Milestone 4

Group D

2024-10-21

R Markdown

Loading all the required packages

```
# Load the randomForest package
# This package is used to build random forest models.
library(randomForest)
```

```
## randomForest 4.7-1.2
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
# Load the caret package
# It provides functions for data preprocessing, model training, tuning, evaluation,
# and visualization of performance metrics.
library(caret)
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'ggplot2'
```

```
## The following object is masked from 'package:randomForest':
##
##     margin
```

```
## Loading required package: lattice
```

```
# Load the party package
# The party package provides tools for creating recursive partitioning models like decision trees.
library(party)
```

```
## Loading required package: grid
```

```
## Loading required package: mvtnorm
```

```
## Loading required package: modeltools
```

```
## Loading required package: stats4
```

```
## Loading required package: strucchange
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: sandwich
```

```r
# Load the partykit package
# partykit is an extension of the party package, focused on improving the visualization and
# handling of decision trees and model objects.
library(partykit)
```

```
## Loading required package: libcoin

##
## Attaching package: 'partykit'

## The following objects are masked from 'package:party':
##
##     cforest, ctree, ctree_control, edge_simple, mob, mob_control,
##     node_barplot, node_bivplot, node_boxplot, node_inner, node_surv,
##     node_terminal, varimp
```

```r
# Load the rpart package
# rpart (Recursive Partitioning and Regression Trees) is used to create decision trees
# for classification or regression problems.
library(rpart)

# Load the rpart.plot package
# This package is used to visualize decision trees created using the rpart package.
# It creates detailed and customizable plots of decision trees, making it easier to interpret
# the splits and decision rules.
library(rpart.plot)

#Load library for partial dependence plots
library(pdp)

data <- read.csv("Prepared_Data (1).csv")
```

Creating the "eligibility" column

```r
data$eligibility <- ifelse(data$Annual.Salary > 50000 &
                           data$yrs_residence > 3 & data$Gross_Year_To_Date > 45000, 1, 0)
data$eligibility <- as.factor(data$eligibility)
```

Splitting of the data sets into training and testing, 75% and 25%, respectively

```r
set.seed(123)
trainIndex <- createDataPartition(data$eligibility, p = 0.75, list = FALSE)
trainData <- data[trainIndex, ]
testData <- data[-trainIndex, ]
```

Selecting the relevant features

```r
features <- c("Gross_Pay_Last_Paycheck", "Gross_Year_To_Date",
              "Gross_FRS_Contribution", "household_size",
              "yrs_residence", "Annual.Salary", "marital_status", "Age")
```

Building the random forest model

```r
rf_model <- randomForest(as.factor(eligibility) ~ .,
                         data = trainData[, c(features, "eligibility")],
                         importance = TRUE, ntree = 10)
```

Predictions on the test set

```r
predictions <- predict(rf_model, newdata = testData[, features])
```

Produce the confusion matrix

```r
conf_matrix <- confusionMatrix(predictions, testData$eligibility)
print(conf_matrix)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction     0     1
##          0 34945     4
##          1     0 11314
##
##                Accuracy : 0.9999
##                  95% CI : (0.9998, 1)
##     No Information Rate : 0.7554
##     P-Value [Acc > NIR] : <2e-16
##
##                   Kappa : 0.9998
##
##  Mcnemar's Test P-Value : 0.1336
##
##             Sensitivity : 1.0000
##             Specificity : 0.9996
##          Pos Pred Value : 0.9999
##          Neg Pred Value : 1.0000
##              Prevalence : 0.7554
##          Detection Rate : 0.7554
##    Detection Prevalence : 0.7554
##       Balanced Accuracy : 0.9998
##
##        'Positive' Class : 0
##
```

Classification report (including precision, recall, and F1-score)

```r
precision <- posPredValue(predictions, testData$eligibility)
recall <- sensitivity(predictions, testData$eligibility)
f1_score <- (2 * precision * recall) / (precision + recall)
cat("Precision:", precision, "\nRecall:", recall, "\nF1-score:", f1_score, "\n")
```

```
## Precision: 0.9998855
```

```
## Recall: 1
## F1-score: 0.9999428
```

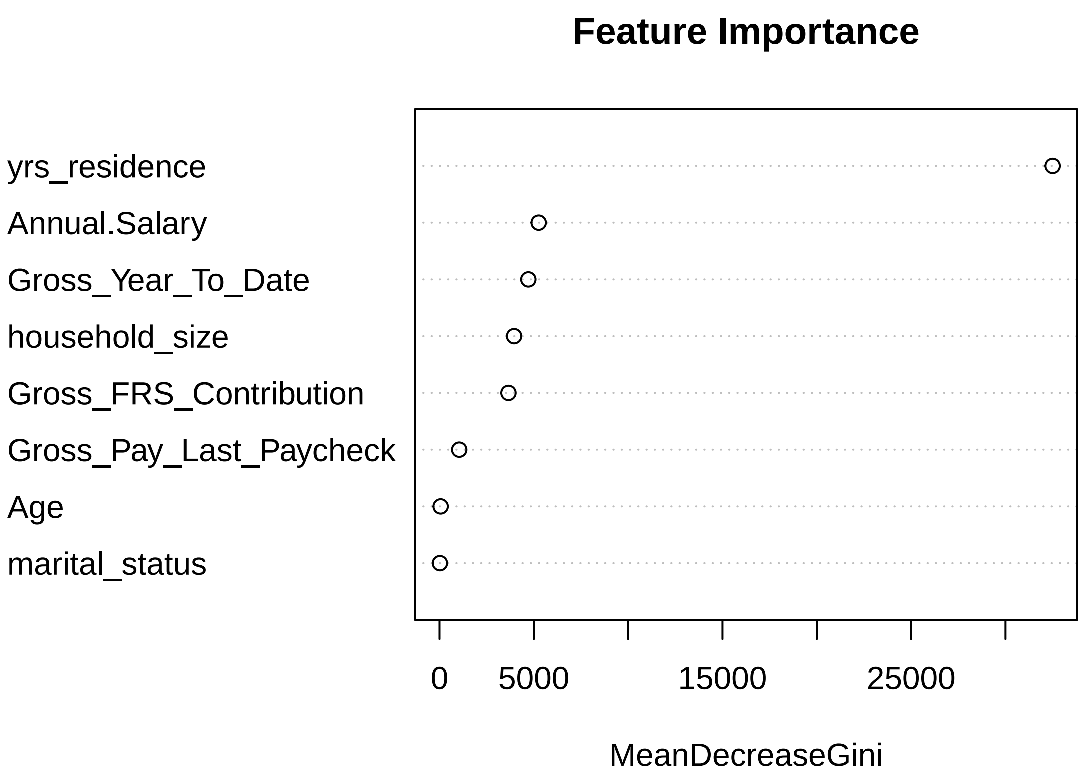
Calculate and rank feature importance

```r
importance_values <- importance(rf_model)
feature_importance <- data.frame(Feature = rownames(importance_values),
                                 Importance = importance_values[, "MeanDecreaseGini"])
feature_importance <- feature_importance[order(-feature_importance$Importance), ]
print(feature_importance)
```

```
##                                          Feature  Importance
## yrs_residence                      yrs_residence 32501.37101
## Annual.Salary                      Annual.Salary  5257.40599
## Gross_Year_To_Date             Gross_Year_To_Date  4709.18636
## household_size                    household_size  3953.27263
## Gross_FRS_Contribution    Gross_FRS_Contribution  3654.69945
## Gross_Pay_Last_Paycheck Gross_Pay_Last_Paycheck  1043.87228
## Age                                          Age    58.77853
## marital_status                    marital_status    19.12357
```

Plotting feature importance

```r
varImpPlot(rf_model,
           main = "Feature Importance",
           n.var = min(10, nrow(importance(rf_model))), # Show top 10 important features
           type = 2) # 'type = 2' is for Mean Decrease in Gini
```



Features with higher bars are more important for the model's decision-making process. They contribute more to reducing uncertainty or impurity when splitting the data.

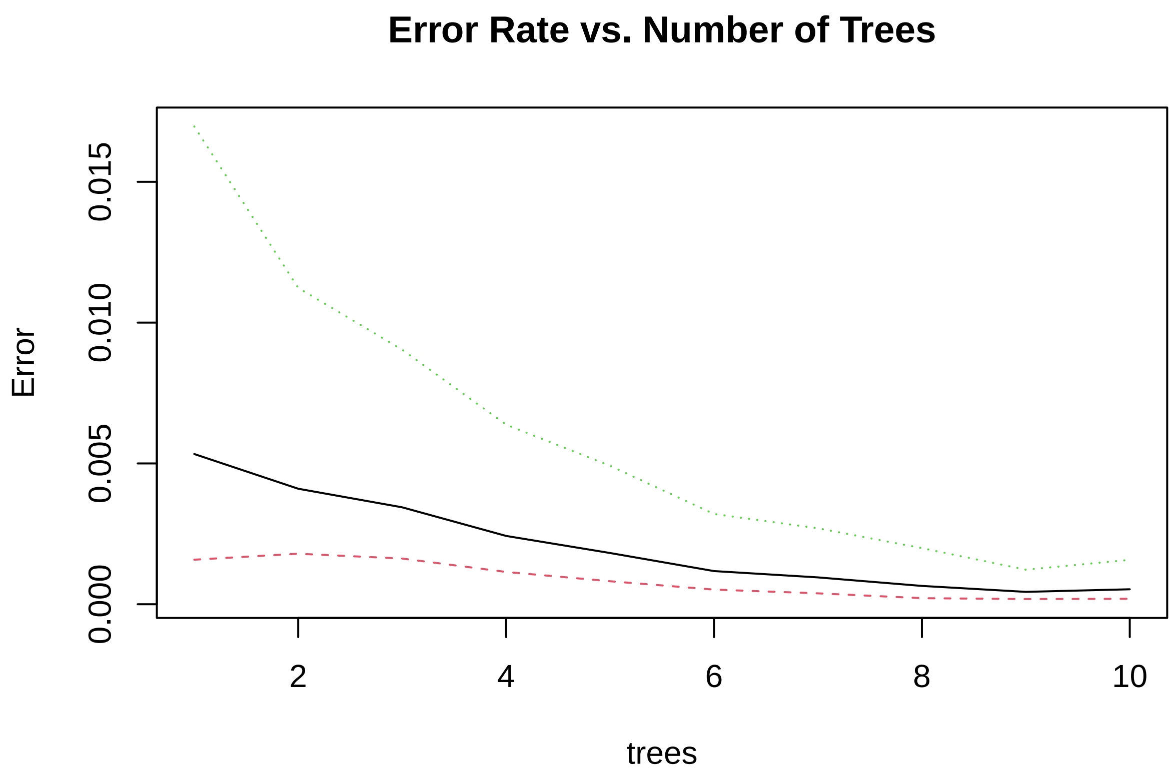
If a feature like Annual.Salary has a much higher importance score than others, it indicates that the Random Forest relies heavily on this feature to classify customers as eligible or not.

Relative importance is key: even if two features have importance values of 10,000 and 5,000 respectively, the one with 10,000 is twice as influential for the model.

This helps identify which features could be prioritized for further analysis or used in simpler models for a similar performance.

Plotting the error rate of the random forest model

```r
plot(rf_model, main = "Error Rate vs. Number of Trees")
```



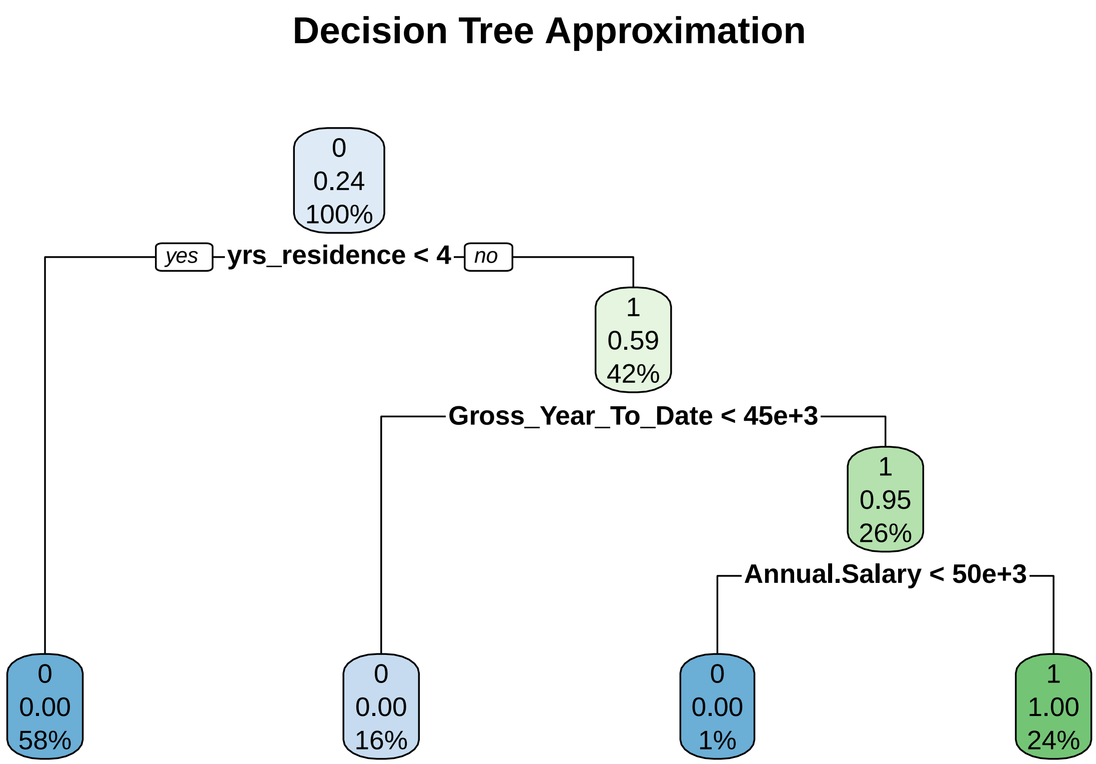
Decreasing Error Rate: If the error rate decreases as more trees are added, it means that the Random Forest is improving its performance. However, this trend usually plateaus after a certain point.

Point of Convergence: The number of trees after which the error rate plateaus is an indicator that adding more trees won't significantly improve performance. For instance, if the error rate levels off around 200 trees, using 500 trees may not provide substantial benefit.

Fit a decision tree using rpart on the same data

```r
tree_model <- rpart(as.factor(eligibility) ~ ., data = trainData[, c(features, "eligibility")])

# Plot the decision tree
rpart.plot(tree_model, main = "Decision Tree Approximation")
```




```r
# Extract and print the structure of the first tree from the random forest
print(getTree(rf_model, k = 1, labelVar = TRUE))
```

```
##    left daughter right daughter            split var split point status
## 1              2              3  Gross_FRS_Contribution   43843.880      1
## 2              4              5 Gross_Pay_Last_Paycheck    2126.000      1
## 3              6              7          marital_status       1.500      1
## 4              8              9          marital_status       2.500      1
## 5             10             11      Gross_Year_To_Date   45025.300      1
## 6             12             13           Annual.Salary   51599.470      1
## 7             14             15          household_size       2.500      1
## 8              0              0                    <NA>       0.000     -1
## 9             16             17 Gross_Pay_Last_Paycheck    1985.665      1
## 10             0              0                    <NA>       0.000     -1
## 11            18             19          household_size       2.500      1
## 12            20             21           Annual.Salary   50046.880      1
## 13            22             23           yrs_residence       3.500      1
## 14            24             25 Gross_Pay_Last_Paycheck    2035.470      1
## 15            26             27           Annual.Salary   49996.050      1
## 16            28             29           yrs_residence       3.500      1
## 17            30             31  Gross_FRS_Contribution   43716.615      1
## 18            32             33           Annual.Salary   61120.540      1
## 19            34             35                     Age      75.000      1
## 20             0              0                    <NA>       0.000     -1
## 21            36             37 Gross_Pay_Last_Paycheck    2634.980      1
## 22             0              0                    <NA>       0.000     -1
## 23             0              0                    <NA>       0.000     -1
## 24            38             39           yrs_residence       3.500      1
## 25            40             41 Gross_Pay_Last_Paycheck    2797.405      1
## 26             0              0                    <NA>       0.000     -1
```

```
## 27               42               43  Gross_FRS_Contribution   44209.575      1
## 28                0                0                    <NA>       0.000     -1
## 29               44               45           Annual.Salary   89438.570      1
## 30                0                0                    <NA>       0.000     -1
## 31               46               47           Annual.Salary   50958.960      1
## 32               48               49           yrs_residence       3.500      1
## 33               50               51           Annual.Salary   62880.870      1
## 34               52               53 Gross_Pay_Last_Paycheck    2793.295      1
## 35                0                0                    <NA>       0.000     -1
## 36               54               55           yrs_residence       3.500      1
## 37                0                0                    <NA>       0.000     -1
## 38                0                0                    <NA>       0.000     -1
## 39               56               57 Gross_Pay_Last_Paycheck    1575.940      1
## 40               58               59           Annual.Salary   50030.370      1
## 41               60               61           Annual.Salary   49996.570      1
## 42               62               63 Gross_Pay_Last_Paycheck    2035.470      1
## 43               64               65                     Age      50.500      1
## 44                0                0                    <NA>       0.000     -1
## 45               66               67           yrs_residence       4.500      1
## 46               68               69      Gross_Year_To_Date   45054.320      1
## 47                0                0                    <NA>       0.000     -1
## 48                0                0                    <NA>       0.000     -1
## 49               70               71  Gross_FRS_Contribution   43750.025      1
## 50                0                0                    <NA>       0.000     -1
## 51                0                0                    <NA>       0.000     -1
## 52               72               73  Gross_FRS_Contribution   43818.050      1
## 53                0                0                    <NA>       0.000     -1
## 54                0                0                    <NA>       0.000     -1
## 55                0                0                    <NA>       0.000     -1
## 56               74               75           Annual.Salary   51062.440      1
## 57                0                0                    <NA>       0.000     -1
## 58                0                0                    <NA>       0.000     -1
## 59               76               77           Annual.Salary   51867.140      1
## 60                0                0                    <NA>       0.000     -1
## 61               78               79 Gross_Pay_Last_Paycheck    9698.370      1
## 62                0                0                    <NA>       0.000     -1
## 63               80               81      Gross_Year_To_Date   44655.830      1
## 64               82               83           Annual.Salary   54545.660      1
## 65                0                0                    <NA>       0.000     -1
## 66               84               85 Gross_Pay_Last_Paycheck    1204.015      1
## 67                0                0                    <NA>       0.000     -1
## 68               86               87          marital_status       3.500      1
## 69                0                0                    <NA>       0.000     -1
## 70               88               89      Gross_Year_To_Date   45031.460      1
## 71               90               91 Gross_Pay_Last_Paycheck    2198.060      1
## 72                0                0                    <NA>       0.000     -1
## 73                0                0                    <NA>       0.000     -1
## 74                0                0                    <NA>       0.000     -1
## 75                0                0                    <NA>       0.000     -1
## 76               92               93          marital_status       3.500      1
## 77               94               95           yrs_residence       3.500      1
## 78               96               97  Gross_FRS_Contribution  120996.670      1
## 79               98               99                     Age      82.500      1
## 80                0                0                    <NA>       0.000     -1
```

```
## 81            0          0                   <NA>      0.000    -1
## 82          100        101      Gross_Year_To_Date  45261.295     1
## 83            0          0                   <NA>      0.000    -1
## 84            0          0                   <NA>      0.000    -1
## 85            0          0                   <NA>      0.000    -1
## 86            0          0                   <NA>      0.000    -1
## 87            0          0                   <NA>      0.000    -1
## 88          102        103                    Age     58.000     1
## 89            0          0                   <NA>      0.000    -1
## 90          104        105         marital_status      3.500     1
## 91          106        107                    Age     62.000     1
## 92          108        109  Gross_FRS_Contribution  84059.910     1
## 93          110        111          Annual.Salary  51376.260     1
## 94            0          0                   <NA>      0.000    -1
## 95          112        113      Gross_Year_To_Date  44984.935     1
## 96          114        115  Gross_FRS_Contribution 114264.715     1
## 97          116        117 Gross_Pay_Last_Paycheck   5030.860     1
## 98          118        119          yrs_residence      3.500     1
## 99          120        121         marital_status      2.500     1
## 100           0          0                   <NA>      0.000    -1
## 101           0          0                   <NA>      0.000    -1
## 102           0          0                   <NA>      0.000    -1
## 103         122        123      Gross_Year_To_Date  45027.230     1
## 104         124        125  Gross_FRS_Contribution  43753.865     1
## 105           0          0                   <NA>      0.000    -1
## 106           0          0                   <NA>      0.000    -1
## 107         126        127 Gross_Pay_Last_Paycheck   2242.020     1
## 108         128        129  Gross_FRS_Contribution  44061.785     1
## 109         130        131                    Age     80.000     1
## 110         132        133      Gross_Year_To_Date  45812.135     1
## 111         134        135          yrs_residence      3.500     1
## 112           0          0                   <NA>      0.000    -1
## 113           0          0                   <NA>      0.000    -1
## 114         136        137  Gross_FRS_Contribution  50760.250     1
## 115         138        139         marital_status      3.500     1
## 116         140        141         marital_status      3.500     1
## 117         142        143                    Age     85.000     1
## 118           0          0                   <NA>      0.000    -1
## 119           0          0                   <NA>      0.000    -1
## 120         144        145          yrs_residence      3.500     1
## 121         146        147          yrs_residence      3.500     1
## 122           0          0                   <NA>      0.000    -1
## 123           0          0                   <NA>      0.000    -1
## 124           0          0                   <NA>      0.000    -1
## 125           0          0                   <NA>      0.000    -1
## 126           0          0                   <NA>      0.000    -1
## 127           0          0                   <NA>      0.000    -1
## 128         148        149         marital_status      2.500     1
## 129         150        151          Annual.Salary  51487.670     1
## 130           0          0                   <NA>      0.000    -1
## 131           0          0                   <NA>      0.000    -1
## 132         152        153                    Age     94.000     1
## 133           0          0                   <NA>      0.000    -1
## 134           0          0                   <NA>      0.000    -1
```

```
## 135          154          155  Gross_FRS_Contribution   44313.225      1
## 136          156          157  Gross_FRS_Contribution   50604.480      1
## 137          158          159                     Age     101.500      1
## 138          160          161           yrs_residence       3.500      1
## 139          162          163           yrs_residence       3.500      1
## 140          164          165           yrs_residence       3.500      1
## 141            0            0                    <NA>       0.000     -1
## 142          166          167           yrs_residence       3.500      1
## 143            0            0                    <NA>       0.000     -1
## 144            0            0                    <NA>       0.000     -1
## 145            0            0                    <NA>       0.000     -1
## 146            0            0                    <NA>       0.000     -1
## 147            0            0                    <NA>       0.000     -1
## 148          168          169 Gross_Pay_Last_Paycheck    2334.380      1
## 149          170          171                     Age      55.000      1
## 150          172          173                     Age      96.500      1
## 151          174          175           yrs_residence       3.500      1
## 152            0            0                    <NA>       0.000     -1
## 153            0            0                    <NA>       0.000     -1
## 154          176          177                     Age      72.500      1
## 155            0            0                    <NA>       0.000     -1
## 156          178          179           yrs_residence       3.500      1
## 157          180          181 Gross_Pay_Last_Paycheck    3737.525      1
## 158          182          183          marital_status       2.500      1
## 159          184          185           yrs_residence       3.500      1
## 160            0            0                    <NA>       0.000     -1
## 161            0            0                    <NA>       0.000     -1
## 162            0            0                    <NA>       0.000     -1
## 163            0            0                    <NA>       0.000     -1
## 164            0            0                    <NA>       0.000     -1
## 165            0            0                    <NA>       0.000     -1
## 166            0            0                    <NA>       0.000     -1
## 167            0            0                    <NA>       0.000     -1
## 168            0            0                    <NA>       0.000     -1
## 169          186          187 Gross_Pay_Last_Paycheck    2415.545      1
## 170            0            0                    <NA>       0.000     -1
## 171            0            0                    <NA>       0.000     -1
## 172          188          189                     Age      38.000      1
## 173            0            0                    <NA>       0.000     -1
## 174            0            0                    <NA>       0.000     -1
## 175            0            0                    <NA>       0.000     -1
## 176            0            0                    <NA>       0.000     -1
## 177            0            0                    <NA>       0.000     -1
## 178            0            0                    <NA>       0.000     -1
## 179            0            0                    <NA>       0.000     -1
## 180          190          191  Gross_FRS_Contribution   50703.810      1
## 181            0            0                    <NA>       0.000     -1
## 182          192          193                     Age      83.500      1
## 183          194          195           Annual.Salary   50460.800      1
## 184            0            0                    <NA>       0.000     -1
## 185            0            0                    <NA>       0.000     -1
## 186            0            0                    <NA>       0.000     -1
## 187            0            0                    <NA>       0.000     -1
## 188            0            0                    <NA>       0.000     -1
```

```
## 189          196          197 Gross_Pay_Last_Paycheck    2534.520       1
## 190          198          199                     Age      74.500       1
## 191            0            0                    <NA>       0.000      -1
## 192          200          201           Annual.Salary  132045.030       1
## 193          202          203  Gross_FRS_Contribution  112887.560       1
## 194          204          205           yrs_residence       3.500       1
## 195          206          207           Annual.Salary  145809.690       1
## 196          208          209       Gross_Year_To_Date   50168.620       1
## 197          210          211           yrs_residence       3.500       1
## 198          212          213                     Age      72.000       1
## 199            0            0                    <NA>       0.000      -1
## 200          214          215                     Age      81.500       1
## 201          216          217  Gross_FRS_Contribution  106766.025       1
## 202          218          219  Gross_FRS_Contribution  112108.800       1
## 203          220          221                     Age      91.500       1
## 204            0            0                    <NA>       0.000      -1
## 205            0            0                    <NA>       0.000      -1
## 206          222          223           yrs_residence       3.500       1
## 207          224          225  Gross_FRS_Contribution   96391.345       1
## 208          226          227       Gross_Year_To_Date   49791.595       1
## 209          228          229       Gross_Year_To_Date   51087.680       1
## 210            0            0                    <NA>       0.000      -1
## 211            0            0                    <NA>       0.000      -1
## 212          230          231           yrs_residence       3.500       1
## 213            0            0                    <NA>       0.000      -1
## 214          232          233 Gross_Pay_Last_Paycheck    9248.450       1
## 215          234          235 Gross_Pay_Last_Paycheck    2828.930       1
## 216          236          237           yrs_residence       3.500       1
## 217          238          239                     Age      82.000       1
## 218          240          241       Gross_Year_To_Date   69889.030       1
## 219            0            0                    <NA>       0.000      -1
## 220          242          243  Gross_FRS_Contribution  113594.045       1
## 221          244          245 Gross_Pay_Last_Paycheck    5011.430       1
## 222            0            0                    <NA>       0.000      -1
## 223            0            0                    <NA>       0.000      -1
## 224          246          247       Gross_Year_To_Date   80548.110       1
## 225          248          249           yrs_residence       3.500       1
## 226          250          251                     Age      41.500       1
## 227            0            0                    <NA>       0.000      -1
## 228          252          253  Gross_FRS_Contribution   49464.575       1
## 229          254          255           yrs_residence       3.500       1
## 230            0            0                    <NA>       0.000      -1
## 231            0            0                    <NA>       0.000      -1
## 232          256          257                     Age      62.500       1
## 233          258          259           yrs_residence       3.500       1
## 234            0            0                    <NA>       0.000      -1
## 235          260          261           Annual.Salary  130624.260       1
## 236            0            0                    <NA>       0.000      -1
## 237            0            0                    <NA>       0.000      -1
## 238          262          263           yrs_residence       3.500       1
## 239            0            0                    <NA>       0.000      -1
## 240          264          265           Annual.Salary   83875.480       1
## 241          266          267           Annual.Salary  112704.670       1
## 242            0            0                    <NA>       0.000      -1
```

```
## 243   0   0                  <NA>      0.000   -1
## 244   0   0                  <NA>      0.000   -1
## 245 268 269                   Age     96.500    1
## 246   0   0                  <NA>      0.000   -1
## 247 270 271         yrs_residence      3.500    1
## 248   0   0                  <NA>      0.000   -1
## 249   0   0                  <NA>      0.000   -1
## 250   0   0                  <NA>      0.000   -1
## 251 272 273 Gross_Pay_Last_Paycheck  2340.540    1
## 252 274 275         Annual.Salary  51412.790    1
## 253   0   0                  <NA>      0.000   -1
## 254   0   0                  <NA>      0.000   -1
## 255   0   0                  <NA>      0.000   -1
## 256 276 277  Gross_FRS_Contribution  61162.370    1
## 257 278 279         yrs_residence      3.500    1
## 258   0   0                  <NA>      0.000   -1
## 259   0   0                  <NA>      0.000   -1
## 260 280 281         Annual.Salary 113597.510    1
## 261   0   0                  <NA>      0.000   -1
## 262   0   0                  <NA>      0.000   -1
## 263   0   0                  <NA>      0.000   -1
## 264 282 283      Gross_Year_To_Date  65698.770    1
## 265 284 285  Gross_FRS_Contribution  66611.580    1
## 266 286 287                   Age     85.500    1
## 267 288 289         yrs_residence      3.500    1
## 268 290 291         yrs_residence      3.500    1
## 269   0   0                  <NA>      0.000   -1
## 270   0   0                  <NA>      0.000   -1
## 271   0   0                  <NA>      0.000   -1
## 272 292 293         yrs_residence      3.500    1
## 273 294 295         Annual.Salary  51435.800    1
## 274 296 297         yrs_residence      3.500    1
## 275   0   0                  <NA>      0.000   -1
## 276 298 299                   Age     41.500    1
## 277 300 301         yrs_residence      3.500    1
## 278   0   0                  <NA>      0.000   -1
## 279   0   0                  <NA>      0.000   -1
## 280 302 303  Gross_FRS_Contribution  57058.210    1
## 281 304 305         yrs_residence      3.500    1
## 282 306 307      Gross_Year_To_Date  55292.560    1
## 283   0   0                  <NA>      0.000   -1
## 284 308 309 Gross_Pay_Last_Paycheck  4661.865    1
## 285 310 311         yrs_residence      3.000    1
## 286 312 313 Gross_Pay_Last_Paycheck  4665.545    1
## 287 314 315  Gross_FRS_Contribution  94782.680    1
## 288   0   0                  <NA>      0.000   -1
## 289   0   0                  <NA>      0.000   -1
## 290   0   0                  <NA>      0.000   -1
## 291   0   0                  <NA>      0.000   -1
## 292   0   0                  <NA>      0.000   -1
## 293   0   0                  <NA>      0.000   -1
## 294 316 317         yrs_residence      3.000    1
## 295   0   0                  <NA>      0.000   -1
## 296   0   0                  <NA>      0.000   -1
```

```
## 297          0          0                  <NA>    0.000   -1
## 298          0          0                  <NA>    0.000   -1
## 299        318        319         Annual.Salary 63240.450    1
## 300          0          0                  <NA>    0.000   -1
## 301          0          0                  <NA>    0.000   -1
## 302          0          0                  <NA>    0.000   -1
## 303        320        321         yrs_residence    3.500    1
## 304          0          0                  <NA>    0.000   -1
## 305          0          0                  <NA>    0.000   -1
## 306        322        323         yrs_residence    3.500    1
## 307        324        325                   Age   90.500    1
## 308        326        327         yrs_residence    3.000    1
## 309          0          0                  <NA>    0.000   -1
## 310          0          0                  <NA>    0.000   -1
## 311          0          0                  <NA>    0.000   -1
## 312        328        329         Annual.Salary 87569.560    1
## 313        330        331  Gross_FRS_Contribution 94707.685    1
## 314        332        333 Gross_Pay_Last_Paycheck  5026.180    1
## 315          0          0                  <NA>    0.000   -1
## 316          0          0                  <NA>    0.000   -1
## 317          0          0                  <NA>    0.000   -1
## 318        334        335                   Age   61.500    1
## 319        336        337  Gross_FRS_Contribution 60971.450    1
## 320          0          0                  <NA>    0.000   -1
## 321          0          0                  <NA>    0.000   -1
## 322          0          0                  <NA>    0.000   -1
## 323          0          0                  <NA>    0.000   -1
## 324        338        339  Gross_FRS_Contribution 63627.365    1
## 325        340        341  Gross_FRS_Contribution 62201.665    1
## 326          0          0                  <NA>    0.000   -1
## 327          0          0                  <NA>    0.000   -1
## 328        342        343  Gross_FRS_Contribution 68842.030    1
## 329        344        345         yrs_residence    3.500    1
## 330        346        347         Annual.Salary 71954.350    1
## 331        348        349         yrs_residence    3.000    1
## 332        350        351  Gross_FRS_Contribution 93889.445    1
## 333        352        353                   Age   95.500    1
## 334        354        355                   Age   57.500    1
## 335          0          0                  <NA>    0.000   -1
## 336        356        357      Gross_Year_To_Date 53073.210    1
## 337          0          0                  <NA>    0.000   -1
## 338        358        359  Gross_FRS_Contribution 60426.265    1
## 339          0          0                  <NA>    0.000   -1
## 340        360        361  Gross_FRS_Contribution 59376.980    1
## 341        362        363      Gross_Year_To_Date 65293.230    1
## 342          0          0                  <NA>    0.000   -1
## 343        364        365         Annual.Salary 85776.470    1
## 344          0          0                  <NA>    0.000   -1
## 345          0          0                  <NA>    0.000   -1
## 346        366        367         Annual.Salary 58577.220    1
## 347        368        369      Gross_Year_To_Date 85186.085    1
## 348          0          0                  <NA>    0.000   -1
## 349          0          0                  <NA>    0.000   -1
## 350        370        371      Gross_Year_To_Date 93892.470    1
```

```
## 351            372            373           yrs_residence     3.500      1
## 352              0              0                    <NA>     0.000     -1
## 353            374            375 Gross_Pay_Last_Paycheck  6180.520      1
## 354            376            377      Gross_Year_To_Date 58207.840      1
## 355            378            379           yrs_residence     3.500      1
## 356              0              0                    <NA>     0.000     -1
## 357            380            381           yrs_residence     3.500      1
## 358            382            383                     Age    88.500      1
## 359            384            385           yrs_residence     3.500      1
## 360            386            387      Gross_Year_To_Date 60840.085      1
## 361            388            389 Gross_Pay_Last_Paycheck  3627.365      1
## 362            390            391  Gross_FRS_Contribution 62437.615      1
## 363              0              0                    <NA>     0.000     -1
## 364            392            393           Annual.Salary 80629.510      1
## 365              0              0                    <NA>     0.000     -1
## 366              0              0                    <NA>     0.000     -1
## 367              0              0                    <NA>     0.000     -1
## 368              0              0                    <NA>     0.000     -1
## 369            394            395      Gross_Year_To_Date 88154.080      1
## 370            396            397  Gross_FRS_Contribution 71811.690      1
## 371            398            399           yrs_residence     3.500      1
## 372              0              0                    <NA>     0.000     -1
## 373              0              0                    <NA>     0.000     -1
## 374              0              0                    <NA>     0.000     -1
## 375            400            401           yrs_residence     3.500      1
## 376            402            403  Gross_FRS_Contribution 55778.140      1
## 377            404            405      Gross_Year_To_Date 60015.545      1
## 378              0              0                    <NA>     0.000     -1
## 379              0              0                    <NA>     0.000     -1
## 380              0              0                    <NA>     0.000     -1
## 381              0              0                    <NA>     0.000     -1
## 382            406            407 Gross_Pay_Last_Paycheck  3169.610      1
## 383            408            409      Gross_Year_To_Date 62149.890      1
## 384              0              0                    <NA>     0.000     -1
## 385              0              0                    <NA>     0.000     -1
## 386            410            411           Annual.Salary 61280.050      1
## 387              0              0                    <NA>     0.000     -1
## 388              0              0                    <NA>     0.000     -1
## 389            412            413  Gross_FRS_Contribution 61084.445      1
## 390              0              0                    <NA>     0.000     -1
## 391            414            415  Gross_FRS_Contribution 62737.580      1
## 392            416            417      Gross_Year_To_Date 73481.455      1
## 393            418            419 Gross_Pay_Last_Paycheck  3258.595      1
## 394              0              0                    <NA>     0.000     -1
## 395              0              0                    <NA>     0.000     -1
## 396            420            421      Gross_Year_To_Date 72054.115      1
## 397            422            423      Gross_Year_To_Date 74651.915      1
## 398              0              0                    <NA>     0.000     -1
## 399              0              0                    <NA>     0.000     -1
## 400              0              0                    <NA>     0.000     -1
## 401              0              0                    <NA>     0.000     -1
## 402            424            425      Gross_Year_To_Date 56618.225      1
## 403            426            427           Annual.Salary 59375.030      1
## 404            428            429           Annual.Salary 50648.000      1
```

```
## 405          430          431 Gross_Pay_Last_Paycheck    2995.020      1
## 406          432          433      Gross_Year_To_Date   56870.900      1
## 407          434          435           Annual.Salary   55089.710      1
## 408          436          437 Gross_Pay_Last_Paycheck    2837.520      1
## 409            0            0                    <NA>       0.000     -1
## 410            0            0                    <NA>       0.000     -1
## 411            0            0                    <NA>       0.000     -1
## 412            0            0                    <NA>       0.000     -1
## 413          438          439 Gross_Pay_Last_Paycheck    4565.915      1
## 414            0            0                    <NA>       0.000     -1
## 415          440          441           yrs_residence       3.500      1
## 416          442          443           yrs_residence       3.500      1
## 417          444          445           yrs_residence       3.500      1
## 418            0            0                    <NA>       0.000     -1
## 419          446          447      Gross_Year_To_Date   71181.450      1
## 420          448          449           Annual.Salary   72511.400      1
## 421          450          451           yrs_residence       3.500      1
## 422          452          453 Gross_Pay_Last_Paycheck    3322.385      1
## 423          454          455  Gross_FRS_Contribution   72999.805      1
## 424          456          457           Annual.Salary   52486.720      1
## 425            0            0                    <NA>       0.000     -1
## 426            0            0                    <NA>       0.000     -1
## 427          458          459           yrs_residence       3.500      1
## 428            0            0                    <NA>       0.000     -1
## 429          460          461           yrs_residence       3.500      1
## 430          462          463                     Age      45.500      1
## 431          464          465  Gross_FRS_Contribution   58632.685      1
## 432            0            0                    <NA>       0.000     -1
## 433          466          467           yrs_residence       3.500      1
## 434          468          469      Gross_Year_To_Date   59388.930      1
## 435          470          471           yrs_residence       3.000      1
## 436            0            0                    <NA>       0.000     -1
## 437          472          473  Gross_FRS_Contribution   58362.450      1
## 438            0            0                    <NA>       0.000     -1
## 439            0            0                    <NA>       0.000     -1
## 440            0            0                    <NA>       0.000     -1
## 441            0            0                    <NA>       0.000     -1
## 442            0            0                    <NA>       0.000     -1
## 443            0            0                    <NA>       0.000     -1
## 444            0            0                    <NA>       0.000     -1
## 445            0            0                    <NA>       0.000     -1
## 446            0            0                    <NA>       0.000     -1
## 447          474          475           Annual.Salary   82130.100      1
## 448          476          477           yrs_residence       3.500      1
## 449          478          479           yrs_residence       3.500      1
## 450            0            0                    <NA>       0.000     -1
## 451            0            0                    <NA>       0.000     -1
## 452            0            0                    <NA>       0.000     -1
## 453          480          481           yrs_residence       3.500      1
## 454          482          483                     Age      90.000      1
## 455          484          485           yrs_residence       3.500      1
## 456            0            0                    <NA>       0.000     -1
## 457          486          487           yrs_residence       3.500      1
## 458            0            0                    <NA>       0.000     -1
```

```
## 459             0              0                    <NA>       0.000     -1
## 460             0              0                    <NA>       0.000     -1
## 461             0              0                    <NA>       0.000     -1
## 462             0              0                    <NA>       0.000     -1
## 463           488            489                     Age      51.500      1
## 464             0              0                    <NA>       0.000     -1
## 465             0              0                    <NA>       0.000     -1
## 466             0              0                    <NA>       0.000     -1
## 467             0              0                    <NA>       0.000     -1
## 468             0              0                    <NA>       0.000     -1
## 469             0              0                    <NA>       0.000     -1
## 470             0              0                    <NA>       0.000     -1
## 471             0              0                    <NA>       0.000     -1
## 472           490            491           Annual.Salary   65683.020      1
## 473           492            493  Gross_FRS_Contribution   59791.460      1
## 474             0              0                    <NA>       0.000     -1
## 475           494            495      Gross_Year_To_Date   76735.995      1
## 476             0              0                    <NA>       0.000     -1
## 477             0              0                    <NA>       0.000     -1
## 478             0              0                    <NA>       0.000     -1
## 479             0              0                    <NA>       0.000     -1
## 480             0              0                    <NA>       0.000     -1
## 481             0              0                    <NA>       0.000     -1
## 482             0              0                    <NA>       0.000     -1
## 483           496            497      Gross_Year_To_Date   75030.355      1
## 484             0              0                    <NA>       0.000     -1
## 485             0              0                    <NA>       0.000     -1
## 486             0              0                    <NA>       0.000     -1
## 487             0              0                    <NA>       0.000     -1
## 488           498            499           Annual.Salary   57847.920      1
## 489           500            501           yrs_residence       3.500      1
## 490           502            503           Annual.Salary   58908.460      1
## 491           504            505 Gross_Pay_Last_Paycheck    2879.065      1
## 492             0              0                    <NA>       0.000     -1
## 493             0              0                    <NA>       0.000     -1
## 494           506            507                     Age      84.500      1
## 495           508            509                     Age      84.500      1
## 496             0              0                    <NA>       0.000     -1
## 497             0              0                    <NA>       0.000     -1
## 498             0              0                    <NA>       0.000     -1
## 499             0              0                    <NA>       0.000     -1
## 500             0              0                    <NA>       0.000     -1
## 501             0              0                    <NA>       0.000     -1
## 502             0              0                    <NA>       0.000     -1
## 503             0              0                    <NA>       0.000     -1
## 504             0              0                    <NA>       0.000     -1
## 505             0              0                    <NA>       0.000     -1
## 506           510            511 Gross_Pay_Last_Paycheck    3316.540      1
## 507           512            513  Gross_FRS_Contribution   71103.510      1
## 508           514            515           Annual.Salary   84507.020      1
## 509           516            517 Gross_Pay_Last_Paycheck    4459.710      1
## 510             0              0                    <NA>       0.000     -1
## 511           518            519 Gross_Pay_Last_Paycheck    3325.470      1
## 512           520            521           Annual.Salary   84312.410      1
```

```
## 513             0             0                   <NA>     0.000     -1
## 514           522           523           yrs_residence     3.000      1
## 515             0             0                   <NA>     0.000     -1
## 516             0             0                   <NA>     0.000     -1
## 517             0             0                   <NA>     0.000     -1
## 518             0             0                   <NA>     0.000     -1
## 519             0             0                   <NA>     0.000     -1
## 520             0             0                   <NA>     0.000     -1
## 521             0             0                   <NA>     0.000     -1
## 522             0             0                   <NA>     0.000     -1
## 523             0             0                   <NA>     0.000     -1
##     prediction
## 1         <NA>
## 2         <NA>
## 3         <NA>
## 4         <NA>
## 5         <NA>
## 6         <NA>
## 7         <NA>
## 8            0
## 9         <NA>
## 10           0
## 11        <NA>
## 12        <NA>
## 13        <NA>
## 14        <NA>
## 15        <NA>
## 16        <NA>
## 17        <NA>
## 18        <NA>
## 19        <NA>
## 20           0
## 21        <NA>
## 22           0
## 23           1
## 24        <NA>
## 25        <NA>
## 26           0
## 27        <NA>
## 28           0
## 29        <NA>
## 30           0
## 31        <NA>
## 32        <NA>
## 33        <NA>
## 34        <NA>
## 35           1
## 36        <NA>
## 37           0
## 38           0
## 39        <NA>
## 40        <NA>
## 41        <NA>
## 42        <NA>
```

```
## 43        <NA>
## 44           0
## 45        <NA>
## 46        <NA>
## 47           1
## 48           0
## 49        <NA>
## 50           1
## 51           0
## 52        <NA>
## 53           0
## 54           0
## 55           1
## 56        <NA>
## 57           0
## 58           0
## 59        <NA>
## 60           0
## 61        <NA>
## 62           0
## 63        <NA>
## 64        <NA>
## 65           1
## 66        <NA>
## 67           1
## 68        <NA>
## 69           0
## 70        <NA>
## 71        <NA>
## 72           1
## 73           0
## 74           0
## 75           1
## 76        <NA>
## 77        <NA>
## 78        <NA>
## 79        <NA>
## 80           0
## 81           1
## 82        <NA>
## 83           1
## 84           0
## 85           1
## 86           0
## 87           1
## 88        <NA>
## 89           1
## 90        <NA>
## 91        <NA>
## 92        <NA>
## 93        <NA>
## 94           0
## 95        <NA>
## 96        <NA>
```

```
## 97         <NA>
## 98         <NA>
## 99         <NA>
## 100           0
## 101           1
## 102           1
## 103        <NA>
## 104        <NA>
## 105           1
## 106           0
## 107        <NA>
## 108        <NA>
## 109        <NA>
## 110        <NA>
## 111        <NA>
## 112           0
## 113           1
## 114        <NA>
## 115        <NA>
## 116        <NA>
## 117        <NA>
## 118           0
## 119           1
## 120        <NA>
## 121        <NA>
## 122           1
## 123           0
## 124           0
## 125           1
## 126           0
## 127           0
## 128        <NA>
## 129        <NA>
## 130           0
## 131           1
## 132        <NA>
## 133           0
## 134           0
## 135        <NA>
## 136        <NA>
## 137        <NA>
## 138        <NA>
## 139        <NA>
## 140        <NA>
## 141           1
## 142        <NA>
## 143           0
## 144           0
## 145           1
## 146           0
## 147           1
## 148        <NA>
## 149        <NA>
## 150        <NA>
```

```
## 151        <NA>
## 152           1
## 153           0
## 154        <NA>
## 155           1
## 156        <NA>
## 157        <NA>
## 158        <NA>
## 159        <NA>
## 160           0
## 161           1
## 162           0
## 163           1
## 164           0
## 165           1
## 166           0
## 167           1
## 168           1
## 169        <NA>
## 170           0
## 171           1
## 172        <NA>
## 173           1
## 174           0
## 175           1
## 176           1
## 177           0
## 178           0
## 179           1
## 180        <NA>
## 181           0
## 182        <NA>
## 183        <NA>
## 184           0
## 185           1
## 186           0
## 187           1
## 188           1
## 189        <NA>
## 190        <NA>
## 191           0
## 192        <NA>
## 193        <NA>
## 194        <NA>
## 195        <NA>
## 196        <NA>
## 197        <NA>
## 198        <NA>
## 199           0
## 200        <NA>
## 201        <NA>
## 202        <NA>
## 203        <NA>
## 204           0
```

```
## 205            1
## 206         <NA>
## 207         <NA>
## 208         <NA>
## 209         <NA>
## 210            0
## 211            1
## 212         <NA>
## 213            1
## 214         <NA>
## 215         <NA>
## 216         <NA>
## 217         <NA>
## 218         <NA>
## 219            0
## 220         <NA>
## 221         <NA>
## 222            0
## 223            1
## 224         <NA>
## 225         <NA>
## 226         <NA>
## 227            0
## 228         <NA>
## 229         <NA>
## 230            0
## 231            1
## 232         <NA>
## 233         <NA>
## 234            1
## 235         <NA>
## 236            0
## 237            1
## 238         <NA>
## 239            1
## 240         <NA>
## 241         <NA>
## 242            1
## 243            0
## 244            0
## 245         <NA>
## 246            0
## 247         <NA>
## 248            0
## 249            1
## 250            0
## 251         <NA>
## 252         <NA>
## 253            1
## 254            0
## 255            1
## 256         <NA>
## 257         <NA>
## 258            0
```

```
## 259           1
## 260        <NA>
## 261           1
## 262           0
## 263           1
## 264        <NA>
## 265        <NA>
## 266        <NA>
## 267        <NA>
## 268        <NA>
## 269           1
## 270           0
## 271           1
## 272        <NA>
## 273        <NA>
## 274        <NA>
## 275           1
## 276        <NA>
## 277        <NA>
## 278           0
## 279           1
## 280        <NA>
## 281        <NA>
## 282        <NA>
## 283           0
## 284        <NA>
## 285        <NA>
## 286        <NA>
## 287        <NA>
## 288           0
## 289           1
## 290           0
## 291           1
## 292           0
## 293           1
## 294        <NA>
## 295           1
## 296           0
## 297           1
## 298           1
## 299        <NA>
## 300           0
## 301           1
## 302           0
## 303        <NA>
## 304           0
## 305           1
## 306        <NA>
## 307        <NA>
## 308        <NA>
## 309           0
## 310           0
## 311           1
## 312        <NA>
```

```
## 313       <NA>
## 314       <NA>
## 315          0
## 316          0
## 317          1
## 318       <NA>
## 319       <NA>
## 320          0
## 321          1
## 322          0
## 323          1
## 324       <NA>
## 325       <NA>
## 326          0
## 327          1
## 328       <NA>
## 329       <NA>
## 330       <NA>
## 331       <NA>
## 332       <NA>
## 333       <NA>
## 334       <NA>
## 335          1
## 336       <NA>
## 337          1
## 338       <NA>
## 339          0
## 340       <NA>
## 341       <NA>
## 342          0
## 343       <NA>
## 344          0
## 345          1
## 346       <NA>
## 347       <NA>
## 348          0
## 349          1
## 350       <NA>
## 351       <NA>
## 352          0
## 353       <NA>
## 354       <NA>
## 355       <NA>
## 356          1
## 357       <NA>
## 358       <NA>
## 359       <NA>
## 360       <NA>
## 361       <NA>
## 362       <NA>
## 363          0
## 364       <NA>
## 365          1
## 366          0
```

```
## 367          1
## 368          0
## 369       <NA>
## 370       <NA>
## 371       <NA>
## 372          0
## 373          1
## 374          0
## 375       <NA>
## 376       <NA>
## 377       <NA>
## 378          0
## 379          1
## 380          0
## 381          1
## 382       <NA>
## 383       <NA>
## 384          0
## 385          1
## 386       <NA>
## 387          1
## 388          0
## 389       <NA>
## 390          1
## 391       <NA>
## 392       <NA>
## 393       <NA>
## 394          1
## 395          0
## 396       <NA>
## 397       <NA>
## 398          0
## 399          1
## 400          0
## 401          1
## 402       <NA>
## 403       <NA>
## 404       <NA>
## 405       <NA>
## 406       <NA>
## 407       <NA>
## 408       <NA>
## 409          0
## 410          0
## 411          0
## 412          0
## 413       <NA>
## 414          0
## 415       <NA>
## 416       <NA>
## 417       <NA>
## 418          0
## 419       <NA>
## 420       <NA>
```

```
## 421        <NA>
## 422        <NA>
## 423        <NA>
## 424        <NA>
## 425           0
## 426           1
## 427        <NA>
## 428           1
## 429        <NA>
## 430        <NA>
## 431        <NA>
## 432           0
## 433        <NA>
## 434        <NA>
## 435        <NA>
## 436           1
## 437        <NA>
## 438           1
## 439           0
## 440           0
## 441           1
## 442           0
## 443           1
## 444           0
## 445           1
## 446           1
## 447        <NA>
## 448        <NA>
## 449        <NA>
## 450           0
## 451           1
## 452           1
## 453        <NA>
## 454        <NA>
## 455        <NA>
## 456           0
## 457        <NA>
## 458           0
## 459           1
## 460           0
## 461           1
## 462           0
## 463        <NA>
## 464           1
## 465           0
## 466           0
## 467           1
## 468           1
## 469           0
## 470           0
## 471           1
## 472        <NA>
## 473        <NA>
## 474           0
```

```
## 475       <NA>
## 476          0
## 477          1
## 478          0
## 479          1
## 480          0
## 481          1
## 482          0
## 483       <NA>
## 484          0
## 485          1
## 486          0
## 487          1
## 488       <NA>
## 489       <NA>
## 490       <NA>
## 491       <NA>
## 492          0
## 493          1
## 494       <NA>
## 495       <NA>
## 496          0
## 497          1
## 498          1
## 499          0
## 500          0
## 501          1
## 502          1
## 503          0
## 504          0
## 505          1
## 506       <NA>
## 507       <NA>
## 508       <NA>
## 509       <NA>
## 510          0
## 511       <NA>
## 512       <NA>
## 513          0
## 514       <NA>
## 515          0
## 516          1
## 517          0
## 518          1
## 519          0
## 520          1
## 521          0
## 522          0
## 523          1
```

Imagine you are a bank manager, and you want to decide which customers can get a loan. Instead of doing this manually, you build a model (the decision tree) to make the decision for you. The decision tree works by asking a series of yes/no questions about each customer. For example:

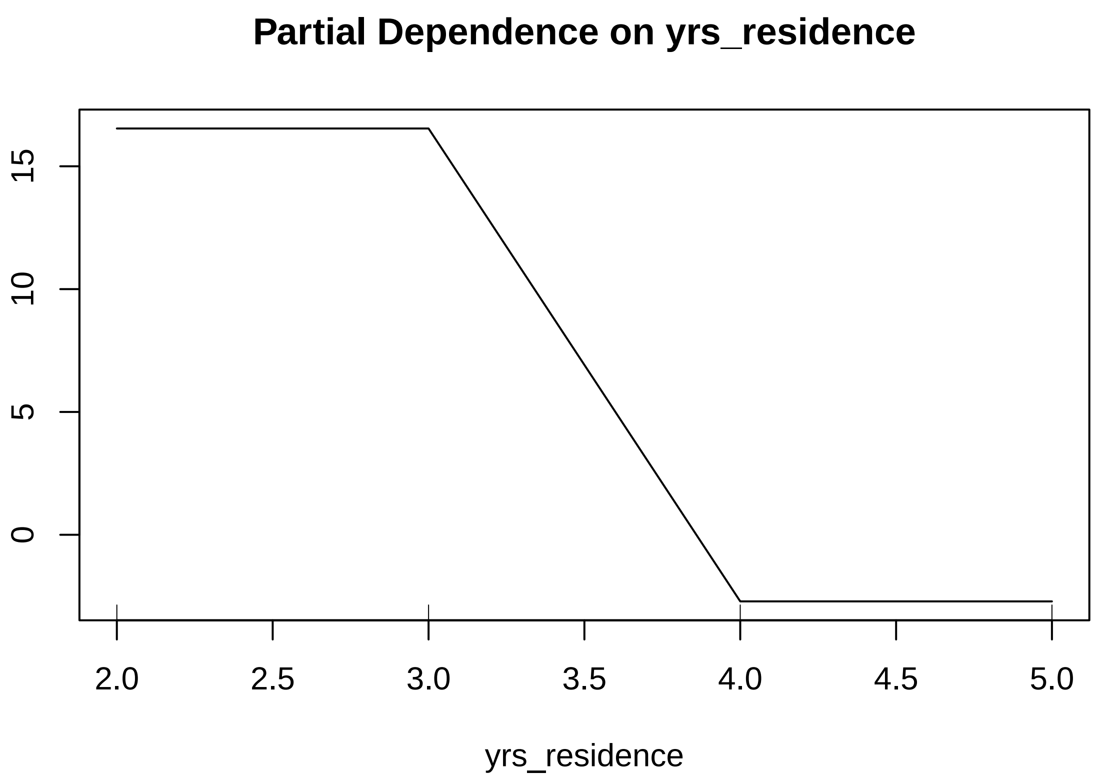
```
1.  Does the customer earn more than R50,000 annually? If yes, they move to the left branch; if no, the
```

```
2.  Has the customer lived in their current residence for more than 3 years? Another yes/no question.
3.  Is their annual gross income more than R45,000? Another split.
```

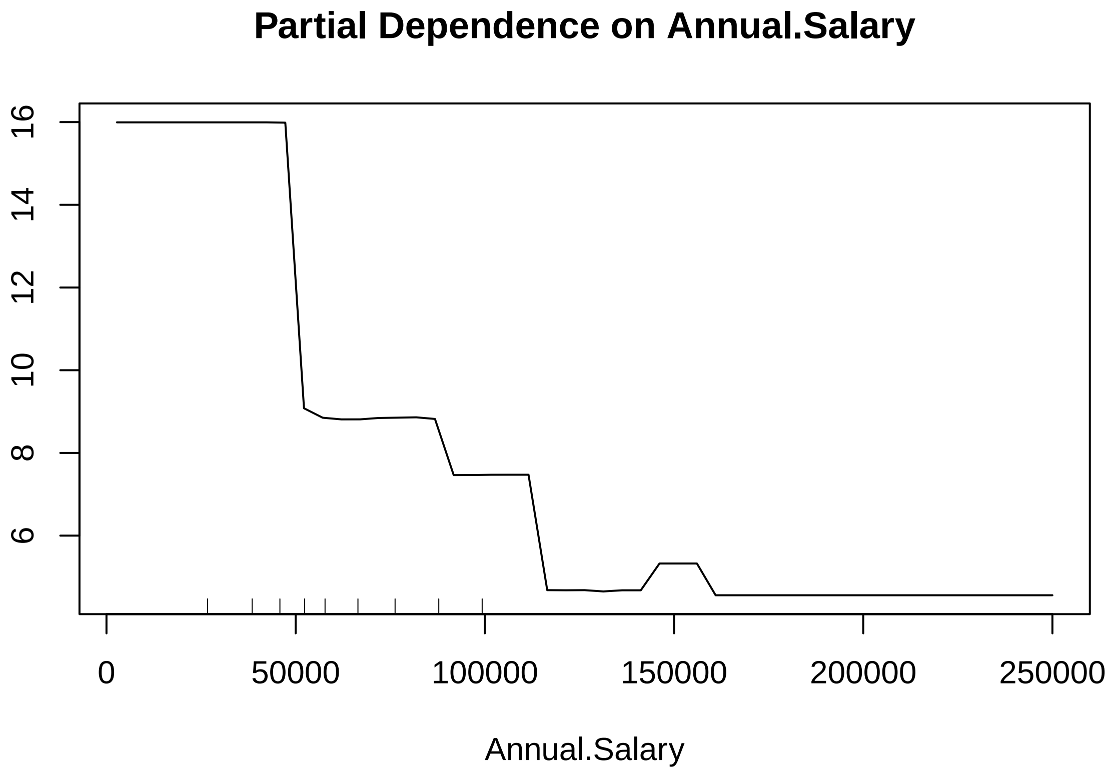
At the end of this question-and-answer game, the tree makes a decision: "eligible" or "not eligible." Each split is chosen because it helps the model make the best decision based on past data.

Create a partial dependence plot

```r
partialPlot(rf_model, trainData, yrs_residence, main = "Partial Dependence on yrs_residence")
```




```r
# Create a partial dependence plot for 'Annual.Salary'
partialPlot(rf_model, trainData, Annual.Salary, main = "Partial Dependence on Annual.Salary")
```



PDPs illustrate how changes in a feature impact the predicted probability of an outcome while holding other features constant

a PDP for yrs_residence might show how the probability of being classified as "eligible" changes as the yrs_residence increases.

If the plot slopes upward, it means that as yrs_residence increases, the probability of a customer being eligible increases.

A flat section in the plot suggests that beyond a certain point, changes in that feature don't significantly impact the prediction.

PDPs can also reveal non-linear effects. For example, eligibility might increase sharply with Annual.Salary up to a certain threshold, then level off.