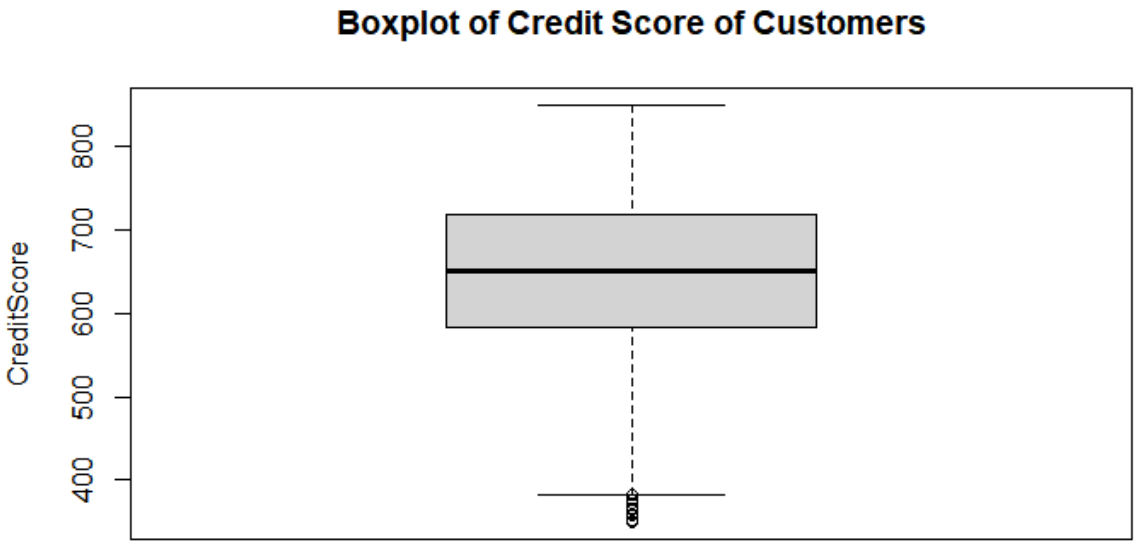
Summary Statistics

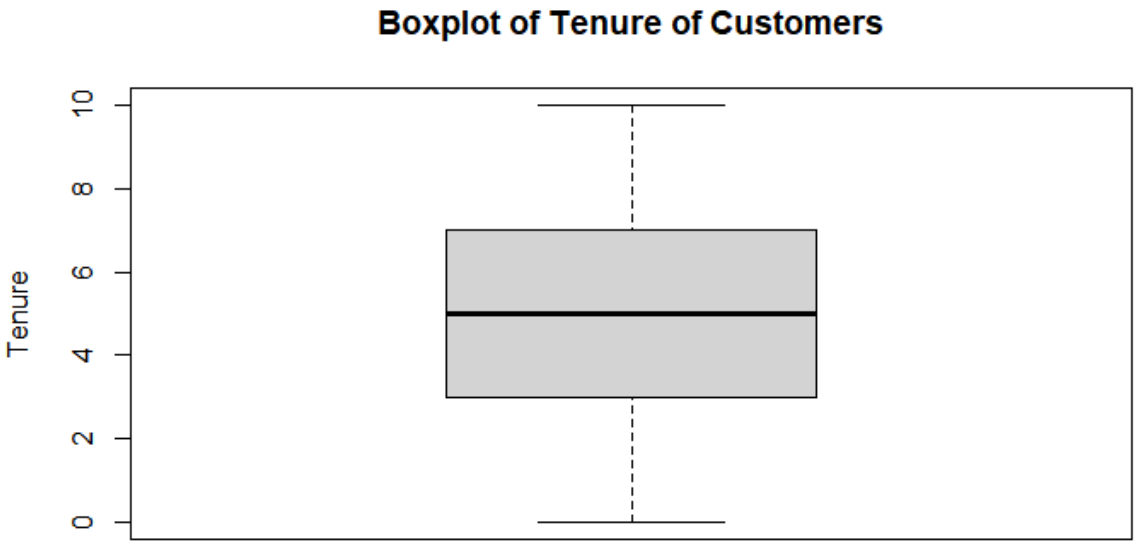
```
  CreditScore      Geography            Gender                Age
 Min.   :350.0   Length:10000       Length:10000       Min.   :18.00
 1st Qu.:584.0   Class :character   Class :character   1st Qu.:32.00
 Median :652.0   Mode  :character   Mode  :character   Median :37.00
 Mean   :650.5                                         Mean   :38.92
 3rd Qu.:718.0                                         3rd Qu.:44.00
 Max.   :850.0                                         Max.   :92.00
     Tenure          Balance       NumOfProducts  HasCrCard IsActiveMember
 Min.   : 0.000   Min.   :     0   Min.   :1.00   0:2945    0:4849
 1st Qu.: 3.000   1st Qu.:     0   1st Qu.:1.00   1:7055    1:5151
 Median : 5.000   Median : 97199   Median :1.00
 Mean   : 5.013   Mean   : 76486   Mean   :1.53
 3rd Qu.: 7.000   3rd Qu.:127644   3rd Qu.:2.00
 Max.   :10.000   Max.   :250898   Max.   :4.00
 EstimatedSalary        Exited
 Min.   :    11.58   Min.   :0.0000
 1st Qu.: 51002.11   1st Qu.:0.0000
 Median :100193.91   Median :0.0000
 Mean   :100090.24   Mean   :0.2037
 3rd Qu.:149388.25   3rd Qu.:0.0000
 Max.   :199992.48   Max.   :1.0000
```

Boxplot of Age of Customers

Boxplot of Credit Score

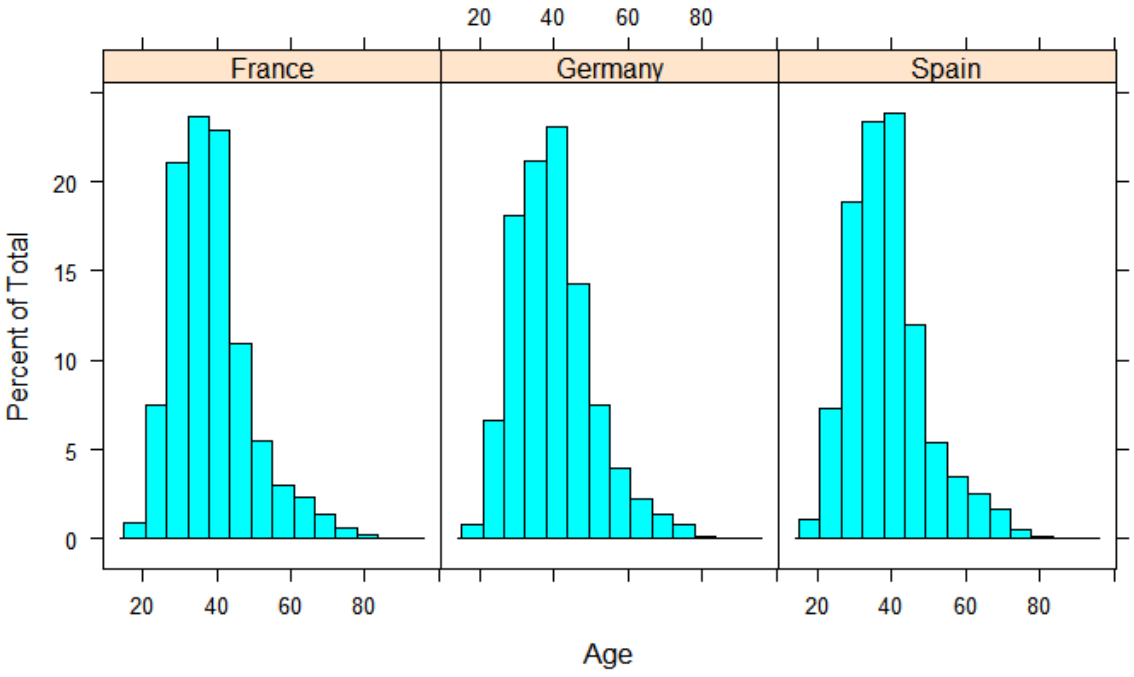


Boxplot of Tenure of Customers

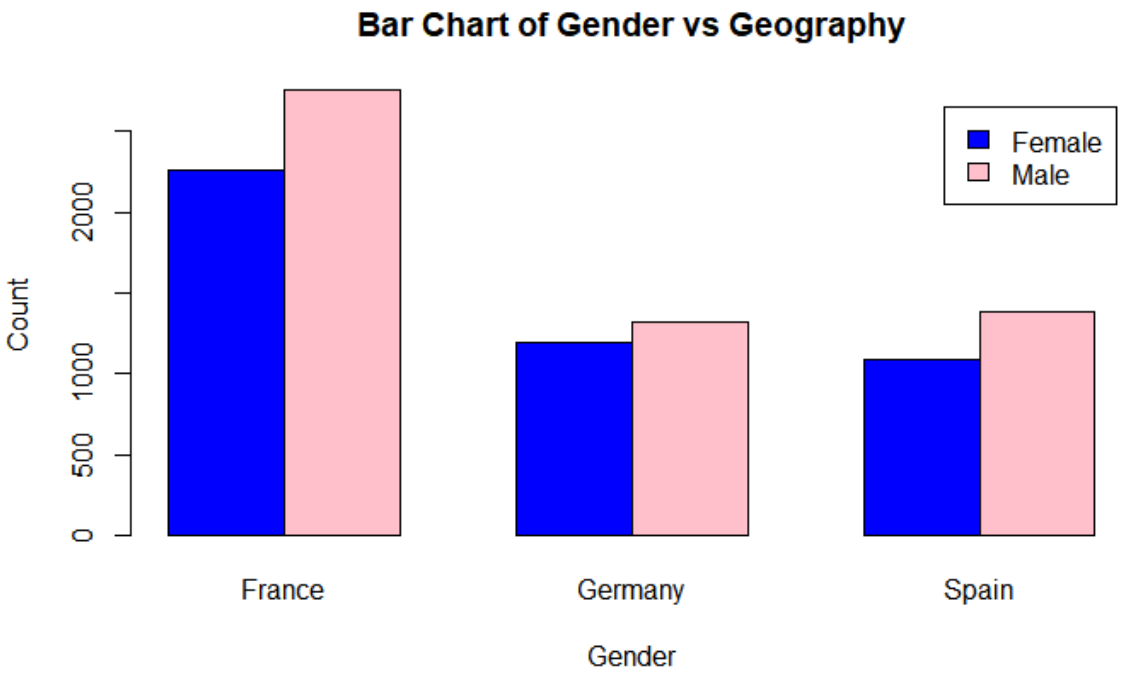


Visualizing Categorical Data

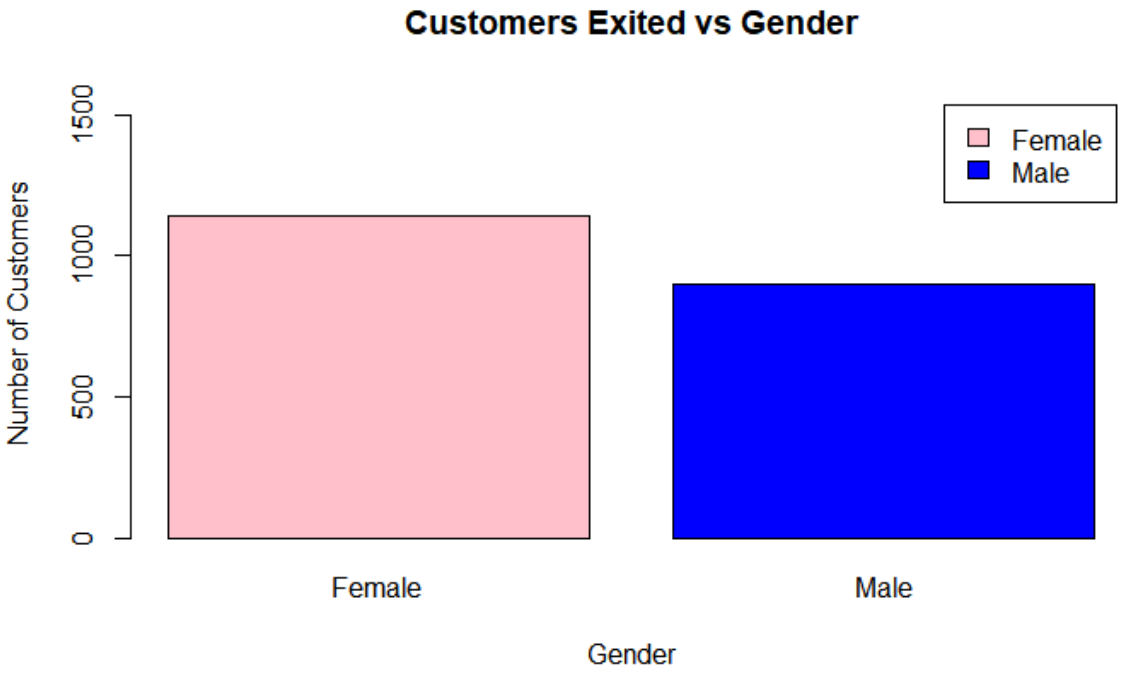
Histogram, Age versus Geography



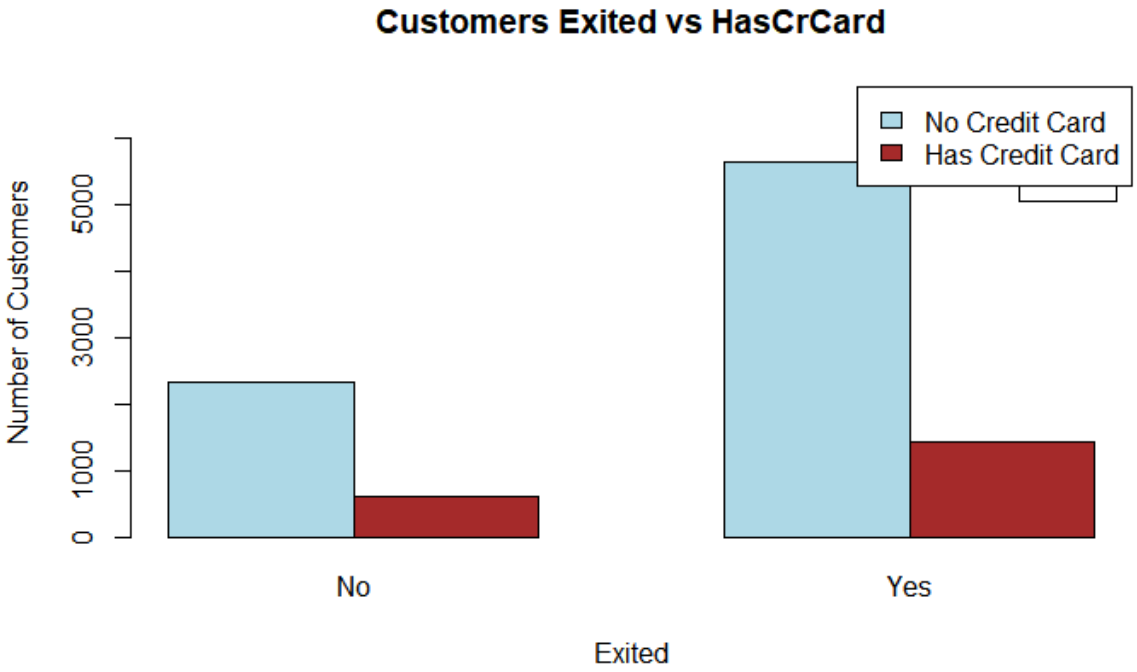
Visualizing Gender and Geography Exited



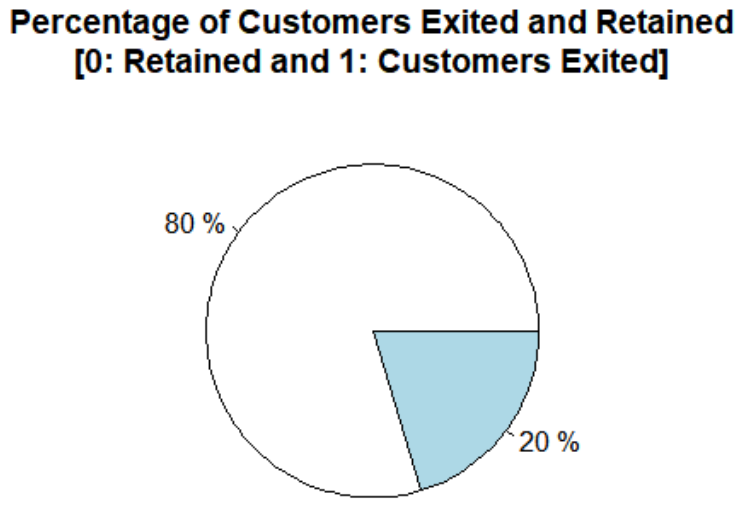
Visualizing Gender versus Exited



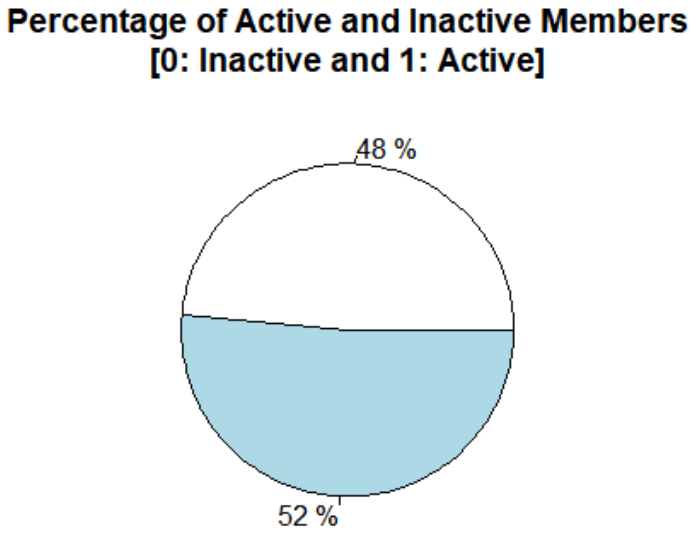
Visualizing HasCrCard Versus Exited



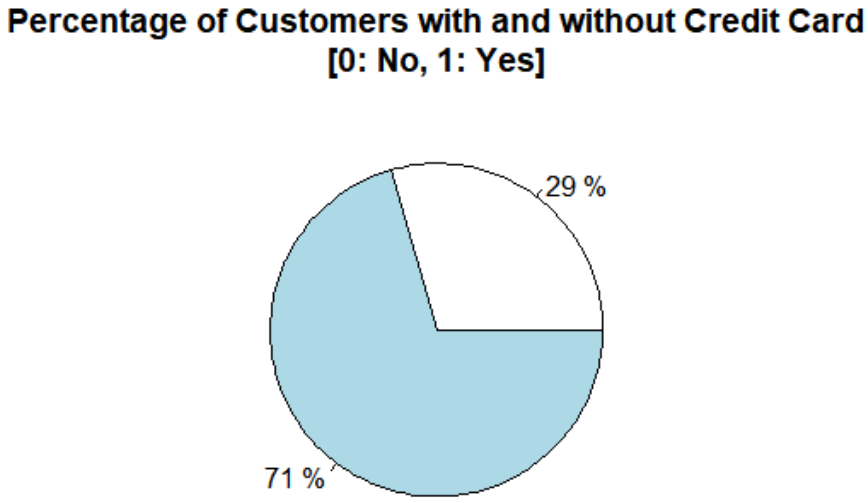
Visualizing Customers Exited using pie chart



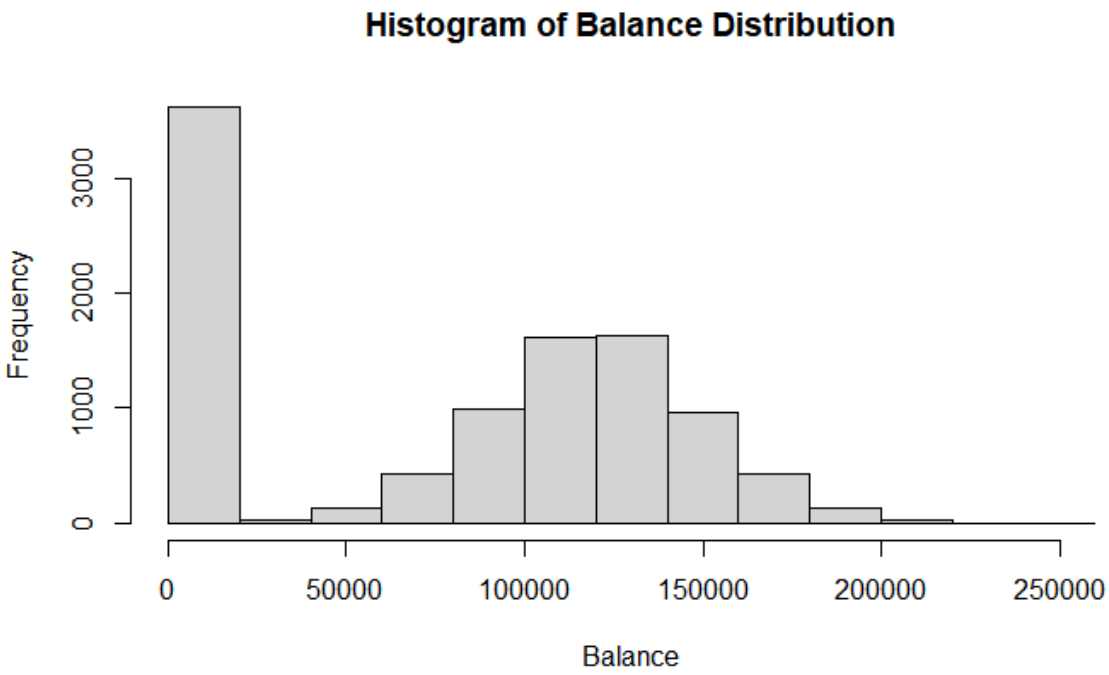
Visualizing IsActiveMember in a pie chart



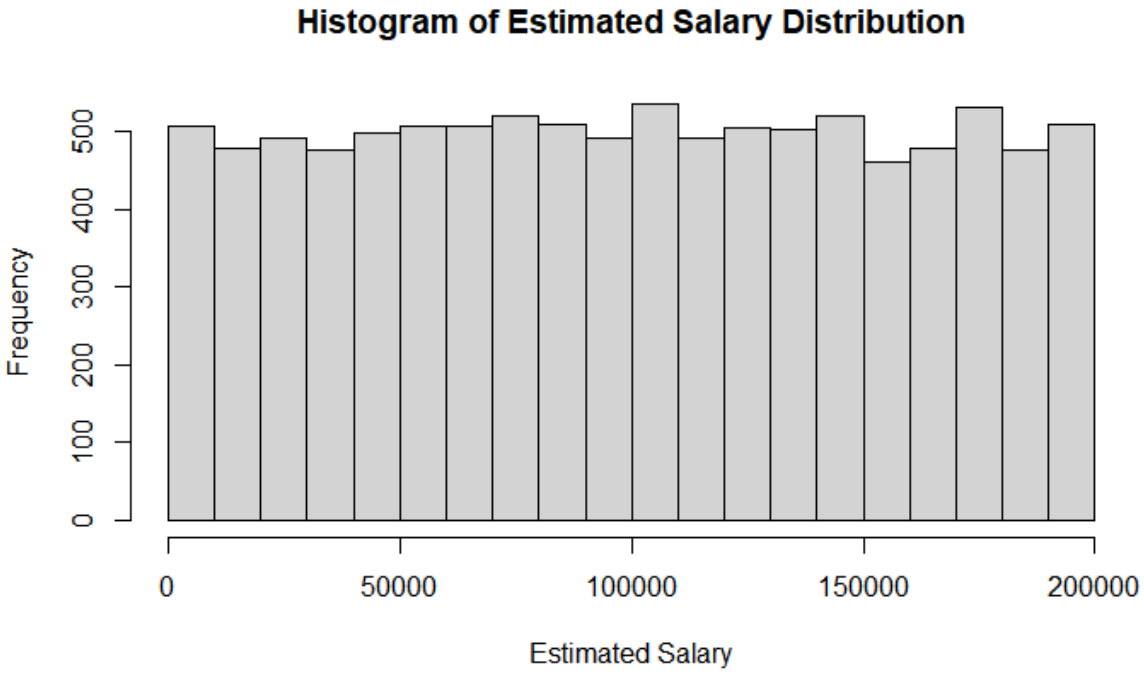
Visualizing HasCrCard in a pie chart



Visualizing Continuous Variables; Balance and EstimatedSalary

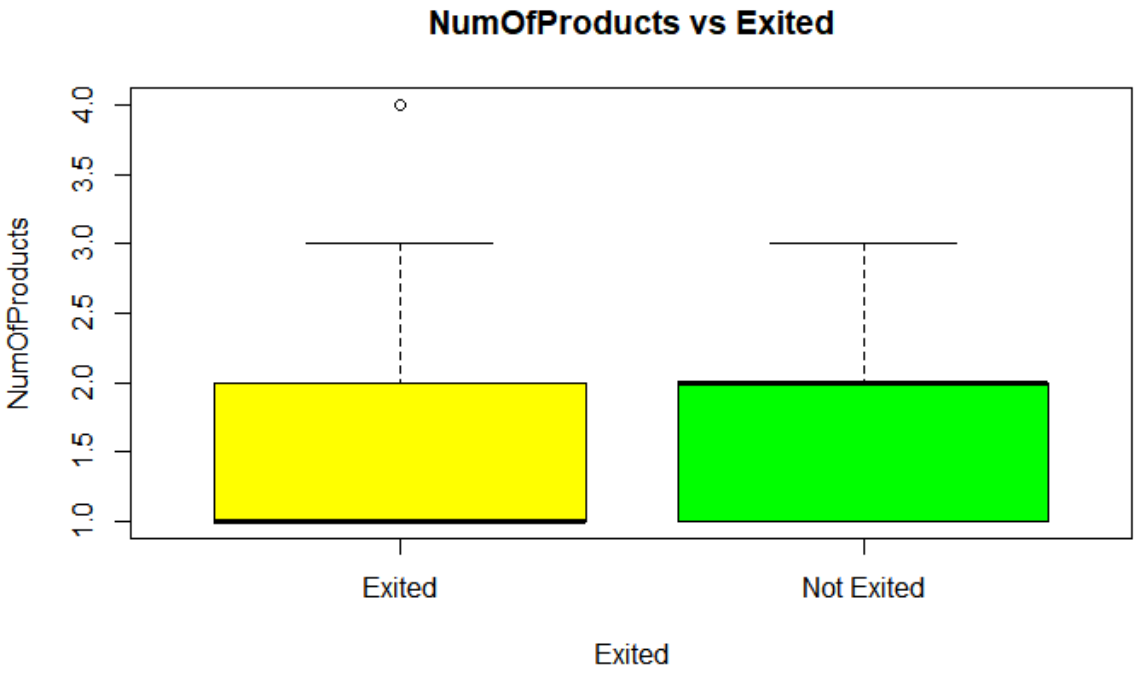


Histogram of Estimated Salary distribution

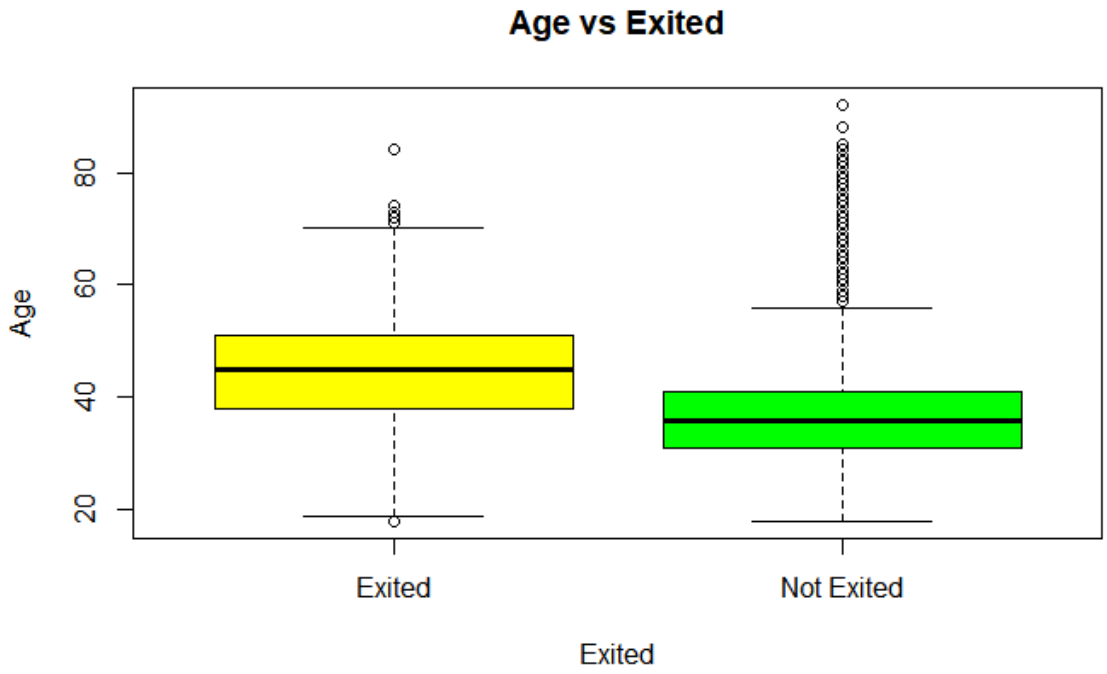


Visualizing Numeric Variables versus Categorical Variables

Visualizing NumofProducts versus Exited

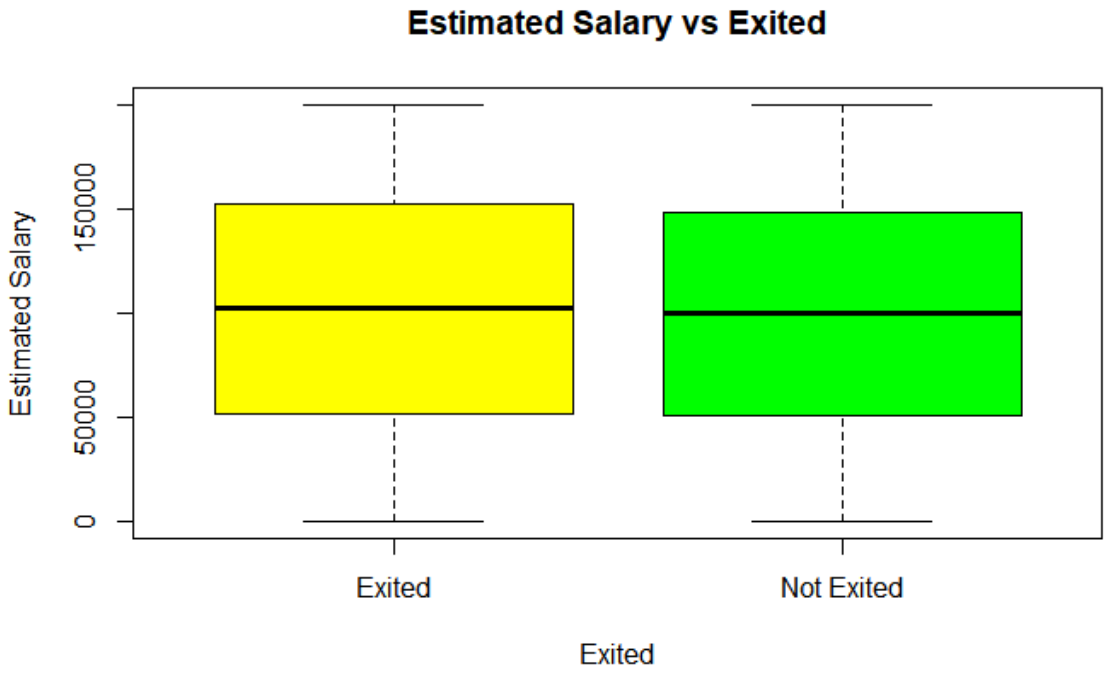


Visualizing Age versus Exited

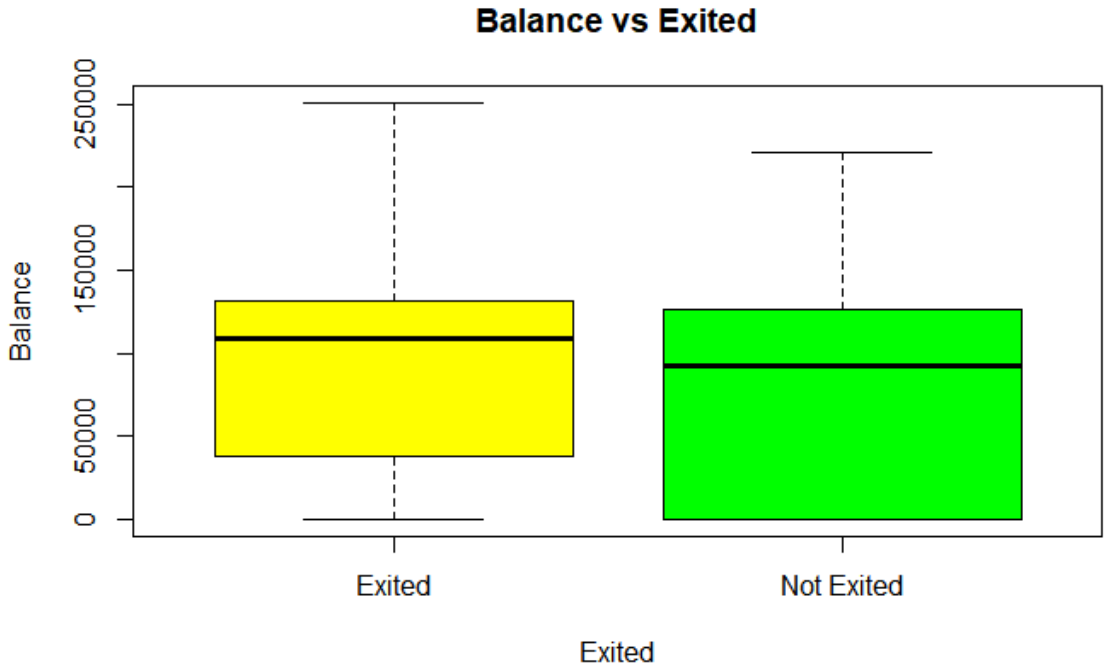


Visualizing Continuous Variables versus Categorical Variables

Visualizing EstimatedSalary versus Exited



Visualizing Balance versus Exited



Building Machine Language Model, Decision Tree

Check the proportion of class variable

```
> plot(churn_train_model)
> prop.table(summary(churn_train$Exited))
        0         1
0.7948889 0.2051111
> prop.table(summary(churn_test$Exited))
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.0000  0.0000  0.0000  0.1756  0.0000  0.8244
>
```

Summary of the Train Model

```
Call:
C5.0.default(x = churn_train[-11], y = churn_train$Exited, control
 = C5.0Control(minCases = 400))


C5.0 [Release 2.07 GPL Edition]         Sun Jun 23 14:00:45 2024
-------------------------------

Class specified by attribute `outcome'

Read 9000 cases (11 attributes) from undefined.data

Decision tree:

IsActiveMember = 1: 0 (4613/653)
IsActiveMember = 0:
:...Age <= 44: 0 (3482/584)
    Age > 44: 1 (905/296)

Evaluation on training data (9000 cases):

            Decision Tree
          ----------------
          Size      Errors

             3 1533(17.0%)   <<


           (a)   (b)    <-classified as
          ----  ----
          6858   296    (a): class 0
          1237   609    (b): class 1


        Attribute usage:

        100.00% IsActiveMember
         48.74% Age


Time: 0.1 secs
```

Display simple facts about the Tree

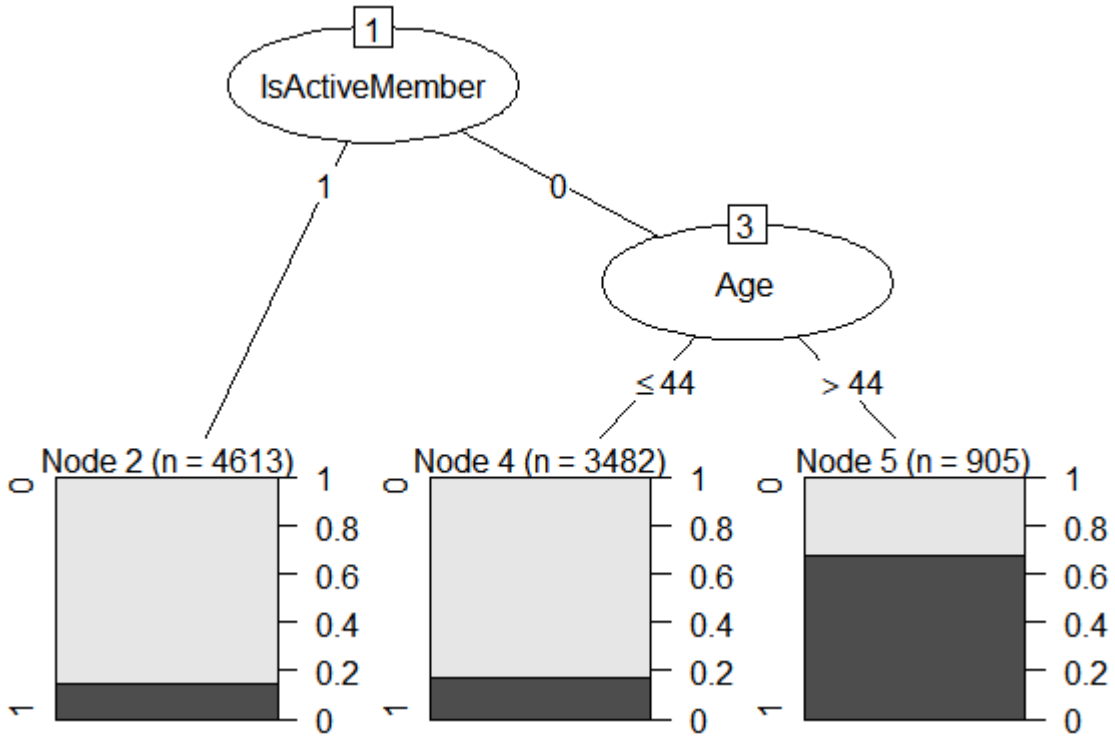
```
Call:
C5.0.default(x = churn_train[-11], y = churn_train$Exited, control
 = C5.0Control(minCases = 400))

Classification Tree
Number of samples: 9000
Number of predictors: 10

Tree size: 3

Non-standard options: attempt to group attributes, minimum number of cases: 400
```

Plot the tree



Evaluating model performance

```
   Cell Contents
|-------------------------|
|                       N |
|           N / Row Total |
|         N / Table Total |
|-------------------------|


Total Observations in Table:  1000


               | predicted Exited
actual Exited |         0 |         1 | Row Total |
--------------|-----------|-----------|-----------|
            0 |       776 |        33 |       809 |
              |     0.959 |     0.041 |     0.809 |
              |     0.776 |     0.033 |           |
--------------|-----------|-----------|-----------|
            1 |       127 |        64 |       191 |
              |     0.665 |     0.335 |     0.191 |
              |     0.127 |     0.064 |           |
--------------|-----------|-----------|-----------|
 Column Total |       903 |        97 |      1000 |
--------------|-----------|-----------|-----------|


> |
```