06  
## 629 3312 2219.737 1092.263047 1.193039e+06  
## 632 2468 2219.737 248.263047 6.163454e+04  
## 633 3633 6233.408 -2600.408248 6.762123e+06  
## 635 2552 2219.737 332.263047 1.103987e+05  
## 639 19665 18274.422 1390.577865 1.933707e+06  
## 642 4558 6233.408 -1675.408248 2.806993e+06  
## 643 1483 2219.737 -736.736953 5.427813e+05  
## 648 5661 6233.408 -572.408248 3.276512e+05  
## 651 2238 2219.737 18.263047 3.335389e+02  
## 653 9380 10247.080 -867.079544 7.518269e+05  
## 655 4477 2219.737 2257.263047 5.095236e+06  
## 661 2476 2219.737 256.263047 6.567075e+04  
## 674 8463 6233.408 2229.591752 4.971079e+06  
## 677 8865 10247.080 -1382.079544 1.910144e+06  
## 678 4960 6233.408 -1273.408248 1.621569e+06  
## 682 16606 14260.751 2345.249161 5.500194e+06  
## 689 10761 10247.080 513.920456 2.641142e+05  
## 693 13225 14260.751 -1035.750839 1.072780e+06  
## 697 6502 6233.408 268.591752 7.214153e+04  
## 707 8380 10247.080 -1867.079544 3.485986e+06  
## 710 8686 6233.408 2452.591752 6.015206e+06  
## 713 4221 6233.408 -2012.408248 4.049787e+06  
## 720 6274 6233.408 40.591752 1.647690e+03  
## 722 4999 6233.408 -1234.408248 1.523764e+06  
## 733 5154 6233.408 -1079.408248 1.165122e+06  
## 735 5647 6233.408 -586.408248 3.438746e+05  
## 736 7847 6233.408 1613.591752 2.603678e+06  
## 738 9738 10247.080 -509.079544 2.591620e+05  
## 743 5405 6233.408 -828.408248 6.862602e+05  
## 744 4163 6233.408 -2070.408248 4.286590e+06  
## 750 6929 6233.408 695.591752 4.838479e+05  
## 755 8412 10247.080 -1835.079544 3.367517e+06  
## 756 2728 2219.737 508.263047 2.583313e+05  
## 758 5486 6233.408 -747.408248 5.586191e+05  
## 759 2322 2219.737 102.263047 1.045773e+04  
## 762 6500 6233.408 266.591752 7.107116e+04  
## 766 15427 14260.751 1166.249161 1.360137e+06  
## 778 5441 6233.408 -792.408248 6.279108e+05  
## 785 5673 6233.408 -560.408248 3.140574e+05  
## 788 10596 10247.080 348.920456 1.217455e+05  
## 790 5343 6233.408 -890.408248 7.928268e+05  
## 791 2610 2219.737 390.263047 1.523052e+05  
## 792 7525 10247.080 -2722.079544 7.409717e+06  
## 793 5006 6233.408 -1227.408248 1.506531e+06  
## 805 4779 6233.408 -1454.408248 2.115303e+06  
## 812 5454 6233.408 -779.408248 6.074772e+05  
## 818 4312 6233.408 -1921.408248 3.691810e+06  
## 819 5220 6233.408 -1013.408248 1.026996e+06  
## 821 10368 10247.080 120.920456 1.462176e+04  
## 826 18789 18274.422 514.577865 2.647904e+05  
## 827 2329 2219.737 109.263047 1.193841e+04  
## 832 10377 10247.080 129.920456 1.687932e+04  
## 833 6230 6233.408 -3.408248 1.161616e+01  
## 834 14118 14260.751 -142.750839 2.037780e+04  
## 836 7587 10247.080 -2660.079544 7.076023e+06  
## 841 5171 6233.408 -1062.408248 1.128711e+06  
## 845 4907 6233.408 -1326.408248 1.759359e+06  
## 846 6151 6233.408 -82.408248 6.791119e+03  
## 847 6347 6233.408 113.591752 1.290309e+04  
## 848 10932 10247.080 684.920456 4.691160e+05  
## 859 2899 2219.737 679.263047 4.613983e+05  
## 862 4538 6233.408 -1695.408248 2.874409e+06  
## 864 5337 6233.408 -896.408248 8.035477e+05  
## 865 6029 6233.408 -204.408248 4.178273e+04  
## 867 10231 10247.080 -16.079544 2.585517e+02

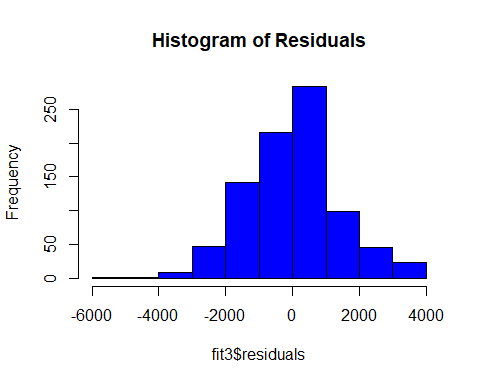
mean(MSPE$SquaredResidual)

## [1] 1887539

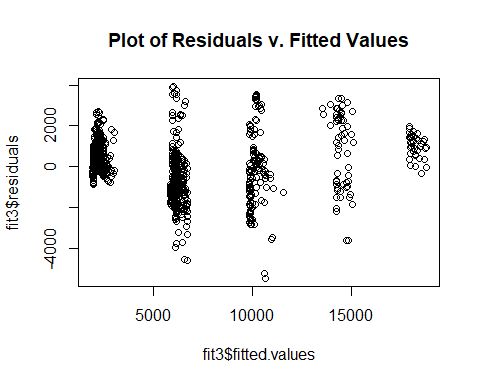
### Chosen model based on RMSE and visualization of assumptions  
### Multiple Linear Regression between MonthlyIncome and JobLevel + TotalWorkingYears  
fit3 <- lm(MonthlyIncome ~ JobLevel + TotalWorkingYears, data = Case2)  
summary(fit3)

##   
## Call:  
## lm(formula = MonthlyIncome ~ JobLevel + TotalWorkingYears, data = Case2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5469.9 -876.8 64.5 728.3 3937.5   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1798.38 99.98 -17.987 < 2e-16 \*\*\*  
## JobLevel 3714.12 69.21 53.664 < 2e-16 \*\*\*  
## TotalWorkingYears 55.66 10.04 5.544 3.94e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1390 on 867 degrees of freedom  
## Multiple R-squared: 0.9088, Adjusted R-squared: 0.9086   
## F-statistic: 4322 on 2 and 867 DF, p-value: < 2.2e-16

hist(fit3$residuals, col = "blue", main = "Histogram of Residuals")



plot(fit3$fitted.values,fit3$residuals, main = "Plot of Residuals v. Fitted Values")



Model3\_Preds = predict(fit3, newdata = test)  
LRM = as.data.frame(Model3\_Preds)  
LRM

## Model3\_Preds  
## 14 2472.385  
## 16 2361.058  
## 21 2472.385  
## 23 2528.049  
## 24 2361.058  
## 28 2138.402  
## 29 2194.066  
## 30 2305.394  
## 34 2249.730  
## 37 2917.697  
## 38 2138.402  
## 44 11570.548  
## 47 2249.730  
## 49 2194.066  
## 52 2305.394  
## 53 1971.410  
## 57 2361.058  
## 61 2472.385  
## 65 2472.385  
## 66 9844.966  
## 76 1971.410  
## 82 5963.852  
## 83 10178.950  
## 85 6520.491  
## 87 1915.746  
## 99 2249.730  
## 100 2082.738  
## 104 1971.410  
## 108 6297.836  
## 113 1971.410  
## 124 2194.066  
## 130 2082.738  
## 131 10123.286  
## 135 5963.852  
## 138 6075.180  
## 144 18720.473  
## 150 2361.058  
## 153 2138.402  
## 156 2249.730  
## 161 1971.410  
## 162 2305.394  
## 163 2361.058  
## 168 18553.481  
## 170 17941.178  
## 172 2249.730  
## 174 2305.394  
## 175 6075.180  
## 180 2249.730  
## 183 6186.508  
## 187 2472.385  
## 190 2862.033  
## 193 2583.713  
## 197 2472.385  
## 201 2472.385  
## 203 9900.630  
## 208 2194.066  
## 210 2583.713  
## 211 2249.730  
## 216 6353.499  
## 222 6576.155  
## 226 2194.066  
## 232 14505.375  
## 236 2639.377  
## 237 6409.163  
## 240 6743.147  
## 242 6019.516  
## 243 6130.844  
## 246 14895.022  
## 247 2027.074  
## 248 10067.622  
## 250 2361.058  
## 251 2249.730  
## 253 6242.172  
## 255 10345.941  
## 257 6186.508  
## 258 6186.508  
## 259 14561.039  
## 260 13614.752  
## 262 2528.049  
## 265 2027.074  
## 268 9844.966  
## 269 2361.058  
## 272 18330.825  
## 276 1971.410  
## 281 6409.163  
## 282 2361.058  
## 287 2194.066  
## 293 1971.410  
## 295 2583.713  
## 296 1971.410  
## 297 2138.402  
## 300 18108.169  
## 303 9900.630  
## 306 6186.508  
## 310 1971.410  
## 312 10234.613  
## 313 2305.394  
## 315 2249.730  
## 316 2249.730  
## 318 6130.844  
## 329 2082.738  
## 331 6242.172  
## 332 2305.394  
## 336 2361.058  
## 343 14282.719  
## 345 5852.524  
## 352 1971.410  
## 353 18052.505  
## 355 6743.147  
## 357 9844.966  
## 358 6019.516  
## 359 1971.410  
## 360 2249.730  
## 363 1971.410  
## 364 5908.188  
## 366 2249.730  
## 374 18664.809  
## 375 17996.842  
## 383 14839.359  
## 386 10178.950  
## 387 6075.180  
## 393 10234.613  
## 395 6242.172  
## 397 9844.966  
## 401 14282.719  
## 404 2583.713  
## 405 2138.402  
## 406 6409.163  
## 407 9900.630  
## 411 1971.410  
## 413 6353.499  
## 415 1971.410  
## 419 2416.722  
## 421 2862.033  
## 422 6687.483  
## 424 6186.508  
## 427 14227.055  
## 428 2194.066  
## 433 2249.730  
## 434 2194.066  
## 441 10234.613  
## 442 5963.852  
## 445 2973.361  
## 446 2194.066  
## 447 6186.508  
## 448 17996.842  
## 449 14449.711  
## 450 9956.294  
## 451 2361.058  
## 459 10178.950  
## 473 6186.508  
## 476 6409.163  
## 480 6186.508  
## 481 10401.605  
## 492 2249.730  
## 495 6019.516  
## 496 2305.394  
## 497 6186.508  
## 502 6130.844  
## 505 2361.058  
## 514 9844.966  
## 517 2361.058  
## 518 6186.508  
## 519 6186.508  
## 522 9844.966  
## 528 10624.261  
## 529 9844.966  
## 535 2249.730  
## 545 6075.180  
## 546 5963.852  
## 549 2082.738  
## 550 6186.508  
## 551 15062.014  
## 552 10234.613  
## 553 2194.066  
## 563 6409.163  
## 566 18163.833  
## 567 6409.163  
## 579 5908.188  
## 582 5852.524  
## 587 2027.074  
## 588 6409.163  
## 589 2249.730  
## 590 2138.402  
## 591 14839.359  
## 595 6743.147  
## 597 14505.375  
## 598 6186.508  
## 608 18664.809  
## 609 2027.074  
## 615 14227.055  
## 616 6353.499  
## 617 14227.055  
## 620 2027.074  
## 623 6297.836  
## 624 2305.394  
## 629 2249.730  
## 632 2416.722  
## 633 6130.844  
## 635 1971.410  
## 639 18386.489  
## 642 6186.508  
## 643 1971.410  
## 648 6130.844  
## 651 2305.394  
## 653 9900.630  
## 655 2305.394  
## 661 1971.410  
## 674 5963.852  
## 677 10624.261  
## 678 6743.147  
## 682 14338.383  
## 689 10345.941  
## 693 14449.711  
## 697 6019.516  
## 707 9900.630  
## 710 6297.836  
## 713 5908.188  
## 720 5963.852  
## 722 5852.524  
## 733 6186.508  
## 735 6242.172  
## 736 6186.508  
## 738 9900.630  
## 743 6186.508  
## 744 6130.844  
## 750 6186.508  
## 755 9900.630  
## 756 2027.074  
## 758 6464.827  
## 759 2082.738  
## 762 5963.852  
## 766 14783.695  
## 778 6242.172  
## 785 6186.508  
## 788 10123.286  
## 790 6186.508  
## 791 2082.738  
## 792 11013.908  
## 793 6130.844  
## 805 6075.180  
## 812 6130.844  
## 818 6520.491  
## 819 6297.836  
## 821 10067.622  
## 826 18219.497  
## 827 2639.377  
## 832 10234.613  
## 833 6520.491  
## 834 14839.359  
## 836 9900.630  
## 841 6353.499  
## 845 5963.852  
## 846 6687.483  
## 847 6687.483  
## 848 10457.269  
## 859 2082.738  
## 862 5852.524  
## 864 6186.508  
## 865 5963.852  
## 867 10624.261

# Model 3 - run statistics and write to csv file  
MSPE = data.frame(Observed = test$MonthlyIncome, Predicted = Model3\_Preds)  
MSPE$Residual = MSPE$Observed - MSPE$Predicted  
MSPE$SquaredResidual = MSPE$Residual^2  
MSPE

## Observed Predicted Residual SquaredResidual  
## 14 3919 2472.385 1.446615e+03 2.092694e+06  
## 16 2404 2361.058 4.294241e+01 1.844050e+03  
## 21 2216 2472.385 -2.563854e+02 6.573350e+04  
## 23 2362 2528.049 -1.660494e+02 2.757239e+04  
## 24 2436 2361.058 7.494241e+01 5.616364e+03  
## 28 2321 2138.402 1.825981e+02 3.334207e+04  
## 29 3464 2194.066 1.269934e+03 1.612733e+06  
## 30 2479 2305.394 1.736063e+02 3.013916e+04  
## 34 2546 2249.730 2.962703e+02 8.777607e+04  
## 37 2461 2917.697 -4.566969e+02 2.085720e+05  
## 38 3730 2138.402 1.591598e+03 2.533185e+06  
## 44 10312 11570.548 -1.258548e+03 1.583942e+06  
## 47 2028 2249.730 -2.217297e+02 4.916408e+04  
## 49 2587 2194.066 3.929342e+02 1.543973e+05  
## 52 2119 2305.394 -1.863937e+02 3.474260e+04  
## 53 2693 1971.410 7.215899e+02 5.206920e+05  
## 57 2561 2361.058 1.999424e+02 3.997697e+04  
## 61 2166 2472.385 -3.063854e+02 9.387204e+04  
## 65 2022 2472.385 -4.503854e+02 2.028470e+05  
## 66 7553 9844.966 -2.291966e+03 5.253108e+06  
## 76 2398 1971.410 4.265899e+02 1.819789e+05  
## 82 8224 5963.852 2.260148e+03 5.108269e+06  
## 83 10448 10178.950 2.690504e+02 7.238814e+04  
## 85 6134 6520.491 -3.864912e+02 1.493755e+05  
## 87 1878 1915.746 -3.774618e+01 1.424774e+03  
## 99 4400 2249.730 2.150270e+03 4.623662e+06  
## 100 2785 2082.738 7.022620e+02 4.931720e+05  
## 104 1091 1971.410 -8.804101e+02 7.751220e+05  
## 108 6172 6297.836 -1.258355e+02 1.583458e+04  
## 113 1081 1971.410 -8.904101e+02 7.928302e+05  
## 124 3041 2194.066 8.469342e+02 7.172975e+05  
## 130 2302 2082.738 2.192620e+02 4.807584e+04  
## 131 7314 10123.286 -2.809286e+03 7.892086e+06  
## 135 9854 5963.852 3.890148e+03 1.513325e+07  
## 138 6380 6075.180 3.048202e+02 9.291533e+04  
## 144 19636 18720.473 9.155274e+02 8.381905e+05  
## 150 2942 2361.058 5.809424e+02 3.374941e+05  
## 153 2696 2138.402 5.575981e+02 3.109157e+05  
## 156 4936 2249.730 2.686270e+03 7.216048e+06  
## 161 2804 1971.410 8.325899e+02 6.932059e+05  
## 162 2267 2305.394 -3.839367e+01 1.474074e+03  
## 163 2109 2361.058 -2.520576e+02 6.353303e+04  
## 168 18200 18553.481 -3.534808e+02 1.249487e+05  
## 170 19626 17941.178 1.684822e+03 2.838627e+06  
## 172 3760 2249.730 1.510270e+03 2.280916e+06  
## 174 2127 2305.394 -1.783937e+02 3.182430e+04  
## 175 5063 6075.180 -1.012180e+03 1.024508e+06  
## 180 2932 2249.730 6.822703e+02 4.654927e+05  
## 183 5679 6186.508 -5.075077e+02 2.575641e+05  
## 187 2258 2472.385 -2.143854e+02 4.596112e+04  
## 190 2587 2862.033 -2.750329e+02 7.564311e+04  
## 193 3673 2583.713 1.089287e+03 1.186546e+06  
## 197 2011 2472.385 -4.613854e+02 2.128765e+05  
## 201 4723 2472.385 2.250615e+03 5.065266e+06  
## 203 7094 9900.630 -2.806630e+03 7.877172e+06  
## 208 2194 2194.066 -6.581222e-02 4.331249e-03  
## 210 2782 2583.713 1.982867e+02 3.931762e+04  
## 211 4766 2249.730 2.516270e+03 6.331616e+06  
## 216 5974 6353.499 -3.794995e+02 1.440198e+05  
## 222 5163 6576.155 -1.413155e+03 1.997008e+06  
## 226 2819 2194.066 6.249342e+02 3.905427e+05  
## 232 17099 14505.375 2.593625e+03 6.726891e+06  
## 236 2451 2639.377 -1.883772e+02 3.548598e+04  
## 237 6687 6409.163 2.778366e+02 7.719318e+04  
## 240 4081 6743.147 -2.662147e+03 7.087026e+06  
## 242 5295 6019.516 -7.245159e+02 5.249233e+05  
## 243 5257 6130.844 -8.738438e+02 7.636029e+05  
## 246 13402 14895.022 -1.493022e+03 2.229116e+06  
## 247 2372 2027.074 3.449260e+02 1.189739e+05  
## 248 10845 10067.622 7.773783e+02 6.043170e+05  
## 250 2756 2361.058 3.949424e+02 1.559795e+05  
## 251 3597 2249.730 1.347270e+03 1.815137e+06  
## 253 6870 6242.172 6.278284e+02 3.941685e+05  
## 255 10820 10345.941 4.740587e+02 2.247316e+05  
## 257 5126 6186.508 -1.060508e+03 1.124677e+06  
## 258 6545 6186.508 3.584923e+02 1.285167e+05  
## 259 16413 14561.039 1.851961e+03 3.429760e+06  
## 260 16184 13614.752 2.569248e+03 6.601034e+06  
## 262 2514 2528.049 -1.404937e+01 1.973849e+02  
## 265 2070 2027.074 4.292597e+01 1.842639e+03  
## 268 8376 9844.966 -1.468966e+03 2.157861e+06  
## 269 3755 2361.058 1.393942e+03 1.943075e+06  
## 272 19144 18330.825 8.131749e+02 6.612534e+05  
## 276 2723 1971.410 7.515899e+02 5.648874e+05  
## 281 6162 6409.163 -2.471634e+02 6.108974e+04  
## 282 2559 2361.058 1.979424e+02 3.918120e+04  
## 287 2045 2194.066 -1.490658e+02 2.222062e+04  
## 293 1563 1971.410 -4.084101e+02 1.667988e+05  
## 295 3748 2583.713 1.164287e+03 1.355564e+06  
## 296 1951 1971.410 -2.041011e+01 4.165724e+02  
## 297 2328 2138.402 1.895981e+02 3.594744e+04  
## 300 18880 18108.169 7.718306e+02 5.957225e+05  
## 303 7406 9900.630 -2.494630e+03 6.223178e+06  
## 306 5056 6186.508 -1.130508e+03 1.278048e+06  
## 310 2008 1971.410 3.658989e+01 1.338820e+03  
## 312 13757 10234.613 3.522387e+03 1.240721e+07  
## 313 2438 2305.394 1.326063e+02 1.758444e+04  
## 315 2774 2249.730 5.242703e+02 2.748593e+05  
## 316 2426 2249.730 1.762703e+02 3.107120e+04  
## 318 5968 6130.844 -1.628438e+02 2.651809e+04  
## 329 2661 2082.738 5.782620e+02 3.343870e+05  
## 331 5674 6242.172 -5.681716e+02 3.228190e+05  
## 332 2279 2305.394 -2.639367e+01 6.966256e+02  
## 336 3537 2361.058 1.175942e+03 1.382841e+06  
## 343 13237 14282.719 -1.045719e+03 1.093529e+06  
## 345 4617 5852.524 -1.235524e+03 1.526520e+06  
## 352 2318 1971.410 3.465899e+02 1.201246e+05  
## 353 19049 18052.505 9.964945e+02 9.930014e+05  
## 355 2133 6743.147 -4.610147e+03 2.125345e+07  
## 357 7412 9844.966 -2.432966e+03 5.919324e+06  
## 358 7756 6019.516 1.736484e+03 3.015377e+06  
## 359 2341 1971.410 3.695899e+02 1.365967e+05  
## 360 3419 2249.730 1.169270e+03 1.367193e+06  
## 363 2083 1971.410 1.115899e+02 1.245230e+04  
## 364 4033 5908.188 -1.875188e+03 3.516330e+06  
## 366 2290 2249.730 4.027026e+01 1.621694e+03  
## 374 19999 18664.809 1.334191e+03 1.780067e+06  
## 375 19436 17996.842 1.439158e+03 2.071177e+06  
## 383 14411 14839.359 -4.283586e+02 1.834911e+05  
## 386 13582 10178.950 3.403050e+03 1.158075e+07  
## 387 4262 6075.180 -1.813180e+03 3.287621e+06  
## 393 10209 10234.613 -2.561349e+01 6.560507e+02  
## 395 5265 6242.172 -9.771716e+02 9.548644e+05  
## 397 8008 9844.966 -1.836966e+03 3.374444e+06  
## 401 17159 14282.719 2.876281e+03 8.272991e+06  
## 404 2564 2583.713 -1.971330e+01 3.886142e+02  
## 405 3688 2138.402 1.549598e+03 2.401254e+06  
## 406 6474 6409.163 6.483661e+01 4.203786e+03  
## 407 11935 9900.630 2.034370e+03 4.138662e+06  
## 411 1274 1971.410 -6.974101e+02 4.863809e+05  
## 413 5957 6353.499 -3.964995e+02 1.572118e+05  
## 415 2109 1971.410 1.375899e+02 1.893098e+04  
## 419 2974 2416.722 5.572785e+02 3.105593e+05  
## 421 3420 2862.033 5.579671e+02 3.113272e+05  
## 422 4735 6687.483 -1.952483e+03 3.812190e+06  
## 424 2406 6186.508 -3.780508e+03 1.429224e+07  
## 427 15402 14227.055 1.174945e+03 1.380495e+06  
## 428 2066 2194.066 -1.280658e+02 1.640085e+04  
## 433 3944 2249.730 1.694270e+03 2.870552e+06  
## 434 3038 2194.066 8.439342e+02 7.122249e+05  
## 441 13744 10234.613 3.509387e+03 1.231579e+07  
## 442 5373 5963.852 -5.908520e+02 3.491061e+05  
## 445 4257 2973.361 1.283639e+03 1.647730e+06  
## 446 2911 2194.066 7.169342e+02 5.139946e+05  
## 447 6142 6186.508 -4.450769e+01 1.980934e+03  
## 448 18665 17996.842 6.681585e+02 4.464357e+05  
## 449 16880 14449.711 2.430289e+03 5.906304e+06  
## 450 8606 9956.294 -1.350294e+03 1.823293e+06  
## 451 2654 2361.058 2.929424e+02 8.581525e+04  
## 459 13458 10178.950 3.279050e+03 1.075217e+07  
## 473 3491 6186.508 -2.695508e+03 7.265762e+06  
## 476 6553 6409.163 1.438366e+02 2.068897e+04  
## 480 5094 6186.508 -1.092508e+03 1.193573e+06  
## 481 12061 10401.605 1.659395e+03 2.753591e+06  
## 492 2436 2249.730 1.862703e+02 3.469661e+04  
## 495 5993 6019.516 -2.651591e+01 7.030933e+02  
## 496 3162 2305.394 8.566063e+02 7.337744e+05  
## 497 5042 6186.508 -1.144508e+03 1.309898e+06  
## 502 6465 6130.844 3.341562e+02 1.116604e+05  
## 505 3629 2361.058 1.267942e+03 1.607678e+06  
## 514 11416 9844.966 1.571034e+03 2.468148e+06  
## 517 3578 2361.058 1.216942e+03 1.480949e+06  
## 518 5363 6186.508 -8.235077e+02 6.781649e+05  
## 519 6811 6186.508 6.244923e+02 3.899906e+05  
## 522 9991 9844.966 1.460340e+02 2.132593e+04  
## 528 10976 10624.261 3.517390e+02 1.237203e+05  
## 529 8621 9844.966 -1.223966e+03 1.498093e+06  
## 535 2725 2249.730 4.752703e+02 2.258818e+05  
## 545 6032 6075.180 -4.317983e+01 1.864498e+03  
## 546 6499 5963.852 5.351480e+02 2.863834e+05  
## 549 3917 2082.738 1.834262e+03 3.364517e+06  
## 550 6294 6186.508 1.074923e+02 1.155460e+04  
## 551 17779 15062.014 2.716986e+03 7.382012e+06  
## 552 13664 10234.613 3.429387e+03 1.176069e+07  
## 553 3517 2194.066 1.322934e+03 1.750155e+06  
## 563 6781 6409.163 3.718366e+02 1.382625e+05  
## 566 19081 18163.833 9.171667e+02 8.411947e+05  
## 567 6854 6409.163 4.448366e+02 1.978796e+05  
## 579 4898 5908.188 -1.010188e+03 1.020480e+06  
## 582 4765 5852.524 -1.087524e+03 1.182709e+06  
## 587 2296 2027.074 2.689260e+02 7.232118e+04  
## 588 5343 6409.163 -1.066163e+03 1.136704e+06  
## 589 3907 2249.730 1.657270e+03 2.746545e+06  
## 590 2500 2138.402 3.615981e+02 1.307532e+05  
## 591 11245 14839.359 -3.594359e+03 1.291941e+07  
## 595 6651 6743.147 -9.214695e+01 8.491061e+03  
## 597 16752 14505.375 2.246625e+03 5.047324e+06  
## 598 5769 6186.508 -4.175077e+02 1.743127e+05  
## 608 19038 18664.809 3.731914e+02 1.392718e+05  
## 609 2096 2027.074 6.892597e+01 4.750789e+03  
## 615 13496 14227.055 -7.310554e+02 5.344419e+05  
## 616 6842 6353.499 4.885005e+02 2.386328e+05  
## 617 17068 14227.055 2.840945e+03 8.070966e+06  
## 620 2323 2027.074 2.959260e+02 8.757218e+04  
## 623 3072 6297.836 -3.225836e+03 1.040601e+07  
## 624 3229 2305.394 9.236063e+02 8.530487e+05  
## 629 3312 2249.730 1.062270e+03 1.128418e+06  
## 632 2468 2416.722 5.127848e+01 2.629483e+03  
## 633 3633 6130.844 -2.497844e+03 6.239223e+06  
## 635 2552 1971.410 5.805899e+02 3.370846e+05  
## 639 19665 18386.489 1.278511e+03 1.634590e+06  
## 642 4558 6186.508 -1.628508e+03 2.652037e+06  
## 643 1483 1971.410 -4.884101e+02 2.385444e+05  
## 648 5661 6130.844 -4.698438e+02 2.207532e+05  
## 651 2238 2305.394 -6.739367e+01 4.541906e+03  
## 653 9380 9900.630 -5.206299e+02 2.710555e+05  
## 655 4477 2305.394 2.171606e+03 4.715874e+06  
## 661 2476 1971.410 5.045899e+02 2.546110e+05  
## 674 8463 5963.852 2.499148e+03 6.245741e+06  
## 677 8865 10624.261 -1.759261e+03 3.094999e+06  
## 678 4960 6743.147 -1.783147e+03 3.179613e+06  
## 682 16606 14338.383 2.267617e+03 5.142086e+06  
## 689 10761 10345.941 4.150587e+02 1.722737e+05  
## 693 13225 14449.711 -1.224711e+03 1.499917e+06  
## 697 6502 6019.516 4.824841e+02 2.327909e+05  
## 707 8380 9900.630 -1.520630e+03 2.312315e+06  
## 710 8686 6297.836 2.388164e+03 5.703329e+06  
## 713 4221 5908.188 -1.687188e+03 2.846604e+06  
## 720 6274 5963.852 3.101480e+02 9.619179e+04  
## 722 4999 5852.524 -8.535241e+02 7.285034e+05  
## 733 5154 6186.508 -1.032508e+03 1.066072e+06  
## 735 5647 6242.172 -5.951716e+02 3.542292e+05  
## 736 7847 6186.508 1.660492e+03 2.757235e+06  
## 738 9738 9900.630 -1.626299e+02 2.644849e+04  
## 743 5405 6186.508 -7.815077e+02 6.107543e+05  
## 744 4163 6130.844 -1.967844e+03 3.872409e+06  
## 750 6929 6186.508 7.424923e+02 5.512948e+05  
## 755 8412 9900.630 -1.488630e+03 2.216019e+06  
## 756 2728 2027.074 7.009260e+02 4.912972e+05  
## 758 5486 6464.827 -9.788273e+02 9.581029e+05  
## 759 2322 2082.738 2.392620e+02 5.724632e+04  
## 762 6500 5963.852 5.361480e+02 2.874547e+05  
## 766 15427 14783.695 6.433054e+02 4.138418e+05  
## 778 5441 6242.172 -8.011716e+02 6.418760e+05  
## 785 5673 6186.508 -5.135077e+02 2.636901e+05  
## 788 10596 10123.286 4.727144e+02 2.234589e+05  
## 790 5343 6186.508 -8.435077e+02 7.115052e+05  
## 791 2610 2082.738 5.272620e+02 2.780053e+05  
## 792 7525 11013.908 -3.488908e+03 1.217248e+07  
## 793 5006 6130.844 -1.124844e+03 1.265273e+06  
## 805 4779 6075.180 -1.296180e+03 1.680082e+06  
## 812 5454 6130.844 -6.768438e+02 4.581175e+05  
## 818 4312 6520.491 -2.208491e+03 4.877434e+06  
## 819 5220 6297.836 -1.077836e+03 1.161729e+06  
## 821 10368 10067.622 3.003783e+02 9.022712e+04  
## 826 18789 18219.497 5.695028e+02 3.243334e+05  
## 827 2329 2639.377 -3.103772e+02 9.633402e+04  
## 832 10377 10234.613 1.423865e+02 2.027392e+04  
## 833 6230 6520.491 -2.904912e+02 8.438516e+04  
## 834 14118 14839.359 -7.213586e+02 5.203582e+05  
## 836 7587 9900.630 -2.313630e+03 5.352883e+06  
## 841 5171 6353.499 -1.182499e+03 1.398305e+06  
## 845 4907 5963.852 -1.056852e+03 1.116936e+06  
## 846 6151 6687.483 -5.364830e+02 2.878140e+05  
## 847 6347 6687.483 -3.404830e+02 1.159287e+05  
## 848 10932 10457.269 4.747308e+02 2.253693e+05  
## 859 2899 2082.738 8.162620e+02 6.662837e+05  
## 862 4538 5852.524 -1.314524e+03 1.727974e+06  
## 864 5337 6186.508 -8.495077e+02 7.216633e+05  
## 865 6029 5963.852 6.514802e+01 4.244265e+03  
## 867 10231 10624.261 -3.932610e+02 1.546542e+05

mean(MSPE$SquaredResidual)

## [1] 1836897

PredJRLRM = cbind(rownames(LRM),LRM)  
rownames(PredJRLRM) = NULL  
colnames(PredJRLRM) = c("ID", "Predicted Salary")  
PredJRLRM

## ID Predicted Salary  
## 1 14 2472.385  
## 2 16 2361.058  
## 3 21 2472.385  
## 4 23 2528.049  
## 5 24 2361.058  
## 6 28 2138.402  
## 7 29 2194.066  
## 8 30 2305.394  
## 9 34 2249.730  
## 10 37 2917.697  
## 11 38 2138.402  
## 12 44 11570.548  
## 13 47 2249.730  
## 14 49 2194.066  
## 15 52 2305.394  
## 16 53 1971.410  
## 17 57 2361.058  
## 18 61 2472.385  
## 19 65 2472.385  
## 20 66 9844.966  
## 21 76 1971.410  
## 22 82 5963.852  
## 23 83 10178.950  
## 24 85 6520.491  
## 25 87 1915.746  
## 26 99 2249.730  
## 27 100 2082.738  
## 28 104 1971.410  
## 29 108 6297.836  
## 30 113 1971.410  
## 31 124 2194.066  
## 32 130 2082.738  
## 33 131 10123.286  
## 34 135 5963.852  
## 35 138 6075.180  
## 36 144 18720.473  
## 37 150 2361.058  
## 38 153 2138.402  
## 39 156 2249.730  
## 40 161 1971.410  
## 41 162 2305.394  
## 42 163 2361.058  
## 43 168 18553.481  
## 44 170 17941.178  
## 45 172 2249.730  
## 46 174 2305.394  
## 47 175 6075.180  
## 48 180 2249.730  
## 49 183 6186.508  
## 50 187 2472.385  
## 51 190 2862.033  
## 52 193 2583.713  
## 53 197 2472.385  
## 54 201 2472.385  
## 55 203 9900.630  
## 56 208 2194.066  
## 57 210 2583.713  
## 58 211 2249.730  
## 59 216 6353.499  
## 60 222 6576.155  
## 61 226 2194.066  
## 62 232 14505.375  
## 63 236 2639.377  
## 64 237 6409.163  
## 65 240 6743.147  
## 66 242 6019.516  
## 67 243 6130.844  
## 68 246 14895.022  
## 69 247 2027.074  
## 70 248 10067.622  
## 71 250 2361.058  
## 72 251 2249.730  
## 73 253 6242.172  
## 74 255 10345.941  
## 75 257 6186.508  
## 76 258 6186.508  
## 77 259 14561.039  
## 78 260 13614.752  
## 79 262 2528.049  
## 80 265 2027.074  
## 81 268 9844.966  
## 82 269 2361.058  
## 83 272 18330.825  
## 84 276 1971.410  
## 85 281 6409.163  
## 86 282 2361.058  
## 87 287 2194.066  
## 88 293 1971.410  
## 89 295 2583.713  
## 90 296 1971.410  
## 91 297 2138.402  
## 92 300 18108.169  
## 93 303 9900.630  
## 94 306 6186.508  
## 95 310 1971.410  
## 96 312 10234.613  
## 97 313 2305.394  
## 98 315 2249.730  
## 99 316 2249.730  
## 100 318 6130.844  
## 101 329 2082.738  
## 102 331 6242.172  
## 103 332 2305.394  
## 104 336 2361.058  
## 105 343 14282.719  
## 106 345 5852.524  
## 107 352 1971.410  
## 108 353 18052.505  
## 109 355 6743.147  
## 110 357 9844.966  
## 111 358 6019.516  
## 112 359 1971.410  
## 113 360 2249.730  
## 114 363 1971.410  
## 115 364 5908.188  
## 116 366 2249.730  
## 117 374 18664.809  
## 118 375 17996.842  
## 119 383 14839.359  
## 120 386 10178.950  
## 121 387 6075.180  
## 122 393 10234.613  
## 123 395 6242.172  
## 124 397 9844.966  
## 125 401 14282.719  
## 126 404 2583.713  
## 127 405 2138.402  
## 128 406 6409.163  
## 129 407 9900.630  
## 130 411 1971.410  
## 131 413 6353.499  
## 132 415 1971.410  
## 133 419 2416.722  
## 134 421 2862.033  
## 135 422 6687.483  
## 136 424 6186.508  
## 137 427 14227.055  
## 138 428 2194.066  
## 139 433 2249.730  
## 140 434 2194.066  
## 141 441 10234.613  
## 142 442 5963.852  
## 143 445 2973.361  
## 144 446 2194.066  
## 145 447 6186.508  
## 146 448 17996.842  
## 147 449 14449.711  
## 148 450 9956.294  
## 149 451 2361.058  
## 150 459 10178.950  
## 151 473 6186.508  
## 152 476 6409.163  
## 153 480 6186.508  
## 154 481 10401.605  
## 155 492 2249.730  
## 156 495 6019.516  
## 157 496 2305.394  
## 158 497 6186.508  
## 159 502 6130.844  
## 160 505 2361.058  
## 161 514 9844.966  
## 162 517 2361.058  
## 163 518 6186.508  
## 164 519 6186.508  
## 165 522 9844.966  
## 166 528 10624.261  
## 167 529 9844.966  
## 168 535 2249.730  
## 169 545 6075.180  
## 170 546 5963.852  
## 171 549 2082.738  
## 172 550 6186.508  
## 173 551 15062.014  
## 174 552 10234.613  
## 175 553 2194.066  
## 176 563 6409.163  
## 177 566 18163.833  
## 178 567 6409.163  
## 179 579 5908.188  
## 180 582 5852.524  
## 181 587 2027.074  
## 182 588 6409.163  
## 183 589 2249.730  
## 184 590 2138.402  
## 185 591 14839.359  
## 186 595 6743.147  
## 187 597 14505.375  
## 188 598 6186.508  
## 189 608 18664.809  
## 190 609 2027.074  
## 191 615 14227.055  
## 192 616 6353.499  
## 193 617 14227.055  
## 194 620 2027.074  
## 195 623 6297.836  
## 196 624 2305.394  
## 197 629 2249.730  
## 198 632 2416.722  
## 199 633 6130.844  
## 200 635 1971.410  
## 201 639 18386.489  
## 202 642 6186.508  
## 203 643 1971.410  
## 204 648 6130.844  
## 205 651 2305.394  
## 206 653 9900.630  
## 207 655 2305.394  
## 208 661 1971.410  
## 209 674 5963.852  
## 210 677 10624.261  
## 211 678 6743.147  
## 212 682 14338.383  
## 213 689 10345.941  
## 214 693 14449.711  
## 215 697 6019.516  
## 216 707 9900.630  
## 217 710 6297.836  
## 218 713 5908.188  
## 219 720 5963.852  
## 220 722 5852.524  
## 221 733 6186.508  
## 222 735 6242.172  
## 223 736 6186.508  
## 224 738 9900.630  
## 225 743 6186.508  
## 226 744 6130.844  
## 227 750 6186.508  
## 228 755 9900.630  
## 229 756 2027.074  
## 230 758 6464.827  
## 231 759 2082.738  
## 232 762 5963.852  
## 233 766 14783.695  
## 234 778 6242.172  
## 235 785 6186.508  
## 236 788 10123.286  
## 237 790 6186.508  
## 238 791 2082.738  
## 239 792 11013.908  
## 240 793 6130.844  
## 241 805 6075.180  
## 242 812 6130.844  
## 243 818 6520.491  
## 244 819 6297.836  
## 245 821 10067.622  
## 246 826 18219.497  
## 247 827 2639.377  
## 248 832 10234.613  
## 249 833 6520.491  
## 250 834 14839.359  
## 251 836 9900.630  
## 252 841 6353.499  
## 253 845 5963.852  
## 254 846 6687.483  
## 255 847 6687.483  
## 256 848 10457.269  
## 257 859 2082.738  
## 258 862 5852.524  
## 259 864 6186.508  
## 260 865 5963.852  
## 261 867 10624.261

write.csv(PredJRLRM,"C:/Users/Team Reed/OneDrive/JEFF/SMU/Doing Data Science/Project 2/Case2PredictionsReedSalary.csv")

### Executive Summary:

### Key Insights

### Notable variables related to Attrition

### Top 3 = Job Level, Job Involvement, Job Satisfaction

### Any rating of 1 within these variables should be noted

### Higher Total Working Years is also associated with less attrition

### Sales Representatives, Human Resources, Lab Technicians have highest attrition

### Notable variables related to Monthly Income / Salary

### Recommendations

### Employees that have a rating of 1 for notable variables and/or have minimal work experience

### Proactively engage to “raise” rating / retain such employees

### Recruit employees with more work experience

### Especially for employees in the following roles: Sales Representatives, Human Resources, Lab Technicians

### Include in survey ratings the option to include “why” behind their ratings to glean insight between different roles and departments

### Note individual scores of 1 and seek to incorporate work interaction with others in their department with higher ratings on satisfaction, involvement and level.