

# Eric\_Hirsch\_621\_Assignment\_4

## Predicting Insurance Claims

Eric Hirsch

4/7/2022

### Contents

1. Data Exploration . . . . .	2
A. Summary Statistics . . . . .	2
B. Distributions . . . . .	5
C. Multicollinearity . . . . .	6
2. Data Preparation . . . . .	6
A. Create Dummy Variables . . . . .	6
B. Address Missing Values . . . . .	6
2. Predicting TARGET_FLAG . . . . .	10
A. Explore relationships . . . . .	10
4. Select model . . . . .	712
5. Predict TARGET_AMT . . . . .	712
10. Conclusion . . . . .	712

```
##          checking for file 'C:\Users\Eric\AppData\Local\Temp\RtmpYLf6dE\remotes2b5c52da347f\ericonsi
##          - preparing 'EHData':
##          checking DESCRIPTION meta-information ...          checking DESCRIPTION meta-information ...    v  check
##          - checking for LF line-endings in source and make files and shell scripts
## - checking for empty or unneeded directories
## - creating default NAMESPACE file
## - building 'EHData_0.1.0.tar.gz'
##
##
```

We examine records of car insurance customers to build two predictive models: one for whether the customer would have crashed, and second, the \$ amount paid for the crash.

The main issue in the dataset are outliers. Without transformation, the distribution of residuals is not normal, and there are two many influential points to create reliable models.

## 1. Data Exploration

**A. Summary Statistics** We first examine the data. The dataset consists of 8161 observations and 26 variables (including two target variables, TARGET\_FLAG and TARGET\_AMT). 14 of the predictor variables are numeric. Approximately 27% of customers had an accident - the rest did not. TARGET\_AMT appears to be highly skewed.

```
##      INDEX      TARGET_FLAG      TARGET_AMT      KIDSDRIV
## Min.      : 1      Min.      :0.0000      Min.      : 0      Min.      :0.0000
## 1st Qu.: 2559      1st Qu.:0.0000      1st Qu.: 0      1st Qu.:0.0000
## Median : 5133      Median :0.0000      Median : 0      Median :0.0000
## Mean    : 5152      Mean    :0.2638      Mean    : 1504      Mean    :0.1711
## 3rd Qu.: 7745      3rd Qu.:1.0000      3rd Qu.: 1036      3rd Qu.:0.0000
## Max.    :10302      Max.    :1.0000      Max.    :107586      Max.    :4.0000
##
##      AGE      HOMEKIDS      YOJ      INCOME
## Min.      :16.00      Min.      :0.0000      Min.      : 0.0      Min.      : 0
## 1st Qu.:39.00      1st Qu.:0.0000      1st Qu.: 9.0      1st Qu.: 28097
## Median :45.00      Median :0.0000      Median :11.0      Median : 54028
## Mean     :44.79      Mean     :0.7212      Mean     :10.5      Mean     : 61898
## 3rd Qu.:51.00      3rd Qu.:1.0000      3rd Qu.:13.0      3rd Qu.: 85986
## Max.     :81.00      Max.     :5.0000      Max.     :23.0      Max.     :367030
## NA's      :6      NA's      :454      NA's      :445
##      PARENT1      HOME_VAL      MSTATUS      SEX
## Length:8161      Min.      : 0      Length:8161      Length:8161
## Class :character      1st Qu.: 0      Class :character      Class :character
## Mode  :character      Median :161160      Mode  :character      Mode  :character
##                               Mean  :154867
##                               3rd Qu.:238724
##                               Max.   :885282
##                               NA's    :464
##      EDUCATION      JOB      TRAVTIME      CAR_USE
## Length:8161      Length:8161      Min.      : 5.00      Length:8161
## Class :character      Class :character      1st Qu.: 22.00      Class :character
## Mode  :character      Mode  :character      Median : 33.00      Mode  :character
##                               Mean    : 33.49
##                               3rd Qu.: 44.00
##                               Max.    :142.00
##
##      BLUEBOOK      TIF      CAR_TYPE      RED_CAR
## Min.      : 1500      Min.      : 1.000      Length:8161      Length:8161
## 1st Qu.: 9280      1st Qu.: 1.000      Class :character      Class :character
## Median :14440      Median : 4.000      Mode  :character      Mode  :character
## Mean     :15710      Mean     : 5.351
## 3rd Qu.:20850      3rd Qu.: 7.000
## Max.     :69740      Max.     :25.000
##
##      OLDCLAIM      CLM_FREQ      REVOKED      MVR_PTS
## Min.      : 0      Min.      :0.0000      Length:8161      Min.      : 0.000
## 1st Qu.: 0      1st Qu.:0.0000      Class :character      1st Qu.: 0.000
## Median : 0      Median :0.0000      Mode  :character      Median : 1.000
## Mean     : 4037      Mean     :0.7986      Mean     : 1.696
## 3rd Qu.: 4636      3rd Qu.:2.0000      3rd Qu.: 3.000
## Max.     :57037      Max.     :5.0000      Max.     :13.000
```

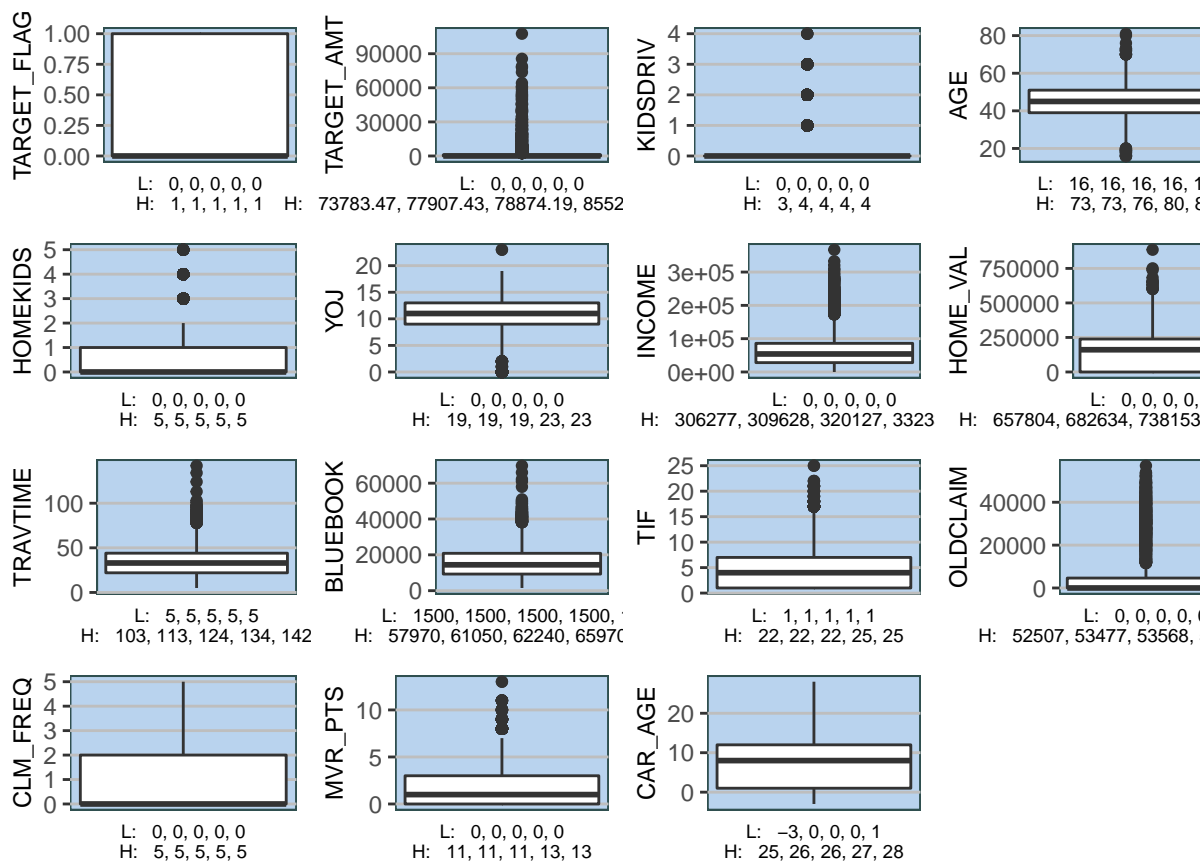
```

##
##      CAR_AGE      URBANICITY
##  Min.      :-3.000   Length:8161
##  1st Qu.: 1.000   Class :character
##  Median : 8.000   Mode  :character
##  Mean    : 8.328
##  3rd Qu.:12.000
##  Max.    :28.000
##  NA's    :510

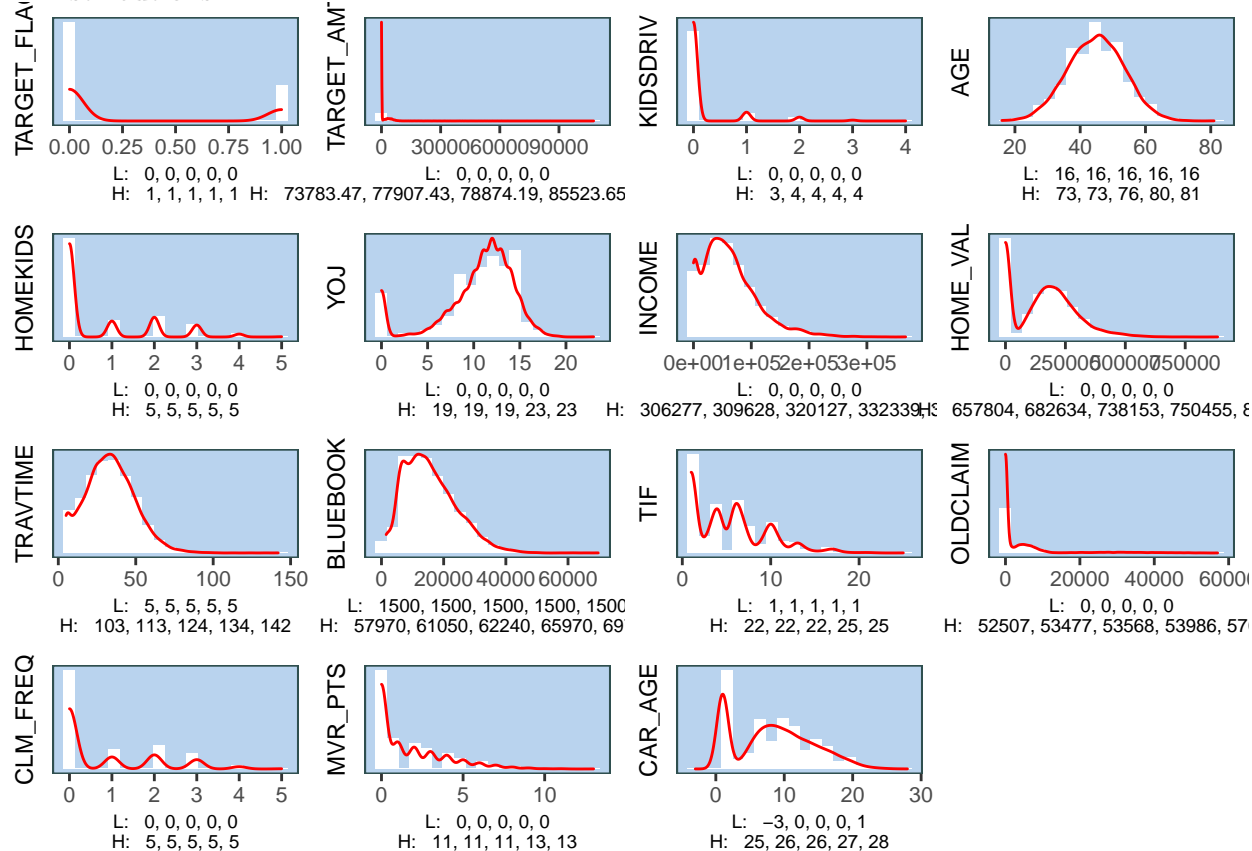
## 'data.frame':      8161 obs. of  26 variables:
##  $ INDEX      : int   1 2 4 5 6 7 8 11 12 13 ...
##  $ TARGET_FLAG: int   0 0 0 0 0 1 0 1 1 0 ...
##  $ TARGET_AMT : num   0 0 0 0 0 ...
##  $ KIDSDRIV   : int   0 0 0 0 0 0 0 1 0 0 ...
##  $ AGE        : int  60 43 35 51 50 34 54 37 34 50 ...
##  $ HOMEKIDS   : int   0 0 1 0 0 1 0 2 0 0 ...
##  $ YOJ        : int  11 11 10 14 NA 12 NA NA 10 7 ...
##  $ INCOME     : num  67349 91449 16039 NA 114986 ...
##  $ PARENT1    : chr   "No" "No" "No" "No" ...
##  $ HOME_VAL   : num   0 257252 124191 306251 243925 ...
##  $ MSTATUS    : chr   "z_No" "z_No" "Yes" "Yes" ...
##  $ SEX        : chr   "M" "M" "z_F" "M" ...
##  $ EDUCATION  : chr   "PhD" "z_High School" "z_High School" "<High School" ...
##  $ JOB        : chr   "Professional" "z_Blue Collar" "Clerical" "z_Blue Collar" ...
##  $ TRAVTIME   : int  14 22 5 32 36 46 33 44 34 48 ...
##  $ CAR_USE    : chr   "Private" "Commercial" "Private" "Private" ...
##  $ BLUEBOOK   : num  14230 14940 4010 15440 18000 ...
##  $ TIF        : int  11 1 4 7 1 1 1 1 1 7 ...
##  $ CAR_TYPE   : chr   "Minivan" "Minivan" "z_SUV" "Minivan" ...
##  $ RED_CAR    : chr   "yes" "yes" "no" "yes" ...
##  $ OLDCLAIM   : num  4461 0 38690 0 19217 ...
##  $ CLM_FREQ   : int   2 0 2 0 2 0 0 1 0 0 ...
##  $ REVOKED    : chr   "No" "No" "No" "No" ...
##  $ MVR_PTS    : int   3 0 3 0 3 0 0 10 0 1 ...
##  $ CAR_AGE    : int  18 1 10 6 17 7 1 7 1 17 ...
##  $ URBANICITY : chr   "Highly Urban/ Urban" "Highly Urban/ Urban" "Highly Urban/ Urban" "Highly Urban/ Urban" ...

```



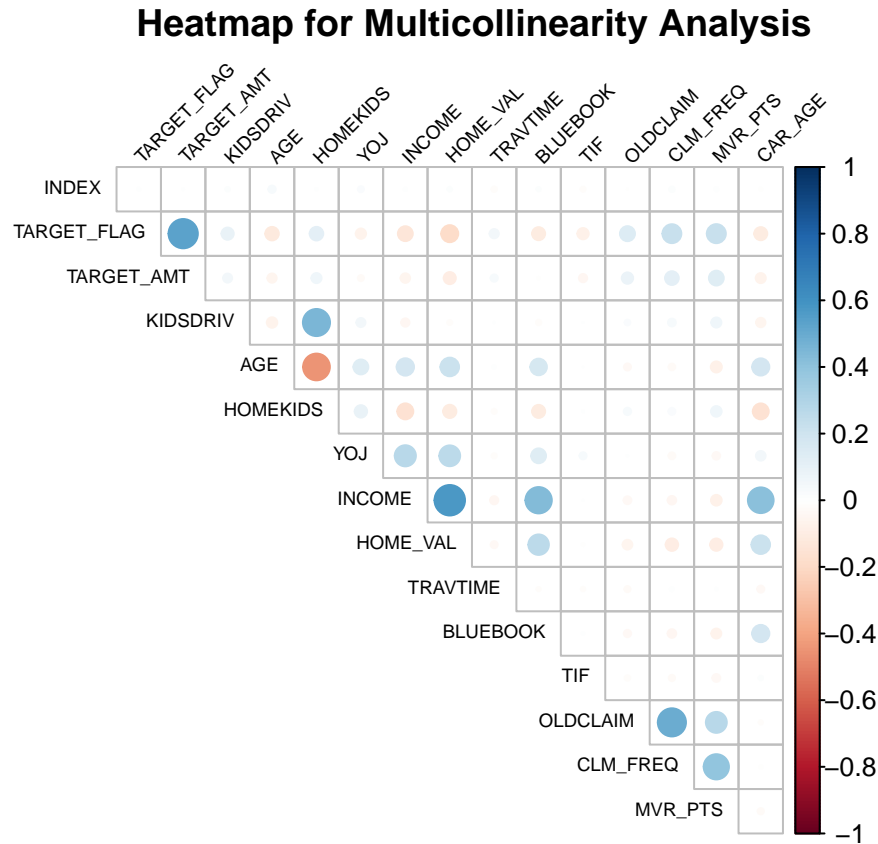


## B. Distributions



Many of the variables are highly skewed, particularly TARGET\_AMT. The level of outliers is very high.

**C. Multicollinearity** The chart below shows multicollinearity for numerical variables only. There are no surprises here - older people tend not to have children at home, income and home value are related, etc. Multicollinearity does not present offhand as a major issue.

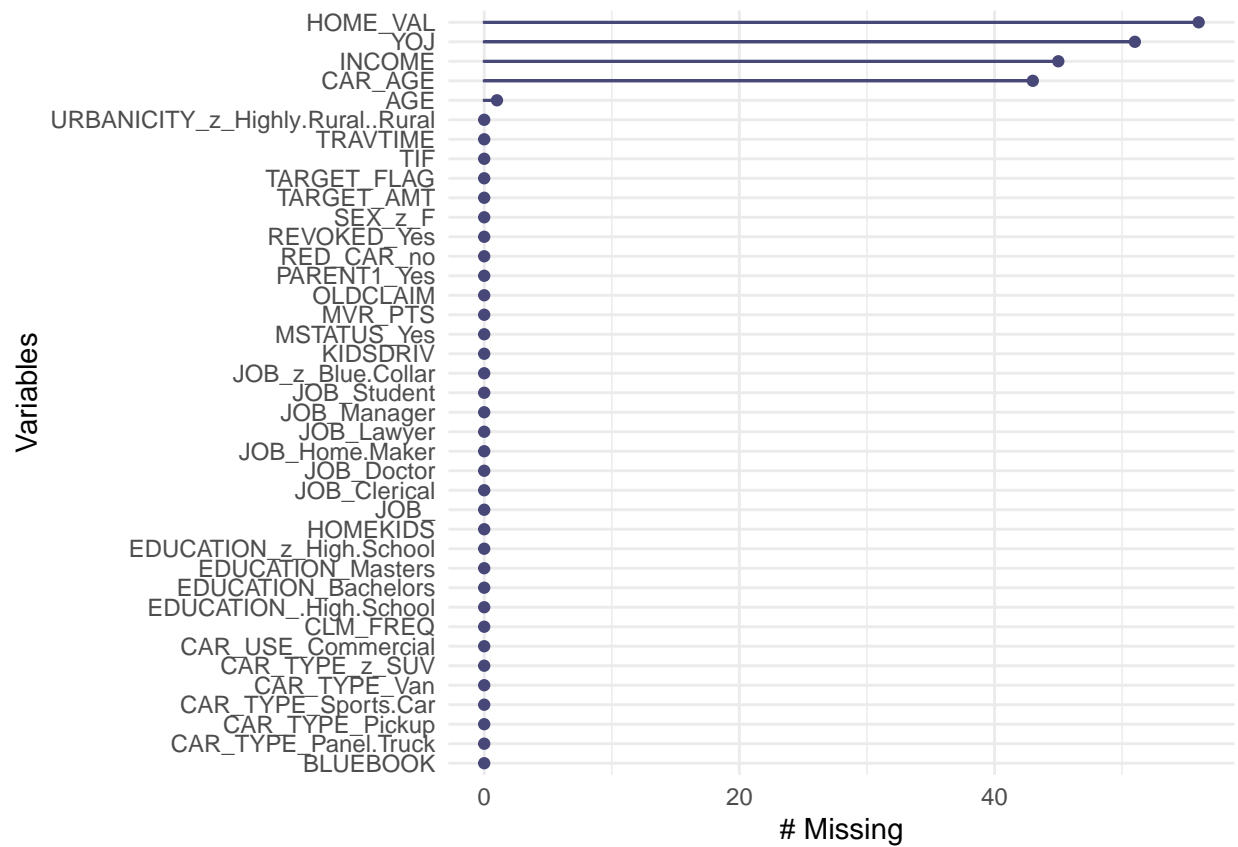


## 2. Data Preparation

**A. Create Dummy Variables** We create dummy variables from the character variables in the database.

**B. Address Missing Values** We consider the missing values. We disregard missing values in character columns because these NAs were isolated out in their own columns when we dummified the data. We convert the 0s in INCOME and HOME\_VAL to NA since 0 is implausible. We create flags to track the NAs for the columns with the most significant NAs - INCOME, HOME\_VAL, CAR\_AGE, and YOJ. Finally we use MICE to populate the missing values.

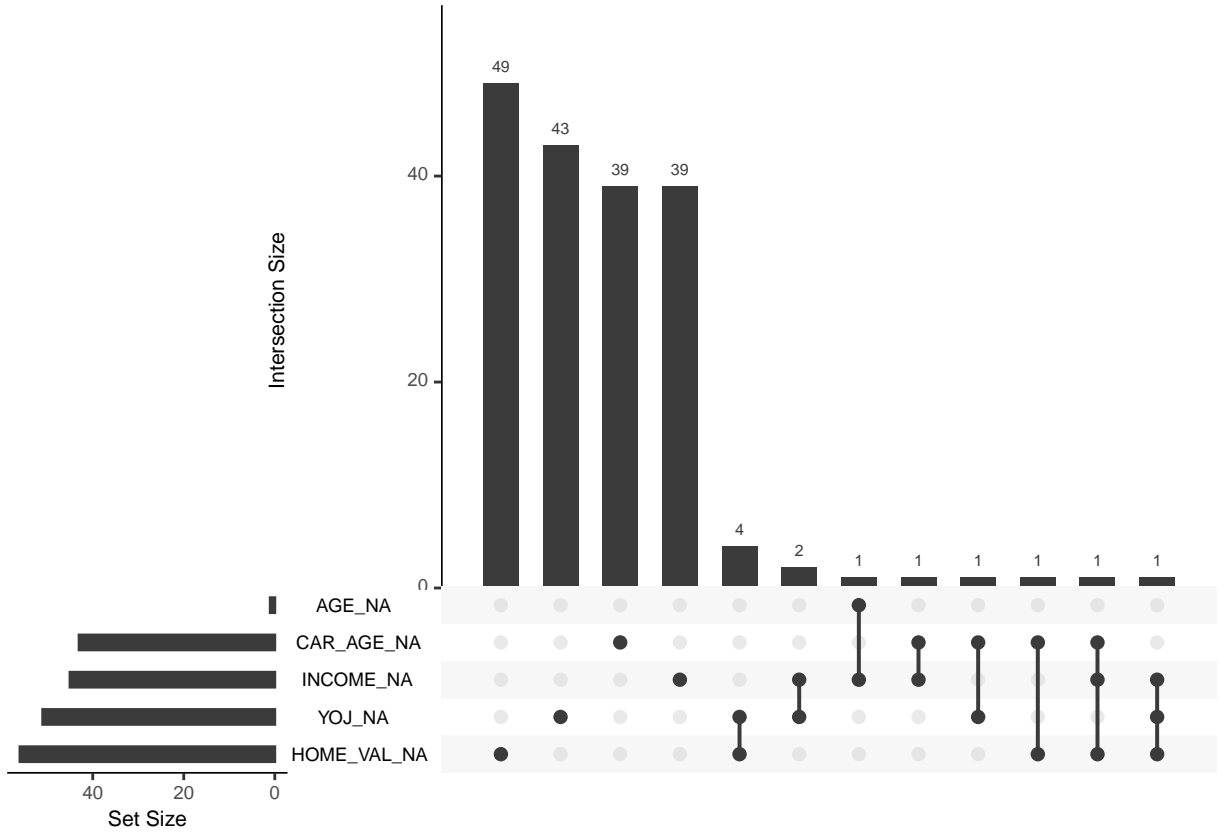
```
## [[1]]
```



```
##
## [[2]]
```







```
##
## iter imp variable
## 1 1 AGE YOJ INCOME HOME_VAL CAR_AGE
## 1 2 AGE YOJ INCOME HOME_VAL CAR_AGE
## 1 3 AGE YOJ INCOME HOME_VAL CAR_AGE
## 1 4 AGE YOJ INCOME HOME_VAL CAR_AGE
## 1 5 AGE YOJ INCOME HOME_VAL CAR_AGE
## 2 1 AGE YOJ INCOME HOME_VAL CAR_AGE
## 2 2 AGE YOJ INCOME HOME_VAL CAR_AGE
## 2 3 AGE YOJ INCOME HOME_VAL CAR_AGE
## 2 4 AGE YOJ INCOME HOME_VAL CAR_AGE
## 2 5 AGE YOJ INCOME HOME_VAL CAR_AGE
## 3 1 AGE YOJ INCOME HOME_VAL CAR_AGE
## 3 2 AGE YOJ INCOME HOME_VAL CAR_AGE
## 3 3 AGE YOJ INCOME HOME_VAL CAR_AGE
## 3 4 AGE YOJ INCOME HOME_VAL CAR_AGE
## 3 5 AGE YOJ INCOME HOME_VAL CAR_AGE
## 4 1 AGE YOJ INCOME HOME_VAL CAR_AGE
## 4 2 AGE YOJ INCOME HOME_VAL CAR_AGE
## 4 3 AGE YOJ INCOME HOME_VAL CAR_AGE
## 4 4 AGE YOJ INCOME HOME_VAL CAR_AGE
## 4 5 AGE YOJ INCOME HOME_VAL CAR_AGE
## 5 1 AGE YOJ INCOME HOME_VAL CAR_AGE
## 5 2 AGE YOJ INCOME HOME_VAL CAR_AGE
## 5 3 AGE YOJ INCOME HOME_VAL CAR_AGE
## 5 4 AGE YOJ INCOME HOME_VAL CAR_AGE
```

```
## 5 5 AGE YOJ INCOME HOME_VAL CAR_AGE
```

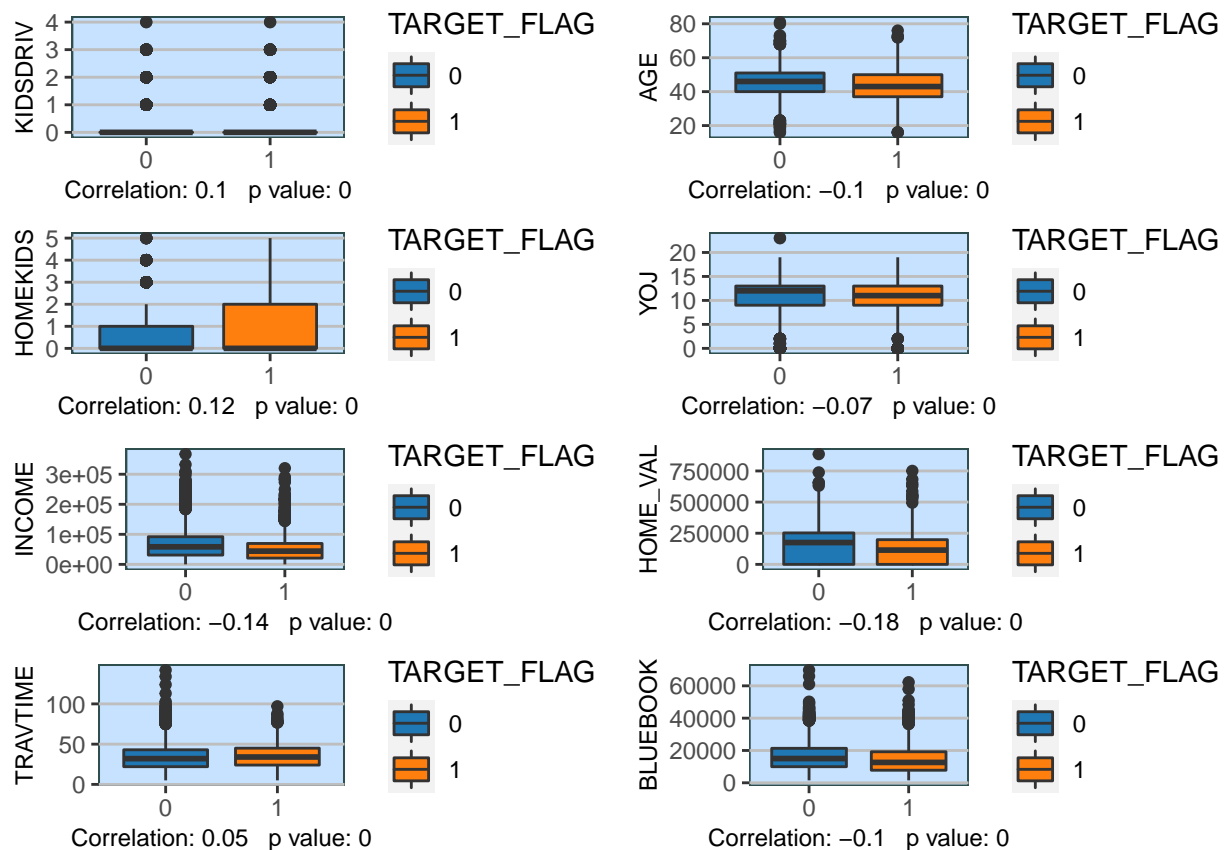
####. C. Perform Transformations

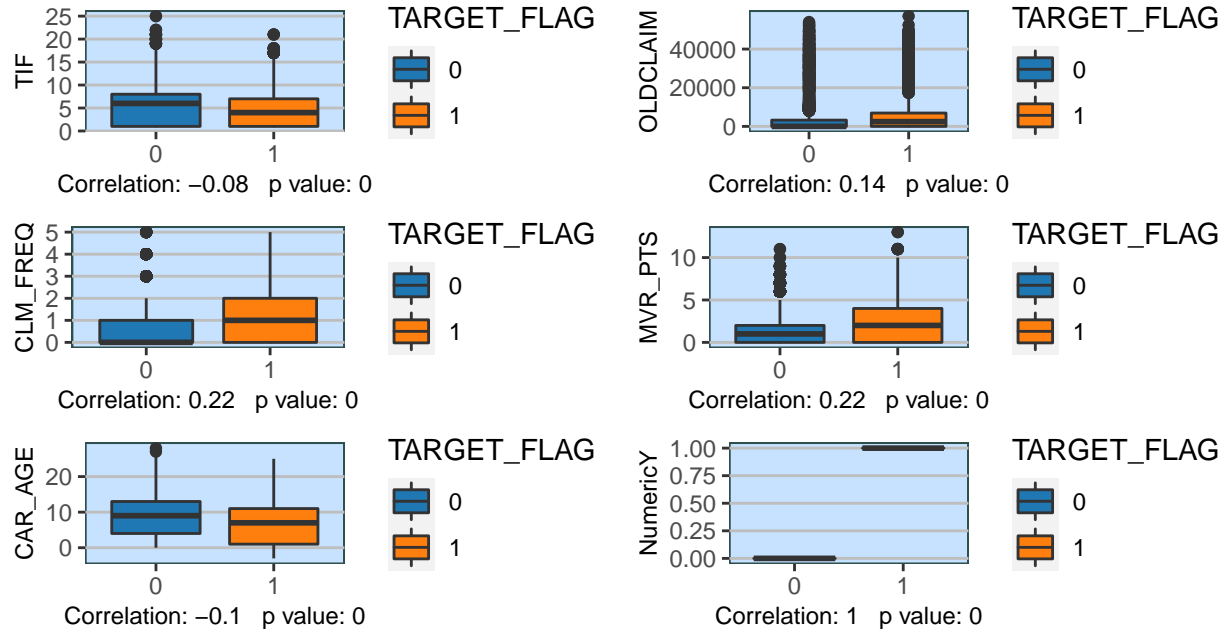
We perform log and other transformations, as well as add an interaction term, to the analysis. These transformations are based on an examination of the distributions of the independent variables. They include:

ageSquared yojSquared income\_log homeval\_log travtime\_log bluebook\_log carage\_log oldclaim\_log  
clm\_freq\_log mvr\_pts\_log tif\_log kidsdriv\_log homekids\_log inter (interaction term = KIDSDRIV\*AGE)

## 2. Predicting TARGET\_FLAG

**A. Explore relationships** We can see from the boxplots run on the original numeric variables against TARGET\_FLAG that the correlations are quite low.





#### B.

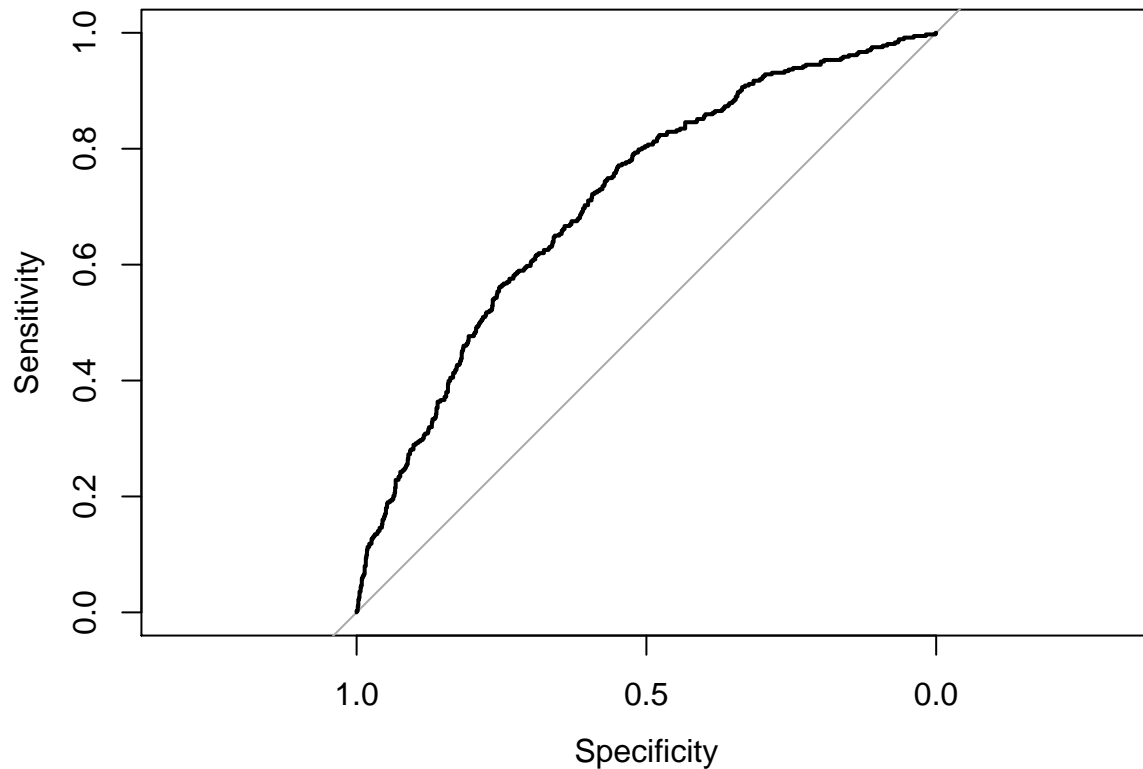
Create Model 1 - the base model with the original numeric variables.

```
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9674  -0.7631  -0.5706   0.9007   2.6031
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.054e-01  2.376e-01  -1.706  0.088016 .
## KIDSDRIV     2.040e-01  7.043e-02   2.896  0.003775 **
## AGE         -1.149e-02  4.573e-03  -2.512  0.012002 *
## HOMEKIDS     7.515e-02  3.781e-02   1.987  0.046895 *
## YOJ         -1.006e-02  8.694e-03  -1.157  0.247220
## INCOME       1.097e-07  1.039e-06   0.106  0.915941
## HOME_VAL    -2.454e-06  3.299e-07  -7.439  1.02e-13 ***
## TRAVTIME     7.986e-03  2.138e-03   3.735  0.000187 ***
## BLUEBOOK    -1.257e-05  4.666e-06  -2.693  0.007071 **
## TIF         -4.489e-02  8.585e-03  -5.229  1.71e-07 ***
## OLDCLAIM     4.734e-06  3.943e-06   1.201  0.229857
## CLM_FREQ     2.707e-01  3.205e-02   8.447  < 2e-16 ***
```

```

## MVR_PTS      1.511e-01  1.577e-02  9.582 < 2e-16 ***
## CAR_AGE      -2.167e-02  6.760e-03 -3.206 0.001346 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 5859.6  on 5075  degrees of freedom
## Residual deviance: 5239.5  on 5062  degrees of freedom
## (1336 observations deleted due to missingness)
## AIC: 5267.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 957 298
##           1  52  65
##
##           Accuracy : 0.7449
##           95% CI : (0.721, 0.7678)
##      No Information Rate : 0.7354
##      P-Value [Acc > NIR] : 0.2227
##
##           Kappa : 0.1629
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9485
##           Specificity : 0.1791
##           Pos Pred Value : 0.7625
##           Neg Pred Value : 0.5556
##           Prevalence : 0.7354
##           Detection Rate : 0.6975
##      Detection Prevalence : 0.9147
##           Balanced Accuracy : 0.5638
##
##           'Positive' Class : 0
##

```

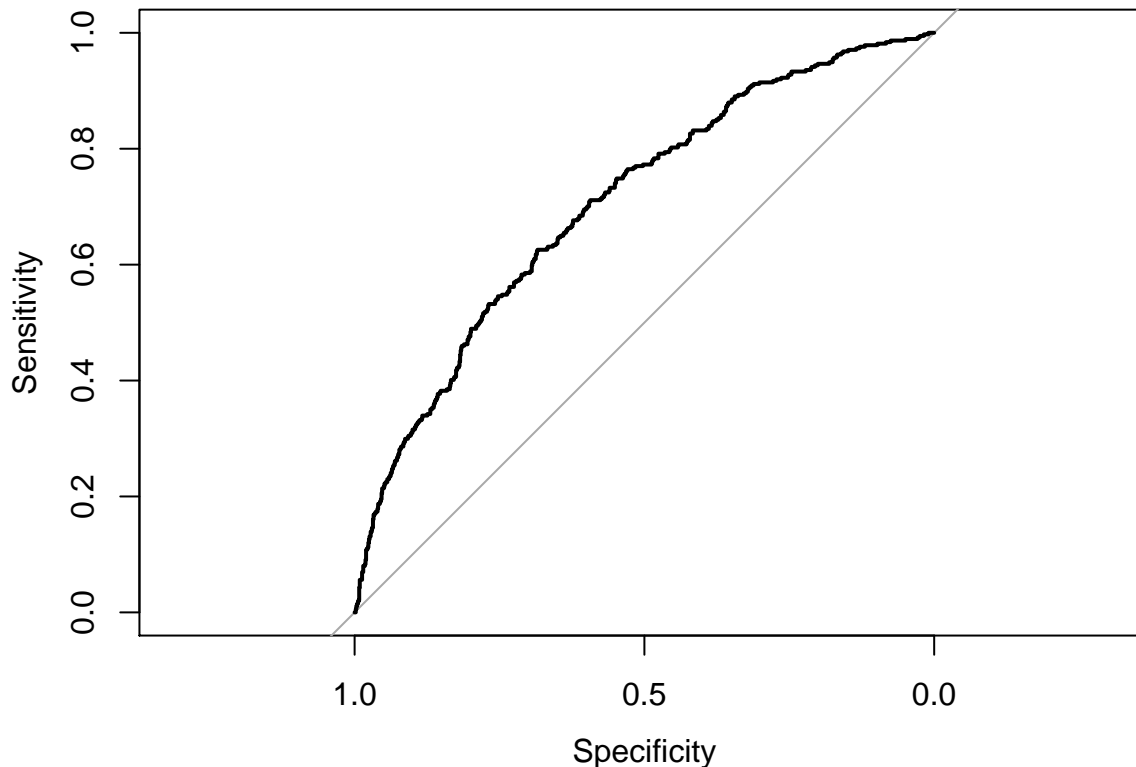


```
## [1] "AUC: 0.711161529703741"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1009 controls (dfPred_raw$class 0) < 363 cases (dfPred_raw$class 1).
## Area under the curve: 0.7112
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7872  -0.7640  -0.5698   0.8744   2.6080
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.568e-01  2.381e-01  -1.919  0.055047 .
## KIDSDRIV     2.890e-01  7.000e-02   4.129  3.65e-05 ***
## AGE          -1.272e-02  4.611e-03  -2.759  0.005801 **
## HOMEKIDS      4.513e-02  3.837e-02   1.176  0.239557
## YOJ           -4.363e-03  8.755e-03  -0.498  0.618218
## INCOME        3.181e-07  1.047e-06   0.304  0.761251
## HOME_VAL      -2.747e-06  3.320e-07  -8.274  < 2e-16 ***
## TRAVTIME      8.218e-03  2.161e-03   3.804  0.000143 ***
## BLUEBOOK      -8.137e-06  4.672e-06  -1.742  0.081547 .
```

```

## TIF          -4.070e-02  8.491e-03  -4.793  1.64e-06 ***
## OLDCLAIM     1.973e-06  4.025e-06   0.490  0.623894
## CLM_FREQ     2.870e-01  3.225e-02   8.897  < 2e-16 ***
## MVR_PTS      1.536e-01  1.597e-02   9.617  < 2e-16 ***
## CAR_AGE      -2.473e-02  6.733e-03  -3.673  0.000240 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5829.0  on 5062  degrees of freedom
## Residual deviance: 5211.8  on 5049  degrees of freedom
##    (1349 observations deleted due to missingness)
## AIC: 5239.8
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  953 288
##           1   58  86
##
##           Accuracy : 0.7502
##           95% CI : (0.7265, 0.7728)
##    No Information Rate : 0.73
##    P-Value [Acc > NIR] : 0.04717
##
##           Kappa : 0.214
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9426
##           Specificity : 0.2299
##           Pos Pred Value : 0.7679
##           Neg Pred Value : 0.5972
##           Prevalence : 0.7300
##           Detection Rate : 0.6881
##           Detection Prevalence : 0.8960
##           Balanced Accuracy : 0.5863
##
##           'Positive' Class : 0
##

```



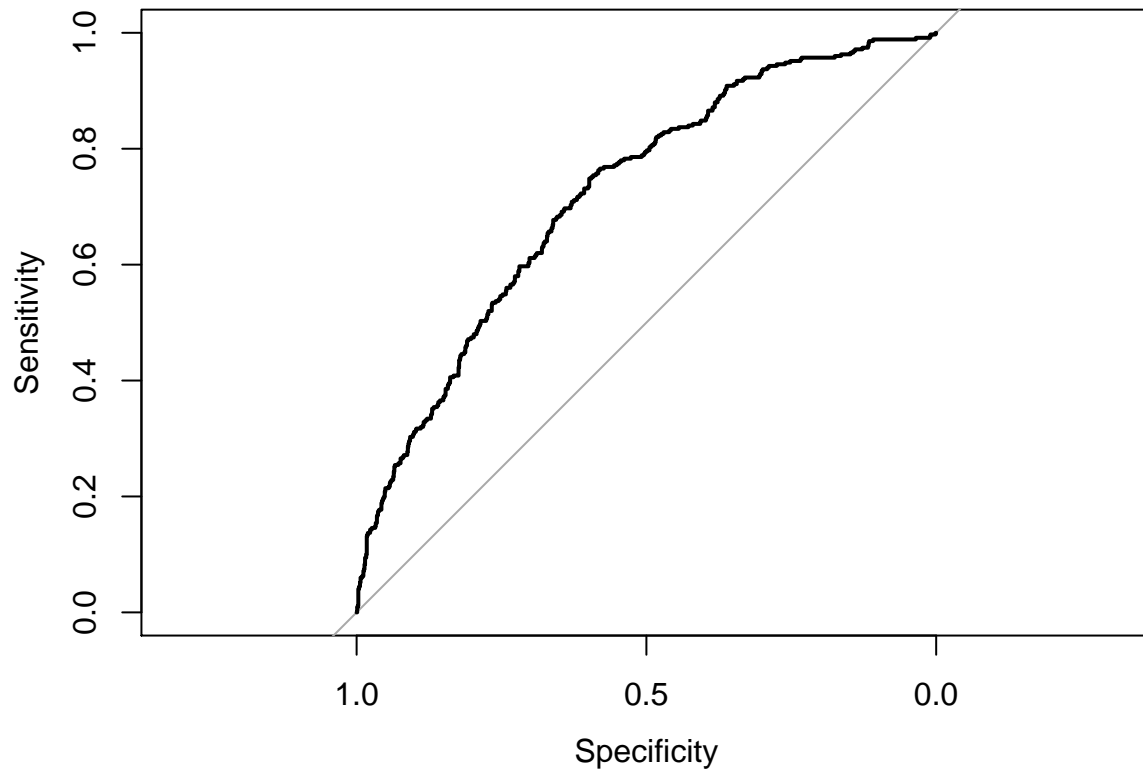
```
## [1] "AUC: 0.70572631534405"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1011 controls (dfPred_raw$class 0) < 374 cases (dfPred_raw$class 1).
## Area under the curve: 0.7057
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7954  -0.7759  -0.5801   0.9403   2.6804
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.539e-01  2.353e-01  -1.504  0.132571
## KIDSDRIV     3.129e-01  6.909e-02   4.529 5.91e-06 ***
## AGE         -1.331e-02  4.572e-03  -2.911 0.003606 **
## HOMEKIDS     5.497e-02  3.749e-02   1.466 0.142575
## YOJ         -4.589e-03  8.675e-03  -0.529 0.596850
## INCOME      -1.166e-06  1.040e-06  -1.121 0.262180
## HOME_VAL    -2.560e-06  3.313e-07  -7.726 1.11e-14 ***
## TRAVTIME     8.040e-03  2.143e-03   3.752 0.000175 ***
## BLUEBOOK    -8.899e-06  4.615e-06  -1.928 0.053811 .
```

```

## TIF          -4.152e-02  8.508e-03  -4.880 1.06e-06 ***
## OLDCLAIM     5.812e-06  3.895e-06   1.492 0.135667
## CLM_FREQ     2.645e-01  3.216e-02   8.224 < 2e-16 ***
## MVR_PTS      1.348e-01  1.573e-02   8.569 < 2e-16 ***
## CAR_AGE      -1.863e-02  6.694e-03  -2.783 0.005378 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5881.7  on 5068  degrees of freedom
## Residual deviance: 5278.2  on 5055  degrees of freedom
## (1343 observations deleted due to missingness)
## AIC: 5306.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 980 281
##           1  49  69
##
##           Accuracy : 0.7607
##           95% CI : (0.7373, 0.783)
##    No Information Rate : 0.7462
##    P-Value [Acc > NIR] : 0.1133
##
##           Kappa : 0.1914
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9524
##           Specificity : 0.1971
##           Pos Pred Value : 0.7772
##           Neg Pred Value : 0.5847
##           Prevalence : 0.7462
##           Detection Rate : 0.7107
##           Detection Prevalence : 0.9144
##           Balanced Accuracy : 0.5748
##
##           'Positive' Class : 0
##

```



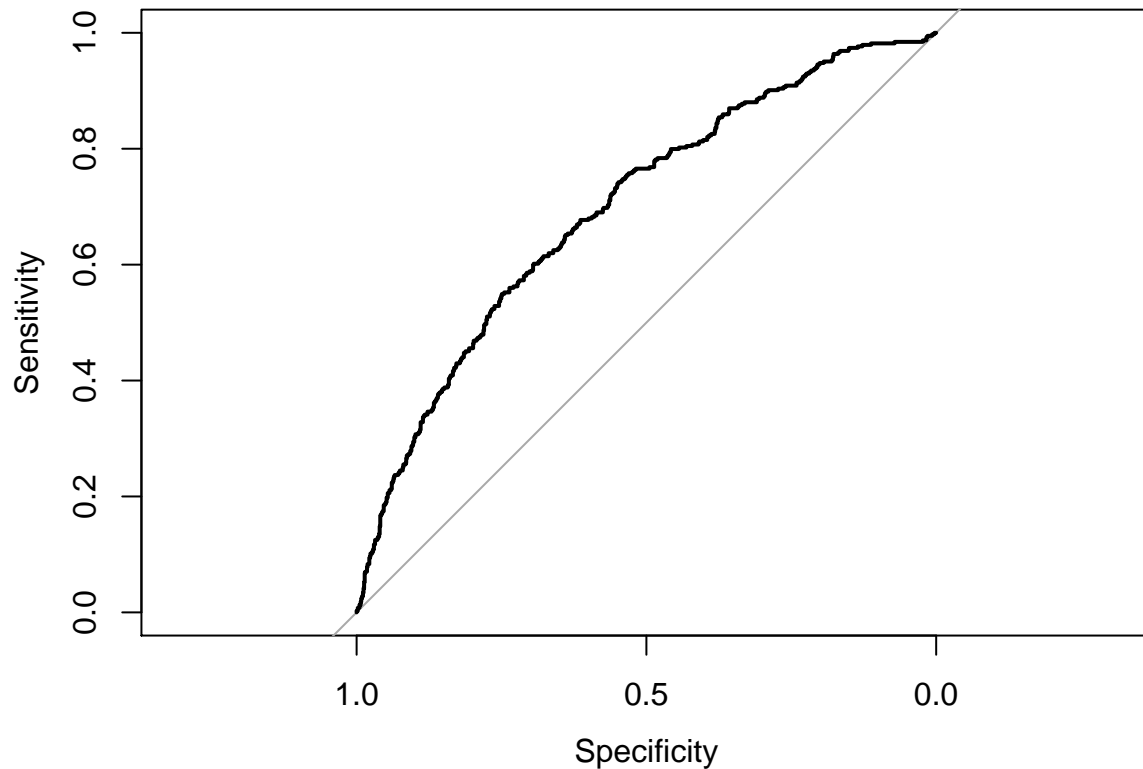


```
## [1] "AUC: 0.720835762876579"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1029 controls (dfPred_raw$class 0) < 350 cases (dfPred_raw$class 1).
## Area under the curve: 0.7208
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0490  -0.7613  -0.5624   0.8621   2.7435
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.456e-01  2.415e-01  -2.259 0.023885 *
## KIDSDRIV     2.515e-01  6.998e-02   3.593 0.000326 ***
## AGE          -8.133e-03  4.607e-03  -1.765 0.077500 .
## HOMEKIDS      7.823e-02  3.838e-02   2.038 0.041514 *
## YOJ          -7.537e-04  8.822e-03  -0.085 0.931918
## INCOME        -4.385e-07  1.054e-06  -0.416 0.677365
## HOME_VAL      -2.458e-06  3.343e-07  -7.353 1.94e-13 ***
## TRAVTIME      7.364e-03  2.143e-03   3.436 0.000591 ***
## BLUEBOOK     -1.605e-05  4.749e-06  -3.380 0.000726 ***
```

```

## TIF          -4.428e-02  8.645e-03  -5.122  3.03e-07 ***
## OLDCLAIM      9.444e-06  4.056e-06   2.328  0.019901 *
## CLM_FREQ      2.824e-01  3.249e-02   8.692  < 2e-16 ***
## MVR_PTS       1.414e-01  1.609e-02   8.788  < 2e-16 ***
## CAR_AGE       -2.736e-02  6.885e-03  -3.974  7.06e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5805.8  on 5058  degrees of freedom
## Residual deviance: 5171.6  on 5045  degrees of freedom
##    (1354 observations deleted due to missingness)
## AIC: 5199.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  946  303
##           1   59   81
##
##               Accuracy : 0.7394
##               95% CI : (0.7154, 0.7623)
##    No Information Rate : 0.7235
##    P-Value [Acc > NIR] : 0.09801
##
##               Kappa : 0.1894
##
## Mcnemar's Test P-Value : < 2e-16
##
##               Sensitivity : 0.9413
##               Specificity : 0.2109
##               Pos Pred Value : 0.7574
##               Neg Pred Value : 0.5786
##               Prevalence : 0.7235
##               Detection Rate : 0.6811
##    Detection Prevalence : 0.8992
##    Balanced Accuracy : 0.5761
##
## 'Positive' Class : 0
##

```

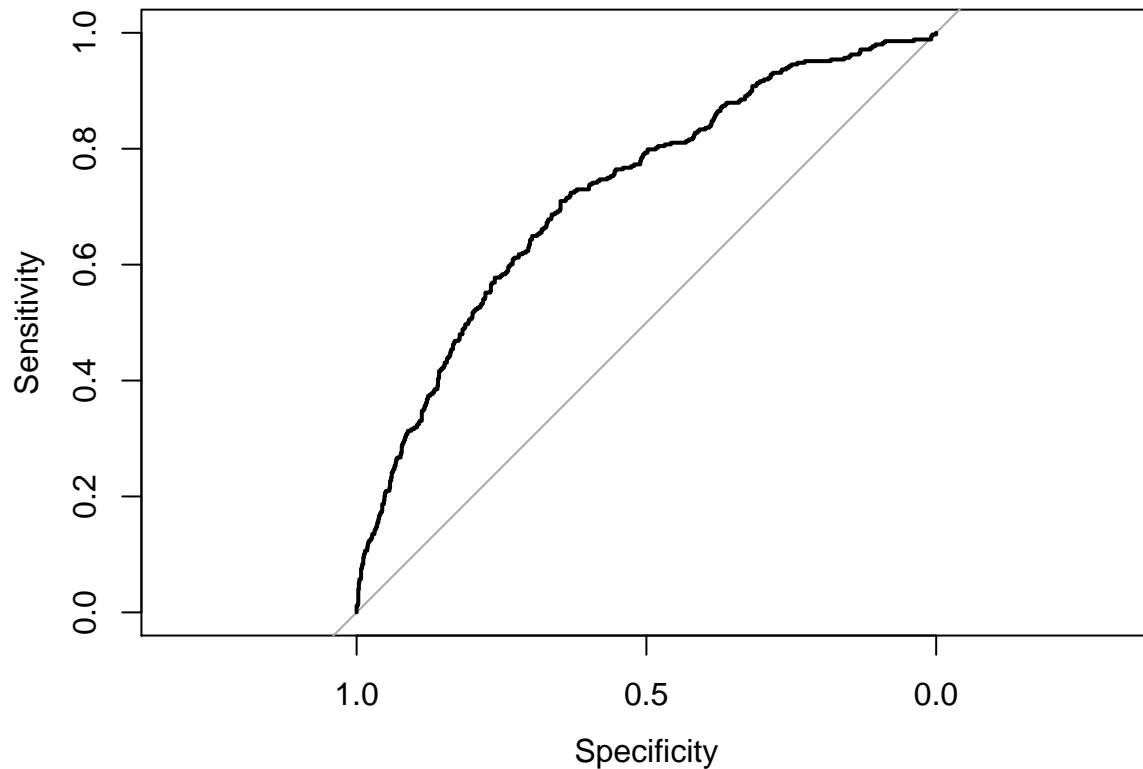


```
## [1] "AUC: 0.697641998341625"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1005 controls (dfPred_raw$class 0) < 384 cases (dfPred_raw$class 1).
## Area under the curve: 0.6976
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0113  -0.7814  -0.5803   0.9353   2.7095
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.925e-01  2.352e-01  -0.818  0.41310
## KIDSDRIV     2.819e-01  6.960e-02   4.051 5.11e-05 ***
## AGE         -1.306e-02  4.538e-03  -2.877  0.00401 **
## HOMEKIDS      5.670e-02  3.785e-02   1.498  0.13413
## YOJ          -5.274e-03  8.672e-03  -0.608  0.54312
## INCOME       -1.168e-06  1.041e-06  -1.122  0.26195
## HOME_VAL     -2.479e-06  3.300e-07  -7.512 5.83e-14 ***
## TRAVTIME      6.771e-03  2.111e-03   3.208  0.00134 **
## BLUEBOOK     -1.315e-05  4.633e-06  -2.838  0.00454 **
```

```

## TIF          -4.778e-02  8.530e-03  -5.601  2.13e-08 ***
## OLDCLAIM      8.081e-06  3.915e-06   2.064  0.03900 *
## CLM_FREQ      2.502e-01  3.240e-02   7.722  1.15e-14 ***
## MVR_PTS       1.283e-01  1.581e-02   8.115  4.84e-16 ***
## CAR_AGE       -2.038e-02  6.697e-03  -3.043  0.00234 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5880.2  on 5059  degrees of freedom
## Residual deviance: 5280.1  on 5046  degrees of freedom
##    (1352 observations deleted due to missingness)
## AIC: 5308.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 989 278
##           1  51  70
##
##           Accuracy : 0.763
##           95% CI : (0.7397, 0.7851)
##    No Information Rate : 0.7493
##    P-Value [Acc > NIR] : 0.1256
##
##           Kappa : 0.1943
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9510
##           Specificity : 0.2011
##           Pos Pred Value : 0.7806
##           Neg Pred Value : 0.5785
##           Prevalence : 0.7493
##           Detection Rate : 0.7125
##           Detection Prevalence : 0.9128
##           Balanced Accuracy : 0.5761
##
##           'Positive' Class : 0
##

```

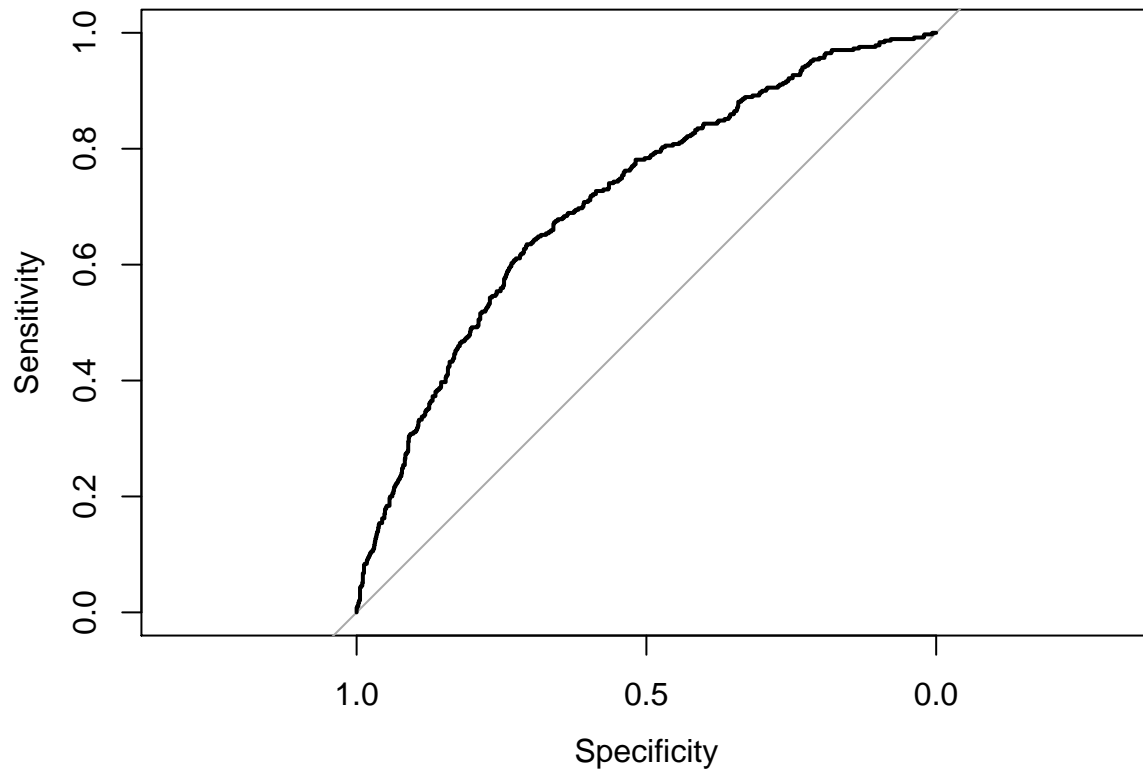


```
## [1] "AUC: 0.722673519009726"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1040 controls (dfPred_raw$class 0) < 348 cases (dfPred_raw$class 1).
## Area under the curve: 0.7227
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0156  -0.7649  -0.5732   0.8942   2.7223
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.270e-01  2.381e-01  -1.793  0.072930 .
## KIDSDRIV     2.725e-01  7.113e-02   3.830  0.000128 ***
## AGE         -1.118e-02  4.541e-03  -2.462  0.013832 *
## HOMEKIDS     4.887e-02  3.858e-02   1.267  0.205277
## YOJ          3.961e-03  8.769e-03   0.452  0.651513
## INCOME      -6.266e-07  1.041e-06  -0.602  0.547336
## HOME_VAL    -2.386e-06  3.314e-07  -7.200  6.00e-13 ***
## TRAVTIME     6.829e-03  2.127e-03   3.210  0.001326 **
## BLUEBOOK    -1.549e-05  4.725e-06  -3.279  0.001041 **
```

```

## TIF          -4.408e-02  8.536e-03  -5.164  2.42e-07 ***
## OLDCLAIM     9.684e-06  3.995e-06   2.424  0.015347 *
## CLM_FREQ     2.718e-01  3.259e-02   8.341  < 2e-16 ***
## MVR_PTS      1.369e-01  1.598e-02   8.568  < 2e-16 ***
## CAR_AGE      -2.305e-02  6.766e-03  -3.407  0.000657 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5833.0  on 5055  degrees of freedom
## Residual deviance: 5221.5  on 5042  degrees of freedom
##    (1356 observations deleted due to missingness)
## AIC: 5249.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  959  295
##           1   63   75
##
##               Accuracy : 0.7428
##               95% CI : (0.719, 0.7656)
##    No Information Rate : 0.7342
##    P-Value [Acc > NIR] : 0.2434
##
##               Kappa : 0.1763
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9384
##           Specificity : 0.2027
##           Pos Pred Value : 0.7648
##           Neg Pred Value : 0.5435
##           Prevalence : 0.7342
##           Detection Rate : 0.6889
##           Detection Prevalence : 0.9009
##           Balanced Accuracy : 0.5705
##
##           'Positive' Class : 0
##

```



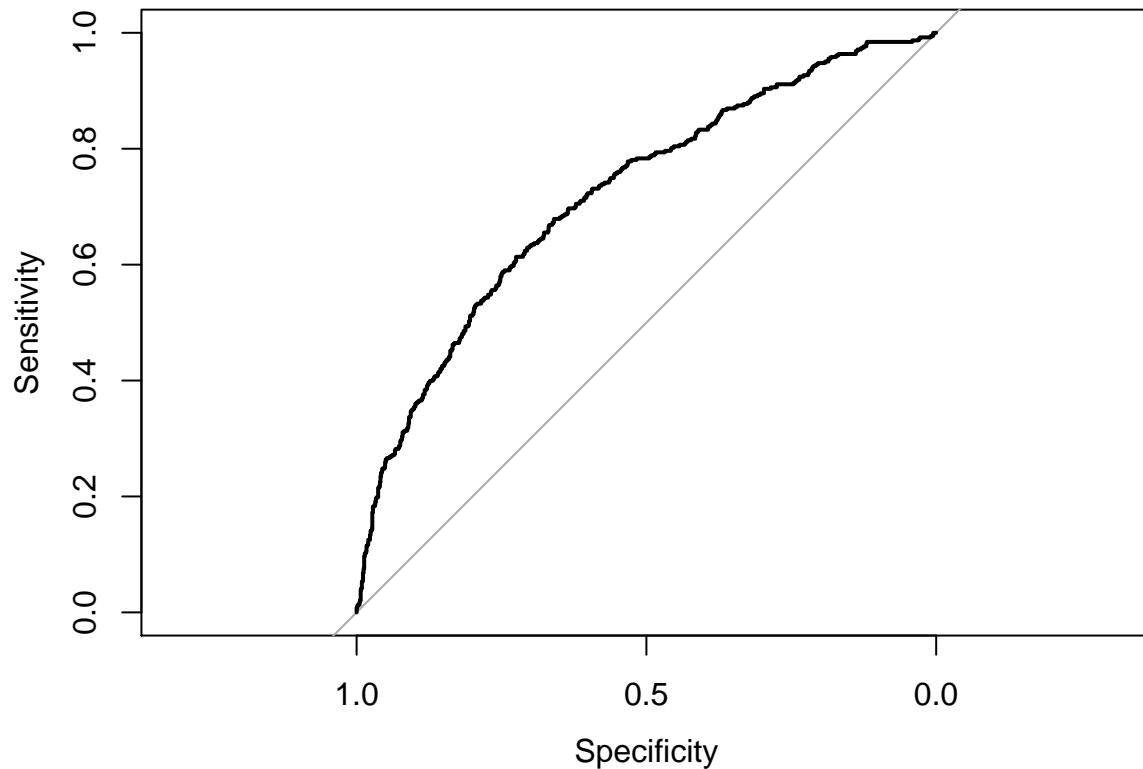
```
## [1] "AUC: 0.712601153012112"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1022 controls (dfPred_raw$class 0) < 370 cases (dfPred_raw$class 1).
## Area under the curve: 0.7126
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9883  -0.7688  -0.5717   0.8869   2.6743
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.248e-01  2.406e-01  -0.518  0.60415
## KIDSDRIV     2.751e-01  6.860e-02   4.010 6.07e-05 ***
## AGE         -1.485e-02  4.652e-03  -3.193  0.00141 **
## HOMEKIDS      5.102e-02  3.809e-02   1.339  0.18043
## YOJ          -8.593e-03  8.754e-03  -0.982  0.32633
## INCOME       -3.410e-07  1.051e-06  -0.324  0.74557
## HOME_VAL     -2.482e-06  3.321e-07  -7.473 7.83e-14 ***
## TRAVTIME      6.293e-03  2.138e-03   2.943  0.00325 **
## BLUEBOOK     -1.114e-05  4.654e-06  -2.394  0.01667 *
```

```

## TIF          -4.762e-02  8.556e-03  -5.565  2.62e-08 ***
## OLDCLAIM     4.338e-06  4.032e-06   1.076  0.28197
## CLM_FREQ     2.685e-01  3.243e-02   8.278  < 2e-16 ***
## MVR_PTS      1.338e-01  1.611e-02   8.306  < 2e-16 ***
## CAR_AGE      -2.692e-02  6.753e-03  -3.986  6.72e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5814.0  on 5068  degrees of freedom
## Residual deviance: 5223.7  on 5055  degrees of freedom
##    (1343 observations deleted due to missingness)
## AIC: 5251.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  954  292
##           1   42   91
##
##           Accuracy : 0.7578
##           95% CI : (0.7343, 0.7802)
##    No Information Rate : 0.7223
##    P-Value [Acc > NIR] : 0.001578
##
##           Kappa : 0.2446
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.9578
##           Specificity : 0.2376
##           Pos Pred Value : 0.7657
##           Neg Pred Value : 0.6842
##           Prevalence : 0.7223
##           Detection Rate : 0.6918
##           Detection Prevalence : 0.9036
##           Balanced Accuracy : 0.5977
##
##           'Positive' Class : 0
##

```



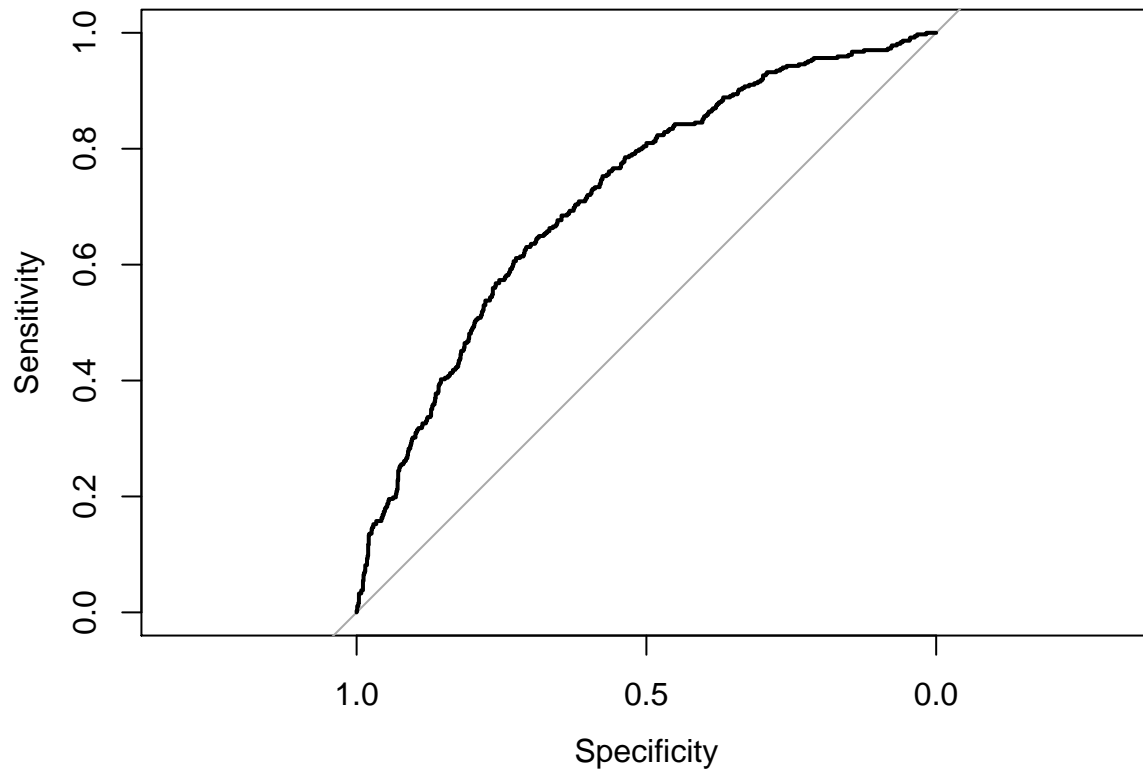


```
## [1] "AUC: 0.720532259586649"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 996 controls (dfPred_raw$class 0) < 383 cases (dfPred_raw$class 1).
## Area under the curve: 0.7205
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8990  -0.7618  -0.5739   0.9092   2.6067
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.832e-01  2.391e-01  -1.603  0.109008
## KIDSDRIV     1.450e-01  7.134e-02   2.032  0.042166 *
## AGE         -1.216e-02  4.572e-03  -2.659  0.007844 **
## HOMEKIDS      8.442e-02  3.768e-02   2.241  0.025049 *
## YOJ         -1.016e-02  8.676e-03  -1.171  0.241597
## INCOME        1.929e-07  1.046e-06   0.184  0.853720
## HOME_VAL     -2.350e-06  3.308e-07  -7.105  1.20e-12 ***
## TRAVTIME      7.750e-03  2.125e-03   3.648  0.000265 ***
## BLUEBOOK     -1.117e-05  4.641e-06  -2.408  0.016056 *
```

```

## TIF          -4.298e-02  8.638e-03  -4.976  6.49e-07  ***
## OLDCLAIM     5.189e-06  3.967e-06   1.308  0.190821
## CLM_FREQ     2.762e-01  3.215e-02   8.591  < 2e-16  ***
## MVR_PTS      1.474e-01  1.595e-02   9.236  < 2e-16  ***
## CAR_AGE      -2.543e-02  6.789e-03  -3.746  0.000180  ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5849.3  on 5075  degrees of freedom
## Residual deviance: 5245.9  on 5062  degrees of freedom
##    (1336 observations deleted due to missingness)
## AIC: 5273.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  948  297
##           1   56   71
##
##           Accuracy : 0.7427
##           95% CI : (0.7187, 0.7657)
##    No Information Rate : 0.7318
##    P-Value [Acc > NIR] : 0.1887
##
##           Kappa : 0.1731
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9442
##           Specificity : 0.1929
##           Pos Pred Value : 0.7614
##           Neg Pred Value : 0.5591
##           Prevalence : 0.7318
##           Detection Rate : 0.6910
##    Detection Prevalence : 0.9074
##           Balanced Accuracy : 0.5686
##
##           'Positive' Class : 0
##

```

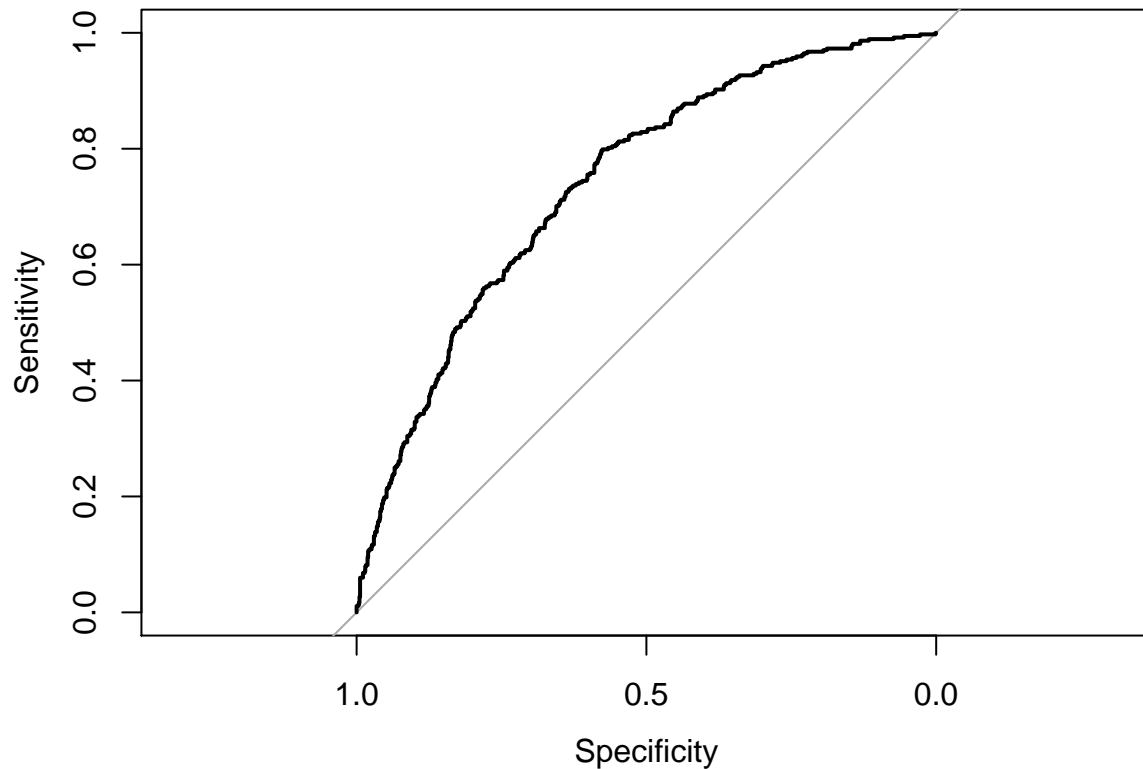


```
## [1] "AUC: 0.719583622033605"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1004 controls (dfPred_raw$class 0) < 368 cases (dfPred_raw$class 1).
## Area under the curve: 0.7196
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0695  -0.7696  -0.5815   0.9258   2.6470
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.029e-01  2.375e-01  -2.539  0.01112 *
## KIDSDRIV     2.959e-01  6.910e-02   4.282 1.85e-05 ***
## AGE         -7.679e-03  4.526e-03  -1.697  0.08978 .
## HOMEKIDS     5.599e-02  3.778e-02   1.482  0.13839
## YOJ         -3.311e-03  8.789e-03  -0.377  0.70637
## INCOME      -9.286e-07  1.045e-06  -0.888  0.37445
## HOME_VAL    -2.220e-06  3.298e-07  -6.731 1.68e-11 ***
## TRAVTIME     8.742e-03  2.119e-03   4.125 3.70e-05 ***
## BLUEBOOK    -1.209e-05  4.659e-06  -2.595  0.00946 **
```

```

## TIF          -4.948e-02  8.621e-03  -5.739  9.52e-09 ***
## OLDCLAIM      9.826e-06  3.878e-06   2.534  0.01129 *
## CLM_FREQ      2.583e-01  3.226e-02   8.006  1.18e-15 ***
## MVR_PTS       1.274e-01  1.573e-02   8.096  5.68e-16 ***
## CAR_AGE       -2.104e-02  6.710e-03  -3.135  0.00172 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5847.5  on 5072  degrees of freedom
## Residual deviance: 5274.9  on 5059  degrees of freedom
##    (1339 observations deleted due to missingness)
## AIC: 5302.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 963 299
##           1  44  69
##
##           Accuracy : 0.7505
##           95% CI : (0.7268, 0.7732)
##    No Information Rate : 0.7324
##    P-Value [Acc > NIR] : 0.06701
##
##           Kappa : 0.1843
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9563
##           Specificity : 0.1875
##           Pos Pred Value : 0.7631
##           Neg Pred Value : 0.6106
##           Prevalence : 0.7324
##           Detection Rate : 0.7004
##           Detection Prevalence : 0.9178
##           Balanced Accuracy : 0.5719
##
##           'Positive' Class : 0
##

```

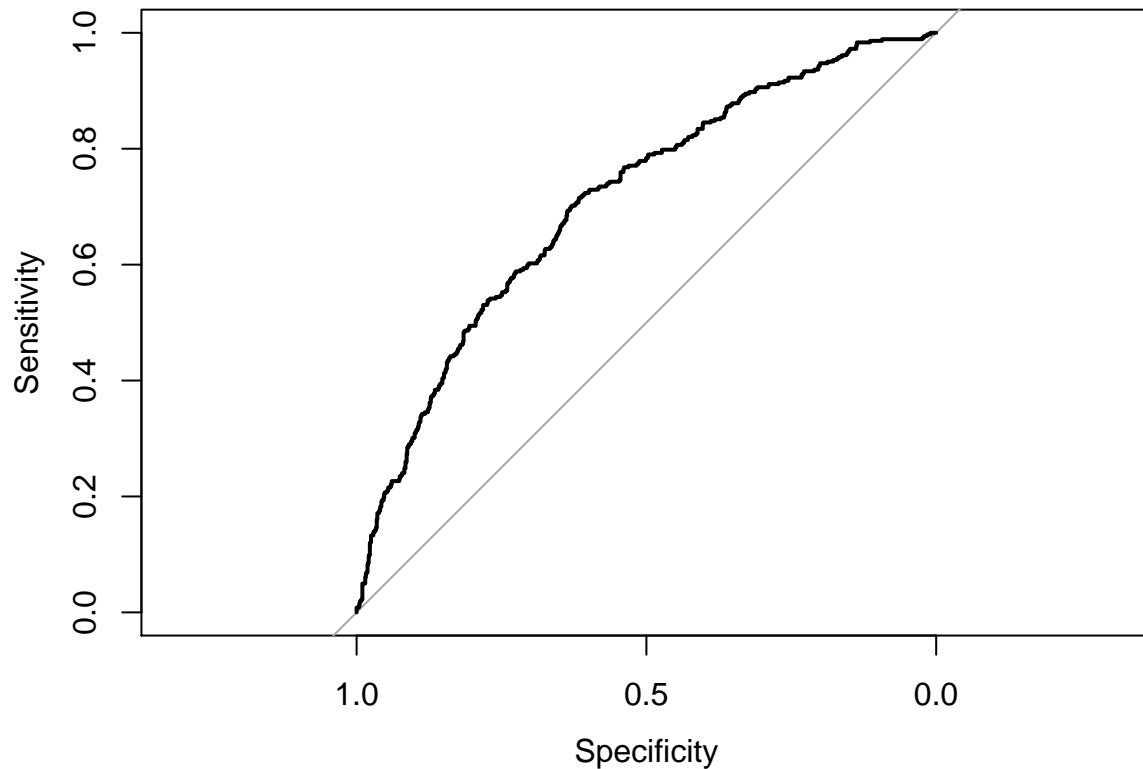


```
## [1] "AUC: 0.740423017140883"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1007 controls (dfPred_raw$class 0) < 368 cases (dfPred_raw$class 1).
## Area under the curve: 0.7404
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0841  -0.7707  -0.5707   0.9017   2.7026
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.736e-01  2.399e-01  -1.974 0.048328 *
## KIDSDRIV     2.530e-01  6.984e-02   3.622 0.000292 ***
## AGE         -7.837e-03  4.579e-03  -1.712 0.086976 .
## HOMEKIDS     8.085e-02  3.769e-02   2.145 0.031947 *
## YOJ         -7.158e-03  8.837e-03  -0.810 0.417928
## INCOME      -1.402e-07  1.043e-06  -0.134 0.893032
## HOME_VAL    -2.467e-06  3.297e-07  -7.483 7.26e-14 ***
## TRAVTIME     9.000e-03  2.127e-03   4.231 2.33e-05 ***
## BLUEBOOK    -1.681e-05  4.685e-06  -3.588 0.000333 ***
```

```

## TIF          -5.135e-02  8.678e-03  -5.917  3.27e-09 ***
## OLDCLAIM      7.733e-06  3.916e-06   1.975  0.048315 *
## CLM_FREQ      2.741e-01  3.240e-02   8.461  < 2e-16 ***
## MVR_PTS       1.299e-01  1.603e-02   8.106  5.25e-16 ***
## CAR_AGE       -2.501e-02  6.767e-03  -3.696  0.000219 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5862.9  on 5077  degrees of freedom
## Residual deviance: 5245.9  on 5064  degrees of freedom
##    (1334 observations deleted due to missingness)
## AIC: 5273.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  946  280
##           1   62   82
##
##           Accuracy : 0.7504
##           95% CI : (0.7266, 0.7731)
##    No Information Rate : 0.7358
##    P-Value [Acc > NIR] : 0.1157
##
##           Kappa : 0.2045
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9385
##           Specificity : 0.2265
##           Pos Pred Value : 0.7716
##           Neg Pred Value : 0.5694
##           Prevalence : 0.7358
##           Detection Rate : 0.6905
##           Detection Prevalence : 0.8949
##           Balanced Accuracy : 0.5825
##
##           'Positive' Class : 0
##

```



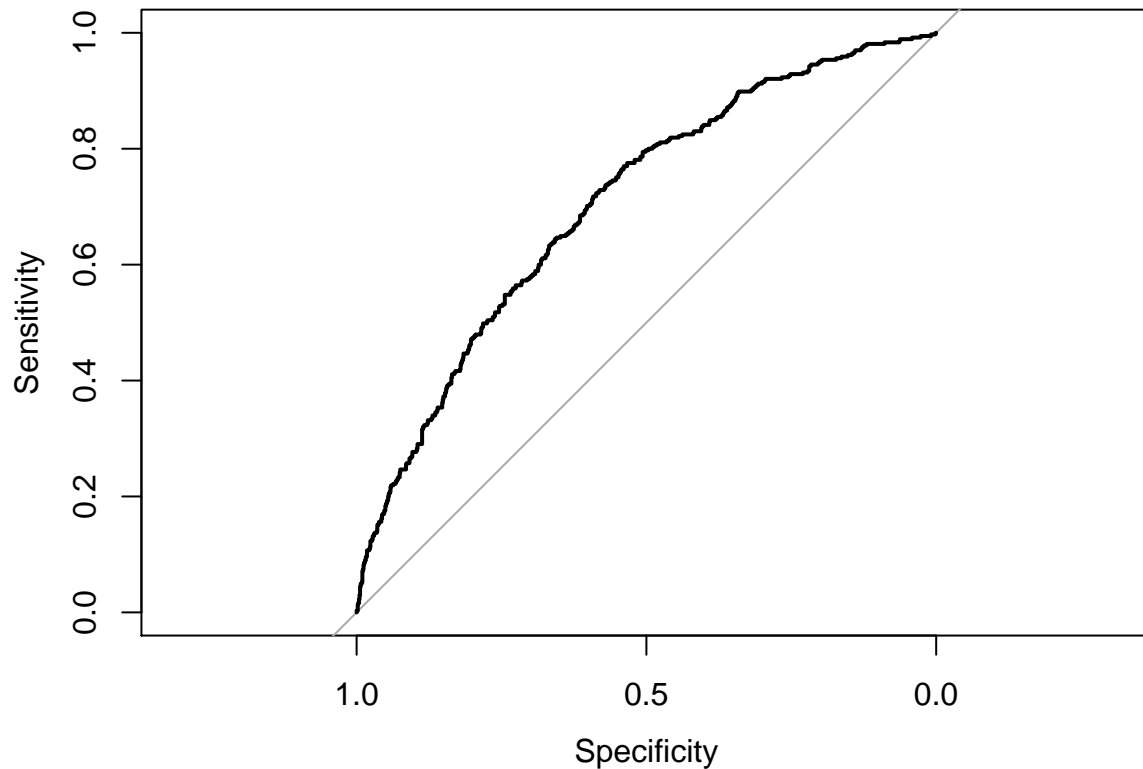
```
## [1] "AUC: 0.711394479522933"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1008 controls (dfPred_raw$class 0) < 362 cases (dfPred_raw$class 1).
## Area under the curve: 0.7114
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9802  -0.7637  -0.5671   0.9051   2.6312
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.162e-01  2.393e-01  -1.321  0.186401
## KIDSDRIV     2.098e-01  6.913e-02   3.035  0.002402 **
## AGE         -1.176e-02  4.601e-03  -2.555  0.010605 *
## HOMEKIDS      8.224e-02  3.764e-02   2.185  0.028912 *
## YOJ          -1.299e-02  8.727e-03  -1.489  0.136581
## INCOME        1.596e-07  1.037e-06   0.154  0.877724
## HOME_VAL     -2.434e-06  3.299e-07  -7.378  1.60e-13 ***
## TRAVTIME      7.691e-03  2.130e-03   3.610  0.000306 ***
## BLUEBOOK     -1.413e-05  4.664e-06  -3.030  0.002446 **
```

```

## TIF          -4.895e-02  8.650e-03  -5.659 1.53e-08 ***
## OLDCLAIM      4.961e-06  3.954e-06   1.255 0.209643
## CLM_FREQ      2.863e-01  3.201e-02   8.945 < 2e-16 ***
## MVR_PTS       1.392e-01  1.582e-02   8.803 < 2e-16 ***
## CAR_AGE       -2.227e-02  6.770e-03  -3.289 0.001004 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5854.3  on 5073  degrees of freedom
## Residual deviance: 5227.8  on 5060  degrees of freedom
## (1338 observations deleted due to missingness)
## AIC: 5255.8
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 958 297
##           1  51  68
##
##           Accuracy : 0.7467
##           95% CI : (0.7229, 0.7695)
##    No Information Rate : 0.7344
##    P-Value [Acc > NIR] : 0.1567
##
##           Kappa : 0.173
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9495
##           Specificity : 0.1863
##           Pos Pred Value : 0.7633
##           Neg Pred Value : 0.5714
##           Prevalence : 0.7344
##           Detection Rate : 0.6972
##           Detection Prevalence : 0.9134
##           Balanced Accuracy : 0.5679
##
##           'Positive' Class : 0
##

```



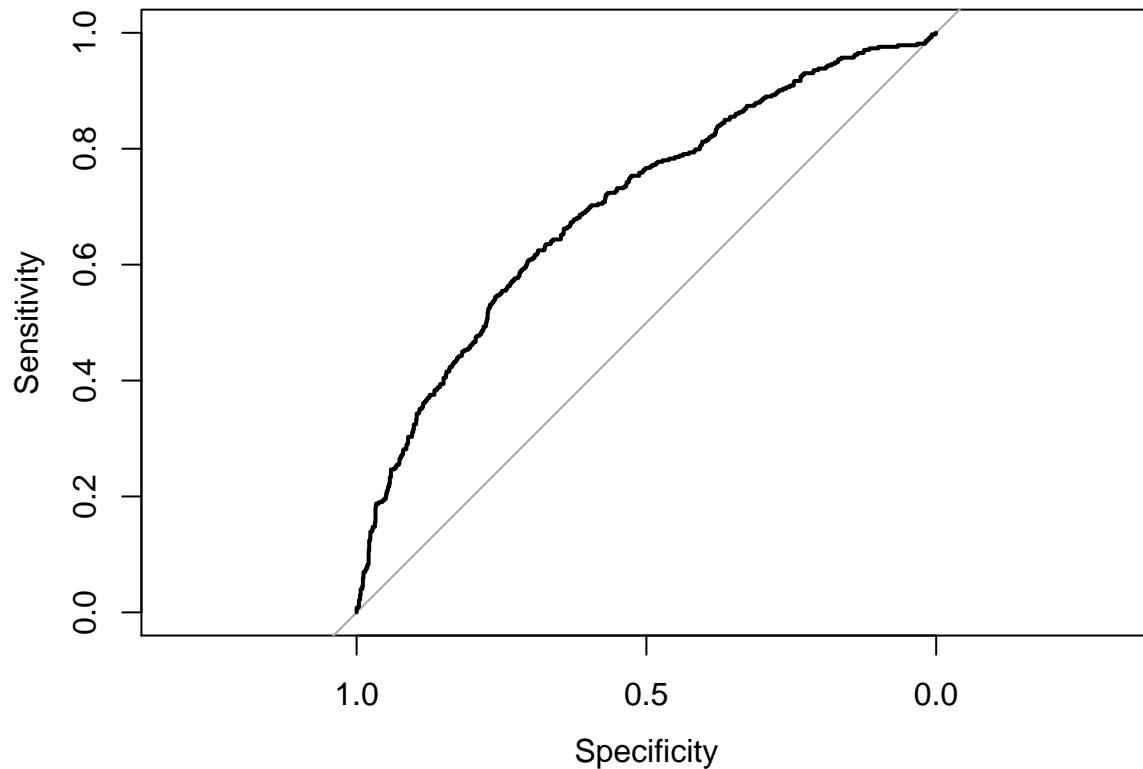


```
## [1] "AUC: 0.704728674803481"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1009 controls (dfPred_raw$class 0) < 365 cases (dfPred_raw$class 1).
## Area under the curve: 0.7047
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0153  -0.7684  -0.5685   0.8926   2.7669
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.012e-01  2.398e-01  -1.256  0.209117
## KIDSDRIV     2.418e-01  6.957e-02   3.476  0.000510 ***
## AGE         -1.082e-02  4.621e-03  -2.341  0.019235 *
## HOMEKIDS      7.654e-02  3.798e-02   2.015  0.043885 *
## YOJ          -4.869e-03  8.763e-03  -0.556  0.578409
## INCOME       -6.805e-07  1.053e-06  -0.646  0.518042
## HOME_VAL    -2.584e-06  3.336e-07  -7.744  9.64e-15 ***
## TRAVTIME      7.051e-03  2.141e-03   3.293  0.000991 ***
## BLUEBOOK    -1.514e-05  4.686e-06  -3.230  0.001237 **
```

```

## TIF          -4.480e-02  8.618e-03  -5.199 2.01e-07 ***
## OLDCLAIM      8.140e-06  4.045e-06   2.013 0.044167 *
## CLM_FREQ      2.529e-01  3.250e-02   7.782 7.11e-15 ***
## MVR_PTS       1.428e-01  1.610e-02   8.873 < 2e-16 ***
## CAR_AGE       -2.823e-02  6.864e-03  -4.113 3.91e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5830.5  on 5061  degrees of freedom
## Residual deviance: 5202.1  on 5048  degrees of freedom
## (1351 observations deleted due to missingness)
## AIC: 5230.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 960 295
##           1  53  78
##
##           Accuracy : 0.7489
##           95% CI : (0.7252, 0.7716)
##    No Information Rate : 0.7309
##    P-Value [Acc > NIR] : 0.06815
##
##           Kappa : 0.1972
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9477
##           Specificity : 0.2091
##           Pos Pred Value : 0.7649
##           Neg Pred Value : 0.5954
##           Prevalence : 0.7309
##           Detection Rate : 0.6926
##           Detection Prevalence : 0.9055
##           Balanced Accuracy : 0.5784
##
##           'Positive' Class : 0
##

```

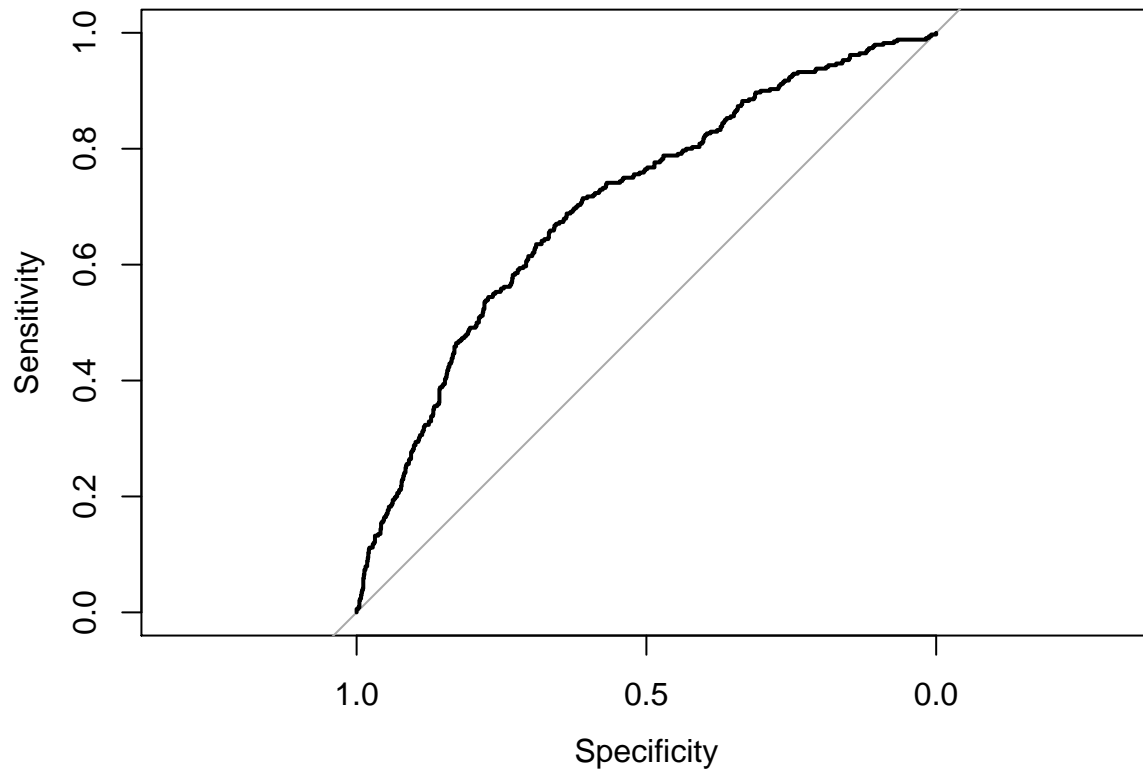


```
## [1] "AUC: 0.700727009996057"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1013 controls (dfPred_raw$class 0) < 373 cases (dfPred_raw$class 1).
## Area under the curve: 0.7007
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0767  -0.7742  -0.5712   0.9259   2.7013
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.922e-01  2.359e-01  -1.663 0.096389 .
## KIDSDRIV     2.639e-01  7.103e-02   3.716 0.000203 ***
## AGE         -9.782e-03  4.530e-03  -2.160 0.030807 *
## HOMEKIDS     8.306e-02  3.806e-02   2.183 0.029065 *
## YOJ         -7.352e-03  8.708e-03  -0.844 0.398514
## INCOME      -5.042e-07  1.042e-06  -0.484 0.628605
## HOME_VAL    -2.575e-06  3.296e-07  -7.813 5.58e-15 ***
## TRAVTIME     7.932e-03  2.116e-03   3.749 0.000178 ***
## BLUEBOOK    -1.542e-05  4.688e-06  -3.290 0.001001 **
```

```

## TIF          -4.837e-02  8.566e-03  -5.647 1.63e-08 ***
## OLDCLAIM     6.603e-06  3.887e-06   1.699 0.089349 .
## CLM_FREQ     2.877e-01  3.235e-02   8.892 < 2e-16 ***
## MVR_PTS      1.309e-01  1.588e-02   8.240 < 2e-16 ***
## CAR_AGE      -1.956e-02  6.717e-03  -2.912 0.003588 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5899.3  on 5064  degrees of freedom
## Residual deviance: 5260.2  on 5051  degrees of freedom
##    (1347 observations deleted due to missingness)
## AIC: 5288.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 972 272
##           1  71  68
##
##           Accuracy : 0.752
##           95% CI : (0.7283, 0.7746)
##    No Information Rate : 0.7542
##    P-Value [Acc > NIR] : 0.5885
##
##           Kappa : 0.1648
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9319
##           Specificity : 0.2000
##           Pos Pred Value : 0.7814
##           Neg Pred Value : 0.4892
##           Prevalence : 0.7542
##           Detection Rate : 0.7028
##           Detection Prevalence : 0.8995
##           Balanced Accuracy : 0.5660
##
##           'Positive' Class : 0
##

```

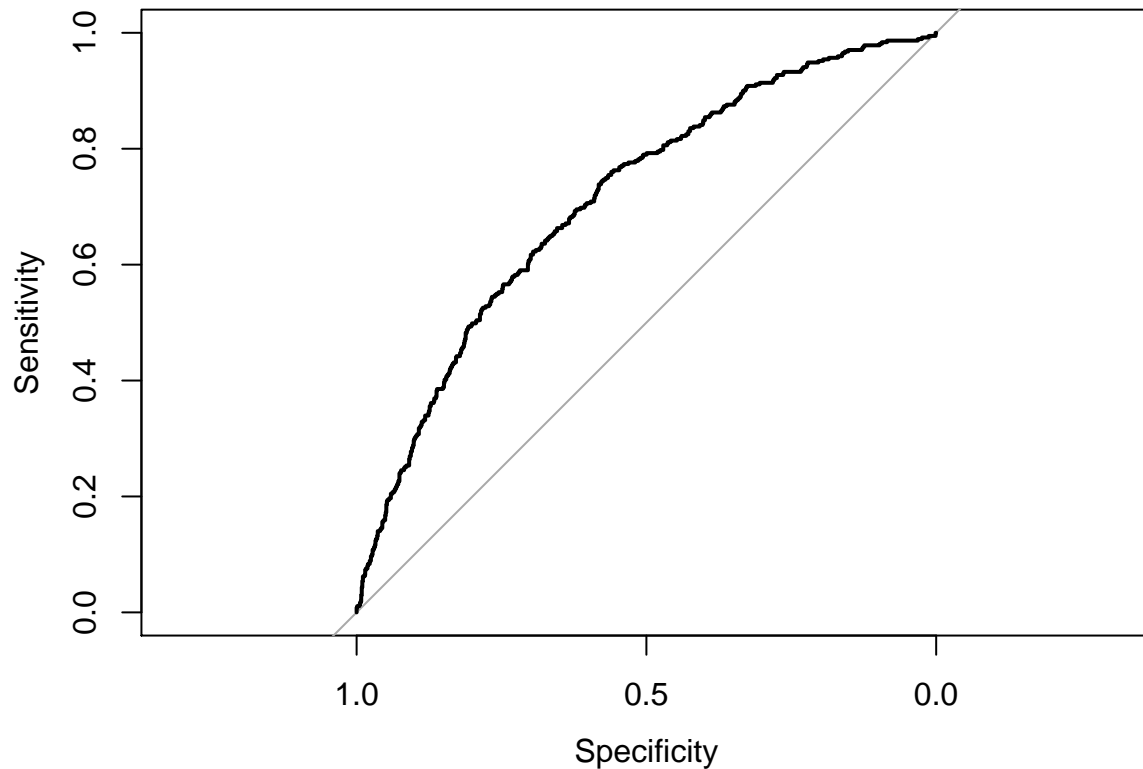


```
## [1] "AUC: 0.703538999492414"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1043 controls (dfPred_raw$class 0) < 340 cases (dfPred_raw$class 1).
## Area under the curve: 0.7035
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0653  -0.7630  -0.5708   0.8914   2.5198
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.390e-01  2.383e-01  -2.681 0.007345 **
## KIDSDRIV     2.670e-01  7.016e-02   3.806 0.000141 ***
## AGE        -7.204e-03  4.561e-03  -1.580 0.114199
## HOMEKIDS     7.863e-02  3.827e-02   2.055 0.039898 *
## YOJ        -3.231e-03  8.752e-03  -0.369 0.711988
## INCOME      -8.132e-07  1.057e-06  -0.769 0.441627
## HOME_VAL    -2.434e-06  3.333e-07  -7.301 2.86e-13 ***
## TRAVTIME     7.914e-03  2.136e-03   3.704 0.000212 ***
## BLUEBOOK    -1.267e-05  4.717e-06  -2.687 0.007214 **
```

```

## TIF          -4.366e-02  8.606e-03  -5.073  3.92e-07 ***
## OLDCLAIM      8.716e-06  3.971e-06   2.195  0.028179 *
## CLM_FREQ      2.744e-01  3.232e-02   8.491  < 2e-16 ***
## MVR_PTS       1.434e-01  1.586e-02   9.042  < 2e-16 ***
## CAR_AGE       -2.298e-02  6.821e-03  -3.369  0.000754 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5834.6  on 5061  degrees of freedom
## Residual deviance: 5213.2  on 5048  degrees of freedom
##    (1351 observations deleted due to missingness)
## AIC: 5241.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 962 302
##           1  53  69
##
##               Accuracy : 0.7439
##               95% CI : (0.72, 0.7667)
##    No Information Rate : 0.7323
##    P-Value [Acc > NIR] : 0.1737
##
##               Kappa : 0.17
##
## Mcnemar's Test P-Value : <2e-16
##
##               Sensitivity : 0.9478
##               Specificity : 0.1860
##               Pos Pred Value : 0.7611
##               Neg Pred Value : 0.5656
##               Prevalence : 0.7323
##               Detection Rate : 0.6941
##    Detection Prevalence : 0.9120
##    Balanced Accuracy : 0.5669
##
##    'Positive' Class : 0
##

```



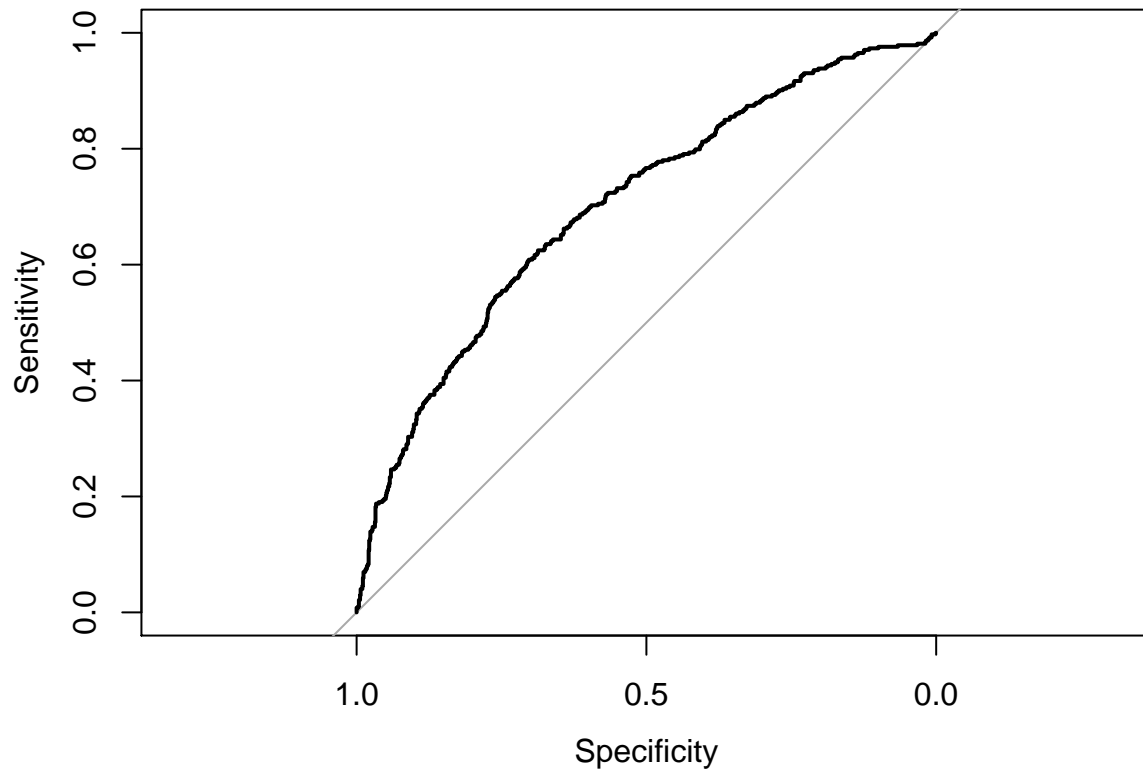
```
## [1] "AUC: 0.711900999827387"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1015 controls (dfPred_raw$class 0) < 371 cases (dfPred_raw$class 1).
## Area under the curve: 0.7119
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0153  -0.7684  -0.5685   0.8926   2.7669
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.012e-01  2.398e-01  -1.256  0.209117
## KIDSDRIV     2.418e-01  6.957e-02   3.476  0.000510 ***
## AGE         -1.082e-02  4.621e-03  -2.341  0.019235 *
## HOMEKIDS      7.654e-02  3.798e-02   2.015  0.043885 *
## YOJ          -4.869e-03  8.763e-03  -0.556  0.578409
## INCOME       -6.805e-07  1.053e-06  -0.646  0.518042
## HOME_VAL    -2.584e-06  3.336e-07  -7.744  9.64e-15 ***
## TRAVTIME      7.051e-03  2.141e-03   3.293  0.000991 ***
## BLUEBOOK    -1.514e-05  4.686e-06  -3.230  0.001237 **
```

```

## TIF          -4.480e-02  8.618e-03  -5.199 2.01e-07 ***
## OLDCLAIM      8.140e-06  4.045e-06   2.013 0.044167 *
## CLM_FREQ      2.529e-01  3.250e-02   7.782 7.11e-15 ***
## MVR_PTS       1.428e-01  1.610e-02   8.873 < 2e-16 ***
## CAR_AGE       -2.823e-02  6.864e-03  -4.113 3.91e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5830.5  on 5061  degrees of freedom
## Residual deviance: 5202.1  on 5048  degrees of freedom
##    (1351 observations deleted due to missingness)
## AIC: 5230.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 960 295
##           1  53  78
##
##           Accuracy : 0.7489
##           95% CI : (0.7252, 0.7716)
##    No Information Rate : 0.7309
##    P-Value [Acc > NIR] : 0.06815
##
##           Kappa : 0.1972
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9477
##           Specificity : 0.2091
##           Pos Pred Value : 0.7649
##           Neg Pred Value : 0.5954
##           Prevalence : 0.7309
##           Detection Rate : 0.6926
##           Detection Prevalence : 0.9055
##           Balanced Accuracy : 0.5784
##
##           'Positive' Class : 0
##

```



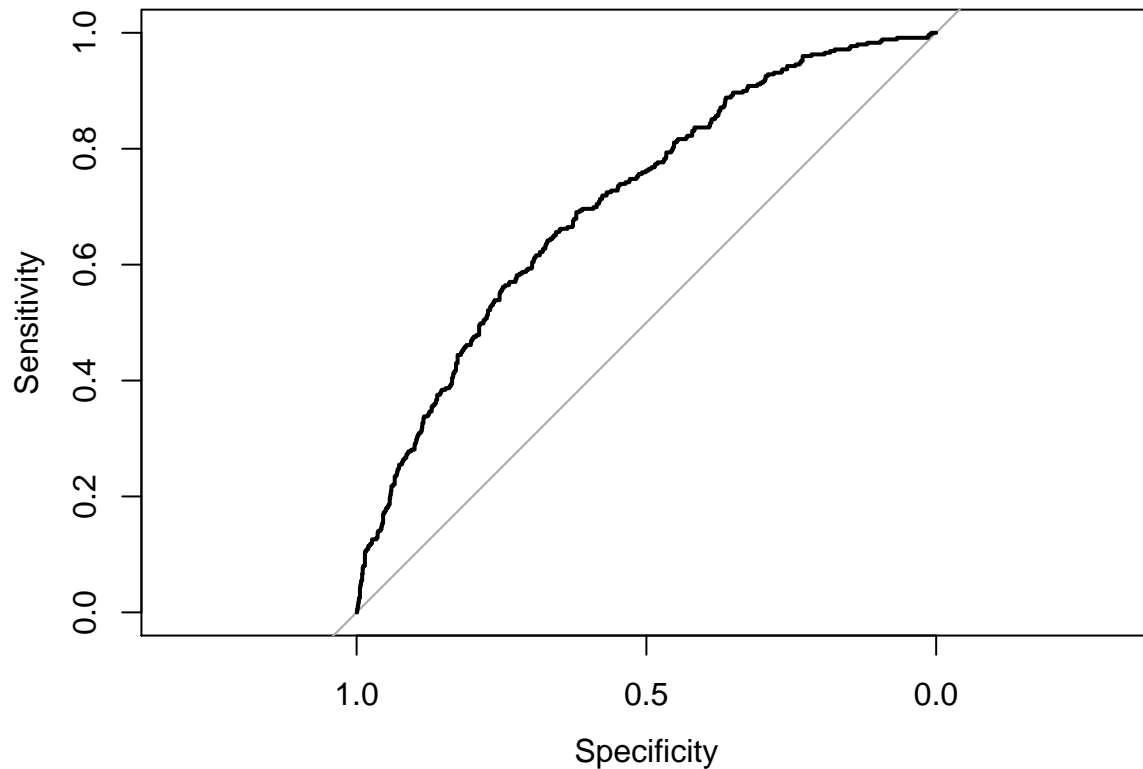


```
## [1] "AUC: 0.700727009996057"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1013 controls (dfPred_raw$class 0) < 373 cases (dfPred_raw$class 1).
## Area under the curve: 0.7007
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7696  -0.7773  -0.5742   0.9371   2.7093
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.195e-01  2.348e-01  -1.786  0.07405 .
## KIDSDRIV     2.856e-01  7.033e-02   4.061 4.90e-05 ***
## AGE         -1.094e-02  4.527e-03  -2.417  0.01567 *
## HOMEKIDS      6.429e-02  3.782e-02   1.700  0.08917 .
## YOJ          -2.648e-03  8.701e-03  -0.304  0.76088
## INCOME       -6.976e-08  1.043e-06  -0.067  0.94669
## HOME_VAL     -2.746e-06  3.296e-07  -8.331 < 2e-16 ***
## TRAVTIME      7.810e-03  2.121e-03   3.683  0.00023 ***
## BLUEBOOK     -1.412e-05  4.664e-06  -3.027  0.00247 **
```

```

## TIF          -4.124e-02  8.512e-03  -4.844  1.27e-06 ***
## OLDCLAIM     6.914e-06  3.910e-06   1.768  0.07703 .
## CLM_FREQ     2.615e-01  3.246e-02   8.054  7.99e-16 ***
## MVR_PTS      1.396e-01  1.578e-02   8.851  < 2e-16 ***
## CAR_AGE      -2.151e-02  6.704e-03  -3.209  0.00133 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5871.3  on 5048  degrees of freedom
## Residual deviance: 5249.0  on 5035  degrees of freedom
## (1363 observations deleted due to missingness)
## AIC: 5277
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 991 284
##           1  59  65
##
##           Accuracy : 0.7548
##           95% CI : (0.7314, 0.7772)
##    No Information Rate : 0.7505
##    P-Value [Acc > NIR] : 0.3687
##
##           Kappa : 0.1657
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9438
##           Specificity : 0.1862
##           Pos Pred Value : 0.7773
##           Neg Pred Value : 0.5242
##           Prevalence : 0.7505
##           Detection Rate : 0.7084
##           Detection Prevalence : 0.9114
##           Balanced Accuracy : 0.5650
##
##           'Positive' Class : 0
##

```

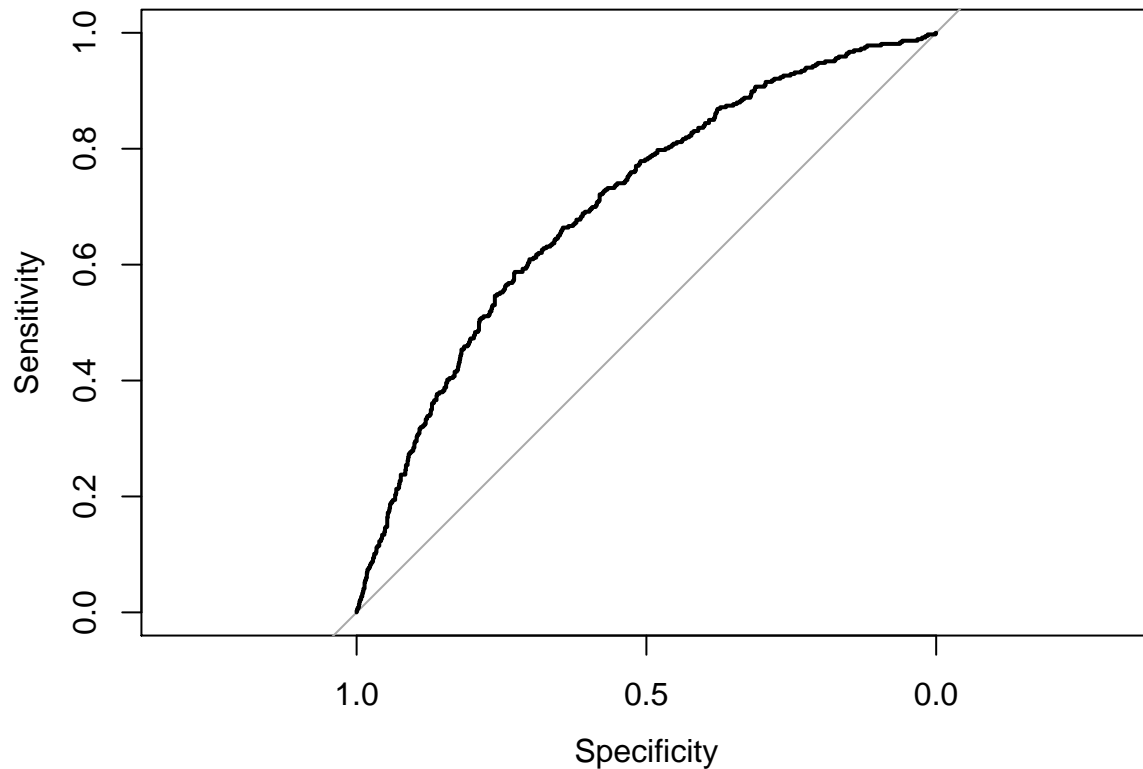


```
## [1] "AUC: 0.708454086505662"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1050 controls (dfPred_raw$class 0) < 349 cases (dfPred_raw$class 1).
## Area under the curve: 0.7085
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9704  -0.7687  -0.5676   0.8913   2.6782
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.788e-01  2.382e-01  -2.430  0.015095 *
## KIDSDRIV     1.896e-01  7.075e-02   2.680  0.007370 **
## AGE          -7.674e-03  4.567e-03  -1.680  0.092950 .
## HOMEKIDS     1.040e-01  3.765e-02   2.762  0.005746 **
## YOJ          -8.897e-03  8.686e-03  -1.024  0.305676
## INCOME        4.594e-07  1.057e-06   0.434  0.663935
## HOME_VAL     -2.677e-06  3.311e-07  -8.086  6.19e-16 ***
## TRAVTIME      7.820e-03  2.125e-03   3.679  0.000234 ***
## BLUEBOOK     -1.410e-05  4.685e-06  -3.010  0.002609 **
```

```

## TIF          -3.846e-02  8.596e-03  -4.475  7.65e-06 ***
## OLDCLAIM      6.571e-06  3.988e-06   1.648  0.099404 .
## CLM_FREQ      2.450e-01  3.236e-02   7.571  3.69e-14 ***
## MVR_PTS       1.606e-01  1.582e-02  10.150  < 2e-16 ***
## CAR_AGE       -2.553e-02  6.830e-03  -3.738  0.000185 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5843.0  on 5058  degrees of freedom
## Residual deviance: 5210.7  on 5045  degrees of freedom
##    (1354 observations deleted due to missingness)
## AIC: 5238.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 969 309
##           1  54  57
##
##           Accuracy : 0.7387
##           95% CI : (0.7147, 0.7616)
##    No Information Rate : 0.7365
##    P-Value [Acc > NIR] : 0.4413
##
##           Kappa : 0.1326
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9472
##           Specificity : 0.1557
##           Pos Pred Value : 0.7582
##           Neg Pred Value : 0.5135
##           Prevalence : 0.7365
##           Detection Rate : 0.6976
##           Detection Prevalence : 0.9201
##           Balanced Accuracy : 0.5515
##
##           'Positive' Class : 0
##

```

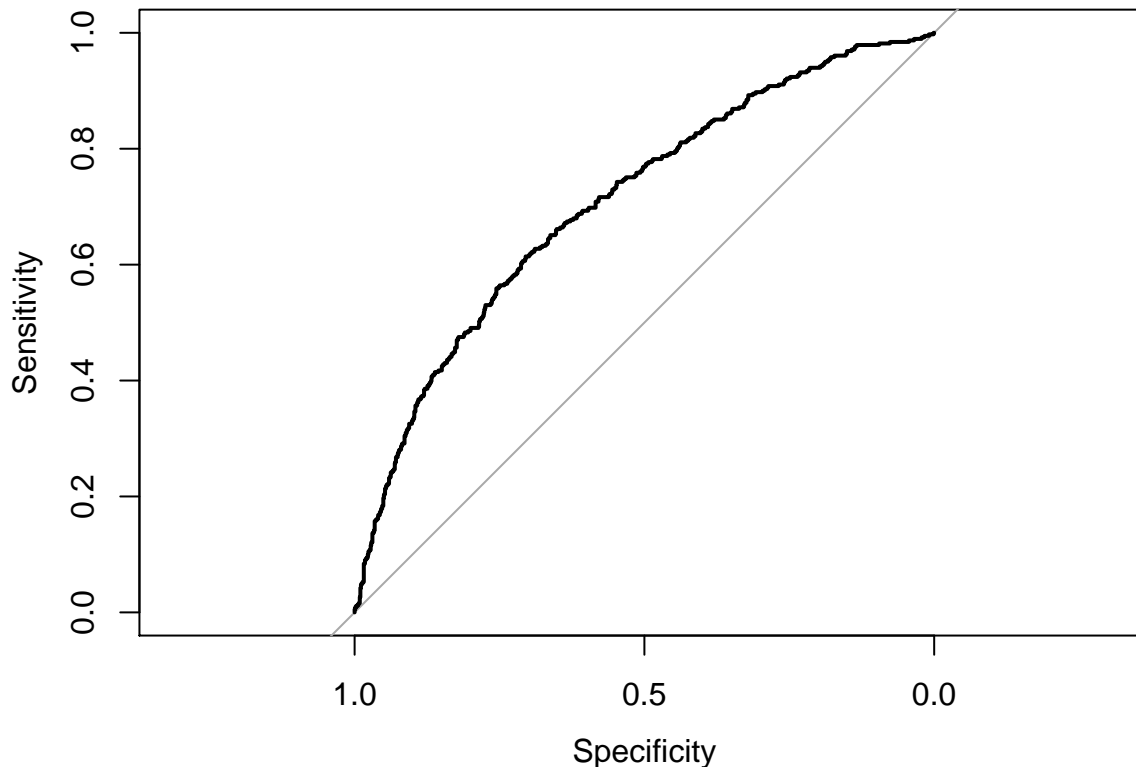


```
## [1] "AUC: 0.703927161621503"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1023 controls (dfPred_raw$class 0) < 366 cases (dfPred_raw$class 1).
## Area under the curve: 0.7039
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0165  -0.7672  -0.5696   0.8732   2.6564
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.047e-01  2.386e-01  -2.115  0.034437 *
## KIDSDRIV     2.367e-01  7.021e-02   3.372  0.000746 ***
## AGE        -7.982e-03  4.596e-03  -1.737  0.082426 .
## HOMEKIDS     8.361e-02  3.833e-02   2.182  0.029141 *
## YOJ        -1.106e-02  8.743e-03  -1.265  0.205778
## INCOME       6.443e-07  1.072e-06   0.601  0.547764
## HOME_VAL    -2.758e-06  3.325e-07  -8.297  < 2e-16 ***
## TRAVTIME     7.637e-03  2.133e-03   3.581  0.000342 ***
## BLUEBOOK    -1.239e-05  4.702e-06  -2.634  0.008442 **
```

```

## TIF          -4.295e-02  8.568e-03  -5.013  5.37e-07 ***
## OLDCLAIM     4.476e-06  4.013e-06   1.115  0.264675
## CLM_FREQ     2.517e-01  3.252e-02   7.741  9.89e-15 ***
## MVR_PTS      1.579e-01  1.598e-02   9.880  < 2e-16 ***
## CAR_AGE      -2.954e-02  6.817e-03  -4.333  1.47e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5810.9  on 5056  degrees of freedom
## Residual deviance: 5193.9  on 5043  degrees of freedom
##    (1356 observations deleted due to missingness)
## AIC: 5221.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 960 306
##           1  50  75
##
##           Accuracy : 0.7441
##           95% CI : (0.7203, 0.7668)
##    No Information Rate : 0.7261
##    P-Value [Acc > NIR] : 0.06963
##
##           Kappa : 0.1863
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9505
##           Specificity : 0.1969
##           Pos Pred Value : 0.7583
##           Neg Pred Value : 0.6000
##           Prevalence : 0.7261
##           Detection Rate : 0.6902
##           Detection Prevalence : 0.9101
##           Balanced Accuracy : 0.5737
##
##           'Positive' Class : 0
##

```



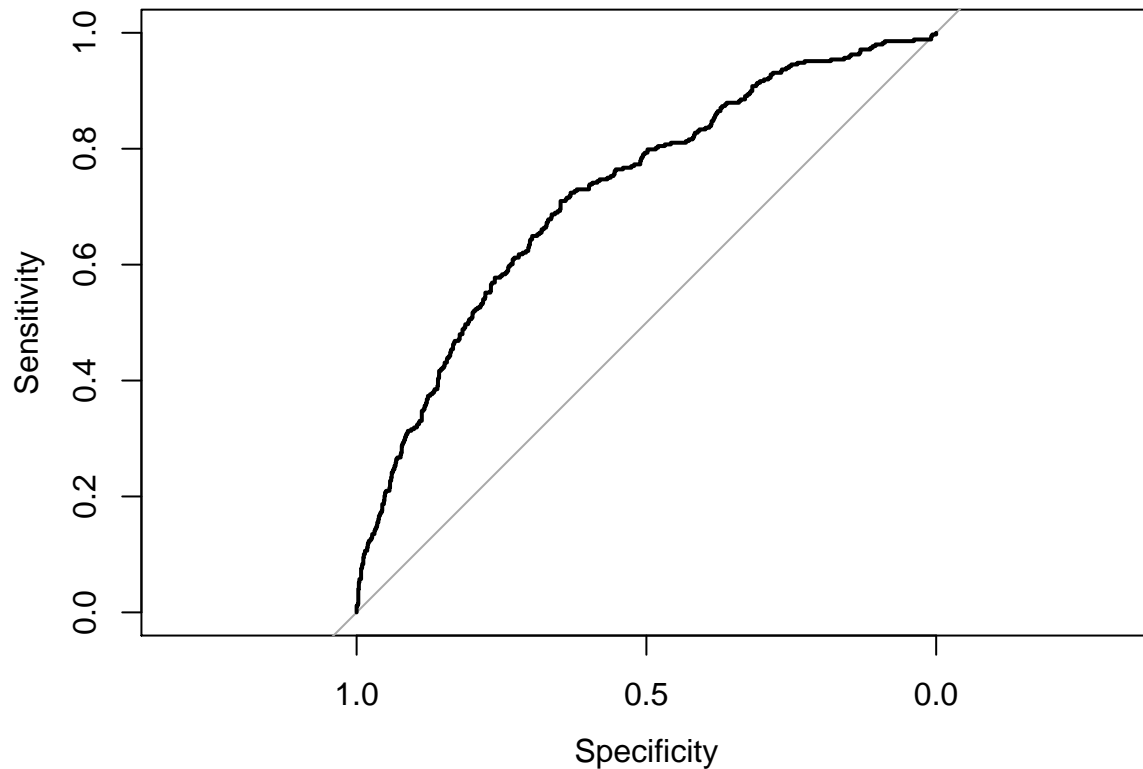
```
## [1] "AUC: 0.708336581689665"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1010 controls (dfPred_raw$class 0) < 381 cases (dfPred_raw$class 1).
## Area under the curve: 0.7083
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0113  -0.7814  -0.5803   0.9353   2.7095
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.925e-01  2.352e-01  -0.818  0.41310
## KIDSDRIV     2.819e-01  6.960e-02   4.051 5.11e-05 ***
## AGE          -1.306e-02  4.538e-03  -2.877  0.00401 **
## HOMEKIDS      5.670e-02  3.785e-02   1.498  0.13413
## YOJ           -5.274e-03  8.672e-03  -0.608  0.54312
## INCOME        -1.168e-06  1.041e-06  -1.122  0.26195
## HOME_VAL      -2.479e-06  3.300e-07  -7.512 5.83e-14 ***
## TRAVTIME      6.771e-03  2.111e-03   3.208  0.00134 **
## BLUEBOOK     -1.315e-05  4.633e-06  -2.838  0.00454 **
```

```

## TIF          -4.778e-02  8.530e-03  -5.601  2.13e-08 ***
## OLDCLAIM     8.081e-06  3.915e-06   2.064  0.03900 *
## CLM_FREQ     2.502e-01  3.240e-02   7.722  1.15e-14 ***
## MVR_PTS      1.283e-01  1.581e-02   8.115  4.84e-16 ***
## CAR_AGE      -2.038e-02  6.697e-03  -3.043  0.00234 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5880.2  on 5059  degrees of freedom
## Residual deviance: 5280.1  on 5046  degrees of freedom
##    (1352 observations deleted due to missingness)
## AIC: 5308.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 989 278
##           1  51  70
##
##           Accuracy : 0.763
##           95% CI : (0.7397, 0.7851)
##    No Information Rate : 0.7493
##    P-Value [Acc > NIR] : 0.1256
##
##           Kappa : 0.1943
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9510
##           Specificity : 0.2011
##           Pos Pred Value : 0.7806
##           Neg Pred Value : 0.5785
##           Prevalence : 0.7493
##           Detection Rate : 0.7125
##           Detection Prevalence : 0.9128
##           Balanced Accuracy : 0.5761
##
##           'Positive' Class : 0
##

```



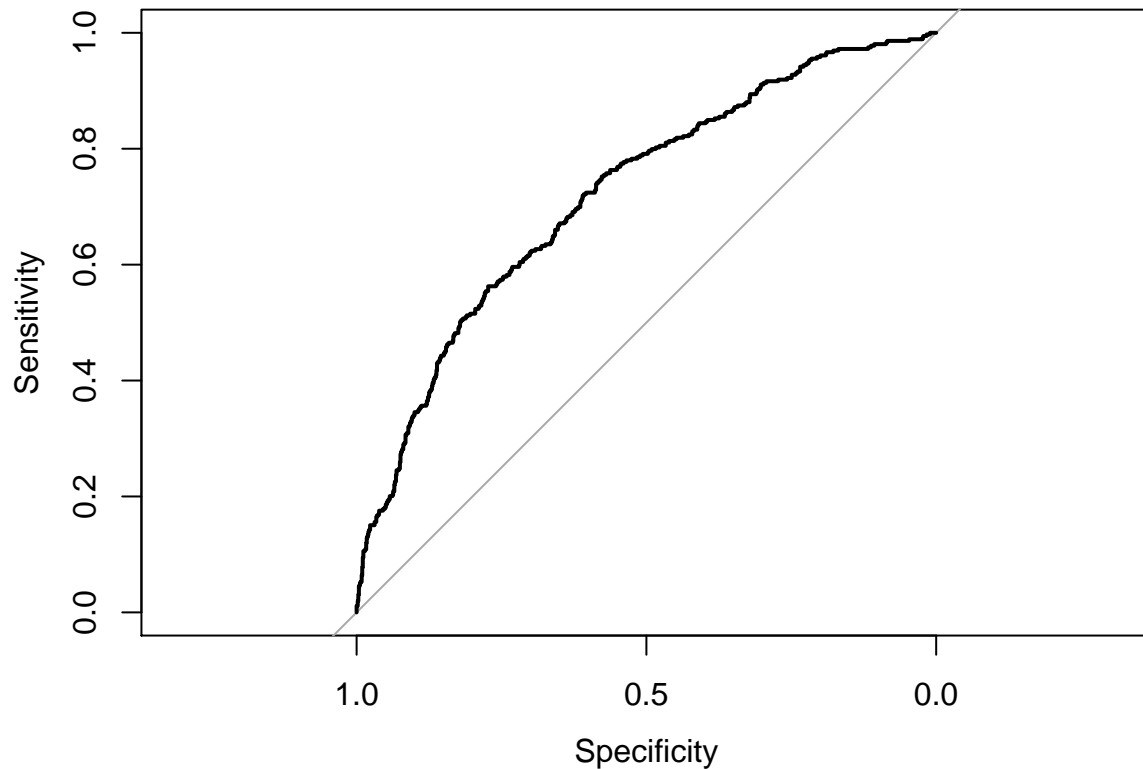


```
## [1] "AUC: 0.722673519009726"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1040 controls (dfPred_raw$class 0) < 348 cases (dfPred_raw$class 1).
## Area under the curve: 0.7227
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0331  -0.7736  -0.5779   0.9146   2.6872
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.286e-01  2.364e-01  -1.813  0.069886 .
## KIDSDRIV     2.533e-01  7.050e-02   3.593  0.000327 ***
## AGE          -9.077e-03  4.514e-03  -2.011  0.044356 *
## HOMEKIDS      5.093e-02  3.811e-02   1.337  0.181377
## YOJ          -1.432e-03  8.798e-03  -0.163  0.870723
## INCOME       -1.512e-07  1.039e-06  -0.146  0.884224
## HOME_VAL     -2.525e-06  3.283e-07  -7.691  1.46e-14 ***
## TRAVTIME      8.419e-03  2.112e-03   3.987  6.70e-05 ***
## BLUEBOOK     -1.575e-05  4.666e-06  -3.375  0.000738 ***
```

```

## TIF          -5.025e-02  8.594e-03  -5.846  5.03e-09 ***
## OLDCLAIM     9.114e-06  3.955e-06   2.304  0.021208 *
## CLM_FREQ     2.414e-01  3.248e-02   7.433  1.06e-13 ***
## MVR_PTS      1.395e-01  1.590e-02   8.770  < 2e-16 ***
## CAR_AGE      -2.552e-02  6.731e-03  -3.792  0.000149 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5856.7  on 5057  degrees of freedom
## Residual deviance: 5258.5  on 5044  degrees of freedom
##    (1354 observations deleted due to missingness)
## AIC: 5286.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 967 287
##           1  64  72
##
##           Accuracy : 0.7475
##           95% CI : (0.7238, 0.7701)
##    No Information Rate : 0.7417
##    P-Value [Acc > NIR] : 0.3243
##
##           Kappa : 0.1736
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9379
##           Specificity : 0.2006
##           Pos Pred Value : 0.7711
##           Neg Pred Value : 0.5294
##           Prevalence : 0.7417
##           Detection Rate : 0.6957
##           Detection Prevalence : 0.9022
##           Balanced Accuracy : 0.5692
##
##           'Positive' Class : 0
##

```

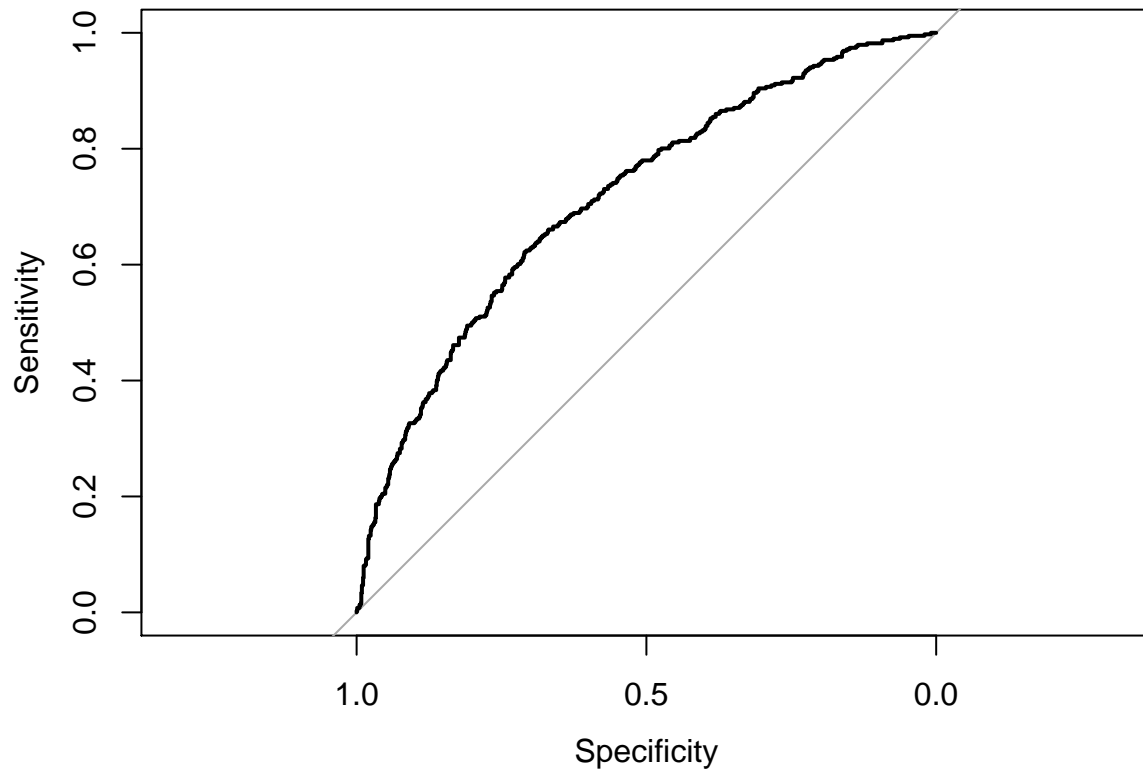


```
## [1] "AUC: 0.72083246651843"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1031 controls (dfPred_raw$class 0) < 359 cases (dfPred_raw$class 1).
## Area under the curve: 0.7208
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9760  -0.7615  -0.5720   0.8659   2.6456
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.239e-01  2.387e-01  -1.357 0.174756
## KIDSDRIV     2.387e-01  7.046e-02   3.388 0.000703 ***
## AGE         -1.149e-02  4.606e-03  -2.494 0.012614 *
## HOMEKIDS      5.839e-02  3.843e-02   1.520 0.128614
## YOJ          -1.068e-02  8.754e-03  -1.220 0.222468
## INCOME        7.521e-07  1.061e-06   0.709 0.478403
## HOME_VAL     -2.578e-06  3.313e-07  -7.783 7.09e-15 ***
## TRAVTIME      7.414e-03  2.142e-03   3.462 0.000537 ***
## BLUEBOOK     -1.346e-05  4.703e-06  -2.863 0.004195 **
```

```

## TIF          -4.582e-02  8.546e-03  -5.362 8.22e-08 ***
## OLDCLAIM      3.771e-06  3.967e-06   0.951 0.341756
## CLM_FREQ      2.546e-01  3.255e-02   7.824 5.14e-15 ***
## MVR_PTS       1.532e-01  1.599e-02   9.576 < 2e-16 ***
## CAR_AGE       -2.868e-02  6.770e-03  -4.237 2.26e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5808.9  on 5070  degrees of freedom
## Residual deviance: 5209.5  on 5057  degrees of freedom
##    (1341 observations deleted due to missingness)
## AIC: 5237.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 942 304
##           1  49  82
##
##           Accuracy : 0.7436
##           95% CI : (0.7197, 0.7665)
##    No Information Rate : 0.7197
##    P-Value [Acc > NIR] : 0.02483
##
##           Kappa : 0.2042
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9506
##           Specificity : 0.2124
##           Pos Pred Value : 0.7560
##           Neg Pred Value : 0.6260
##           Prevalence : 0.7197
##           Detection Rate : 0.6841
##    Detection Prevalence : 0.9049
##           Balanced Accuracy : 0.5815
##
##           'Positive' Class : 0
##

```

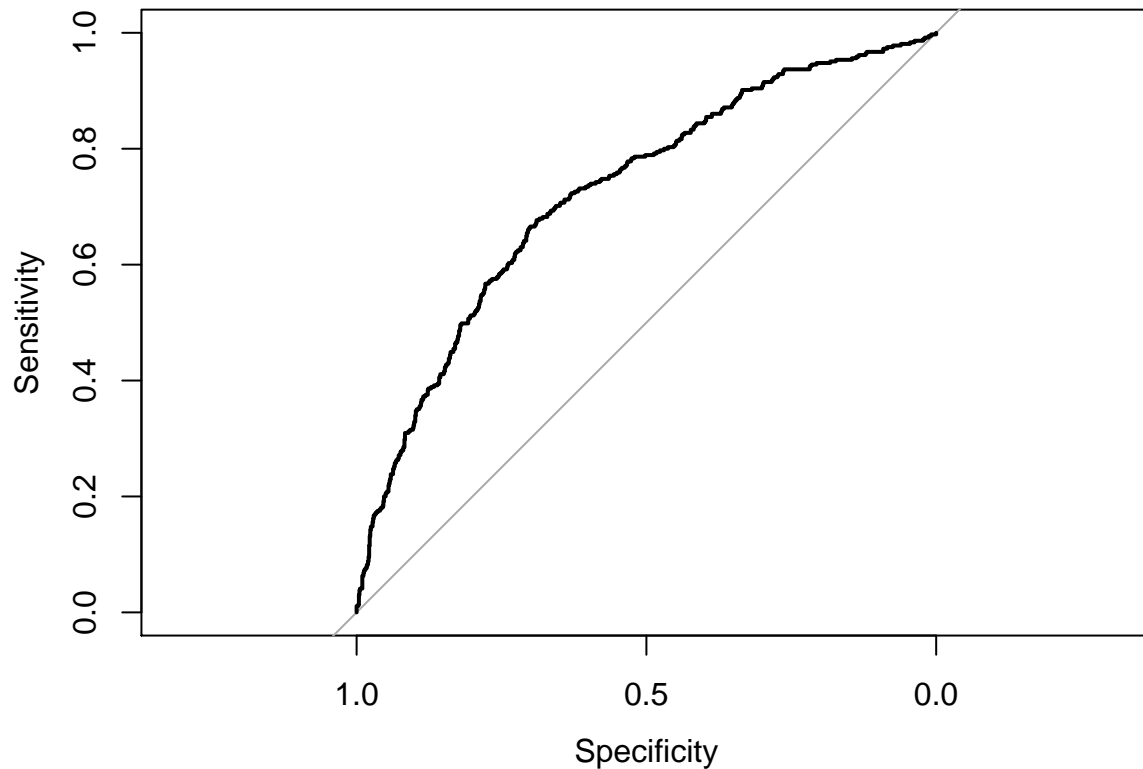


```
## [1] "AUC: 0.71544417895777"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 991 controls (dfPred_raw$class 0) < 386 cases (dfPred_raw$class 1).
## Area under the curve: 0.7154
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9654  -0.7731  -0.5775   0.9236   2.6853
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.709e-01  2.380e-01  -1.558  0.119153
## KIDSDRIV     2.007e-01  7.050e-02   2.847  0.004408 **
## AGE         -1.059e-02  4.571e-03  -2.318  0.020459 *
## HOMEKIDS      8.605e-02  3.773e-02   2.281  0.022570 *
## YOJ          -7.269e-03  8.674e-03  -0.838  0.401981
## INCOME       -9.425e-07  1.063e-06  -0.887  0.375214
## HOME_VAL    -2.453e-06  3.335e-07  -7.357  1.88e-13 ***
## TRAVTIME      7.356e-03  2.120e-03   3.470  0.000521 ***
## BLUEBOOK    -1.071e-05  4.630e-06  -2.313  0.020737 *
```

```

## TIF          -4.201e-02  8.628e-03  -4.870  1.12e-06 ***
## OLDCLAIM     8.105e-06  3.985e-06   2.034  0.041970 *
## CLM_FREQ     2.503e-01  3.243e-02   7.717  1.19e-14 ***
## MVR_PTS      1.404e-01  1.605e-02   8.749  < 2e-16 ***
## CAR_AGE      -2.756e-02  6.828e-03  -4.036  5.44e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5848.8  on 5064  degrees of freedom
## Residual deviance: 5250.7  on 5051  degrees of freedom
## (1348 observations deleted due to missingness)
## AIC: 5278.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 966 290
##           1  52  75
##
##           Accuracy : 0.7527
##           95% CI : (0.7291, 0.7753)
##    No Information Rate : 0.7361
##    P-Value [Acc > NIR] : 0.08424
##
##           Kappa : 0.1952
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9489
##           Specificity : 0.2055
##           Pos Pred Value : 0.7691
##           Neg Pred Value : 0.5906
##           Prevalence : 0.7361
##           Detection Rate : 0.6985
##           Detection Prevalence : 0.9082
##           Balanced Accuracy : 0.5772
##
##           'Positive' Class : 0
##

```



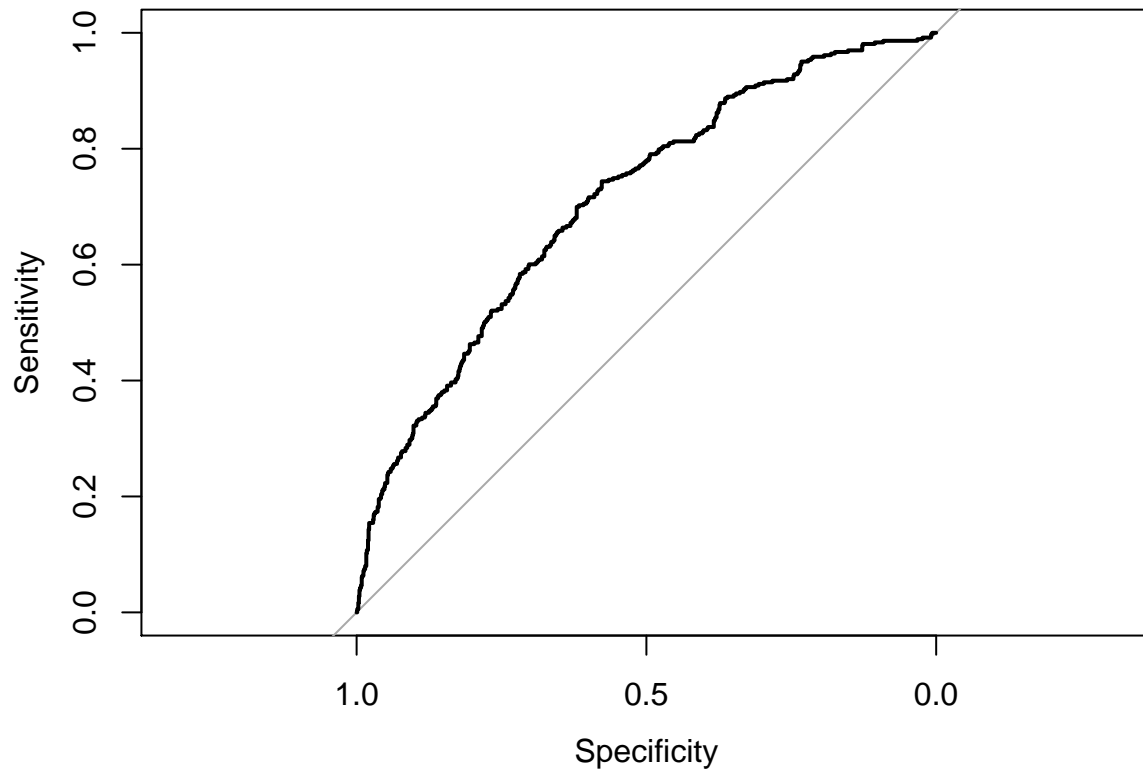
```
## [1] "AUC: 0.723834001668595"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1018 controls (dfPred_raw$class 0) < 365 cases (dfPred_raw$class 1).
## Area under the curve: 0.7238
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7692  -0.7738  -0.5724   0.9180   2.7193
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.588e-01  2.384e-01  -1.086  0.277675
## KIDSDRIV     2.983e-01  6.886e-02   4.331  1.48e-05 ***
## AGE         -1.429e-02  4.620e-03  -3.094  0.001975 **
## HOMEKIDS      5.377e-02  3.762e-02   1.429  0.152922
## YOJ          -2.400e-03  8.742e-03  -0.274  0.783707
## INCOME       -8.095e-07  1.037e-06  -0.780  0.435148
## HOME_VAL     -2.576e-06  3.322e-07  -7.757  8.73e-15 ***
## TRAVTIME      7.441e-03  2.149e-03   3.463  0.000535 ***
## BLUEBOOK     -1.195e-05  4.643e-06  -2.575  0.010029 *
```

```

## TIF          -4.207e-02  8.546e-03  -4.923  8.53e-07 ***
## OLDCLAIM      6.278e-06  3.973e-06   1.580  0.114101
## CLM_FREQ      2.723e-01  3.231e-02   8.428  < 2e-16 ***
## MVR_PTS       1.329e-01  1.596e-02   8.330  < 2e-16 ***
## CAR_AGE       -2.277e-02  6.753e-03  -3.373  0.000744 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5853.5  on 5065  degrees of freedom
## Residual deviance: 5239.2  on 5052  degrees of freedom
##    (1346 observations deleted due to missingness)
## AIC: 5267.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 969 283
##           1  50  80
##
##           Accuracy : 0.759
##           95% CI : (0.7356, 0.7814)
##    No Information Rate : 0.7373
##    P-Value [Acc > NIR] : 0.03479
##
##           Kappa : 0.2159
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9509
##           Specificity : 0.2204
##           Pos Pred Value : 0.7740
##           Neg Pred Value : 0.6154
##           Prevalence : 0.7373
##           Detection Rate : 0.7012
##           Detection Prevalence : 0.9059
##           Balanced Accuracy : 0.5857
##
##           'Positive' Class : 0
##

```



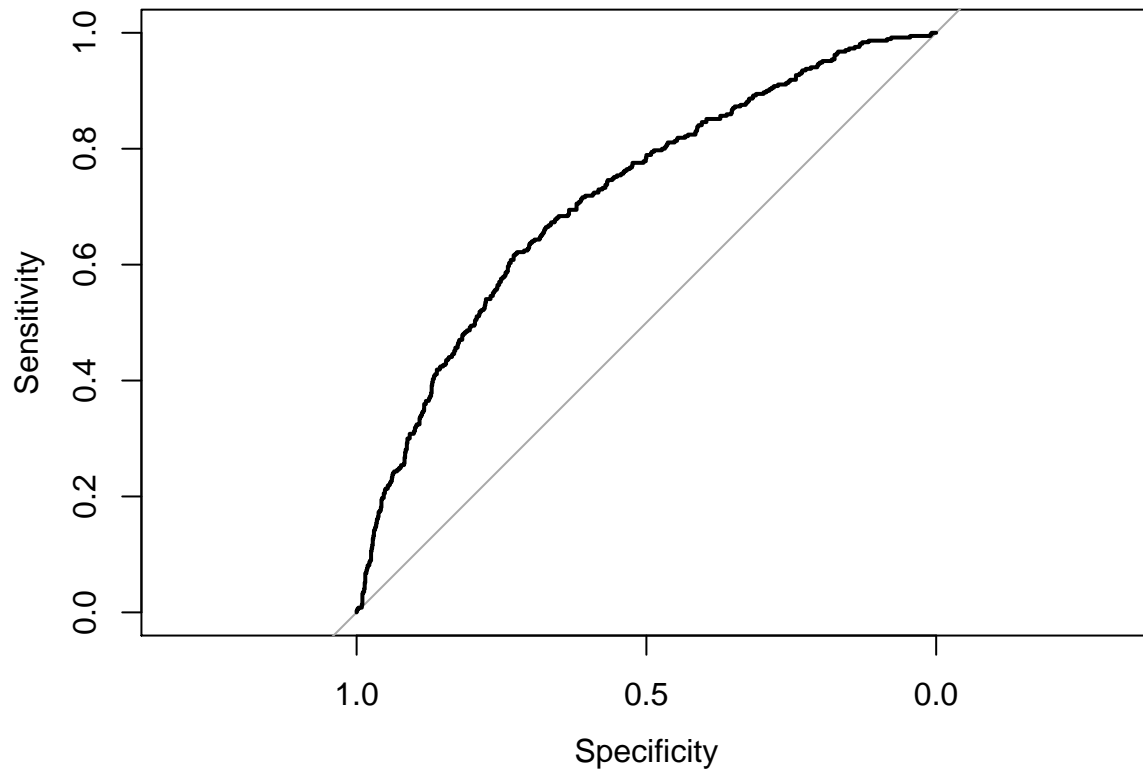


```
## [1] "AUC: 0.709681343725415"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1019 controls (dfPred_raw$class 0) < 363 cases (dfPred_raw$class 1).
## Area under the curve: 0.7097
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0389  -0.7690  -0.5706   0.8777   2.6548
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.239e-01  2.385e-01  -1.777 0.075526 .
## KIDSDRIV     2.596e-01  7.068e-02   3.673 0.000240 ***
## AGE          -1.003e-02  4.595e-03  -2.183 0.029033 *
## HOMEKIDS      8.367e-02  3.806e-02   2.198 0.027925 *
## YOJ           -1.141e-02  8.720e-03  -1.309 0.190549
## INCOME         4.472e-07  1.063e-06   0.421 0.674104
## HOME_VAL      -2.625e-06  3.310e-07  -7.931 2.18e-15 ***
## TRAVTIME       7.704e-03  2.132e-03   3.614 0.000301 ***
## BLUEBOOK      -1.305e-05  4.699e-06  -2.777 0.005490 **
```

```

## TIF          -4.397e-02  8.537e-03  -5.151  2.60e-07 ***
## OLDCLAIM      3.583e-06  3.916e-06   0.915  0.360247
## CLM_FREQ      2.739e-01  3.253e-02   8.421  < 2e-16 ***
## MVR_PTS       1.438e-01  1.599e-02   8.990  < 2e-16 ***
## CAR_AGE       -2.477e-02  6.734e-03  -3.678  0.000235 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5844  on 5073  degrees of freedom
## Residual deviance: 5233  on 5060  degrees of freedom
## (1338 observations deleted due to missingness)
## AIC: 5261
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  948 289
##           1   56  81
##
##           Accuracy : 0.7489
##           95% CI : (0.7251, 0.7716)
##    No Information Rate : 0.7307
##    P-Value [Acc > NIR] : 0.06733
##
##           Kappa : 0.2036
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9442
##           Specificity : 0.2189
##           Pos Pred Value : 0.7664
##           Neg Pred Value : 0.5912
##           Prevalence : 0.7307
##           Detection Rate : 0.6900
##           Detection Prevalence : 0.9003
##           Balanced Accuracy : 0.5816
##
##           'Positive' Class : 0
##

```

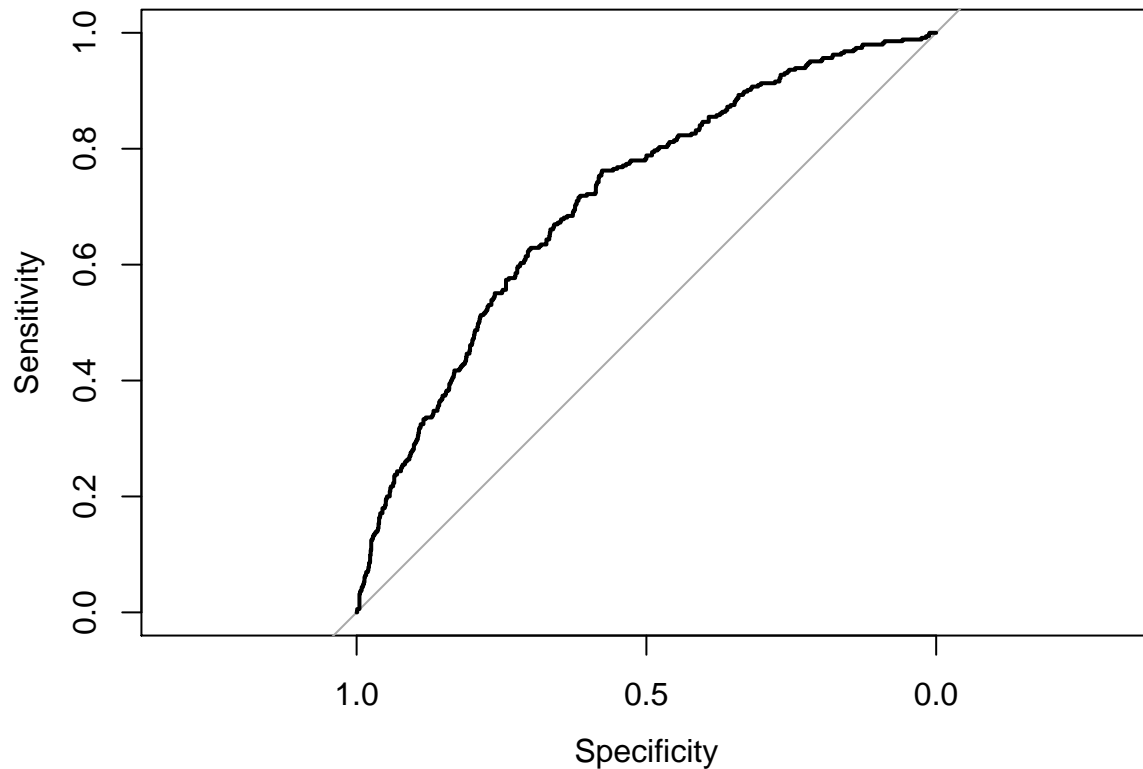


```
## [1] "AUC: 0.716722299989232"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1004 controls (dfPred_raw$class 0) < 370 cases (dfPred_raw$class 1).
## Area under the curve: 0.7167
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8025  -0.7770  -0.5758   0.9315   2.7015
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.457e-01  2.366e-01  -1.884  0.05958 .
## KIDSDRIV     3.181e-01  7.052e-02   4.511 6.45e-06 ***
## AGE         -1.264e-02  4.582e-03  -2.759  0.00580 **
## HOMEKIDS      7.035e-02  3.747e-02   1.877  0.06047 .
## YOJ          -4.969e-04  8.681e-03  -0.057  0.95435
## INCOME       -1.122e-06  1.041e-06  -1.077  0.28127
## HOME_VAL     -2.645e-06  3.320e-07  -7.966 1.64e-15 ***
## TRAVTIME      8.111e-03  2.149e-03   3.775  0.00016 ***
## BLUEBOOK     -9.899e-06  4.643e-06  -2.132  0.03299 *
```

```

## TIF          -3.627e-02  8.483e-03  -4.275  1.91e-05 ***
## OLDCLAIM      5.782e-06  3.915e-06   1.477  0.13968
## CLM_FREQ      2.759e-01  3.238e-02   8.521  < 2e-16 ***
## MVR_PTS       1.363e-01  1.592e-02   8.562  < 2e-16 ***
## CAR_AGE       -1.864e-02  6.714e-03  -2.776  0.00550 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5893.1  on 5070  degrees of freedom
## Residual deviance: 5269.4  on 5057  degrees of freedom
##    (1341 observations deleted due to missingness)
## AIC: 5297.4
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 967 268
##           1  65  77
##
##           Accuracy : 0.7582
##           95% CI : (0.7347, 0.7806)
##    No Information Rate : 0.7495
##    P-Value [Acc > NIR] : 0.238
##
##           Kappa : 0.1992
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9370
##           Specificity : 0.2232
##           Pos Pred Value : 0.7830
##           Neg Pred Value : 0.5423
##           Prevalence : 0.7495
##           Detection Rate : 0.7023
##           Detection Prevalence : 0.8969
##           Balanced Accuracy : 0.5801
##
##           'Positive' Class : 0
##

```

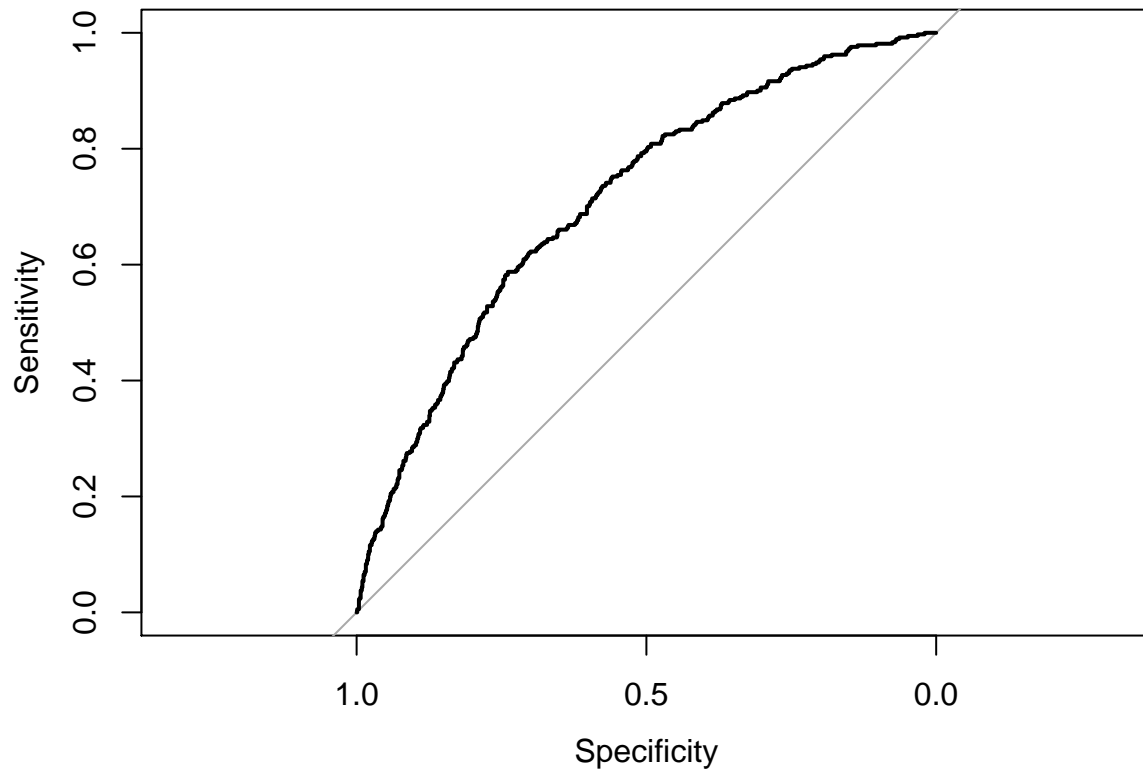


```
## [1] "AUC: 0.711984608470958"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1032 controls (dfPred_raw$class 0) < 345 cases (dfPred_raw$class 1).
## Area under the curve: 0.712
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9294  -0.7652  -0.5708   0.8857   2.6677
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.983e-01  2.382e-01  -1.672  0.094445 .
## KIDSDRIV     1.907e-01  7.103e-02   2.685  0.007260 **
## AGE          -1.113e-02  4.576e-03  -2.433  0.014989 *
## HOMEKIDS      7.974e-02  3.772e-02   2.114  0.034526 *
## YOJ           -8.589e-03  8.694e-03  -0.988  0.323239
## INCOME         5.713e-07  1.047e-06   0.546  0.585309
## HOME_VAL      -2.506e-06  3.300e-07  -7.595  3.08e-14 ***
## TRAVTIME       7.589e-03  2.134e-03   3.557  0.000375 ***
## BLUEBOOK      -1.513e-05  4.685e-06  -3.229  0.001242 **
```

```

## TIF          -4.145e-02  8.572e-03  -4.835  1.33e-06 ***
## OLDCLAIM     5.754e-06  3.939e-06   1.461  0.144075
## CLM_FREQ     2.478e-01  3.237e-02   7.654  1.94e-14 ***
## MVR_PTS      1.559e-01  1.583e-02   9.847  < 2e-16 ***
## CAR_AGE      -2.475e-02  6.783e-03  -3.648  0.000264 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5841.3  on 5072  degrees of freedom
## Residual deviance: 5226.8  on 5059  degrees of freedom
##    (1339 observations deleted due to missingness)
## AIC: 5254.8
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 951 304
##           1  53  67
##
##               Accuracy : 0.7404
##               95% CI : (0.7163, 0.7634)
##    No Information Rate : 0.7302
##    P-Value [Acc > NIR] : 0.2065
##
##               Kappa : 0.1625
##
## Mcnemar's Test P-Value : <2e-16
##
##               Sensitivity : 0.9472
##               Specificity : 0.1806
##    Pos Pred Value : 0.7578
##    Neg Pred Value : 0.5583
##    Prevalence : 0.7302
##    Detection Rate : 0.6916
##    Detection Prevalence : 0.9127
##    Balanced Accuracy : 0.5639
##
##    'Positive' Class : 0
##

```



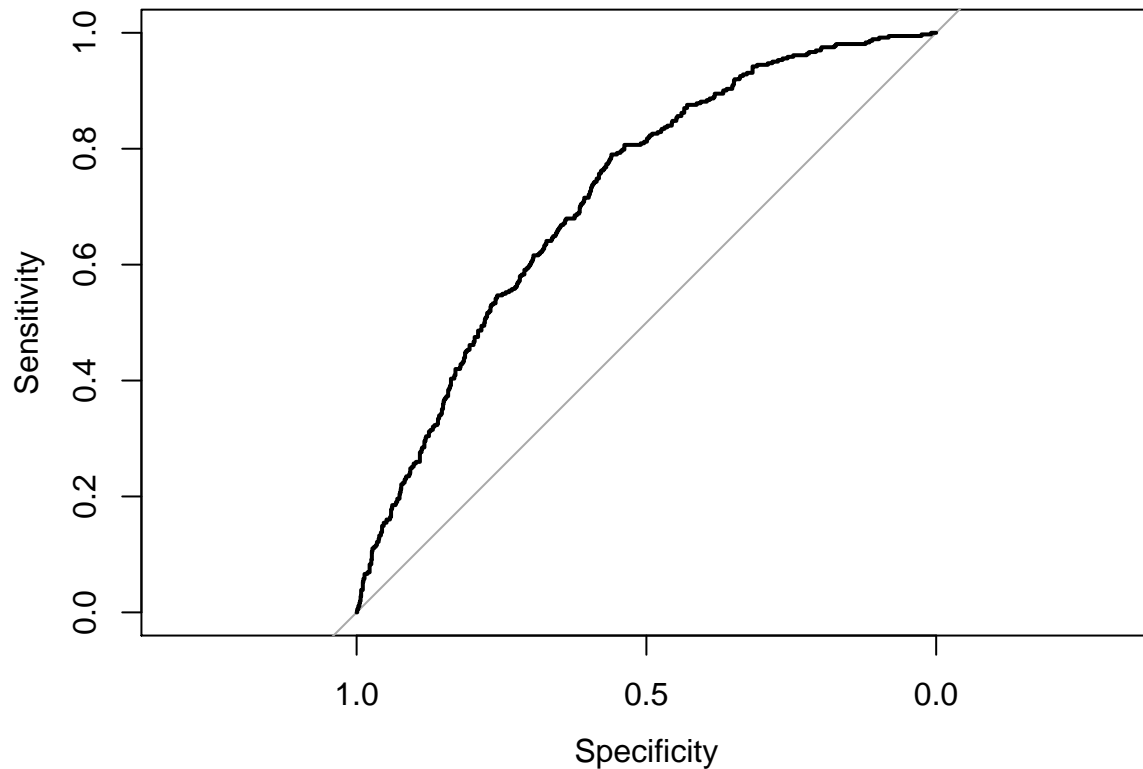
```
## [1] "AUC: 0.711807755500907"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1004 controls (dfPred_raw$class 0) < 371 cases (dfPred_raw$class 1).
## Area under the curve: 0.7118
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7900  -0.7643  -0.5725   0.9156   2.6125
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.768e-01  2.384e-01  -2.839 0.004518 **
## KIDSDRIV     2.887e-01  6.988e-02   4.131 3.61e-05 ***
## AGE          -1.129e-02  4.568e-03  -2.472 0.013441 *
## HOMEKIDS      6.834e-02  3.753e-02   1.821 0.068591 .
## YOJ           2.581e-03  8.688e-03   0.297 0.766372
## INCOME        -5.525e-07  1.035e-06  -0.534 0.593340
## HOME_VAL      -2.502e-06  3.315e-07  -7.549 4.39e-14 ***
## TRAVTIME       7.689e-03  2.140e-03   3.593 0.000328 ***
## BLUEBOOK      -7.972e-06  4.655e-06  -1.713 0.086788 .
```

```

## TIF          -3.322e-02  8.509e-03  -3.904  9.48e-05 ***
## OLDCLAIM     6.465e-06  3.983e-06   1.623  0.104581
## CLM_FREQ     2.902e-01  3.213e-02   9.034  < 2e-16 ***
## MVR_PTS      1.437e-01  1.573e-02   9.136  < 2e-16 ***
## CAR_AGE      -1.699e-02  6.716e-03  -2.530  0.011422 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5851.8  on 5059  degrees of freedom
## Residual deviance: 5239.6  on 5046  degrees of freedom
## (1352 observations deleted due to missingness)
## AIC: 5267.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 966 303
##           1  60  59
##
##           Accuracy : 0.7385
##           95% CI : (0.7145, 0.7614)
##    No Information Rate : 0.7392
##    P-Value [Acc > NIR] : 0.5384
##
##           Kappa : 0.1335
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9415
##           Specificity : 0.1630
##           Pos Pred Value : 0.7612
##           Neg Pred Value : 0.4958
##           Prevalence : 0.7392
##           Detection Rate : 0.6960
##           Detection Prevalence : 0.9143
##           Balanced Accuracy : 0.5523
##
##           'Positive' Class : 0
##

```



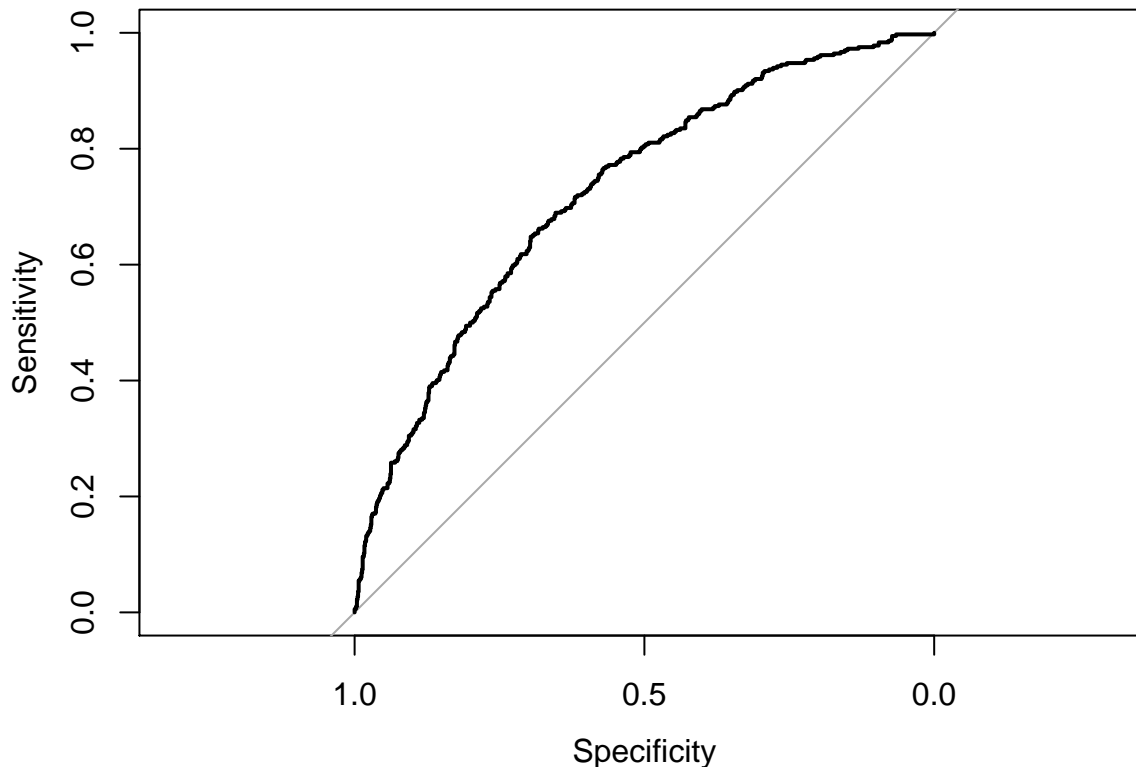


```
## [1] "AUC: 0.718662832649457"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1026 controls (dfPred_raw$class 0) < 362 cases (dfPred_raw$class 1).
## Area under the curve: 0.7187
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9486  -0.7670  -0.5770   0.9273   2.6333
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.083e-01  2.366e-01  -1.303  0.192619
## KIDSDRIV     1.876e-01  7.021e-02   2.672  0.007529 **
## AGE         -1.031e-02  4.545e-03  -2.268  0.023318 *
## HOMEKIDS      8.613e-02  3.766e-02   2.287  0.022186 *
## YOJ          -1.692e-02  8.689e-03  -1.947  0.051530 .
## INCOME       -3.797e-08  1.062e-06  -0.036  0.971492
## HOME_VAL     -2.313e-06  3.308e-07  -6.993  2.70e-12 ***
## TRAVTIME      8.157e-03  2.112e-03   3.861  0.000113 ***
## BLUEBOOK     -1.391e-05  4.644e-06  -2.996  0.002738 **
```

```

## TIF          -5.017e-02  8.683e-03  -5.778 7.58e-09 ***
## OLDCLAIM     6.623e-06  3.825e-06   1.731 0.083391 .
## CLM_FREQ     2.639e-01  3.230e-02   8.170 3.09e-16 ***
## MVR_PTS      1.318e-01  1.583e-02   8.323 < 2e-16 ***
## CAR_AGE      -2.525e-02  6.762e-03  -3.734 0.000188 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5860.0  on 5079  degrees of freedom
## Residual deviance: 5268.1  on 5066  degrees of freedom
##    (1332 observations deleted due to missingness)
## AIC: 5296.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 966 297
##           1  38  67
##
##           Accuracy : 0.7551
##           95% CI : (0.7314, 0.7777)
##    No Information Rate : 0.7339
##    P-Value [Acc > NIR] : 0.03972
##
##           Kappa : 0.1891
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9622
##           Specificity : 0.1841
##           Pos Pred Value : 0.7648
##           Neg Pred Value : 0.6381
##           Prevalence : 0.7339
##           Detection Rate : 0.7061
##           Detection Prevalence : 0.9232
##           Balanced Accuracy : 0.5731
##
##           'Positive' Class : 0
##

```

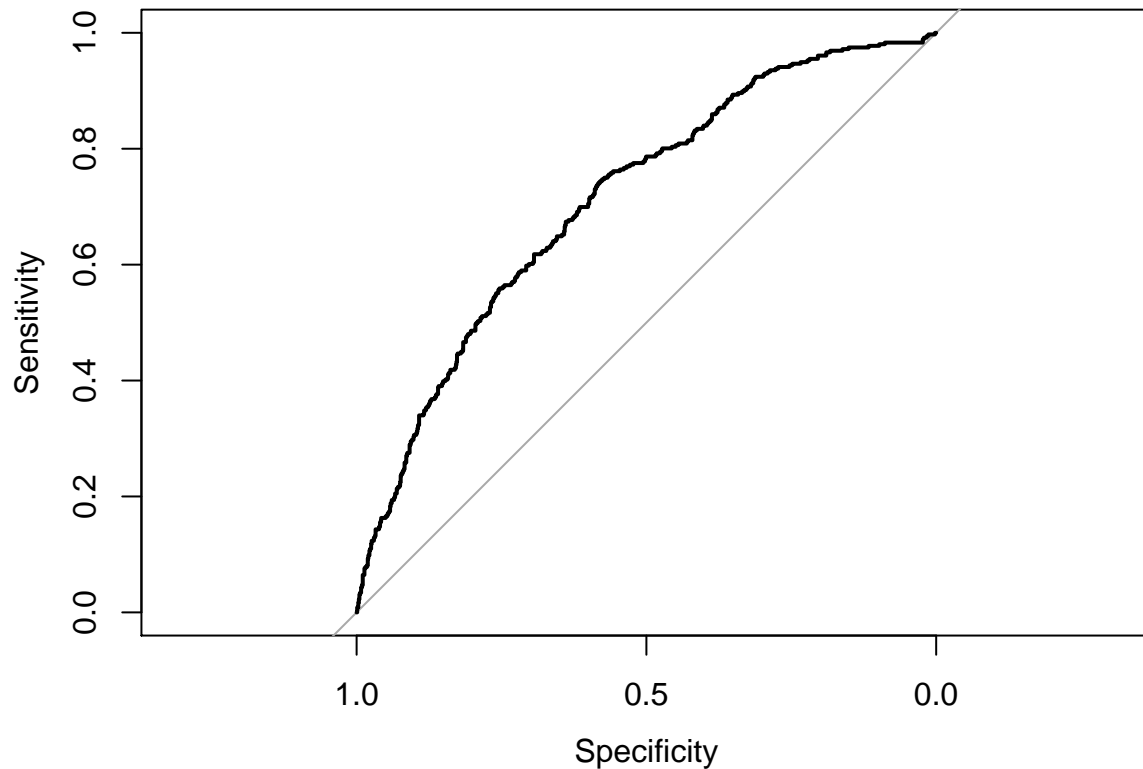


```
## [1] "AUC: 0.725529201873823"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1004 controls (dfPred_raw$class 0) < 364 cases (dfPred_raw$class 1).
## Area under the curve: 0.7255
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8207  -0.7734  -0.5738   0.9151   2.6425
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -7.719e-01  2.367e-01  -3.261 0.001111 **
## KIDSDRIV     2.850e-01  6.969e-02   4.089 4.33e-05 ***
## AGE          -5.880e-03  4.540e-03  -1.295 0.195273
## HOMEKIDS      7.178e-02  3.766e-02   1.906 0.056657 .
## YOJ          -1.174e-03  8.787e-03  -0.134 0.893669
## INCOME        -2.696e-07  1.048e-06  -0.257 0.796941
## HOME_VAL      -2.789e-06  3.308e-07  -8.431 < 2e-16 ***
## TRAVTIME      9.972e-03  2.137e-03   4.667 3.05e-06 ***
## BLUEBOOK     -1.022e-05  4.648e-06  -2.199 0.027872 *
```

```

## TIF          -4.068e-02  8.590e-03  -4.735  2.19e-06 ***
## OLDCLAIM      7.586e-06  3.981e-06   1.905  0.056742 .
## CLM_FREQ      2.528e-01  3.223e-02   7.844  4.37e-15 ***
## MVR_PTS       1.510e-01  1.581e-02   9.548  < 2e-16 ***
## CAR_AGE       -2.466e-02  6.772e-03  -3.641  0.000271 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5859.7  on 5052  degrees of freedom
## Residual deviance: 5239.3  on 5039  degrees of freedom
##    (1360 observations deleted due to missingness)
## AIC: 5267.3
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  973  287
##           1   66   69
##
##           Accuracy : 0.747
##           95% CI : (0.7233, 0.7696)
##    No Information Rate : 0.7448
##    P-Value [Acc > NIR] : 0.4409
##
##           Kappa : 0.1637
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9365
##           Specificity : 0.1938
##           Pos Pred Value : 0.7722
##           Neg Pred Value : 0.5111
##           Prevalence : 0.7448
##           Detection Rate : 0.6975
##           Detection Prevalence : 0.9032
##           Balanced Accuracy : 0.5651
##
##           'Positive' Class : 0
##

```

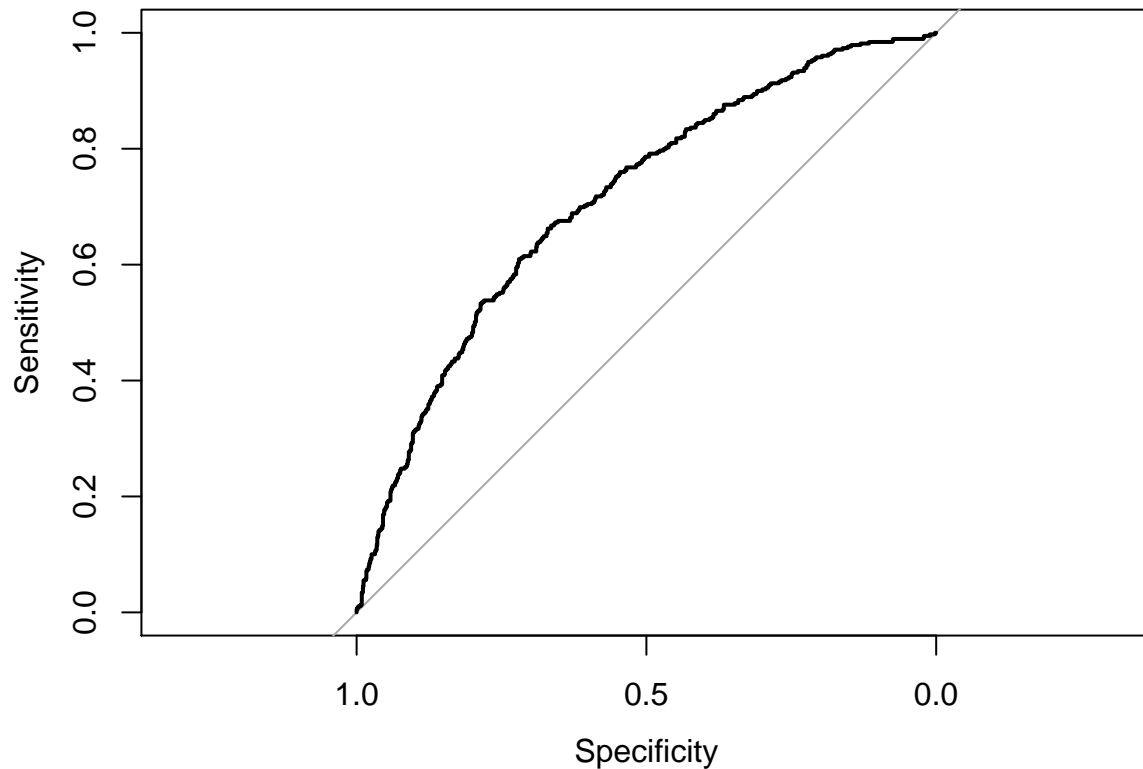


```
## [1] "AUC: 0.712009711152686"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1039 controls (dfPred_raw$class 0) < 356 cases (dfPred_raw$class 1).
## Area under the curve: 0.712
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0264  -0.7645  -0.5675   0.8837   2.7201
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.321e-01  2.388e-01  -2.647 0.008112 **
## KIDSDRIV     2.534e-01  7.070e-02   3.584 0.000339 ***
## AGE         -6.851e-03  4.561e-03  -1.502 0.133067
## HOMEKIDS     8.352e-02  3.816e-02   2.188 0.028634 *
## YOJ         -1.533e-03  8.755e-03  -0.175 0.861047
## INCOME      -3.687e-07  1.065e-06  -0.346 0.729154
## HOME_VAL    -2.480e-06  3.334e-07  -7.439 1.01e-13 ***
## TRAVTIME     7.497e-03  2.132e-03   3.516 0.000437 ***
## BLUEBOOK    -1.540e-05  4.735e-06  -3.252 0.001145 **
```

```

## TIF          -4.001e-02  8.593e-03  -4.656  3.22e-06 ***
## OLDCLAIM     9.841e-06  3.968e-06   2.480  0.013145 *
## CLM_FREQ     2.509e-01  3.263e-02   7.688  1.49e-14 ***
## MVR_PTS      1.479e-01  1.591e-02   9.298  < 2e-16 ***
## CAR_AGE      -2.590e-02  6.843e-03  -3.785  0.000154 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5816.2  on 5058  degrees of freedom
## Residual deviance: 5200.3  on 5045  degrees of freedom
## (1354 observations deleted due to missingness)
## AIC: 5228.3
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  959  310
##           1   51   69
##
##           Accuracy : 0.7401
##           95% CI : (0.7162, 0.763)
##    No Information Rate : 0.7271
##    P-Value [Acc > NIR] : 0.1458
##
##           Kappa : 0.1673
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9495
##           Specificity : 0.1821
##           Pos Pred Value : 0.7557
##           Neg Pred Value : 0.5750
##           Prevalence : 0.7271
##           Detection Rate : 0.6904
##           Detection Prevalence : 0.9136
##           Balanced Accuracy : 0.5658
##
##           'Positive' Class : 0
##

```



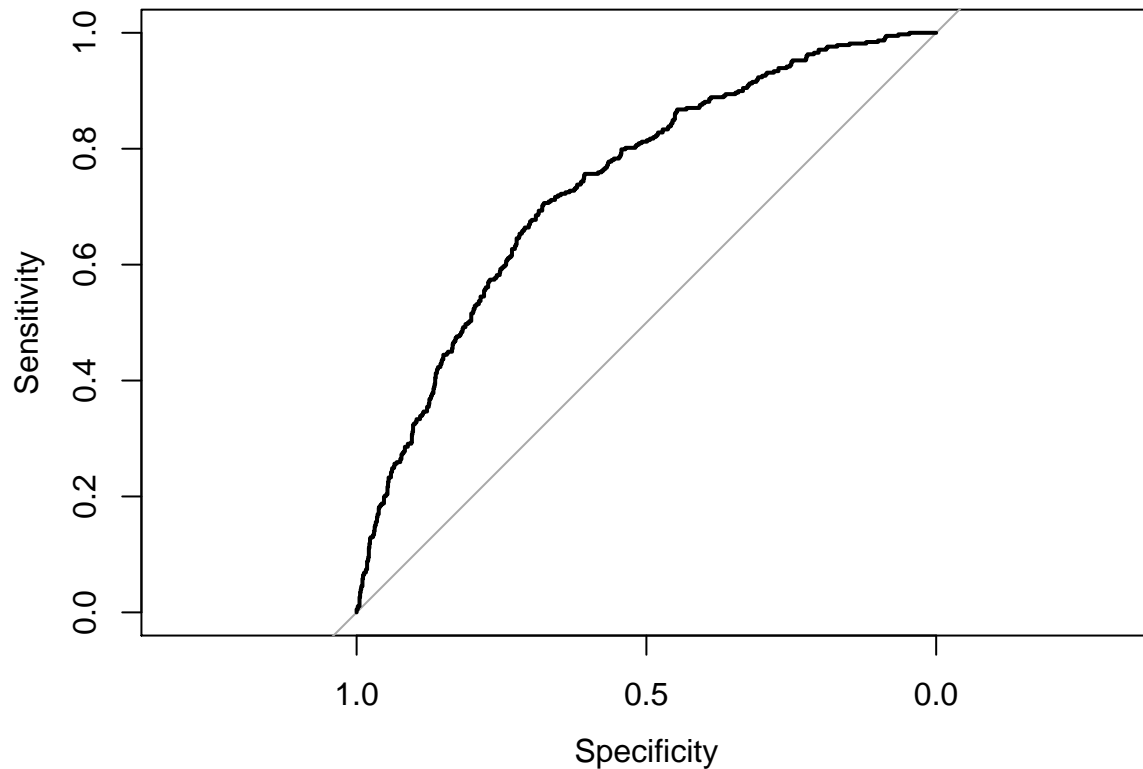
```
## [1] "AUC: 0.712440241385616"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1010 controls (dfPred_raw$class 0) < 379 cases (dfPred_raw$class 1).
## Area under the curve: 0.7124
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9482  -0.7685  -0.5774   0.9129   2.6524
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.388e-01  2.388e-01  -1.838 0.066056 .
## KIDSDRIV     2.164e-01  7.063e-02   3.063 0.002188 **
## AGE          -1.023e-02  4.549e-03  -2.250 0.024479 *
## HOMEKIDS      7.722e-02  3.780e-02   2.043 0.041058 *
## YOJ          -6.593e-03  8.708e-03  -0.757 0.448960
## INCOME        -4.845e-07  1.063e-06  -0.456 0.648529
## HOME_VAL      -2.160e-06  3.323e-07  -6.501 8.00e-11 ***
## TRAVTIME       7.342e-03  2.114e-03   3.473 0.000514 ***
## BLUEBOOK      -1.356e-05  4.676e-06  -2.900 0.003729 **
```

```

## TIF          -4.391e-02  8.635e-03  -5.085  3.67e-07 ***
## OLDCLAIM      9.170e-06  3.877e-06   2.365  0.018020 *
## CLM_FREQ      2.673e-01  3.253e-02   8.216  < 2e-16 ***
## MVR_PTS       1.284e-01  1.593e-02   8.060  7.65e-16 ***
## CAR_AGE       -2.512e-02  6.777e-03  -3.707  0.000210 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5827.4  on 5073  degrees of freedom
## Residual deviance: 5253.0  on 5060  degrees of freedom
##    (1338 observations deleted due to missingness)
## AIC: 5281
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 955 308
##           1  41  70
##
##           Accuracy : 0.746
##           95% CI : (0.7221, 0.7688)
##    No Information Rate : 0.7249
##    P-Value [Acc > NIR] : 0.04172
##
##           Kappa : 0.1844
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9588
##           Specificity : 0.1852
##           Pos Pred Value : 0.7561
##           Neg Pred Value : 0.6306
##           Prevalence : 0.7249
##           Detection Rate : 0.6951
##           Detection Prevalence : 0.9192
##           Balanced Accuracy : 0.5720
##
##           'Positive' Class : 0
##

```



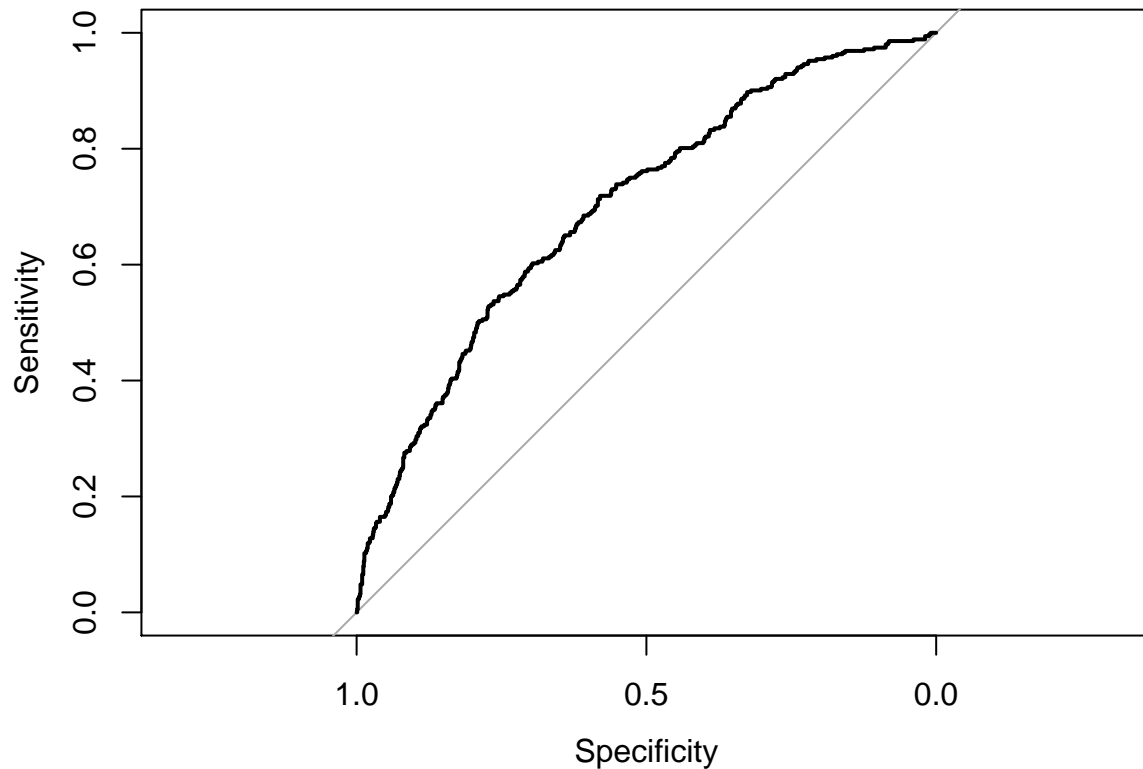


```
## [1] "AUC: 0.738113830985317"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 996 controls (dfPred_raw$class 0) < 378 cases (dfPred_raw$class 1).
## Area under the curve: 0.7381
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8087  -0.7712  -0.5714   0.9176   2.7042
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.210e-01  2.358e-01  -1.785  0.074262 .
## KIDSDRIV    2.807e-01  7.084e-02   3.963  7.39e-05 ***
## AGE        -1.208e-02  4.546e-03  -2.659  0.007845 **
## HOMEKIDS     4.977e-02  3.823e-02   1.302  0.192983
## YOJ         1.123e-03  8.738e-03   0.128  0.897778
## INCOME      -1.690e-07  1.034e-06  -0.163  0.870232
## HOME_VAL    -2.742e-06  3.305e-07  -8.297  < 2e-16 ***
## TRAVTIME     8.040e-03  2.139e-03   3.759  0.000171 ***
## BLUEBOOK    -1.309e-05  4.679e-06  -2.797  0.005150 **
```

```

## TIF          -4.150e-02  8.503e-03  -4.880 1.06e-06 ***
## OLDCLAIM     6.182e-06  3.984e-06   1.552 0.120723
## CLM_FREQ     2.768e-01  3.236e-02   8.554 < 2e-16 ***
## MVR_PTS      1.455e-01  1.591e-02   9.147 < 2e-16 ***
## CAR_AGE      -2.232e-02  6.734e-03  -3.314 0.000919 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5866.5  on 5050  degrees of freedom
## Residual deviance: 5231.1  on 5037  degrees of freedom
##    (1361 observations deleted due to missingness)
## AIC: 5259.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 974 276
##           1  71  76
##
##           Accuracy : 0.7516
##           95% CI : (0.7281, 0.7741)
##    No Information Rate : 0.748
##    P-Value [Acc > NIR] : 0.3926
##
##           Kappa : 0.1834
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9321
##           Specificity : 0.2159
##           Pos Pred Value : 0.7792
##           Neg Pred Value : 0.5170
##           Prevalence : 0.7480
##           Detection Rate : 0.6972
##           Detection Prevalence : 0.8948
##           Balanced Accuracy : 0.5740
##
##           'Positive' Class : 0
##

```

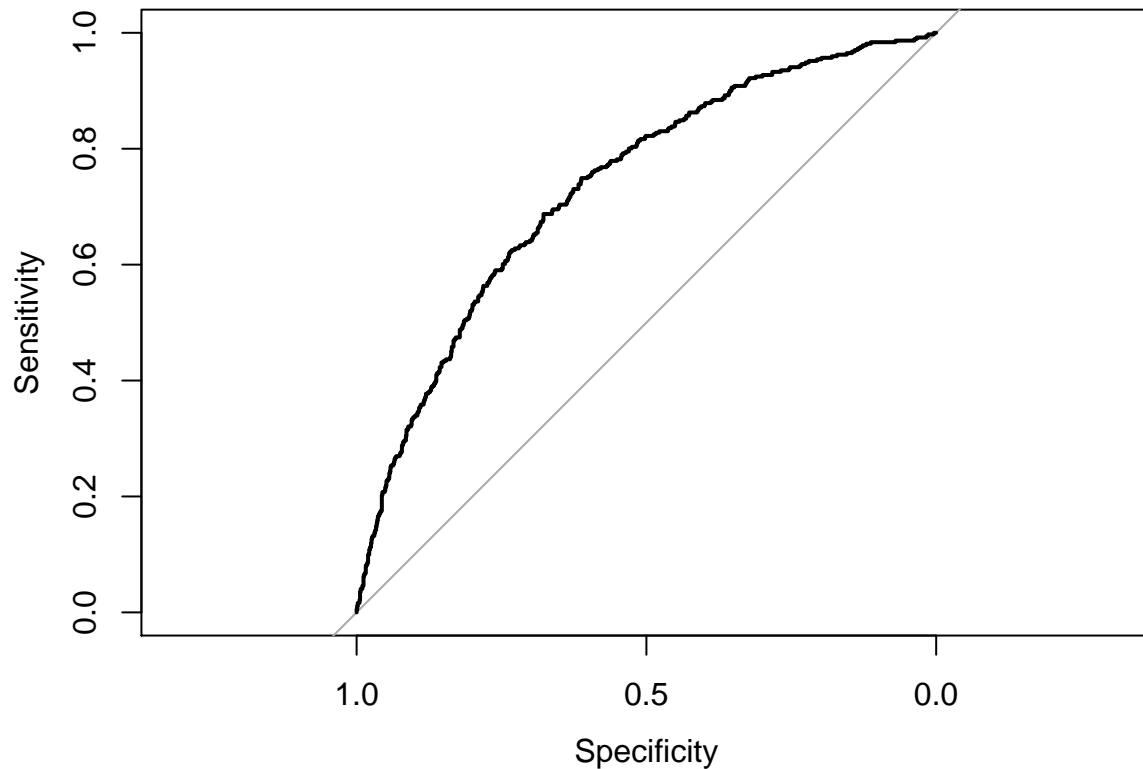


```
## [1] "AUC: 0.699611244019139"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1045 controls (dfPred_raw$class 0) < 352 cases (dfPred_raw$class 1).
## Area under the curve: 0.6996
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9506  -0.7717  -0.5796   0.9075   2.5939
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.949e-01  2.388e-01  -1.234  0.217029
## KIDSDRIV     2.551e-01  6.916e-02   3.688  0.000226 ***
## AGE          -1.468e-02  4.612e-03  -3.182  0.001464 **
## HOMEKIDS      6.470e-02  3.762e-02   1.720  0.085433 .
## YOJ           -7.734e-03  8.641e-03  -0.895  0.370765
## INCOME        -3.088e-07  1.047e-06  -0.295  0.767986
## HOME_VAL      -2.527e-06  3.304e-07  -7.648  2.04e-14 ***
## TRAVTIME      6.326e-03  2.126e-03   2.975  0.002930 **
## BLUEBOOK      -6.492e-06  4.611e-06  -1.408  0.159116
```

```

## TIF          -3.938e-02  8.477e-03  -4.645  3.40e-06 ***
## OLDCLAIM     3.275e-06  4.030e-06   0.813  0.416404
## CLM_FREQ     2.573e-01  3.224e-02   7.980  1.47e-15 ***
## MVR_PTS      1.457e-01  1.586e-02   9.186  < 2e-16 ***
## CAR_AGE      -2.196e-02  6.696e-03  -3.279  0.001042 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5837.0  on 5065  degrees of freedom
## Residual deviance: 5256.9  on 5052  degrees of freedom
##    (1346 observations deleted due to missingness)
## AIC: 5284.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 967 301
##           1  44  70
##
##           Accuracy : 0.7504
##           95% CI : (0.7267, 0.773)
##    No Information Rate : 0.7315
##    P-Value [Acc > NIR] : 0.06001
##
##           Kappa : 0.1859
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9565
##           Specificity : 0.1887
##           Pos Pred Value : 0.7626
##           Neg Pred Value : 0.6140
##           Prevalence : 0.7315
##           Detection Rate : 0.6997
##           Detection Prevalence : 0.9175
##           Balanced Accuracy : 0.5726
##
##           'Positive' Class : 0
##

```

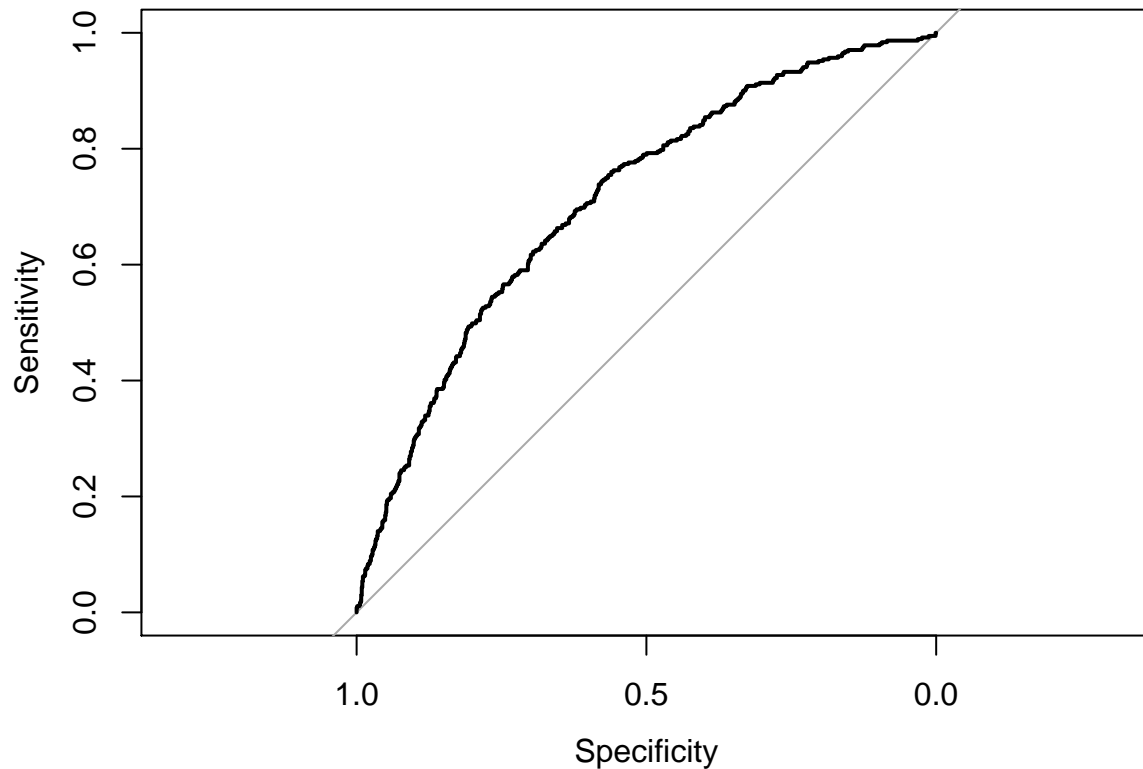


```
## [1] "AUC: 0.734707969745202"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1011 controls (dfPred_raw$class 0) < 371 cases (dfPred_raw$class 1).
## Area under the curve: 0.7347
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0653  -0.7630  -0.5708   0.8914   2.5198
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.390e-01  2.383e-01  -2.681 0.007345 **
## KIDSDRIV     2.670e-01  7.016e-02   3.806 0.000141 ***
## AGE        -7.204e-03  4.561e-03  -1.580 0.114199
## HOMEKIDS     7.863e-02  3.827e-02   2.055 0.039898 *
## YOJ        -3.231e-03  8.752e-03  -0.369 0.711988
## INCOME      -8.132e-07  1.057e-06  -0.769 0.441627
## HOME_VAL    -2.434e-06  3.333e-07  -7.301 2.86e-13 ***
## TRAVTIME     7.914e-03  2.136e-03   3.704 0.000212 ***
## BLUEBOOK    -1.267e-05  4.717e-06  -2.687 0.007214 **
```

```

## TIF          -4.366e-02  8.606e-03  -5.073  3.92e-07 ***
## OLDCLAIM      8.716e-06  3.971e-06   2.195  0.028179 *
## CLM_FREQ      2.744e-01  3.232e-02   8.491  < 2e-16 ***
## MVR_PTS       1.434e-01  1.586e-02   9.042  < 2e-16 ***
## CAR_AGE       -2.298e-02  6.821e-03  -3.369  0.000754 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5834.6  on 5061  degrees of freedom
## Residual deviance: 5213.2  on 5048  degrees of freedom
##    (1351 observations deleted due to missingness)
## AIC: 5241.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 962 302
##           1  53  69
##
##               Accuracy : 0.7439
##               95% CI : (0.72, 0.7667)
##    No Information Rate : 0.7323
##    P-Value [Acc > NIR] : 0.1737
##
##               Kappa : 0.17
##
## Mcnemar's Test P-Value : <2e-16
##
##               Sensitivity : 0.9478
##               Specificity : 0.1860
##               Pos Pred Value : 0.7611
##               Neg Pred Value : 0.5656
##               Prevalence : 0.7323
##               Detection Rate : 0.6941
##    Detection Prevalence : 0.9120
##    Balanced Accuracy : 0.5669
##
##    'Positive' Class : 0
##

```



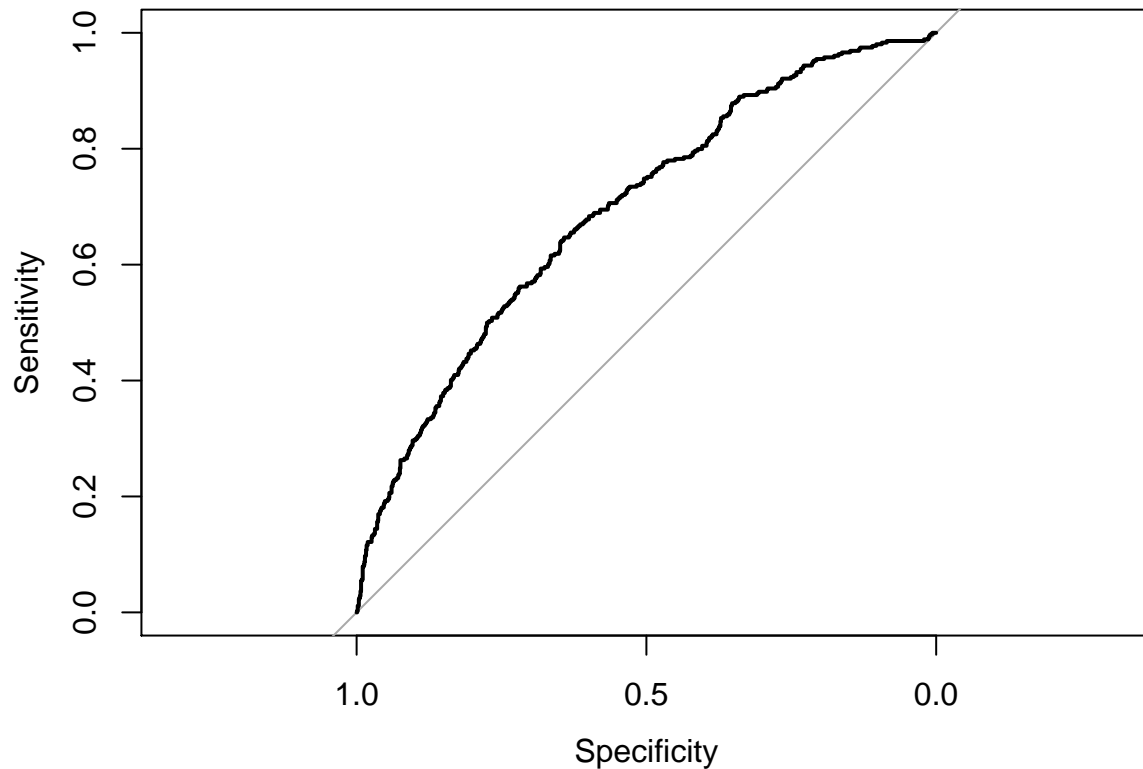
```
## [1] "AUC: 0.711900999827387"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1015 controls (dfPred_raw$class 0) < 371 cases (dfPred_raw$class 1).
## Area under the curve: 0.7119
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7735  -0.7739  -0.5684   0.9226   2.7311
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.329e-01  2.375e-01  -1.402  0.160994
## KIDSDRIV     2.838e-01  6.960e-02   4.078  4.54e-05 ***
## AGE         -1.223e-02  4.572e-03  -2.676  0.007460 **
## HOMEKIDS     5.826e-02  3.804e-02   1.531  0.125674
## YOJ         -2.010e-03  8.771e-03  -0.229  0.818767
## INCOME      -1.375e-07  1.033e-06  -0.133  0.894104
## HOME_VAL    -2.720e-06  3.306e-07  -8.227  < 2e-16 ***
## TRAVTIME     7.667e-03  2.131e-03   3.598  0.000321 ***
## BLUEBOOK    -1.473e-05  4.677e-06  -3.150  0.001631 **
```

```

## TIF          -4.538e-02  8.564e-03  -5.299 1.16e-07 ***
## OLDCLAIM      6.462e-06  3.994e-06   1.618 0.105686
## CLM_FREQ      2.926e-01  3.233e-02   9.052 < 2e-16 ***
## MVR_PTS       1.333e-01  1.596e-02   8.355 < 2e-16 ***
## CAR_AGE       -2.287e-02  6.744e-03  -3.392 0.000695 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5861.2  on 5048  degrees of freedom
## Residual deviance: 5219.8  on 5035  degrees of freedom
##    (1363 observations deleted due to missingness)
## AIC: 5247.8
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 982 278
##           1  63  76
##
##           Accuracy : 0.7563
##           95% CI : (0.7329, 0.7786)
##    No Information Rate : 0.747
##    P-Value [Acc > NIR] : 0.2217
##
##           Kappa : 0.1932
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9397
##           Specificity : 0.2147
##           Pos Pred Value : 0.7794
##           Neg Pred Value : 0.5468
##           Prevalence : 0.7470
##           Detection Rate : 0.7019
##           Detection Prevalence : 0.9006
##           Balanced Accuracy : 0.5772
##
##           'Positive' Class : 0
##

```



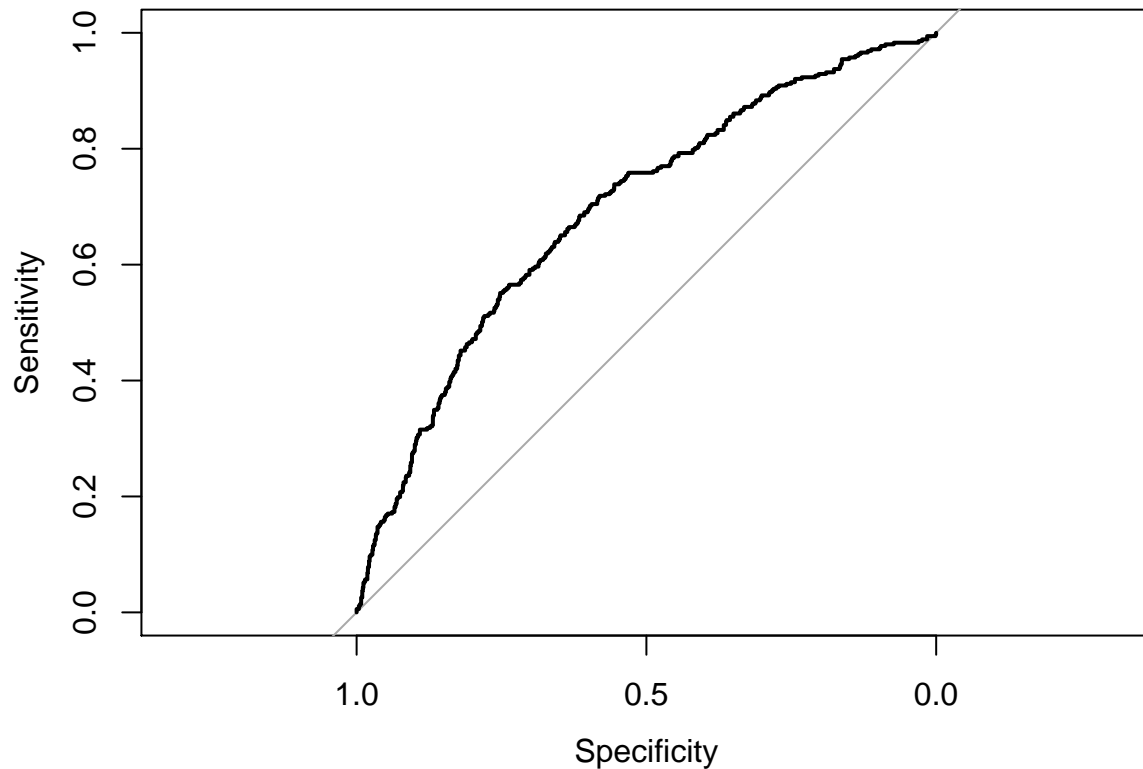


```
## [1] "AUC: 0.693552834320006"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1045 controls (dfPred_raw$class 0) < 354 cases (dfPred_raw$class 1).
## Area under the curve: 0.6936
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0989  -0.7677  -0.5662   0.8903   2.5416
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.003e-01  2.375e-01  -2.527 0.011503 *
## KIDSDRIV     2.380e-01  7.117e-02   3.345 0.000824 ***
## AGE          -6.506e-03  4.567e-03  -1.424 0.154311
## HOMEKIDS      1.047e-01  3.807e-02   2.751 0.005945 **
## YOJ          -9.366e-03  8.734e-03  -1.072 0.283541
## INCOME       -3.447e-07  1.057e-06  -0.326 0.744255
## HOME_VAL     -2.664e-06  3.324e-07  -8.015 1.10e-15 ***
## TRAVTIME      8.821e-03  2.140e-03   4.123 3.74e-05 ***
## BLUEBOOK     -1.432e-05  4.713e-06  -3.039 0.002377 **
```

```

## TIF          -4.484e-02  8.616e-03  -5.204 1.95e-07 ***
## OLDCLAIM      6.041e-06  3.933e-06   1.536 0.124514
## CLM_FREQ      2.823e-01  3.229e-02   8.744 < 2e-16 ***
## MVR_PTS       1.478e-01  1.595e-02   9.267 < 2e-16 ***
## CAR_AGE       -2.307e-02  6.823e-03  -3.381 0.000721 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5878.3  on 5069  degrees of freedom
## Residual deviance: 5221.4  on 5056  degrees of freedom
##    (1343 observations deleted due to missingness)
## AIC: 5249.4
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 957 286
##           1  69  66
##
##           Accuracy : 0.7424
##           95% CI : (0.7184, 0.7653)
##    No Information Rate : 0.7446
##    P-Value [Acc > NIR] : 0.5874
##
##           Kappa : 0.1508
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9327
##           Specificity : 0.1875
##           Pos Pred Value : 0.7699
##           Neg Pred Value : 0.4889
##           Prevalence : 0.7446
##           Detection Rate : 0.6945
##           Detection Prevalence : 0.9020
##           Balanced Accuracy : 0.5601
##
##           'Positive' Class : 0
##

```

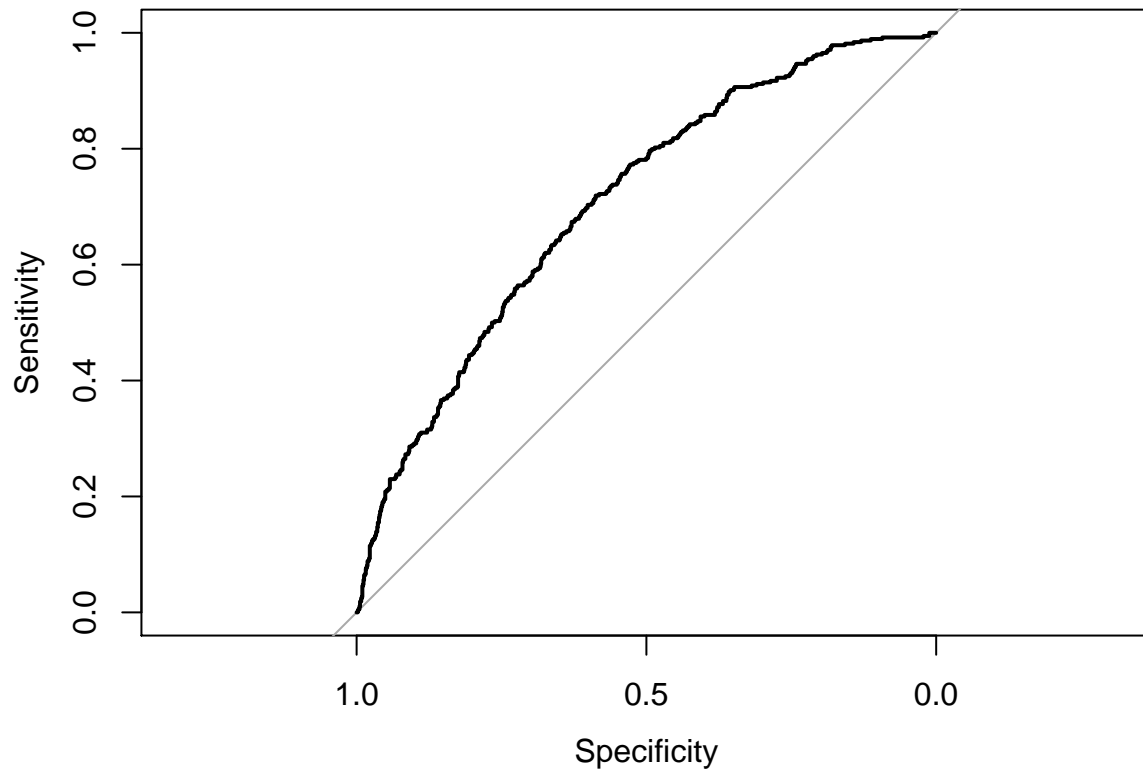


```
## [1] "AUC: 0.692877237285132"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1026 controls (dfPred_raw$class 0) < 352 cases (dfPred_raw$class 1).
## Area under the curve: 0.6929
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7573  -0.7622  -0.5687   0.8928   2.6942
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.026e-01  2.400e-01  -2.094 0.036268 *
## KIDSDRIV     3.071e-01  6.930e-02   4.431 9.37e-06 ***
## AGE          -1.153e-02  4.605e-03  -2.503 0.012297 *
## HOMEKIDS      5.509e-02  3.800e-02   1.450 0.147182
## YOJ           1.542e-03  8.801e-03   0.175 0.860884
## INCOME        -5.857e-07  1.039e-06  -0.564 0.572739
## HOME_VAL      -2.453e-06  3.329e-07  -7.368 1.74e-13 ***
## TRAVTIME       7.696e-03  2.152e-03   3.576 0.000349 ***
## BLUEBOOK      -1.273e-05  4.702e-06  -2.706 0.006805 **
```

```

## TIF          -4.149e-02  8.572e-03  -4.841  1.29e-06 ***
## OLDCLAIM      7.490e-06  3.982e-06   1.881  0.059962 .
## CLM_FREQ      3.016e-01  3.232e-02   9.330  < 2e-16 ***
## MVR_PTS       1.315e-01  1.595e-02   8.245  < 2e-16 ***
## CAR_AGE       -2.197e-02  6.773e-03  -3.243  0.001182 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5829.0  on 5062  degrees of freedom
## Residual deviance: 5207.9  on 5049  degrees of freedom
##    (1349 observations deleted due to missingness)
## AIC: 5235.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 958 295
##           1  53  79
##
##           Accuracy : 0.7487
##           95% CI : (0.725, 0.7714)
##    No Information Rate : 0.73
##    P-Value [Acc > NIR] : 0.06057
##
##           Kappa : 0.1995
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9476
##           Specificity : 0.2112
##           Pos Pred Value : 0.7646
##           Neg Pred Value : 0.5985
##           Prevalence : 0.7300
##           Detection Rate : 0.6917
##           Detection Prevalence : 0.9047
##           Balanced Accuracy : 0.5794
##
##           'Positive' Class : 0
##

```

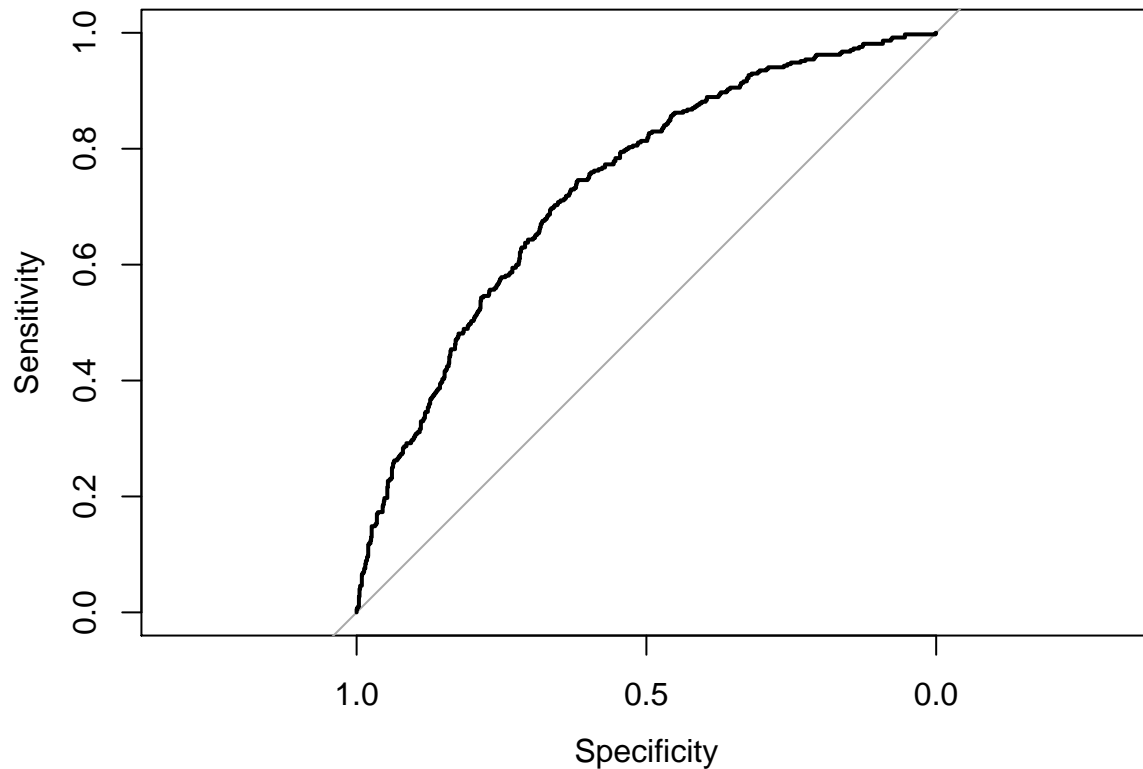


```
## [1] "AUC: 0.706681053862063"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1011 controls (dfPred_raw$class 0) < 374 cases (dfPred_raw$class 1).
## Area under the curve: 0.7067
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9870  -0.7642  -0.5773   0.9232   2.6354
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.430e-01  2.383e-01  -1.859 0.063000 .
## KIDSDRIV     2.305e-01  7.009e-02   3.289 0.001004 **
## AGE          -1.058e-02  4.547e-03  -2.327 0.019992 *
## HOMEKIDS      7.170e-02  3.790e-02   1.892 0.058515 .
## YOJ          -8.193e-03  8.706e-03  -0.941 0.346662
## INCOME       -9.251e-07  1.055e-06  -0.877 0.380740
## HOME_VAL     -2.111e-06  3.324e-07  -6.352 2.13e-10 ***
## TRAVTIME      7.728e-03  2.119e-03   3.648 0.000264 ***
## BLUEBOOK     -1.098e-05  4.660e-06  -2.357 0.018413 *
```

```

## TIF          -4.742e-02  8.652e-03  -5.481 4.24e-08 ***
## OLDCLAIM      8.081e-06  3.880e-06   2.083 0.037286 *
## CLM_FREQ      2.904e-01  3.223e-02   9.010 < 2e-16 ***
## MVR_PTS       1.239e-01  1.589e-02   7.800 6.19e-15 ***
## CAR_AGE       -2.237e-02  6.757e-03  -3.311 0.000930 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5845.8  on 5076  degrees of freedom
## Residual deviance: 5263.9  on 5063  degrees of freedom
##    (1335 observations deleted due to missingness)
## AIC: 5291.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 954 299
##           1  47  71
##
##           Accuracy : 0.7476
##           95% CI : (0.7238, 0.7704)
##    No Information Rate : 0.7301
##    P-Value [Acc > NIR] : 0.07566
##
##           Kappa : 0.1846
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9530
##           Specificity : 0.1919
##           Pos Pred Value : 0.7614
##           Neg Pred Value : 0.6017
##           Prevalence : 0.7301
##           Detection Rate : 0.6958
##    Detection Prevalence : 0.9139
##           Balanced Accuracy : 0.5725
##
##           'Positive' Class : 0
##

```



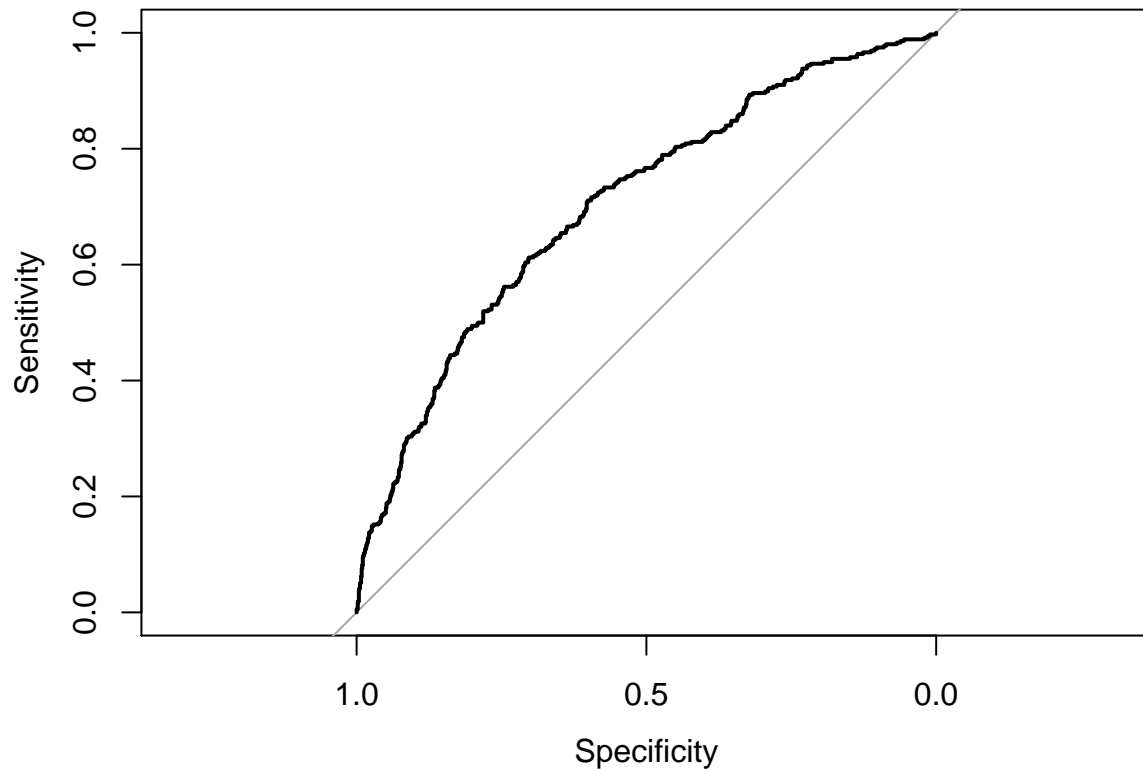
```
## [1] "AUC: 0.733517833517834"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1001 controls (dfPred_raw$class 0) < 370 cases (dfPred_raw$class 1).
## Area under the curve: 0.7335
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0139  -0.7662  -0.5730   0.9116   2.6767
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.927e-01  2.360e-01  -1.240  0.214943
## KIDSDRIV     2.429e-01  7.072e-02   3.435  0.000592 ***
## AGE         -1.126e-02  4.539e-03  -2.480  0.013134 *
## HOMEKIDS      5.774e-02  3.840e-02   1.504  0.132684
## YOJ          -6.407e-03  8.738e-03  -0.733  0.463415
## INCOME       -2.229e-07  1.041e-06  -0.214  0.830392
## HOME_VAL     -2.528e-06  3.298e-07  -7.665  1.79e-14 ***
## TRAVTIME      7.650e-03  2.125e-03   3.600  0.000318 ***
## BLUEBOOK     -1.586e-05  4.692e-06  -3.380  0.000726 ***
```

```

## TIF          -5.018e-02  8.573e-03  -5.853 4.84e-09 ***
## OLDCLAIM      6.909e-06  3.934e-06   1.756 0.079039 .
## CLM_FREQ      2.680e-01  3.234e-02   8.287 < 2e-16 ***
## MVR_PTS       1.399e-01  1.587e-02   8.811 < 2e-16 ***
## CAR_AGE       -2.322e-02  6.748e-03  -3.441 0.000580 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5865.3  on 5061  degrees of freedom
## Residual deviance: 5239.6  on 5048  degrees of freedom
##    (1350 observations deleted due to missingness)
## AIC: 5267.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 970 286
##           1  60  70
##
##           Accuracy : 0.7504
##           95% CI : (0.7267, 0.773)
##    No Information Rate : 0.7431
##    P-Value [Acc > NIR] : 0.2807
##
##           Kappa : 0.1747
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9417
##           Specificity : 0.1966
##           Pos Pred Value : 0.7723
##           Neg Pred Value : 0.5385
##           Prevalence : 0.7431
##           Detection Rate : 0.6999
##           Detection Prevalence : 0.9062
##           Balanced Accuracy : 0.5692
##
##           'Positive' Class : 0
##

```



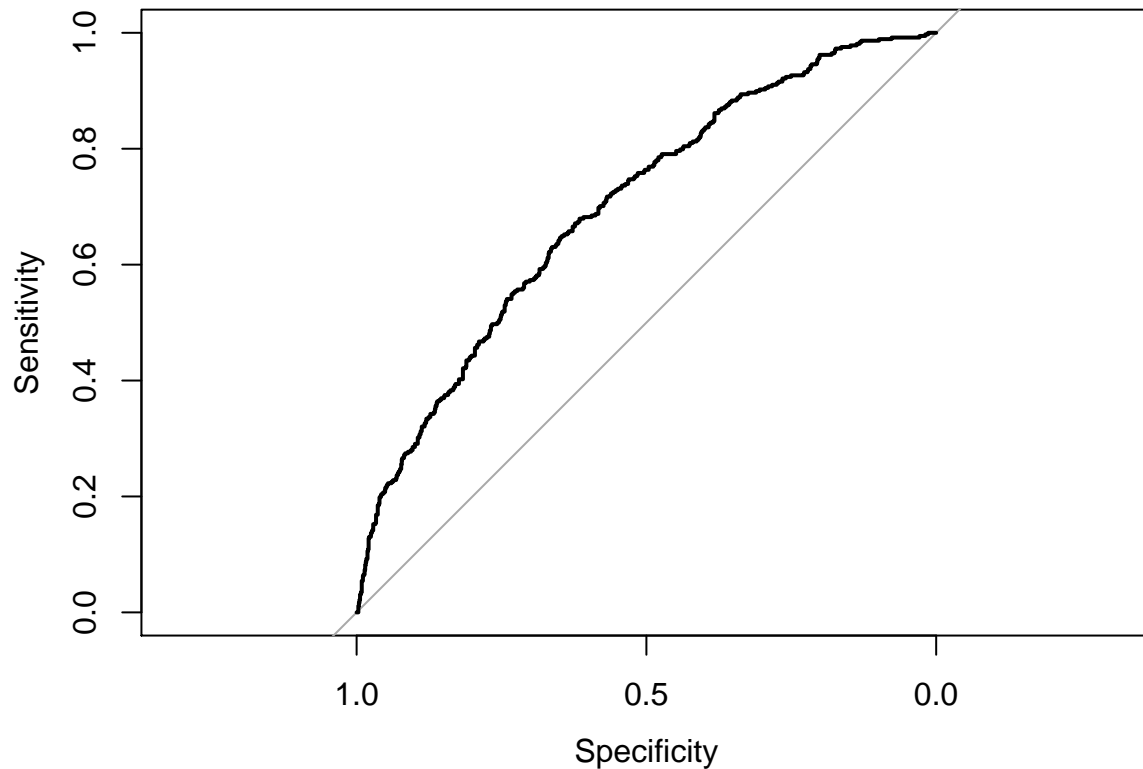


```
## [1] "AUC: 0.704505290716701"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1030 controls (dfPred_raw$class 0) < 356 cases (dfPred_raw$class 1).
## Area under the curve: 0.7045
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7786  -0.7691  -0.5682   0.8993   2.7168
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.603e-01  2.382e-01  -1.513 0.130277
## KIDSDRIV     2.663e-01  6.939e-02   3.838 0.000124 ***
## AGE          -1.136e-02  4.599e-03  -2.470 0.013514 *
## HOMEKIDS      6.815e-02  3.774e-02   1.806 0.070988 .
## YOJ           -7.231e-03  8.768e-03  -0.825 0.409507
## INCOME        2.560e-07  1.045e-06   0.245 0.806440
## HOME_VAL      -2.642e-06  3.313e-07  -7.975 1.53e-15 ***
## TRAVTIME      8.179e-03  2.146e-03   3.812 0.000138 ***
## BLUEBOOK     -1.571e-05  4.687e-06  -3.352 0.000803 ***
```

```

## TIF          -4.421e-02  8.598e-03  -5.142  2.71e-07 ***
## OLDCLAIM      5.839e-06  3.921e-06   1.489  0.136444
## CLM_FREQ      2.753e-01  3.237e-02   8.505  < 2e-16 ***
## MVR_PTS       1.391e-01  1.589e-02   8.752  < 2e-16 ***
## CAR_AGE       -2.485e-02  6.774e-03  -3.669  0.000243 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5843.2  on 5065  degrees of freedom
## Residual deviance: 5214.8  on 5052  degrees of freedom
##    (1346 observations deleted due to missingness)
## AIC: 5242.8
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 964 291
##           1  50  77
##
##           Accuracy : 0.7533
##           95% CI : (0.7296, 0.7758)
##    No Information Rate : 0.7337
##    P-Value [Acc > NIR] : 0.05254
##
##           Kappa : 0.2021
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9507
##           Specificity : 0.2092
##           Pos Pred Value : 0.7681
##           Neg Pred Value : 0.6063
##           Prevalence : 0.7337
##           Detection Rate : 0.6975
##           Detection Prevalence : 0.9081
##           Balanced Accuracy : 0.5800
##
##           'Positive' Class : 0
##

```

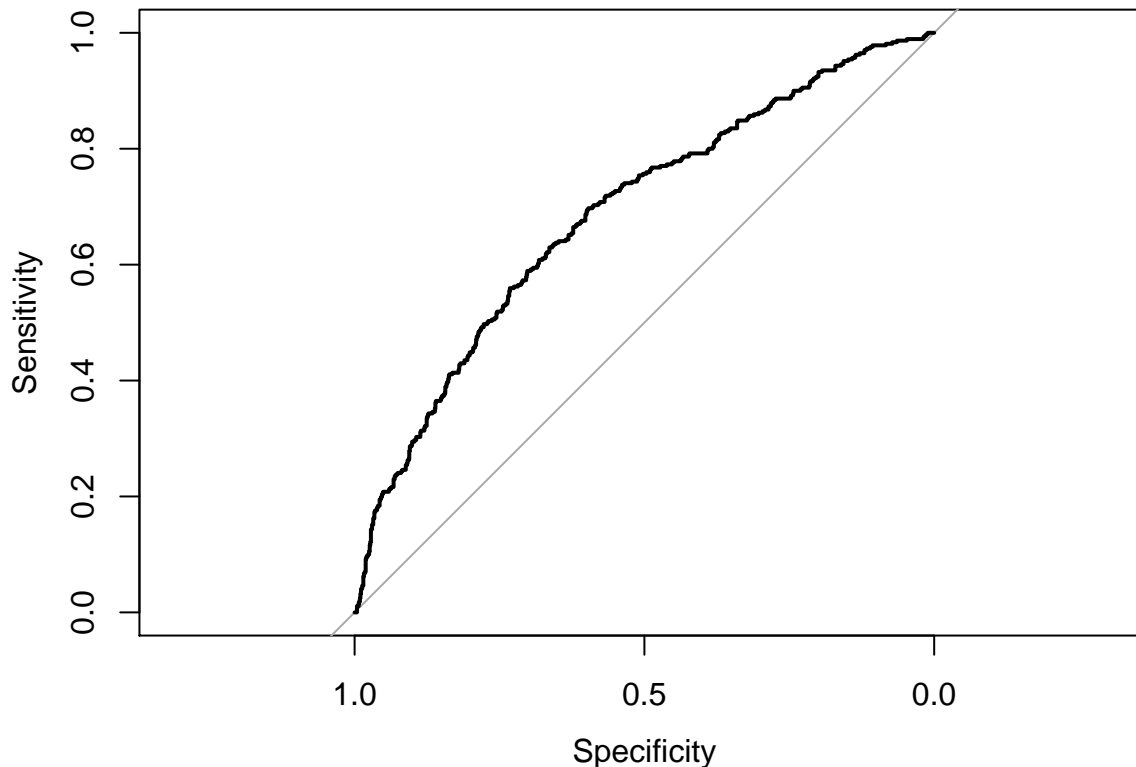


```
## [1] "AUC: 0.699484392419175"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1014 controls (dfPred_raw$class 0) < 368 cases (dfPred_raw$class 1).
## Area under the curve: 0.6995
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0401  -0.7600  -0.5585   0.8682   2.7581
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.235e-01  2.407e-01  -1.344  0.178943
## KIDSDRIV     2.248e-01  7.124e-02   3.155  0.001603 **
## AGE          -1.090e-02  4.623e-03  -2.357  0.018416 *
## HOMEKIDS      7.796e-02  3.829e-02   2.036  0.041762 *
## YOJ          -6.684e-03  8.817e-03  -0.758  0.448445
## INCOME        1.339e-07  1.044e-06   0.128  0.897991
## HOME_VAL     -2.517e-06  3.323e-07  -7.574  3.61e-14 ***
## TRAVTIME      7.981e-03  2.155e-03   3.703  0.000213 ***
## BLUEBOOK     -1.865e-05  4.746e-06  -3.930  8.49e-05 ***
```

```

## TIF          -4.831e-02  8.640e-03  -5.591  2.26e-08 ***
## OLDCLAIM      5.703e-06  3.966e-06   1.438  0.150422
## CLM_FREQ      2.940e-01  3.246e-02   9.056  < 2e-16 ***
## MVR_PTS       1.410e-01  1.619e-02   8.710  < 2e-16 ***
## CAR_AGE       -2.647e-02  6.837e-03  -3.872  0.000108 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5848.2  on 5080  degrees of freedom
## Residual deviance: 5196.2  on 5067  degrees of freedom
##    (1331 observations deleted due to missingness)
## AIC: 5224.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  928  283
##           1   69   87
##
##           Accuracy : 0.7425
##           95% CI : (0.7185, 0.7655)
##    No Information Rate : 0.7293
##    P-Value [Acc > NIR] : 0.1432
##
##           Kappa : 0.2028
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9308
##           Specificity : 0.2351
##           Pos Pred Value : 0.7663
##           Neg Pred Value : 0.5577
##           Prevalence : 0.7293
##           Detection Rate : 0.6789
##           Detection Prevalence : 0.8859
##           Balanced Accuracy : 0.5830
##
##           'Positive' Class : 0
##

```

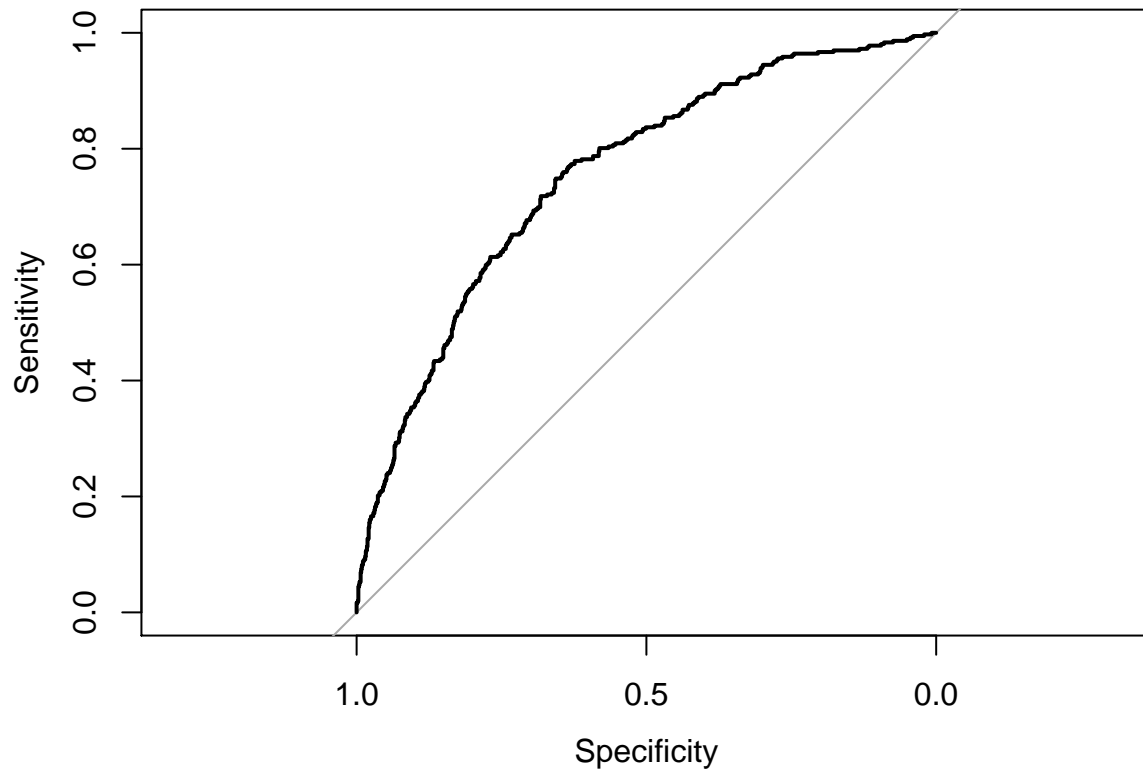


```
## [1] "AUC: 0.686771124183361"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 997 controls (dfPred_raw$class 0) < 370 cases (dfPred_raw$class 1).
## Area under the curve: 0.6868
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9720  -0.7778  -0.5885   0.9558   2.6485
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.351e-01  2.373e-01  -1.412  0.157866
## KIDSDRIV     2.316e-01  6.937e-02   3.338  0.000843 ***
## AGE         -1.107e-02  4.538e-03  -2.440  0.014697 *
## HOMEKIDS      6.375e-02  3.724e-02   1.712  0.086930 .
## YOJ          -7.419e-03  8.706e-03  -0.852  0.394082
## INCOME       -1.072e-06  1.052e-06  -1.019  0.308029
## HOME_VAL     -2.234e-06  3.300e-07  -6.768  1.30e-11 ***
## TRAVTIME      8.165e-03  2.102e-03   3.885  0.000102 ***
## BLUEBOOK     -1.012e-05  4.577e-06  -2.211  0.027008 *
```

```

## TIF          -4.807e-02  8.649e-03  -5.558 2.73e-08 ***
## OLDCLAIM     9.151e-06  3.892e-06   2.351 0.018728 *
## CLM_FREQ     2.348e-01  3.239e-02   7.249 4.20e-13 ***
## MVR_PTS      1.245e-01  1.592e-02   7.820 5.30e-15 ***
## CAR_AGE      -2.532e-02  6.717e-03  -3.770 0.000163 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5861.6  on 5075  degrees of freedom
## Residual deviance: 5313.0  on 5062  degrees of freedom
##    (1336 observations deleted due to missingness)
## AIC: 5341
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 967 287
##           1  43  75
##
##           Accuracy : 0.7595
##           95% CI : (0.736, 0.7819)
##    No Information Rate : 0.7362
##    P-Value [Acc > NIR] : 0.02598
##
##           Kappa : 0.21
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9574
##           Specificity : 0.2072
##           Pos Pred Value : 0.7711
##           Neg Pred Value : 0.6356
##           Prevalence : 0.7362
##           Detection Rate : 0.7048
##           Detection Prevalence : 0.9140
##           Balanced Accuracy : 0.5823
##
##           'Positive' Class : 0
##

```



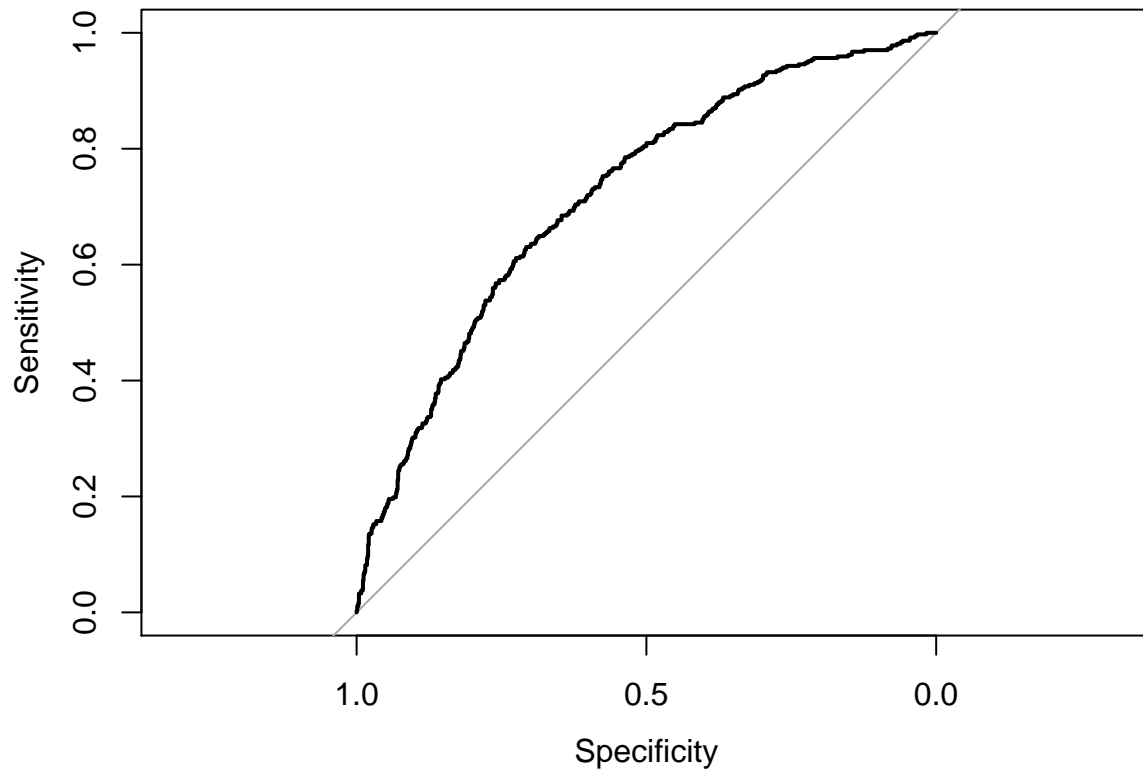
```
## [1] "AUC: 0.754644166074066"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1010 controls (dfPred_raw$class 0) < 362 cases (dfPred_raw$class 1).
## Area under the curve: 0.7546
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8990  -0.7618  -0.5739   0.9092   2.6067
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.832e-01  2.391e-01  -1.603  0.109008
## KIDSDRIV     1.450e-01  7.134e-02   2.032  0.042166 *
## AGE         -1.216e-02  4.572e-03  -2.659  0.007844 **
## HOMEKIDS      8.442e-02  3.768e-02   2.241  0.025049 *
## YOJ          -1.016e-02  8.676e-03  -1.171  0.241597
## INCOME        1.929e-07  1.046e-06   0.184  0.853720
## HOME_VAL     -2.350e-06  3.308e-07  -7.105  1.20e-12 ***
## TRAVTIME      7.750e-03  2.125e-03   3.648  0.000265 ***
## BLUEBOOK     -1.117e-05  4.641e-06  -2.408  0.016056 *
```

```

## TIF          -4.298e-02  8.638e-03  -4.976  6.49e-07  ***
## OLDCLAIM      5.189e-06  3.967e-06   1.308  0.190821
## CLM_FREQ      2.762e-01  3.215e-02   8.591  < 2e-16  ***
## MVR_PTS       1.474e-01  1.595e-02   9.236  < 2e-16  ***
## CAR_AGE       -2.543e-02  6.789e-03  -3.746  0.000180  ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5849.3  on 5075  degrees of freedom
## Residual deviance: 5245.9  on 5062  degrees of freedom
##    (1336 observations deleted due to missingness)
## AIC: 5273.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  948  297
##           1   56   71
##
##           Accuracy : 0.7427
##           95% CI : (0.7187, 0.7657)
##    No Information Rate : 0.7318
##    P-Value [Acc > NIR] : 0.1887
##
##           Kappa : 0.1731
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9442
##           Specificity : 0.1929
##           Pos Pred Value : 0.7614
##           Neg Pred Value : 0.5591
##           Prevalence : 0.7318
##           Detection Rate : 0.6910
##           Detection Prevalence : 0.9074
##           Balanced Accuracy : 0.5686
##
##           'Positive' Class : 0
##

```



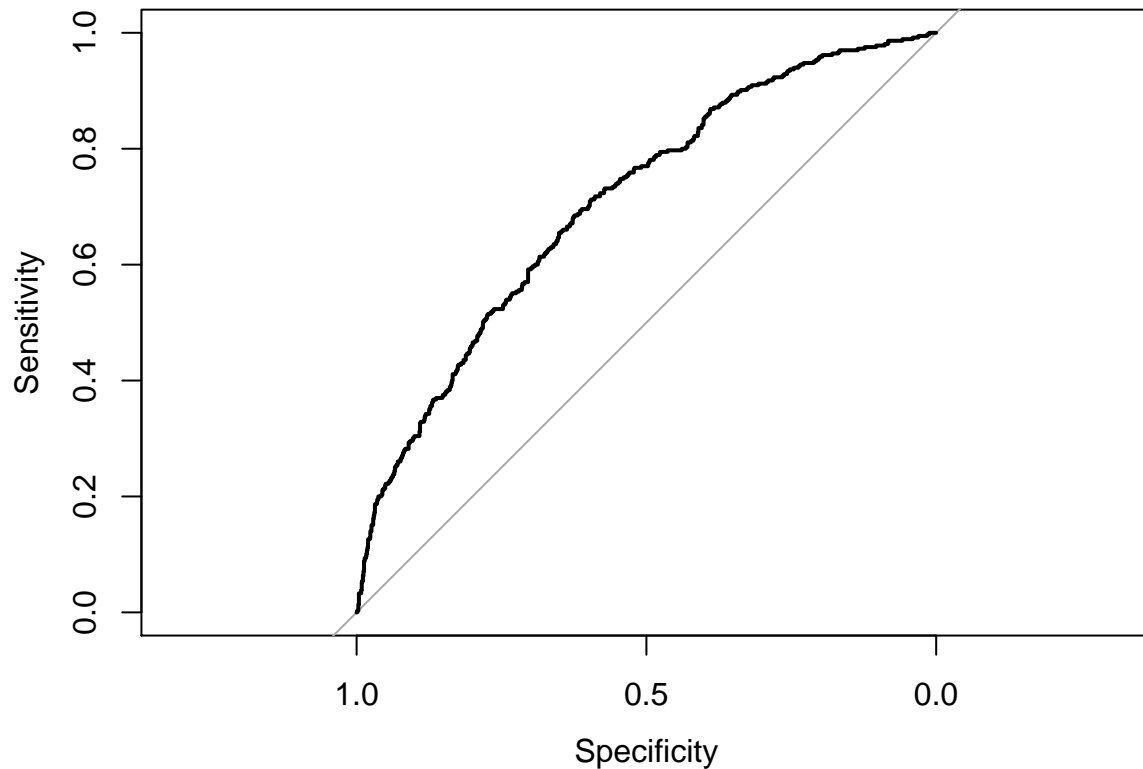


```
## [1] "AUC: 0.719583622033605"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1004 controls (dfPred_raw$class 0) < 368 cases (dfPred_raw$class 1).
## Area under the curve: 0.7196
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7764  -0.7657  -0.5708   0.9195   2.6559
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.441e-01  2.392e-01  -1.439  0.150233
## KIDSDRIV     2.239e-01  6.958e-02   3.218  0.001292 **
## AGE         -1.239e-02  4.596e-03  -2.696  0.007019 **
## HOMEKIDS      7.250e-02  3.769e-02   1.923  0.054441 .
## YOJ          -8.881e-03  8.748e-03  -1.015  0.310001
## INCOME       -9.391e-08  1.045e-06  -0.090  0.928372
## HOME_VAL     -2.485e-06  3.322e-07  -7.481  7.39e-14 ***
## TRAVTIME      8.340e-03  2.136e-03   3.904  9.48e-05 ***
## BLUEBOOK     -1.180e-05  4.644e-06  -2.540  0.011072 *
```

```

## TIF          -4.560e-02  8.666e-03  -5.262 1.43e-07 ***
## OLDCLAIM      5.367e-06  3.947e-06   1.360 0.173927
## CLM_FREQ      3.034e-01  3.217e-02   9.430 < 2e-16 ***
## MVR_PTS       1.302e-01  1.604e-02   8.121 4.64e-16 ***
## CAR_AGE       -2.576e-02  6.782e-03  -3.798 0.000146 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5851.2  on 5068  degrees of freedom
## Residual deviance: 5233.7  on 5055  degrees of freedom
##    (1343 observations deleted due to missingness)
## AIC: 5261.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 958 284
##           1  56  81
##
##           Accuracy : 0.7534
##           95% CI : (0.7298, 0.776)
##    No Information Rate : 0.7353
##    P-Value [Acc > NIR] : 0.06659
##
##           Kappa : 0.2083
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9448
##           Specificity : 0.2219
##           Pos Pred Value : 0.7713
##           Neg Pred Value : 0.5912
##           Prevalence : 0.7353
##           Detection Rate : 0.6947
##    Detection Prevalence : 0.9007
##           Balanced Accuracy : 0.5833
##
##           'Positive' Class : 0
##

```

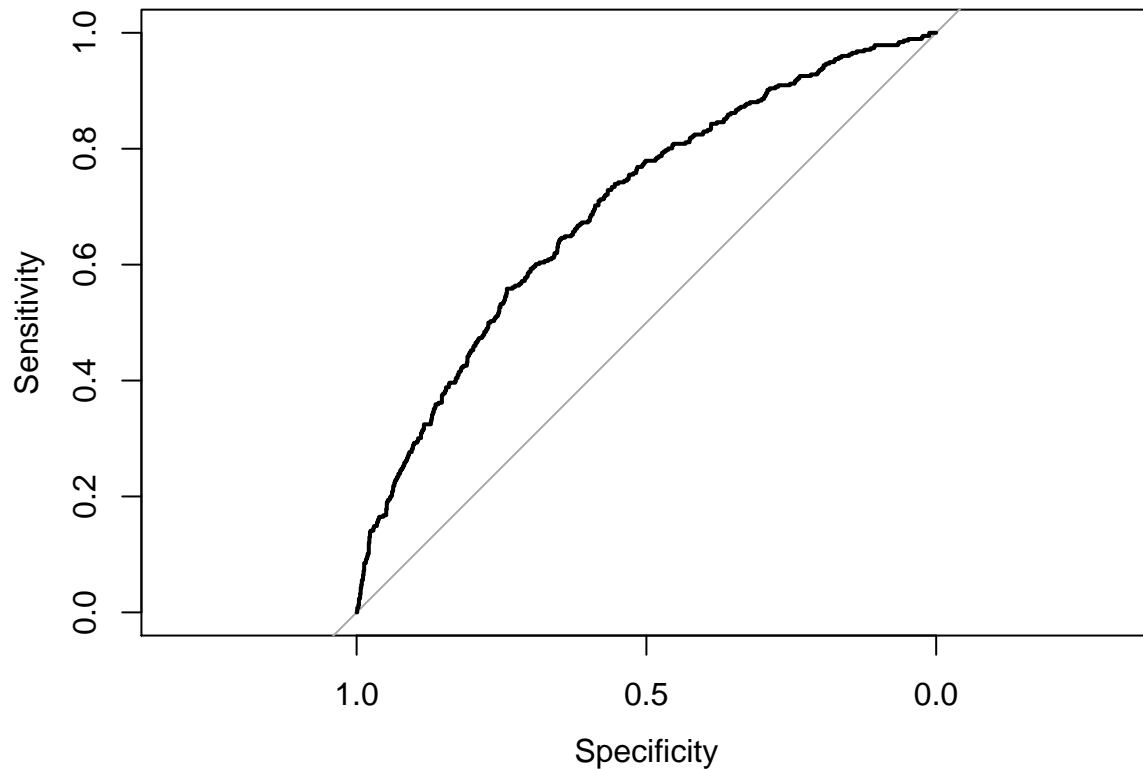


```
## [1] "AUC: 0.707776066574802"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1014 controls (dfPred_raw$class 0) < 365 cases (dfPred_raw$class 1).
## Area under the curve: 0.7078
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9517  -0.7588  -0.5637   0.8829   2.6904
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.043e-01  2.408e-01  -1.264  0.206356
## KIDSDRIV     1.911e-01  7.030e-02   2.718  0.006559 **
## AGE         -1.258e-02  4.621e-03  -2.722  0.006483 **
## HOMEKIDS      7.273e-02  3.794e-02   1.917  0.055272 .
## YOJ          -7.670e-03  8.763e-03  -0.875  0.381465
## INCOME        4.806e-07  1.037e-06   0.464  0.642975
## HOME_VAL     -2.476e-06  3.308e-07  -7.485  7.17e-14 ***
## TRAVTIME      7.417e-03  2.145e-03   3.458  0.000544 ***
## BLUEBOOK     -1.571e-05  4.696e-06  -3.346  0.000820 ***
```

```

## TIF          -4.578e-02  8.628e-03  -5.306  1.12e-07 ***
## OLDCLAIM      5.129e-06  4.025e-06   1.274  0.202559
## CLM_FREQ      2.789e-01  3.221e-02   8.658  < 2e-16 ***
## MVR_PTS       1.497e-01  1.600e-02   9.359  < 2e-16 ***
## CAR_AGE       -2.604e-02  6.822e-03  -3.817  0.000135 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5830.9  on 5072  degrees of freedom
## Residual deviance: 5198.5  on 5059  degrees of freedom
##    (1339 observations deleted due to missingness)
## AIC: 5226.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 938 300
##           1  61  76
##
##               Accuracy : 0.7375
##               95% CI : (0.7133, 0.7605)
##    No Information Rate : 0.7265
##    P-Value [Acc > NIR] : 0.1904
##
##               Kappa : 0.1759
##
## Mcnemar's Test P-Value : <2e-16
##
##               Sensitivity : 0.9389
##               Specificity : 0.2021
##               Pos Pred Value : 0.7577
##               Neg Pred Value : 0.5547
##               Prevalence : 0.7265
##               Detection Rate : 0.6822
##    Detection Prevalence : 0.9004
##    Balanced Accuracy : 0.5705
##
##           'Positive' Class : 0
##

```

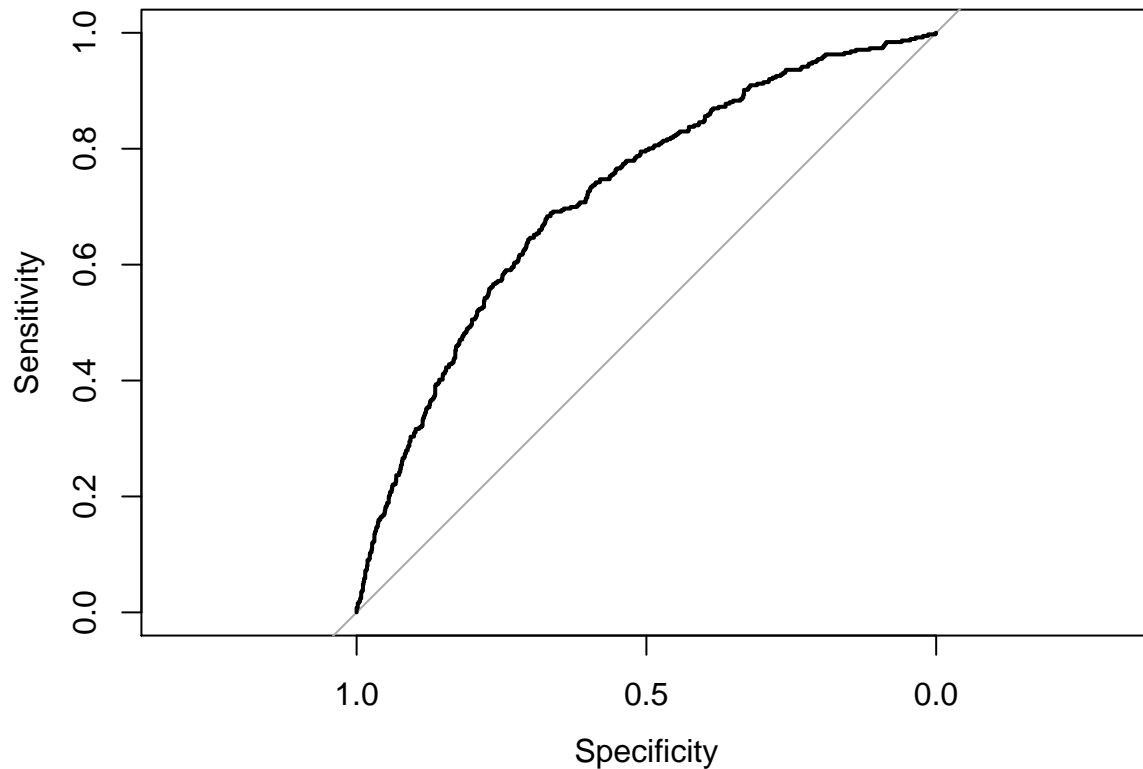


```
## [1] "AUC: 0.696686047749878"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 999 controls (dfPred_raw$class 0) < 376 cases (dfPred_raw$class 1).
## Area under the curve: 0.6967
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9997  -0.7670  -0.5721   0.8951   2.6593
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.178e-01  2.397e-01  -2.577  0.00996 **
## KIDSDRIV     2.101e-01  7.094e-02   2.962  0.00305 **
## AGE         -7.856e-03  4.558e-03  -1.724  0.08479 .
## HOMEKIDS     8.771e-02  3.811e-02   2.301  0.02137 *
## YOJ         -3.130e-03  8.732e-03  -0.358  0.72000
## INCOME      -7.066e-07  1.064e-06  -0.664  0.50674
## HOME_VAL    -2.330e-06  3.342e-07  -6.974 3.09e-12 ***
## TRAVTIME     7.675e-03  2.123e-03   3.616  0.00030 ***
## BLUEBOOK    -1.150e-05  4.691e-06  -2.452  0.01421 *
```

```

## TIF          -4.144e-02  8.652e-03  -4.790  1.67e-06 ***
## OLDCLAIM     9.280e-06  3.996e-06   2.322  0.02023 *
## CLM_FREQ     2.796e-01  3.243e-02   8.621  < 2e-16 ***
## MVR_PTS      1.392e-01  1.604e-02   8.679  < 2e-16 ***
## CAR_AGE      -2.675e-02  6.849e-03  -3.906  9.38e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5824.3  on 5061  degrees of freedom
## Residual deviance: 5220.7  on 5048  degrees of freedom
##    (1351 observations deleted due to missingness)
## AIC: 5248.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 953 300
##           1  57  76
##
##           Accuracy : 0.7424
##           95% CI : (0.7185, 0.7653)
##    No Information Rate : 0.7287
##    P-Value [Acc > NIR] : 0.1316
##
##           Kappa : 0.1828
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9436
##           Specificity : 0.2021
##           Pos Pred Value : 0.7606
##           Neg Pred Value : 0.5714
##           Prevalence : 0.7287
##           Detection Rate : 0.6876
##           Detection Prevalence : 0.9040
##           Balanced Accuracy : 0.5728
##
##           'Positive' Class : 0
##

```



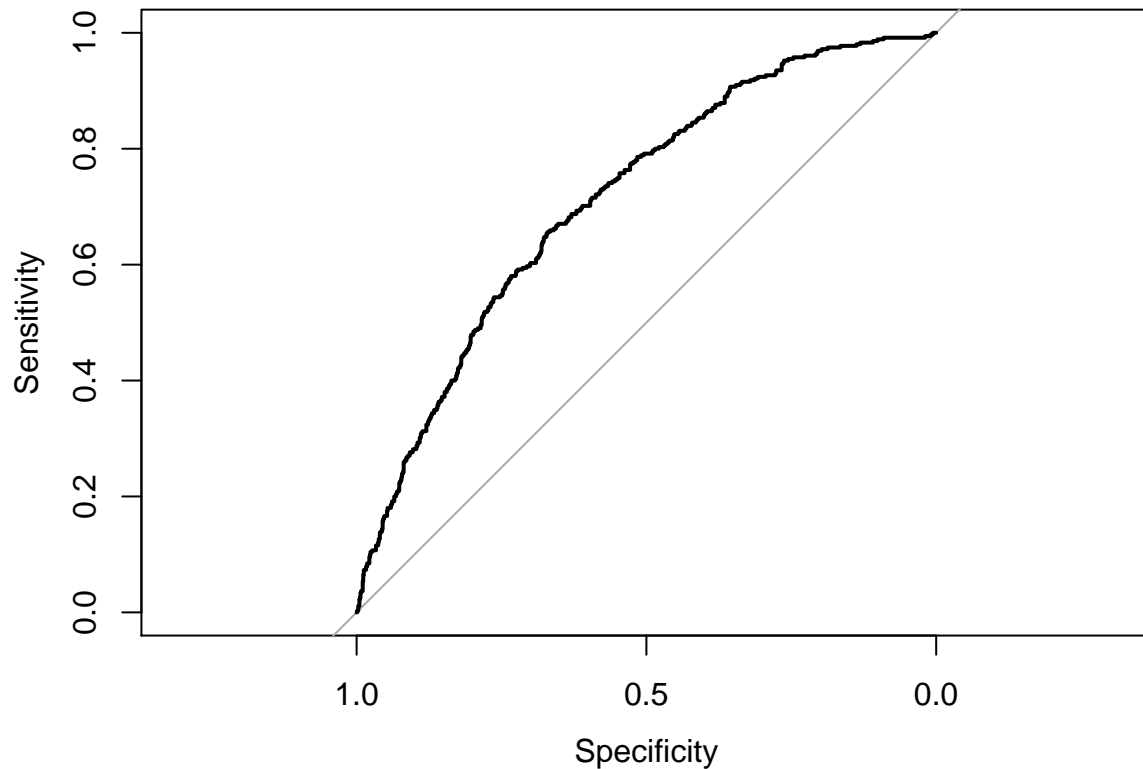
```
## [1] "AUC: 0.720233831893828"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1010 controls (dfPred_raw$class 0) < 376 cases (dfPred_raw$class 1).
## Area under the curve: 0.7202
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7546  -0.7691  -0.5747   0.9263   2.6858
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.626e-01  2.366e-01  -2.378 0.017403 *
## KIDSDRIV     3.273e-01  7.023e-02   4.660 3.16e-06 ***
## AGE          -1.112e-02  4.530e-03  -2.454 0.014137 *
## HOMEKIDS      5.125e-02  3.809e-02   1.345 0.178466
## YOJ           5.965e-03  8.734e-03   0.683 0.494617
## INCOME        -8.953e-07  1.036e-06  -0.864 0.387730
## HOME_VAL      -2.562e-06  3.313e-07  -7.732 1.06e-14 ***
## TRAVTIME       7.360e-03  2.127e-03   3.461 0.000539 ***
## BLUEBOOK      -1.108e-05  4.680e-06  -2.368 0.017876 *
```

```

## TIF          -3.871e-02  8.491e-03  -4.559 5.14e-06 ***
## OLDCLAIM      8.604e-06  3.974e-06   2.165 0.030379 *
## CLM_FREQ      2.885e-01  3.242e-02   8.900 < 2e-16 ***
## MVR_PTS       1.323e-01  1.583e-02   8.358 < 2e-16 ***
## CAR_AGE       -1.864e-02  6.702e-03  -2.781 0.005426 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5857.4  on 5045  degrees of freedom
## Residual deviance: 5240.2  on 5032  degrees of freedom
##    (1366 observations deleted due to missingness)
## AIC: 5268.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 979 287
##           1  68  68
##
##           Accuracy : 0.7468
##           95% CI : (0.7232, 0.7694)
##    No Information Rate : 0.7468
##    P-Value [Acc > NIR] : 0.5143
##
##           Kappa : 0.159
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9351
##           Specificity : 0.1915
##           Pos Pred Value : 0.7733
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7468
##           Detection Rate : 0.6983
##           Detection Prevalence : 0.9030
##           Balanced Accuracy : 0.5633
##
##           'Positive' Class : 0
##

```



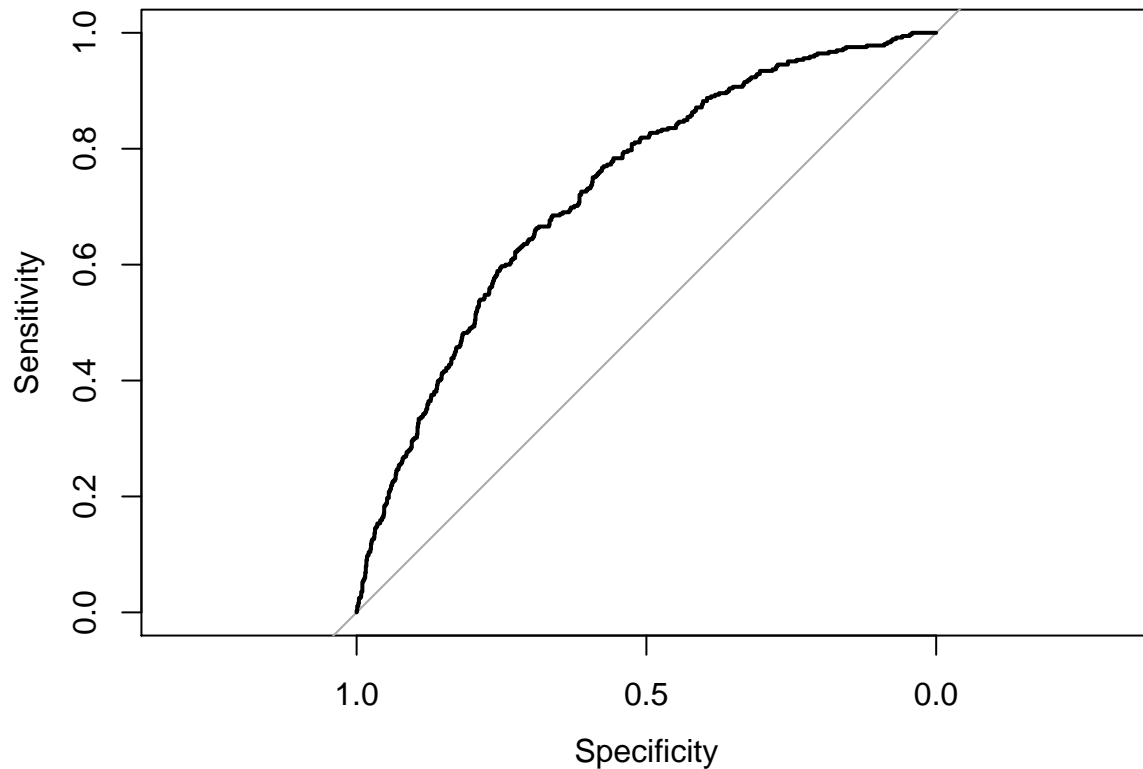


```
## [1] "AUC: 0.713480500961836"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1047 controls (dfPred_raw$class 0) < 355 cases (dfPred_raw$class 1).
## Area under the curve: 0.7135
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8915  -0.7706  -0.5775   0.9276   2.6099
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.867e-01  2.381e-01  -1.624  0.104326
## KIDSDRIV     1.523e-01  7.068e-02   2.155  0.031151 *
## AGE         -1.111e-02  4.555e-03  -2.440  0.014683 *
## HOMEKIDS      9.769e-02  3.730e-02   2.619  0.008822 **
## YOJ          -1.378e-02  8.640e-03  -1.595  0.110810
## INCOME        3.417e-07  1.055e-06   0.324  0.746037
## HOME_VAL     -2.362e-06  3.299e-07  -7.161  7.99e-13 ***
## TRAVTIME      7.663e-03  2.107e-03   3.637  0.000275 ***
## BLUEBOOK     -1.220e-05  4.626e-06  -2.637  0.008354 **
```

```

## TIF          -4.248e-02  8.646e-03  -4.913  8.96e-07 ***
## OLDCLAIM     6.046e-06  3.896e-06   1.552  0.120740
## CLM_FREQ     2.610e-01  3.226e-02   8.091  5.94e-16 ***
## MVR_PTS      1.415e-01  1.582e-02   8.942  < 2e-16 ***
## CAR_AGE      -2.462e-02  6.760e-03  -3.641  0.000271 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5854.3  on 5073  degrees of freedom
## Residual deviance: 5263.3  on 5060  degrees of freedom
##    (1338 observations deleted due to missingness)
## AIC: 5291.3
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 962 305
##           1  47  60
##
##           Accuracy : 0.7438
##           95% CI : (0.7199, 0.7667)
##    No Information Rate : 0.7344
##    P-Value [Acc > NIR] : 0.2232
##
##           Kappa : 0.1521
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9534
##           Specificity : 0.1644
##           Pos Pred Value : 0.7593
##           Neg Pred Value : 0.5607
##           Prevalence : 0.7344
##           Detection Rate : 0.7001
##           Detection Prevalence : 0.9221
##           Balanced Accuracy : 0.5589
##
##           'Positive' Class : 0
##

```

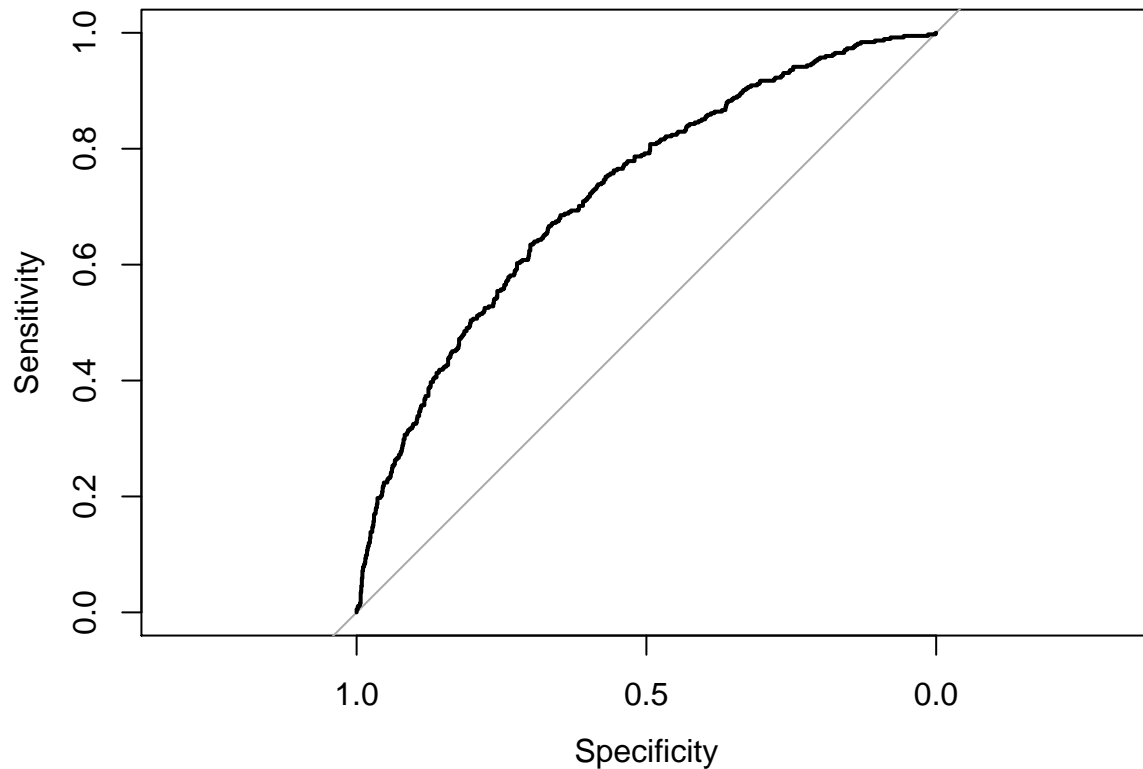


```
## [1] "AUC: 0.730241524905983"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1009 controls (dfPred_raw$class 0) < 365 cases (dfPred_raw$class 1).
## Area under the curve: 0.7302
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0045  -0.7665  -0.5751   0.8900   2.6164
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.353e-01  2.372e-01  -1.414  0.157447
## KIDSDRIV     2.574e-01  6.932e-02   3.713  0.000205 ***
## AGE         -1.070e-02  4.584e-03  -2.335  0.019541 *
## HOMEKIDS      6.793e-02  3.811e-02   1.783  0.074643 .
## YOJ          -1.583e-02  8.720e-03  -1.815  0.069461 .
## INCOME        4.210e-07  1.062e-06   0.396  0.691823
## HOME_VAL     -2.541e-06  3.305e-07  -7.690  1.47e-14 ***
## TRAVTIME      7.660e-03  2.128e-03   3.599  0.000319 ***
## BLUEBOOK     -1.182e-05  4.669e-06  -2.532  0.011335 *
```

```

## TIF          -4.889e-02  8.567e-03  -5.707  1.15e-08 ***
## OLDCLAIM     3.667e-06  3.899e-06   0.940  0.347060
## CLM_FREQ     2.628e-01  3.234e-02   8.127  4.39e-16 ***
## MVR_PTS      1.425e-01  1.581e-02   9.009  < 2e-16 ***
## CAR_AGE      -2.497e-02  6.720e-03  -3.716  0.000202 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5832.4  on 5071  degrees of freedom
## Residual deviance: 5238.3  on 5058  degrees of freedom
##    (1340 observations deleted due to missingness)
## AIC: 5266.3
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 965 301
##           1  36  74
##
##           Accuracy : 0.7551
##           95% CI : (0.7315, 0.7776)
##    No Information Rate : 0.7275
##    P-Value [Acc > NIR] : 0.01099
##
##           Kappa : 0.2071
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9640
##           Specificity : 0.1973
##           Pos Pred Value : 0.7622
##           Neg Pred Value : 0.6727
##           Prevalence : 0.7275
##           Detection Rate : 0.7013
##           Detection Prevalence : 0.9201
##           Balanced Accuracy : 0.5807
##
##           'Positive' Class : 0
##

```

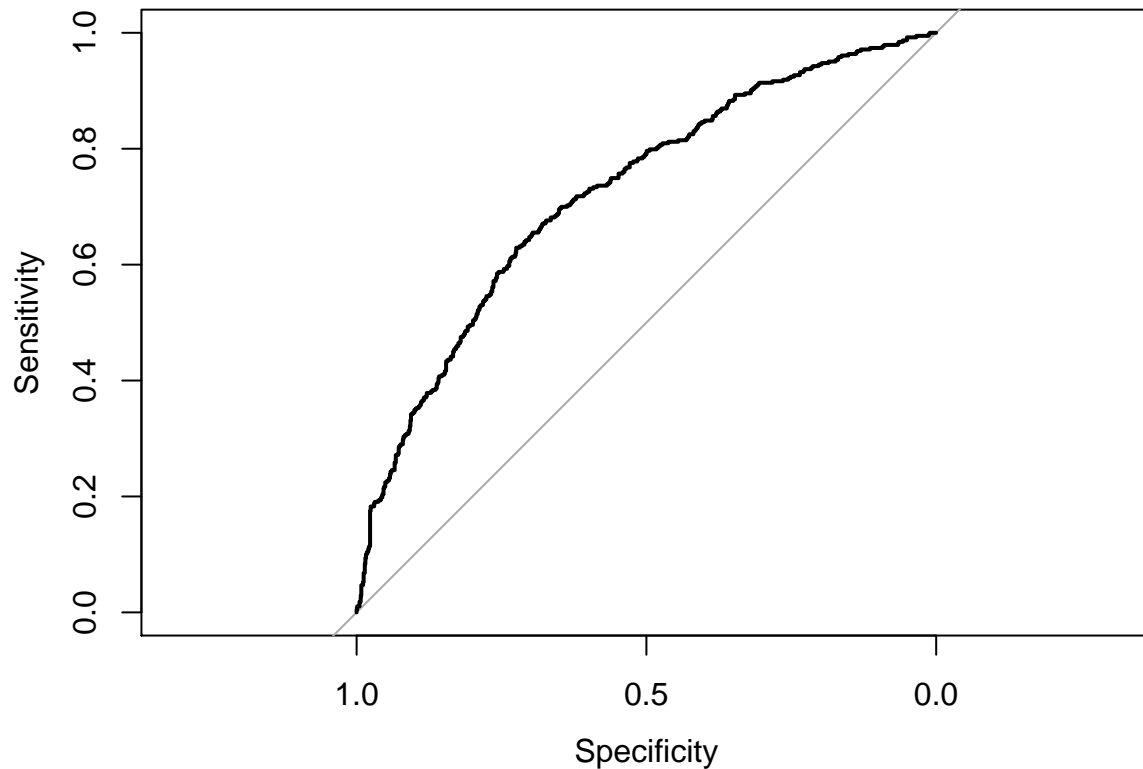


```
## [1] "AUC: 0.722914418914419"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1001 controls (dfPred_raw$class 0) < 375 cases (dfPred_raw$class 1).
## Area under the curve: 0.7229
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9465  -0.7610  -0.5743   0.8832   2.5878
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.052e-01  2.397e-01  -1.273  0.202890
## KIDSDRIV     1.948e-01  7.075e-02   2.753  0.005897 **
## AGE         -1.260e-02  4.603e-03  -2.738  0.006180 **
## HOMEKIDS      6.270e-02  3.838e-02   1.634  0.102302
## YOJ          -1.208e-02  8.735e-03  -1.383  0.166611
## INCOME        3.452e-07  1.061e-06   0.325  0.744836
## HOME_VAL     -2.416e-06  3.323e-07  -7.272  3.55e-13 ***
## TRAVTIME      7.516e-03  2.133e-03   3.524  0.000425 ***
## BLUEBOOK     -9.394e-06  4.659e-06  -2.016  0.043753 *
```

```

## TIF          -4.749e-02  8.611e-03  -5.515  3.48e-08 ***
## OLDCLAIM     3.236e-06  3.994e-06   0.810  0.417827
## CLM_FREQ     2.836e-01  3.233e-02   8.774  < 2e-16 ***
## MVR_PTS      1.443e-01  1.613e-02   8.947  < 2e-16 ***
## CAR_AGE      -2.938e-02  6.776e-03  -4.335  1.45e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5817.0  on 5073  degrees of freedom
## Residual deviance: 5228.4  on 5060  degrees of freedom
## (1338 observations deleted due to missingness)
## AIC: 5256.4
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 936 296
##           1  55  87
##
##           Accuracy : 0.7445
##           95% CI : (0.7206, 0.7674)
##           No Information Rate : 0.7213
##           P-Value [Acc > NIR] : 0.02824
##
##           Kappa : 0.2127
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9445
##           Specificity : 0.2272
##           Pos Pred Value : 0.7597
##           Neg Pred Value : 0.6127
##           Prevalence : 0.7213
##           Detection Rate : 0.6812
##           Detection Prevalence : 0.8967
##           Balanced Accuracy : 0.5858
##
##           'Positive' Class : 0
##

```



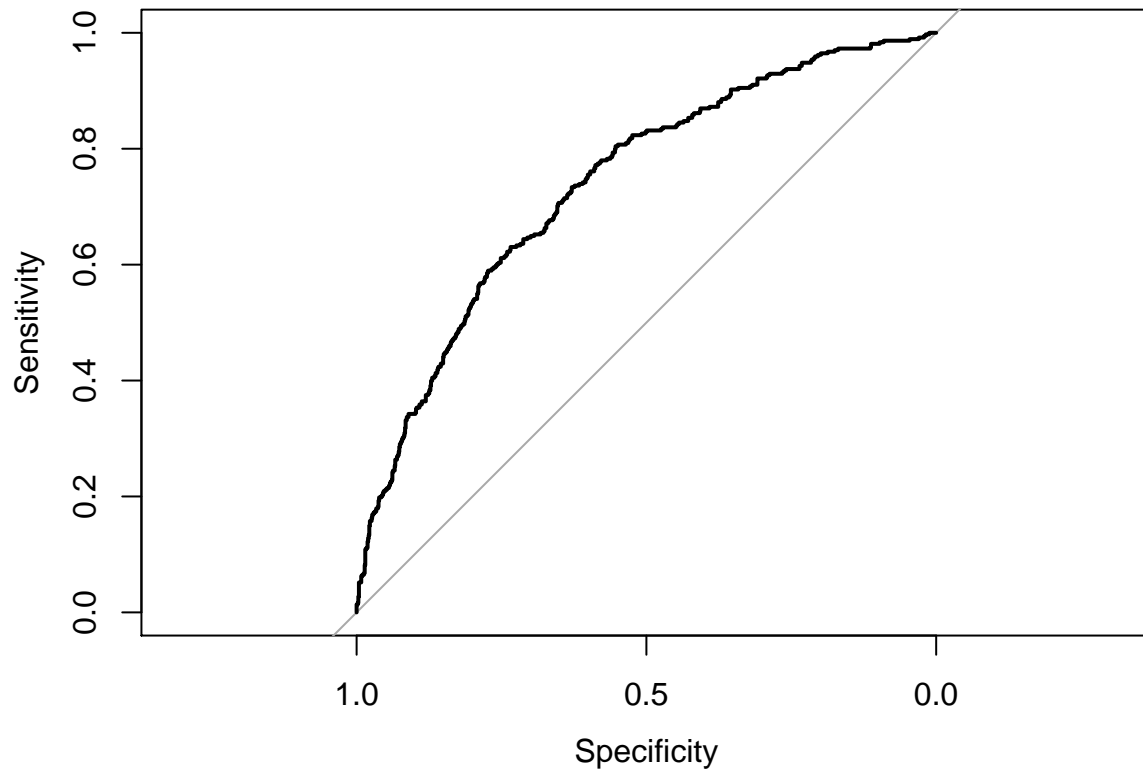
```
## [1] "AUC: 0.722544677554913"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 991 controls (dfPred_raw$class 0) < 383 cases (dfPred_raw$class 1).
## Area under the curve: 0.7225
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0112  -0.7741  -0.5812   0.9194   2.6750
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.542e-01  2.373e-01  -1.493  0.135514
## KIDSDRIV     2.679e-01  6.975e-02   3.841  0.000122 ***
## AGE         -1.109e-02  4.560e-03  -2.431  0.015039 *
## HOMEKIDS     4.623e-02  3.767e-02   1.227  0.219682
## YOJ          -2.119e-03  8.766e-03  -0.242  0.809039
## INCOME       -8.265e-07  1.043e-06  -0.792  0.428227
## HOME_VAL     -2.379e-06  3.300e-07  -7.211  5.56e-13 ***
## TRAVTIME     8.223e-03  2.130e-03   3.862  0.000113 ***
## BLUEBOOK     -1.299e-05  4.632e-06  -2.804  0.005044 **
```

```

## TIF          -4.709e-02  8.575e-03  -5.492  3.98e-08 ***
## OLDCLAIM     9.018e-06  3.936e-06   2.291  0.021950 *
## CLM_FREQ     2.212e-01  3.251e-02   6.805  1.01e-11 ***
## MVR_PTS      1.391e-01  1.591e-02   8.747  < 2e-16 ***
## CAR_AGE      -2.536e-02  6.738e-03  -3.764  0.000167 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5848.7  on 5074  degrees of freedom
## Residual deviance: 5277.0  on 5061  degrees of freedom
##    (1337 observations deleted due to missingness)
## AIC: 5305
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 958 292
##           1  47  76
##
##           Accuracy : 0.7531
##           95% CI : (0.7294, 0.7757)
##    No Information Rate : 0.732
##    P-Value [Acc > NIR] : 0.04036
##
##           Kappa : 0.2025
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9532
##           Specificity : 0.2065
##           Pos Pred Value : 0.7664
##           Neg Pred Value : 0.6179
##           Prevalence : 0.7320
##           Detection Rate : 0.6977
##           Detection Prevalence : 0.9104
##           Balanced Accuracy : 0.5799
##
##           'Positive' Class : 0
##

```



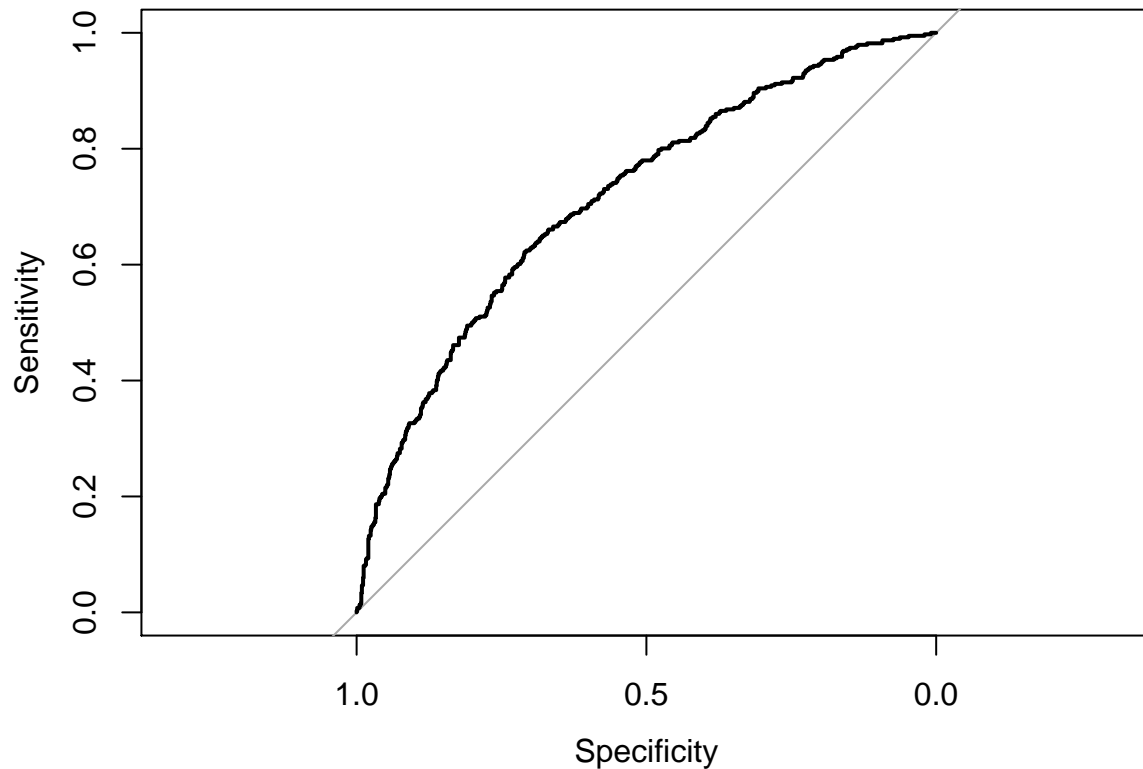


```
## [1] "AUC: 0.738313865455332"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1005 controls (dfPred_raw$class 0) < 368 cases (dfPred_raw$class 1).
## Area under the curve: 0.7383
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9760  -0.7615  -0.5720   0.8659   2.6456
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.239e-01  2.387e-01  -1.357 0.174756
## KIDSDRIV     2.387e-01  7.046e-02   3.388 0.000703 ***
## AGE         -1.149e-02  4.606e-03  -2.494 0.012614 *
## HOMEKIDS     5.839e-02  3.843e-02   1.520 0.128614
## YOJ         -1.068e-02  8.754e-03  -1.220 0.222468
## INCOME       7.521e-07  1.061e-06   0.709 0.478403
## HOME_VAL    -2.578e-06  3.313e-07  -7.783 7.09e-15 ***
## TRAVTIME     7.414e-03  2.142e-03   3.462 0.000537 ***
## BLUEBOOK    -1.346e-05  4.703e-06  -2.863 0.004195 **
```

```

## TIF          -4.582e-02  8.546e-03  -5.362 8.22e-08 ***
## OLDCLAIM     3.771e-06  3.967e-06   0.951 0.341756
## CLM_FREQ     2.546e-01  3.255e-02   7.824 5.14e-15 ***
## MVR_PTS      1.532e-01  1.599e-02   9.576 < 2e-16 ***
## CAR_AGE      -2.868e-02  6.770e-03  -4.237 2.26e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5808.9  on 5070  degrees of freedom
## Residual deviance: 5209.5  on 5057  degrees of freedom
##    (1341 observations deleted due to missingness)
## AIC: 5237.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 942 304
##           1  49  82
##
##           Accuracy : 0.7436
##           95% CI : (0.7197, 0.7665)
##    No Information Rate : 0.7197
##    P-Value [Acc > NIR] : 0.02483
##
##           Kappa : 0.2042
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9506
##           Specificity : 0.2124
##           Pos Pred Value : 0.7560
##           Neg Pred Value : 0.6260
##           Prevalence : 0.7197
##           Detection Rate : 0.6841
##           Detection Prevalence : 0.9049
##           Balanced Accuracy : 0.5815
##
##           'Positive' Class : 0
##

```

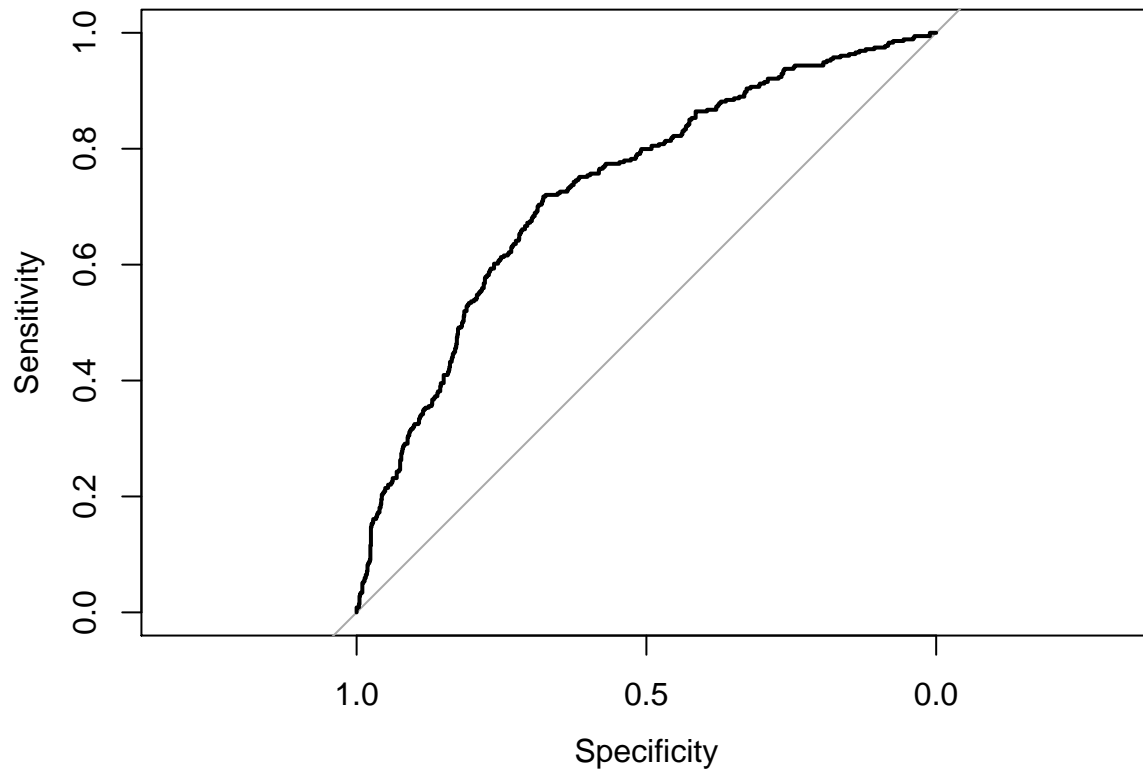


```
## [1] "AUC: 0.71544417895777"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 991 controls (dfPred_raw$class 0) < 386 cases (dfPred_raw$class 1).
## Area under the curve: 0.7154
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9858  -0.7759  -0.5800   0.9339   2.6813
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.973e-01  2.380e-01  -1.249  0.211483
## KIDSDRIV     2.216e-01  7.100e-02   3.120  0.001806 **
## AGE          -1.248e-02  4.570e-03  -2.731  0.006315 **
## HOMEKIDS      8.678e-02  3.750e-02   2.314  0.020641 *
## YOJ          -7.931e-03  8.657e-03  -0.916  0.359618
## INCOME        -1.132e-06  1.055e-06  -1.073  0.283361
## HOME_VAL      -2.322e-06  3.323e-07  -6.988  2.78e-12 ***
## TRAVTIME       7.445e-03  2.120e-03   3.513  0.000444 ***
## BLUEBOOK      -1.125e-05  4.628e-06  -2.432  0.015020 *
```

```

## TIF          -4.271e-02  8.600e-03  -4.966 6.83e-07 ***
## OLDCLAIM     6.994e-06  3.890e-06   1.798 0.072185 .
## CLM_FREQ     2.728e-01  3.245e-02   8.406 < 2e-16 ***
## MVR_PTS      1.263e-01  1.608e-02   7.854 4.04e-15 ***
## CAR_AGE      -2.293e-02  6.751e-03  -3.396 0.000683 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5881.7  on 5081  degrees of freedom
## Residual deviance: 5287.6  on 5068  degrees of freedom
##    (1330 observations deleted due to missingness)
## AIC: 5315.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 954 276
##           1  58  78
##
##           Accuracy : 0.7555
##           95% CI : (0.7318, 0.7781)
##    No Information Rate : 0.7408
##    P-Value [Acc > NIR] : 0.1138
##
##           Kappa : 0.2038
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9427
##           Specificity : 0.2203
##           Pos Pred Value : 0.7756
##           Neg Pred Value : 0.5735
##           Prevalence : 0.7408
##           Detection Rate : 0.6984
##           Detection Prevalence : 0.9004
##           Balanced Accuracy : 0.5815
##
##           'Positive' Class : 0
##

```

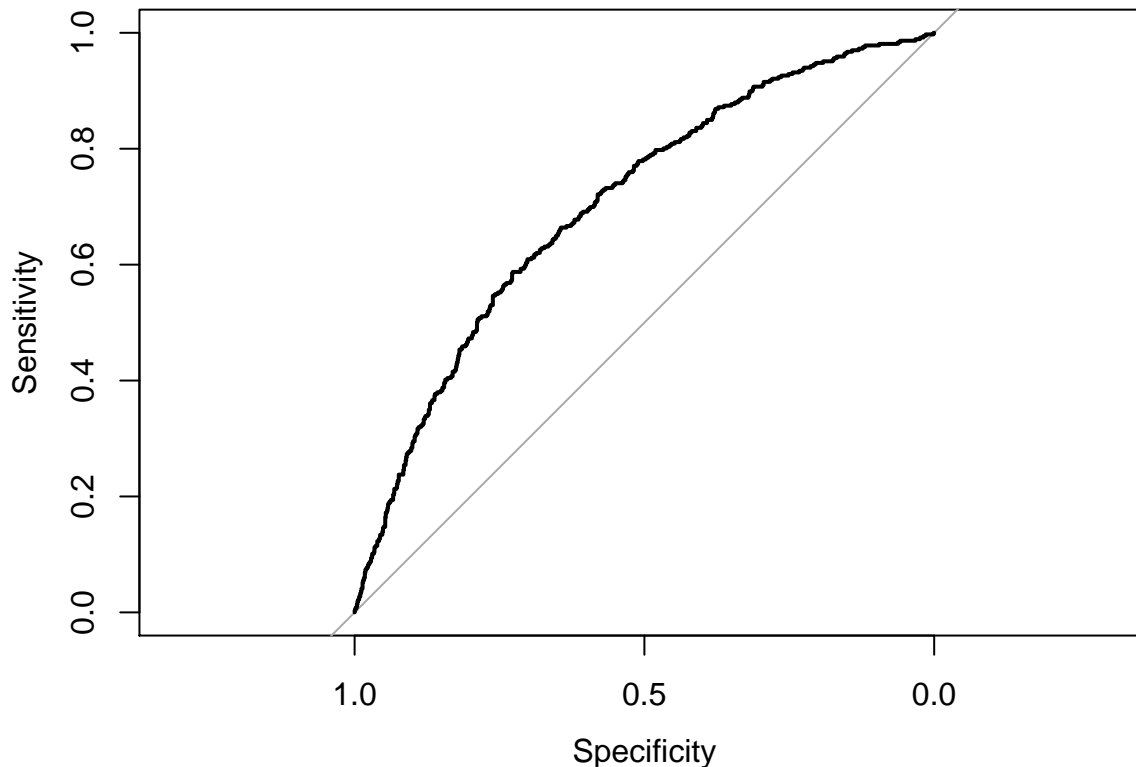


```
## [1] "AUC: 0.731029901074116"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1012 controls (dfPred_raw$class 0) < 354 cases (dfPred_raw$class 1).
## Area under the curve: 0.731
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9704  -0.7687  -0.5676   0.8913   2.6782
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.788e-01  2.382e-01  -2.430  0.015095 *
## KIDSDRIV     1.896e-01  7.075e-02   2.680  0.007370 **
## AGE          -7.674e-03  4.567e-03  -1.680  0.092950 .
## HOMEKIDS     1.040e-01  3.765e-02   2.762  0.005746 **
## YOJ          -8.897e-03  8.686e-03  -1.024  0.305676
## INCOME        4.594e-07  1.057e-06   0.434  0.663935
## HOME_VAL     -2.677e-06  3.311e-07  -8.086  6.19e-16 ***
## TRAVTIME      7.820e-03  2.125e-03   3.679  0.000234 ***
## BLUEBOOK     -1.410e-05  4.685e-06  -3.010  0.002609 **
```

```

## TIF          -3.846e-02  8.596e-03  -4.475  7.65e-06 ***
## OLDCLAIM      6.571e-06  3.988e-06   1.648  0.099404 .
## CLM_FREQ      2.450e-01  3.236e-02   7.571  3.69e-14 ***
## MVR_PTS       1.606e-01  1.582e-02  10.150  < 2e-16 ***
## CAR_AGE       -2.553e-02  6.830e-03  -3.738  0.000185 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5843.0  on 5058  degrees of freedom
## Residual deviance: 5210.7  on 5045  degrees of freedom
##    (1354 observations deleted due to missingness)
## AIC: 5238.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 969 309
##           1  54  57
##
##           Accuracy : 0.7387
##           95% CI : (0.7147, 0.7616)
##    No Information Rate : 0.7365
##    P-Value [Acc > NIR] : 0.4413
##
##           Kappa : 0.1326
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9472
##           Specificity : 0.1557
##           Pos Pred Value : 0.7582
##           Neg Pred Value : 0.5135
##           Prevalence : 0.7365
##           Detection Rate : 0.6976
##           Detection Prevalence : 0.9201
##           Balanced Accuracy : 0.5515
##
##           'Positive' Class : 0
##

```



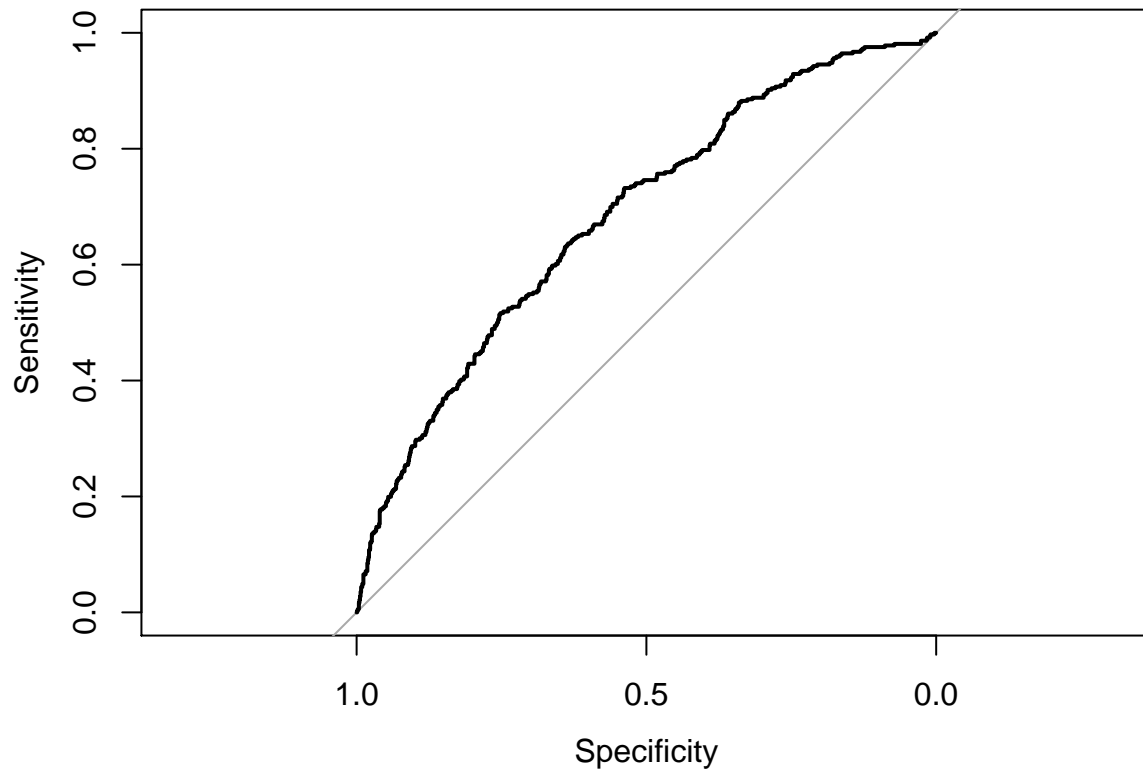
```
## [1] "AUC: 0.703927161621503"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1023 controls (dfPred_raw$class 0) < 366 cases (dfPred_raw$class 1).
## Area under the curve: 0.7039
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8329  -0.7676  -0.5625   0.8821   2.7262
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.424e-01  2.392e-01  -2.268 0.023345 *
## KIDSDRIV     2.586e-01  6.973e-02   3.709 0.000208 ***
## AGE         -8.970e-03  4.609e-03  -1.946 0.051632 .
## HOMEKIDS      7.901e-02  3.806e-02   2.076 0.037888 *
## YOJ         -3.720e-03  8.794e-03  -0.423 0.672309
## INCOME       1.742e-08  1.047e-06   0.017 0.986717
## HOME_VAL    -2.818e-06  3.334e-07  -8.452 < 2e-16 ***
## TRAVTIME     8.600e-03  2.156e-03   3.990 6.62e-05 ***
## BLUEBOOK    -1.368e-05  4.701e-06  -2.910 0.003611 **
```

```

## TIF          -4.161e-02  8.615e-03  -4.830  1.37e-06 ***
## OLDCLAIM     5.892e-06  4.043e-06   1.457  0.145011
## CLM_FREQ     2.876e-01  3.226e-02   8.914  < 2e-16 ***
## MVRPTS       1.501e-01  1.601e-02   9.376  < 2e-16 ***
## CAR_AGE      -2.649e-02  6.852e-03  -3.866  0.000111 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5840.0  on 5053  degrees of freedom
## Residual deviance: 5180.7  on 5040  degrees of freedom
## (1359 observations deleted due to missingness)
## AIC: 5208.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 965 291
##           1  63  75
##
##           Accuracy : 0.7461
##           95% CI : (0.7223, 0.7687)
##           No Information Rate : 0.7374
##           P-Value [Acc > NIR] : 0.2427
##
##           Kappa : 0.1797
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9387
##           Specificity : 0.2049
##           Pos Pred Value : 0.7683
##           Neg Pred Value : 0.5435
##           Prevalence : 0.7374
##           Detection Rate : 0.6923
##           Detection Prevalence : 0.9010
##           Balanced Accuracy : 0.5718
##
##           'Positive' Class : 0
##

```



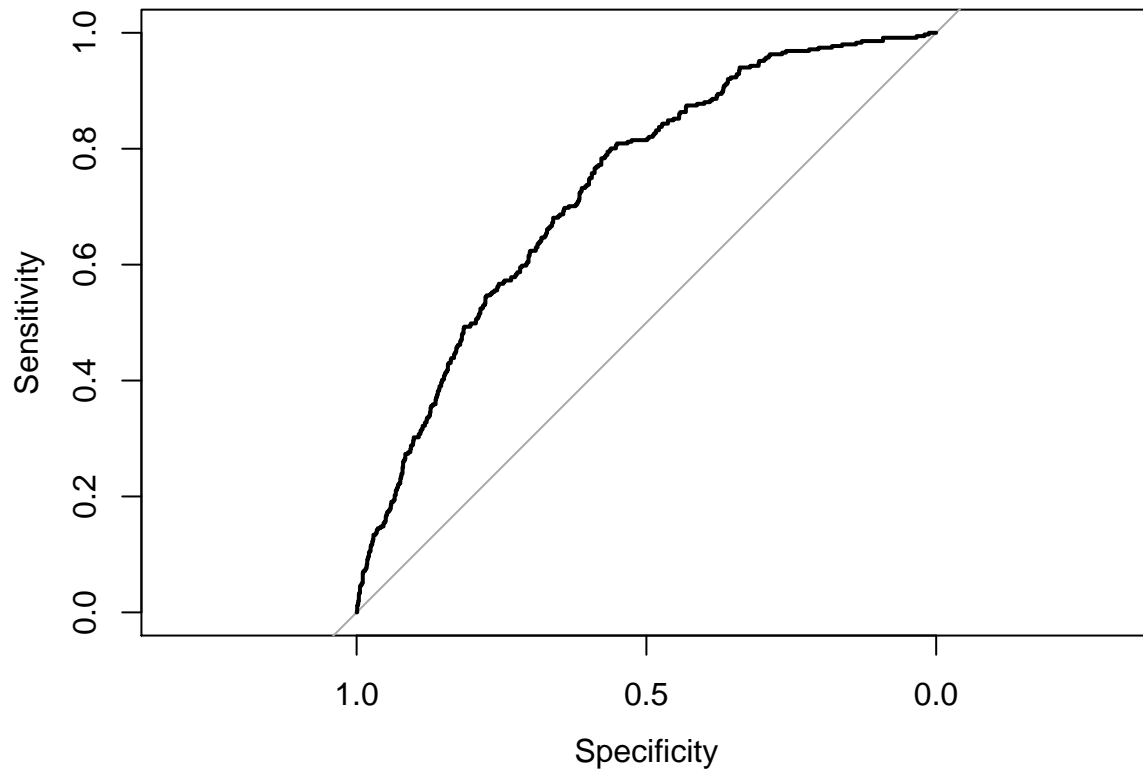


```
## [1] "AUC: 0.683623567434245"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1028 controls (dfPred_raw$class 0) < 366 cases (dfPred_raw$class 1).
## Area under the curve: 0.6836
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8263  -0.7737  -0.5801   0.9387   2.5897
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.743e-01  2.366e-01  -2.850  0.00437 **
## KIDSDRIV     2.702e-01  6.918e-02   3.905  9.41e-05 ***
## AGE         -9.329e-03  4.537e-03  -2.056  0.03976 *
## HOMEKIDS     6.887e-02  3.708e-02   1.857  0.06328 .
## YOJ         -2.917e-03  8.713e-03  -0.335  0.73779
## INCOME      -8.734e-08  1.032e-06  -0.085  0.93256
## HOME_VAL    -2.631e-06  3.281e-07  -8.020  1.06e-15 ***
## TRAVTIME     9.397e-03  2.126e-03   4.421  9.84e-06 ***
## BLUEBOOK    -8.226e-06  4.598e-06  -1.789  0.07360 .
```

```

## TIF          -3.911e-02  8.557e-03  -4.570  4.88e-06 ***
## OLDCLAIM      5.979e-06  3.944e-06   1.516  0.12953
## CLM_FREQ      2.604e-01  3.200e-02   8.138  4.03e-16 ***
## MVR_PTS       1.453e-01  1.564e-02   9.286  < 2e-16 ***
## CAR_AGE       -1.950e-02  6.678e-03  -2.919  0.00351 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5875.4  on 5061  degrees of freedom
## Residual deviance: 5279.1  on 5048  degrees of freedom
##    (1350 observations deleted due to missingness)
## AIC: 5307.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 980 291
##           1  55  60
##
##           Accuracy : 0.7504
##           95% CI : (0.7267, 0.773)
##    No Information Rate : 0.7468
##    P-Value [Acc > NIR] : 0.3923
##
##           Kappa : 0.1514
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9469
##           Specificity : 0.1709
##           Pos Pred Value : 0.7710
##           Neg Pred Value : 0.5217
##           Prevalence : 0.7468
##           Detection Rate : 0.7071
##           Detection Prevalence : 0.9170
##           Balanced Accuracy : 0.5589
##
##           'Positive' Class : 0
##

```

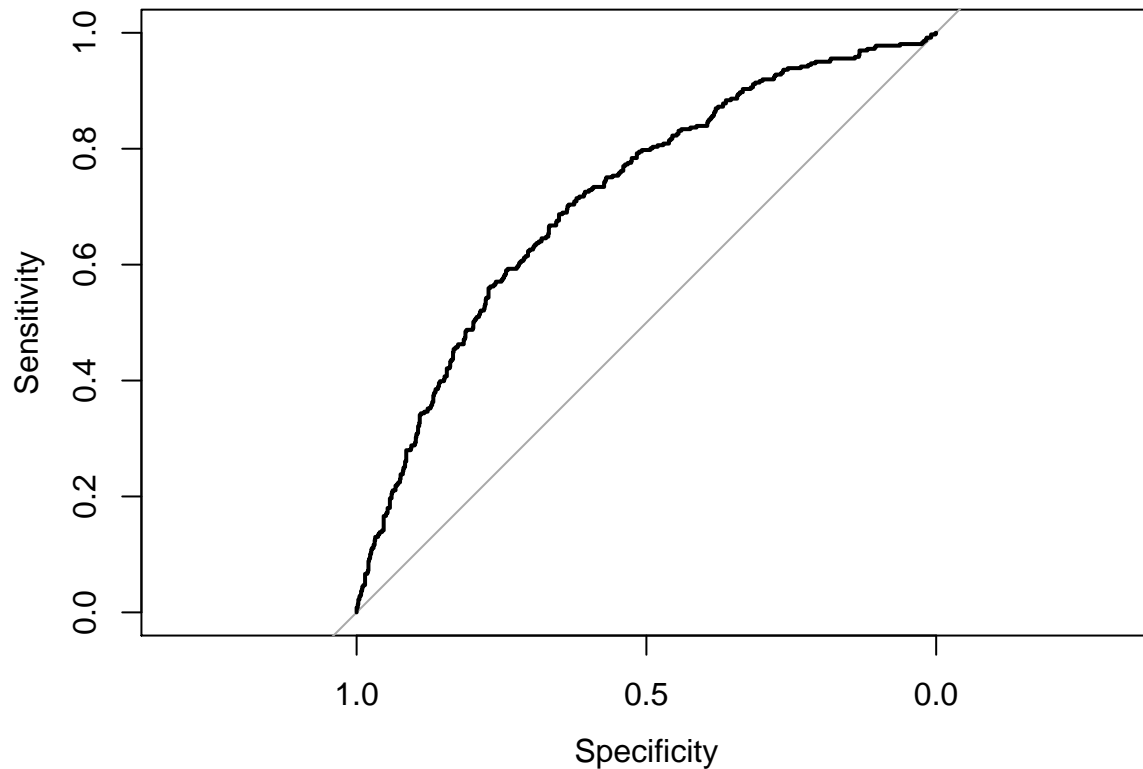


```
## [1] "AUC: 0.731599157686114"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1035 controls (dfPred_raw$class 0) < 351 cases (dfPred_raw$class 1).
## Area under the curve: 0.7316
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9774  -0.7731  -0.5734   0.9123   2.6826
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.729e-01  2.382e-01  -1.986 0.047066 *
## KIDSDRIV     2.245e-01  7.018e-02   3.199 0.001378 **
## AGE          -1.065e-02  4.582e-03  -2.325 0.020073 *
## HOMEKIDS      8.864e-02  3.752e-02   2.363 0.018145 *
## YOJ           -4.149e-03  8.652e-03  -0.479 0.631589
## INCOME        -6.479e-07  1.048e-06  -0.618 0.536558
## HOME_VAL      -2.610e-06  3.319e-07  -7.864 3.73e-15 ***
## TRAVTIME       7.055e-03  2.129e-03   3.314 0.000919 ***
## BLUEBOOK      -1.031e-05  4.640e-06  -2.221 0.026342 *
```

```

## TIF          -3.641e-02  8.544e-03  -4.261  2.03e-05 ***
## OLDCLAIM      7.045e-06  4.044e-06   1.742  0.081525 .
## CLM_FREQ      2.420e-01  3.231e-02   7.490  6.90e-14 ***
## MVR_PTS       1.547e-01  1.587e-02   9.743  < 2e-16 ***
## CAR_AGE       -2.326e-02  6.803e-03  -3.420  0.000627 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5853.2  on 5058  degrees of freedom
## Residual deviance: 5236.7  on 5045  degrees of freedom
##    (1354 observations deleted due to missingness)
## AIC: 5264.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  973  298
##           1   55   63
##
##           Accuracy : 0.7459
##           95% CI : (0.7221, 0.7686)
##    No Information Rate : 0.7401
##    P-Value [Acc > NIR] : 0.3246
##
##           Kappa : 0.1548
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9465
##           Specificity : 0.1745
##           Pos Pred Value : 0.7655
##           Neg Pred Value : 0.5339
##           Prevalence : 0.7401
##           Detection Rate : 0.7005
##           Detection Prevalence : 0.9150
##           Balanced Accuracy : 0.5605
##
##           'Positive' Class : 0
##

```

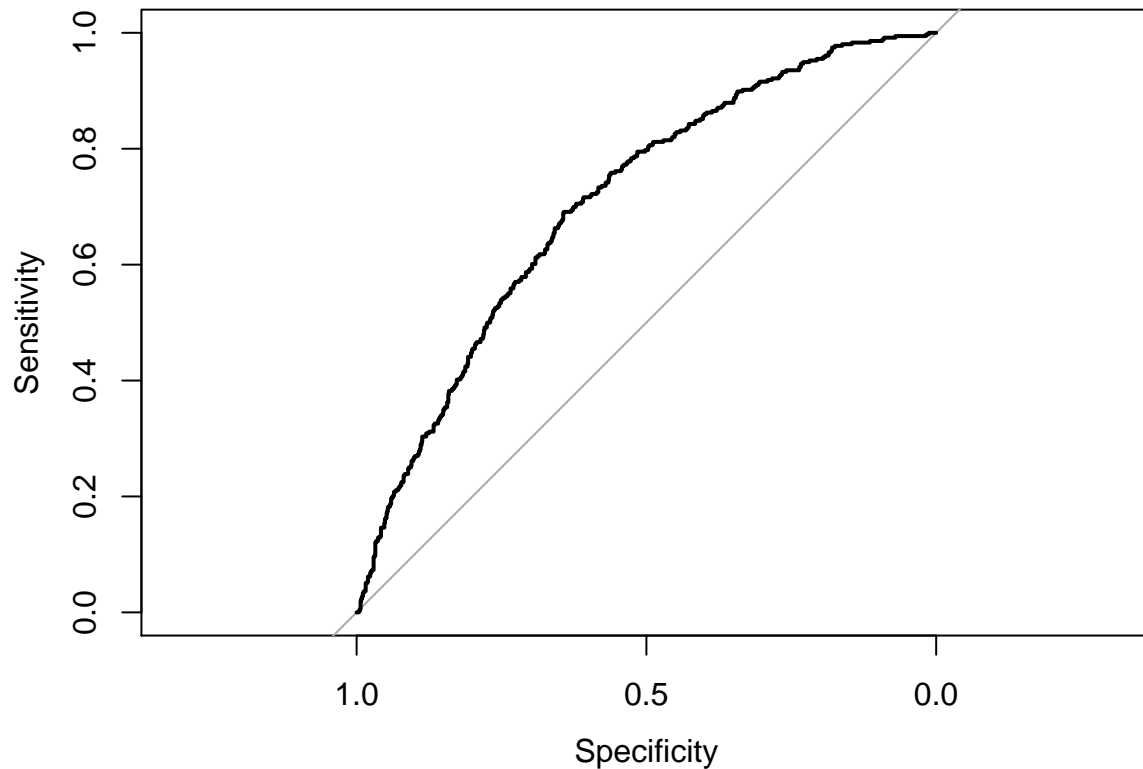


```
## [1] "AUC: 0.716039535660778"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1028 controls (dfPred_raw$class 0) < 361 cases (dfPred_raw$class 1).
## Area under the curve: 0.716
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7849  -0.7660  -0.5694   0.9006   2.6761
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.909e-01  2.383e-01  -2.899  0.003738 **
## KIDSDRIV     3.257e-01  7.092e-02   4.593  4.37e-06 ***
## AGE         -9.849e-03  4.570e-03  -2.155  0.031145 *
## HOMEKIDS      7.217e-02  3.784e-02   1.907  0.056518 .
## YOJ          3.705e-03  8.737e-03   0.424  0.671513
## INCOME       -9.038e-07  1.043e-06  -0.867  0.386139
## HOME_VAL     -2.527e-06  3.329e-07  -7.591  3.18e-14 ***
## TRAVTIME      8.298e-03  2.152e-03   3.857  0.000115 ***
## BLUEBOOK     -1.061e-05  4.702e-06  -2.256  0.024081 *
```

```

## TIF          -3.576e-02  8.511e-03  -4.201  2.66e-05 ***
## OLDCLAIM      6.928e-06  3.924e-06   1.765  0.077490 .
## CLM_FREQ      3.049e-01  3.240e-02   9.409  < 2e-16 ***
## MVR_PTS       1.350e-01  1.592e-02   8.480  < 2e-16 ***
## CAR_AGE       -1.780e-02  6.733e-03  -2.644  0.008200 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5869.0  on 5067  degrees of freedom
## Residual deviance: 5237.4  on 5054  degrees of freedom
##    (1344 observations deleted due to missingness)
## AIC: 5265.4
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  956  282
##           1   68   74
##
##           Accuracy : 0.7464
##           95% CI : (0.7225, 0.7692)
##    No Information Rate : 0.742
##    P-Value [Acc > NIR] : 0.3692
##
##           Kappa : 0.176
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9336
##           Specificity : 0.2079
##           Pos Pred Value : 0.7722
##           Neg Pred Value : 0.5211
##           Prevalence : 0.7420
##           Detection Rate : 0.6928
##           Detection Prevalence : 0.8971
##           Balanced Accuracy : 0.5707
##
##           'Positive' Class : 0
##

```



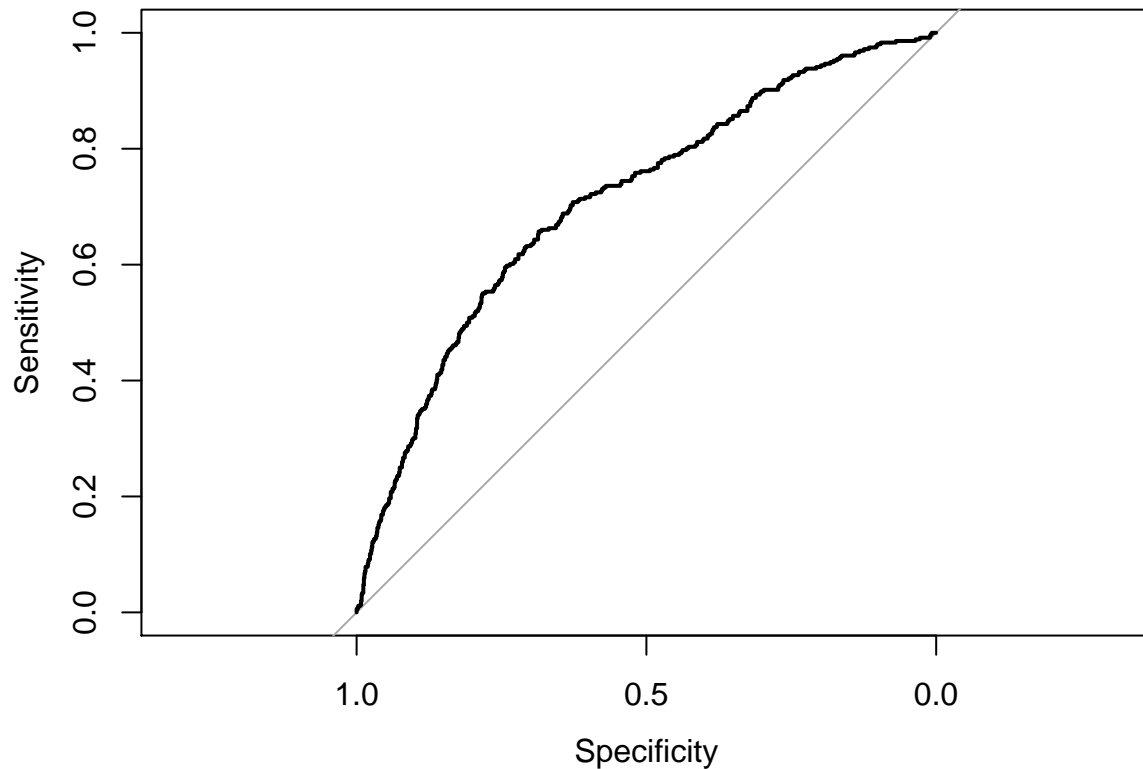
```
## [1] "AUC: 0.707357685217697"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1024 controls (dfPred_raw$class 0) < 356 cases (dfPred_raw$class 1).
## Area under the curve: 0.7074
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0668  -0.7740  -0.5703   0.9065   2.6694
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.934e-01  2.378e-01  -1.654  0.098083 .
## KIDSDRIV     2.776e-01  7.090e-02   3.915  9.03e-05 ***
## AGE         -1.100e-02  4.567e-03  -2.409  0.016008 *
## HOMEKIDS     7.313e-02  3.837e-02   1.906  0.056641 .
## YOJ         -6.024e-03  8.719e-03  -0.691  0.489649
## INCOME       3.701e-08  1.050e-06   0.035  0.971882
## HOME_VAL    -2.701e-06  3.301e-07  -8.182  2.79e-16 ***
## TRAVTIME     7.204e-03  2.118e-03   3.402  0.000670 ***
## BLUEBOOK    -1.201e-05  4.687e-06  -2.562  0.010393 *
```

```

## TIF          -4.518e-02  8.503e-03  -5.314 1.08e-07 ***
## OLDCLAIM     4.087e-06  3.989e-06   1.025 0.305577
## CLM_FREQ     2.917e-01  3.249e-02   8.979 < 2e-16 ***
## MVR_PTS      1.376e-01  1.605e-02   8.573 < 2e-16 ***
## CAR_AGE      -2.282e-02  6.705e-03  -3.403 0.000666 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5862.2  on 5056  degrees of freedom
## Residual deviance: 5239.1  on 5043  degrees of freedom
##    (1355 observations deleted due to missingness)
## AIC: 5267.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 964 275
##           1  71  81
##
##           Accuracy : 0.7513
##           95% CI : (0.7277, 0.7738)
##    No Information Rate : 0.7441
##    P-Value [Acc > NIR] : 0.2808
##
##           Kappa : 0.1957
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9314
##           Specificity : 0.2275
##           Pos Pred Value : 0.7780
##           Neg Pred Value : 0.5329
##           Prevalence : 0.7441
##           Detection Rate : 0.6930
##           Detection Prevalence : 0.8907
##           Balanced Accuracy : 0.5795
##
##           'Positive' Class : 0
##

```



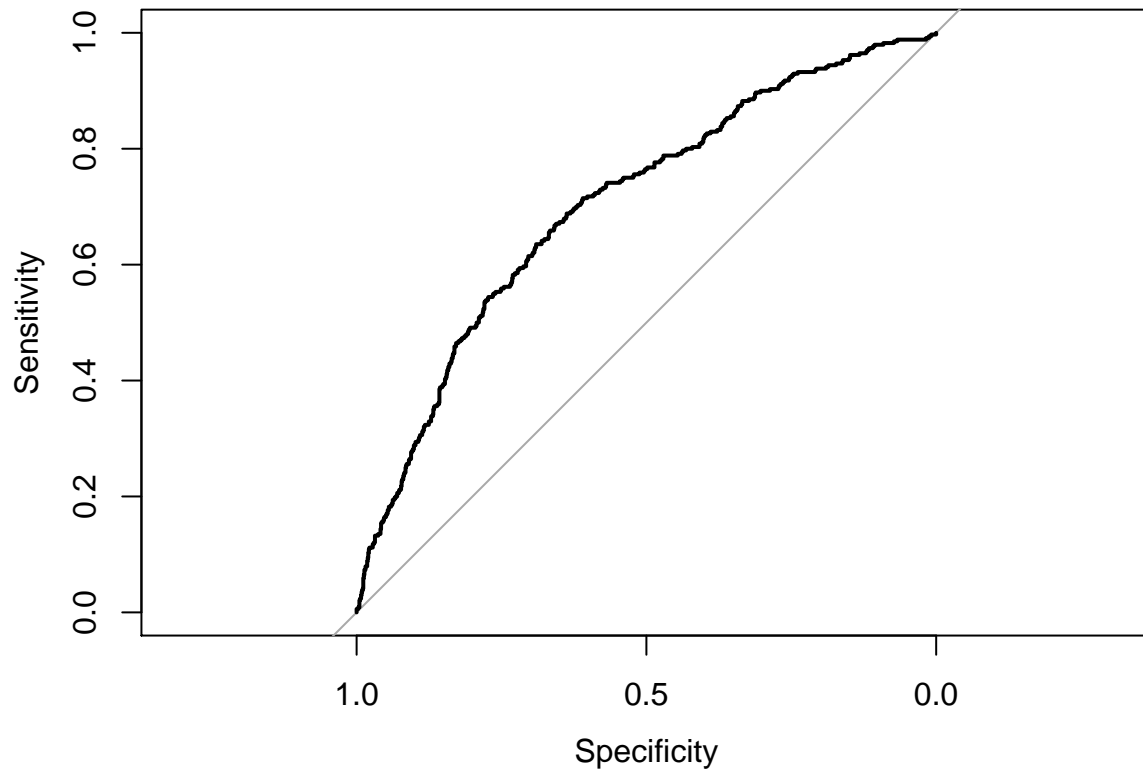


```
## [1] "AUC: 0.710495033382185"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1035 controls (dfPred_raw$class 0) < 356 cases (dfPred_raw$class 1).
## Area under the curve: 0.7105
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0767  -0.7742  -0.5712   0.9259   2.7013
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.922e-01  2.359e-01  -1.663  0.096389 .
## KIDSDRIV     2.639e-01  7.103e-02   3.716  0.000203 ***
## AGE          -9.782e-03  4.530e-03  -2.160  0.030807 *
## HOMEKIDS      8.306e-02  3.806e-02   2.183  0.029065 *
## YOJ          -7.352e-03  8.708e-03  -0.844  0.398514
## INCOME        -5.042e-07  1.042e-06  -0.484  0.628605
## HOME_VAL      -2.575e-06  3.296e-07  -7.813  5.58e-15 ***
## TRAVTIME       7.932e-03  2.116e-03   3.749  0.000178 ***
## BLUEBOOK      -1.542e-05  4.688e-06  -3.290  0.001001 **
```

```

## TIF          -4.837e-02  8.566e-03  -5.647 1.63e-08 ***
## OLDCLAIM     6.603e-06  3.887e-06   1.699 0.089349 .
## CLM_FREQ     2.877e-01  3.235e-02   8.892 < 2e-16 ***
## MVR_PTS      1.309e-01  1.588e-02   8.240 < 2e-16 ***
## CAR_AGE      -1.956e-02  6.717e-03  -2.912 0.003588 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5899.3  on 5064  degrees of freedom
## Residual deviance: 5260.2  on 5051  degrees of freedom
##    (1347 observations deleted due to missingness)
## AIC: 5288.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 972 272
##           1  71  68
##
##           Accuracy : 0.752
##           95% CI : (0.7283, 0.7746)
##    No Information Rate : 0.7542
##    P-Value [Acc > NIR] : 0.5885
##
##           Kappa : 0.1648
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9319
##           Specificity : 0.2000
##           Pos Pred Value : 0.7814
##           Neg Pred Value : 0.4892
##           Prevalence : 0.7542
##           Detection Rate : 0.7028
##           Detection Prevalence : 0.8995
##           Balanced Accuracy : 0.5660
##
##           'Positive' Class : 0
##

```

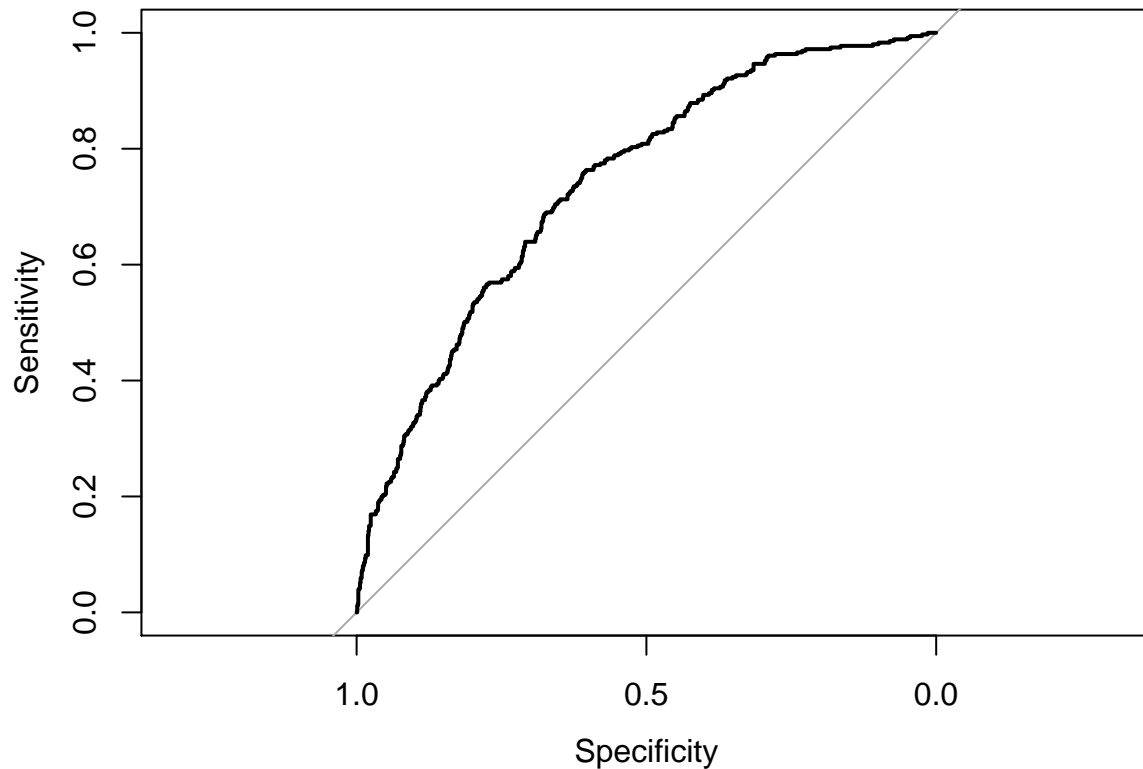


```
## [1] "AUC: 0.703538999492414"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1043 controls (dfPred_raw$class 0) < 340 cases (dfPred_raw$class 1).
## Area under the curve: 0.7035
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7645  -0.7758  -0.5819   0.9509   2.6164
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.784e-01  2.366e-01  -2.445 0.014488 *
## KIDSDRIV     2.476e-01  6.948e-02   3.563 0.000366 ***
## AGE         -9.213e-03  4.525e-03  -2.036 0.041773 *
## HOMEKIDS     6.658e-02  3.730e-02   1.785 0.074242 .
## YOJ         -6.191e-03  8.736e-03  -0.709 0.478503
## INCOME      -3.809e-07  1.046e-06  -0.364 0.715773
## HOME_VAL    -2.463e-06  3.296e-07  -7.472 7.91e-14 ***
## TRAVTIME     9.701e-03  2.118e-03   4.581 4.62e-06 ***
## BLUEBOOK    -8.540e-06  4.591e-06  -1.860 0.062875 .
```

```

## TIF          -4.460e-02  8.638e-03  -5.164  2.42e-07 ***
## OLDCLAIM     7.003e-06  3.890e-06   1.800  0.071808 .
## CLM_FREQ     2.690e-01  3.214e-02   8.370  < 2e-16 ***
## MVR_PTS      1.314e-01  1.583e-02   8.302  < 2e-16 ***
## CAR_AGE      -2.375e-02  6.705e-03  -3.542  0.000397 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5871.0  on 5067  degrees of freedom
## Residual deviance: 5293.9  on 5054  degrees of freedom
##    (1344 observations deleted due to missingness)
## AIC: 5321.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  972  277
##           1   53   78
##
##           Accuracy : 0.7609
##           95% CI : (0.7375, 0.7832)
##    No Information Rate : 0.7428
##    P-Value [Acc > NIR] : 0.06483
##
##           Kappa : 0.2117
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9483
##           Specificity : 0.2197
##           Pos Pred Value : 0.7782
##           Neg Pred Value : 0.5954
##           Prevalence : 0.7428
##           Detection Rate : 0.7043
##           Detection Prevalence : 0.9051
##           Balanced Accuracy : 0.5840
##
##           'Positive' Class : 0
##

```

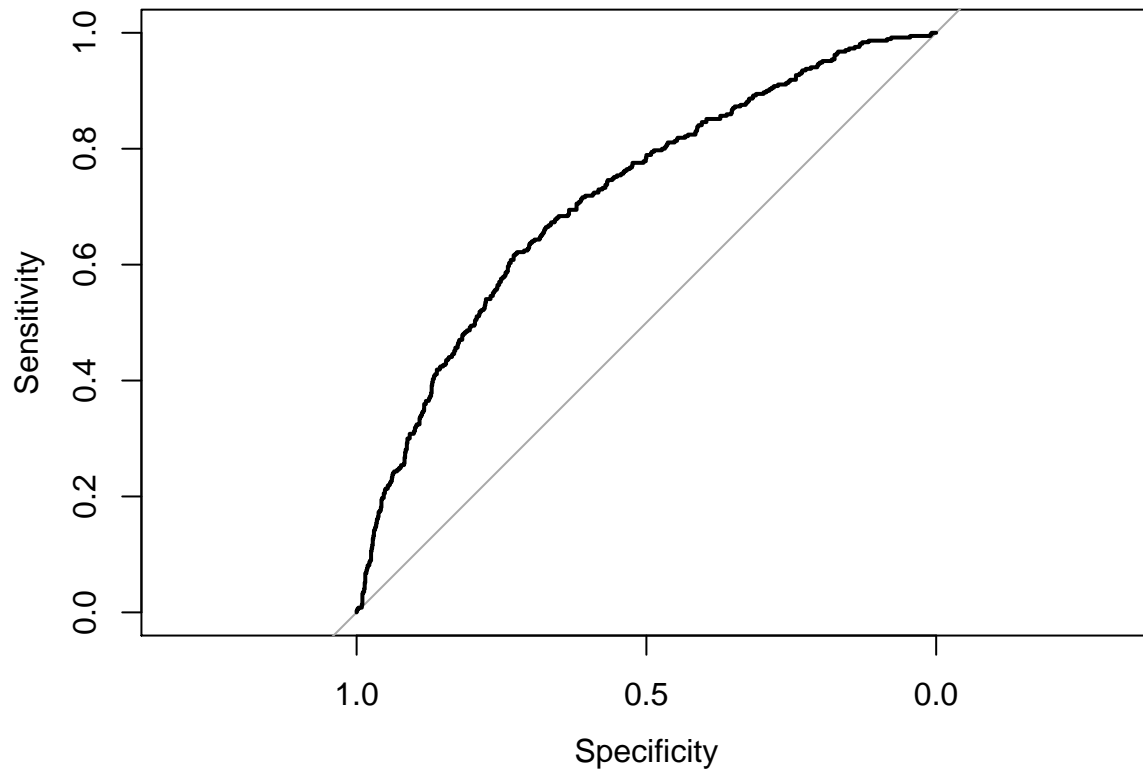


```
## [1] "AUC: 0.739696324287187"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1025 controls (dfPred_raw$class 0) < 355 cases (dfPred_raw$class 1).
## Area under the curve: 0.7397
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0389  -0.7690  -0.5706   0.8777   2.6548
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.239e-01  2.385e-01  -1.777 0.075526 .
## KIDSDRIV     2.596e-01  7.068e-02   3.673 0.000240 ***
## AGE         -1.003e-02  4.595e-03  -2.183 0.029033 *
## HOMEKIDS     8.367e-02  3.806e-02   2.198 0.027925 *
## YOJ         -1.141e-02  8.720e-03  -1.309 0.190549
## INCOME       4.472e-07  1.063e-06   0.421 0.674104
## HOME_VAL    -2.625e-06  3.310e-07  -7.931 2.18e-15 ***
## TRAVTIME     7.704e-03  2.132e-03   3.614 0.000301 ***
## BLUEBOOK    -1.305e-05  4.699e-06  -2.777 0.005490 **
```

```

## TIF          -4.397e-02  8.537e-03  -5.151  2.60e-07 ***
## OLDCLAIM      3.583e-06  3.916e-06   0.915  0.360247
## CLM_FREQ      2.739e-01  3.253e-02   8.421  < 2e-16 ***
## MVR_PTS       1.438e-01  1.599e-02   8.990  < 2e-16 ***
## CAR_AGE       -2.477e-02  6.734e-03  -3.678  0.000235 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5844  on 5073  degrees of freedom
## Residual deviance: 5233  on 5060  degrees of freedom
## (1338 observations deleted due to missingness)
## AIC: 5261
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  948  289
##           1   56   81
##
##           Accuracy : 0.7489
##           95% CI : (0.7251, 0.7716)
##    No Information Rate : 0.7307
##    P-Value [Acc > NIR] : 0.06733
##
##           Kappa : 0.2036
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9442
##           Specificity : 0.2189
##           Pos Pred Value : 0.7664
##           Neg Pred Value : 0.5912
##           Prevalence : 0.7307
##           Detection Rate : 0.6900
##           Detection Prevalence : 0.9003
##           Balanced Accuracy : 0.5816
##
##           'Positive' Class : 0
##

```



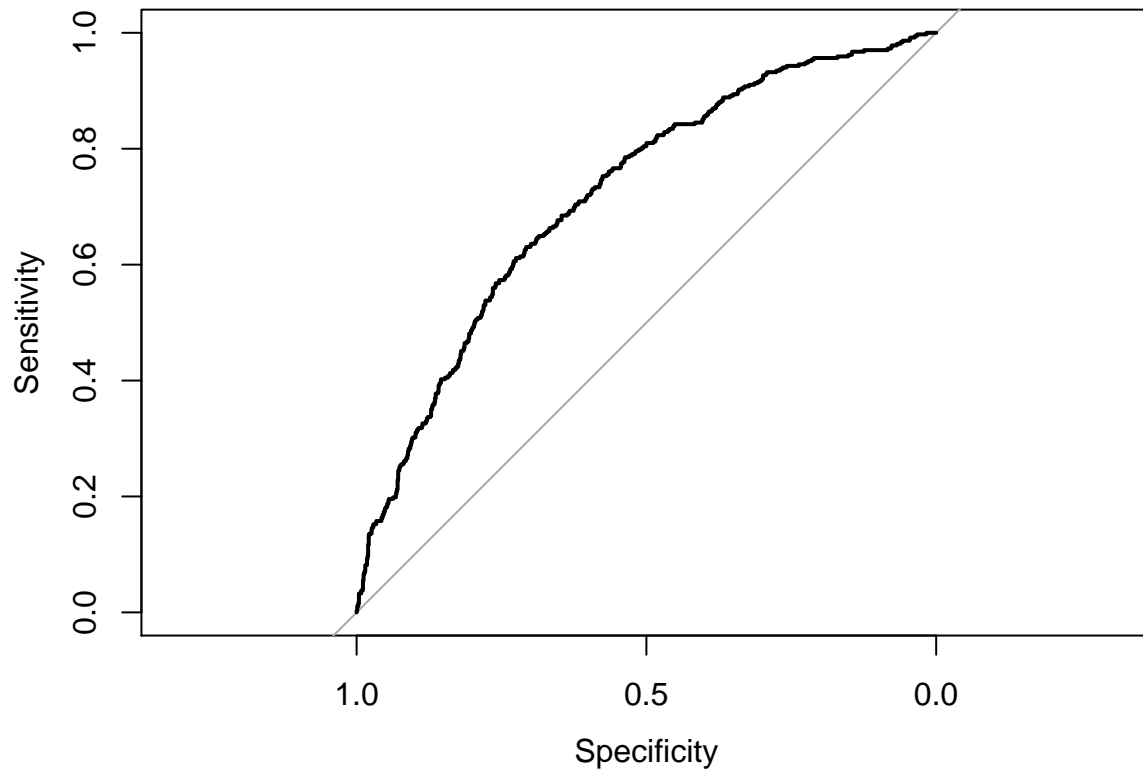
```
## [1] "AUC: 0.716722299989232"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1004 controls (dfPred_raw$class 0) < 370 cases (dfPred_raw$class 1).
## Area under the curve: 0.7167
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8990  -0.7618  -0.5739   0.9092   2.6067
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.832e-01  2.391e-01  -1.603  0.109008
## KIDSDRIV     1.450e-01  7.134e-02   2.032  0.042166 *
## AGE         -1.216e-02  4.572e-03  -2.659  0.007844 **
## HOMEKIDS      8.442e-02  3.768e-02   2.241  0.025049 *
## YOJ          -1.016e-02  8.676e-03  -1.171  0.241597
## INCOME        1.929e-07  1.046e-06   0.184  0.853720
## HOME_VAL     -2.350e-06  3.308e-07  -7.105  1.20e-12 ***
## TRAVTIME      7.750e-03  2.125e-03   3.648  0.000265 ***
## BLUEBOOK     -1.117e-05  4.641e-06  -2.408  0.016056 *
```

```

## TIF          -4.298e-02  8.638e-03  -4.976  6.49e-07  ***
## OLDCLAIM      5.189e-06  3.967e-06   1.308  0.190821
## CLM_FREQ      2.762e-01  3.215e-02   8.591  < 2e-16  ***
## MVR_PTS       1.474e-01  1.595e-02   9.236  < 2e-16  ***
## CAR_AGE       -2.543e-02  6.789e-03  -3.746  0.000180  ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5849.3  on 5075  degrees of freedom
## Residual deviance: 5245.9  on 5062  degrees of freedom
##    (1336 observations deleted due to missingness)
## AIC: 5273.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  948  297
##           1   56   71
##
##           Accuracy : 0.7427
##           95% CI : (0.7187, 0.7657)
##    No Information Rate : 0.7318
##    P-Value [Acc > NIR] : 0.1887
##
##           Kappa : 0.1731
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9442
##           Specificity : 0.1929
##           Pos Pred Value : 0.7614
##           Neg Pred Value : 0.5591
##           Prevalence : 0.7318
##           Detection Rate : 0.6910
##           Detection Prevalence : 0.9074
##           Balanced Accuracy : 0.5686
##
##           'Positive' Class : 0
##

```



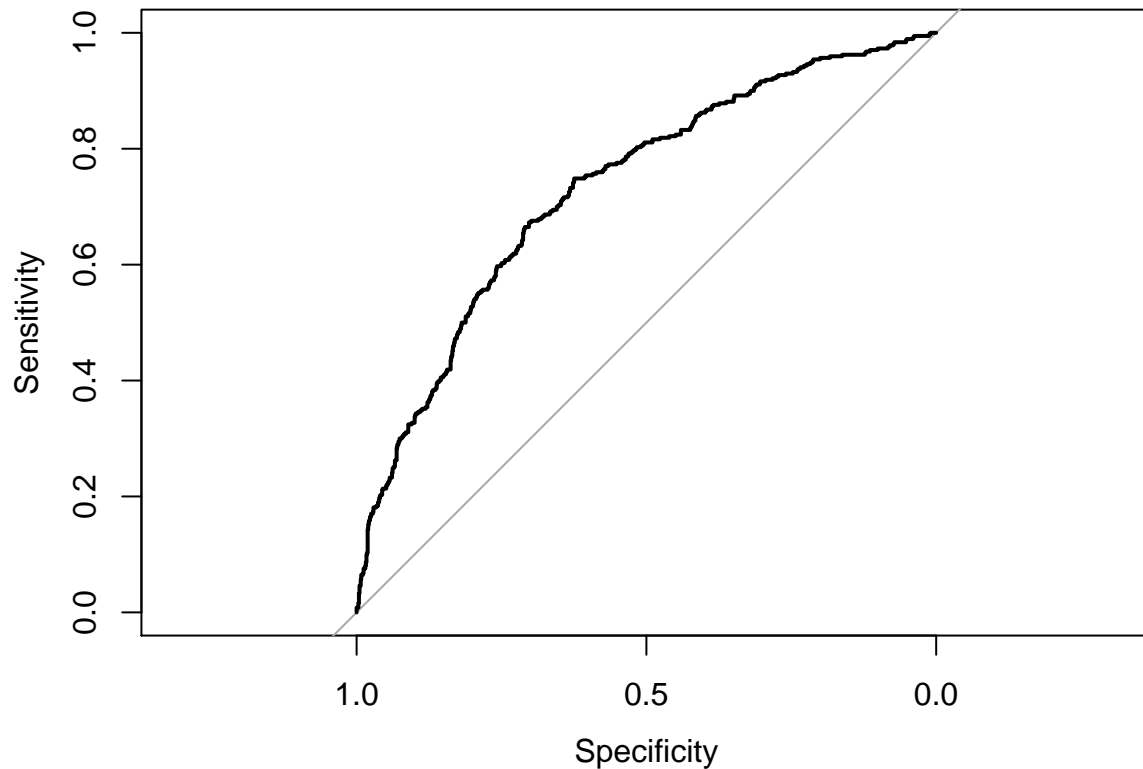


```
## [1] "AUC: 0.719583622033605"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1004 controls (dfPred_raw$class 0) < 368 cases (dfPred_raw$class 1).
## Area under the curve: 0.7196
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9216  -0.7697  -0.5790   0.9227   2.6723
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.914e-01  2.383e-01  -0.803  0.421896
## KIDSDRIV     2.002e-01  7.080e-02   2.828  0.004678 **
## AGE          -1.403e-02  4.583e-03  -3.062  0.002202 **
## HOMEKIDS      6.187e-02  3.782e-02   1.636  0.101929
## YOJ          -7.062e-03  8.687e-03  -0.813  0.416239
## INCOME        -8.371e-07  1.053e-06  -0.795  0.426671
## HOME_VAL      -2.270e-06  3.326e-07  -6.826  8.72e-12 ***
## TRAVTIME       7.119e-03  2.130e-03   3.343  0.000829 ***
## BLUEBOOK      -1.170e-05  4.631e-06  -2.527  0.011511 *
```

```

## TIF          -4.506e-02  8.606e-03  -5.236 1.64e-07 ***
## OLDCLAIM      7.332e-06  3.936e-06   1.863 0.062517 .
## CLM_FREQ      2.529e-01  3.245e-02   7.795 6.46e-15 ***
## MVR_PTS       1.357e-01  1.607e-02   8.445 < 2e-16 ***
## CAR_AGE       -2.671e-02  6.784e-03  -3.937 8.26e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5847.0  on 5078  degrees of freedom
## Residual deviance: 5265.5  on 5065  degrees of freedom
##    (1333 observations deleted due to missingness)
## AIC: 5293.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  949  291
##           1   50   79
##
##               Accuracy : 0.7509
##               95% CI : (0.7271, 0.7736)
##    No Information Rate : 0.7297
##    P-Value [Acc > NIR] : 0.04055
##
##               Kappa : 0.2056
##
## Mcnemar's Test P-Value : < 2e-16
##
##               Sensitivity : 0.9499
##               Specificity : 0.2135
##    Pos Pred Value : 0.7653
##    Neg Pred Value : 0.6124
##    Prevalence : 0.7297
##    Detection Rate : 0.6932
##    Detection Prevalence : 0.9058
##    Balanced Accuracy : 0.5817
##
##    'Positive' Class : 0
##

```

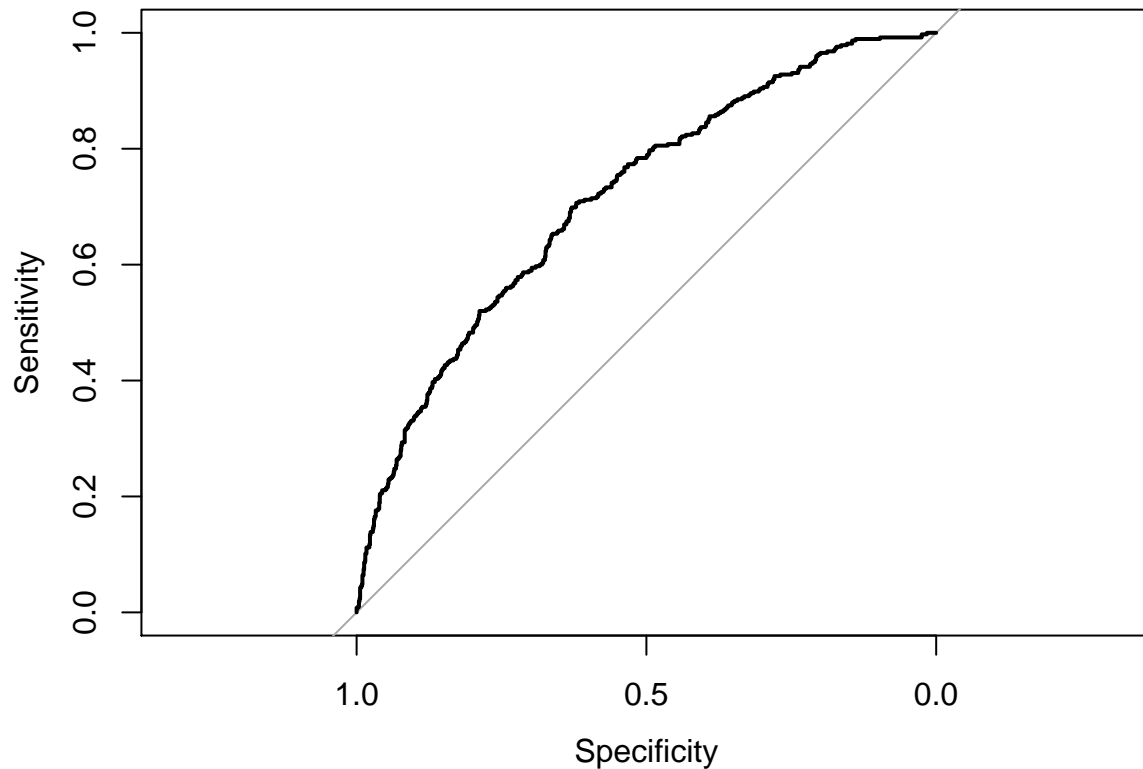


```
## [1] "AUC: 0.731417904390877"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 999 controls (dfPred_raw$class 0) < 370 cases (dfPred_raw$class 1).
## Area under the curve: 0.7314
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0159  -0.7671  -0.5725   0.9082   2.6977
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.728e-01  2.389e-01  -1.560  0.118661
## KIDSDRIV     2.388e-01  6.904e-02   3.460  0.000541 ***
## AGE         -8.256e-03  4.571e-03  -1.806  0.070889 .
## HOMEKIDS     7.042e-02  3.763e-02   1.871  0.061310 .
## YOJ         -9.854e-03  8.832e-03  -1.116  0.264571
## INCOME       3.010e-07  1.049e-06   0.287  0.774164
## HOME_VAL    -2.440e-06  3.292e-07  -7.412  1.24e-13 ***
## TRAVTIME     8.583e-03  2.119e-03   4.050  5.11e-05 ***
## BLUEBOOK    -1.822e-05  4.674e-06  -3.897  9.72e-05 ***
```

```

## TIF          -5.339e-02  8.697e-03  -6.139 8.32e-10 ***
## OLDCLAIM      8.771e-06  3.895e-06   2.252 0.024335 *
## CLM_FREQ      2.393e-01  3.252e-02   7.360 1.84e-13 ***
## MVR_PTS       1.333e-01  1.590e-02   8.385 < 2e-16 ***
## CAR_AGE       -2.815e-02  6.775e-03  -4.155 3.25e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5833  on 5072  degrees of freedom
## Residual deviance: 5239  on 5059  degrees of freedom
## (1339 observations deleted due to missingness)
## AIC: 5267
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 956 298
##           1  44  77
##
##           Accuracy : 0.7513
##           95% CI : (0.7275, 0.7739)
##    No Information Rate : 0.7273
##    P-Value [Acc > NIR] : 0.02375
##
##           Kappa : 0.2047
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9560
##           Specificity : 0.2053
##           Pos Pred Value : 0.7624
##           Neg Pred Value : 0.6364
##           Prevalence : 0.7273
##           Detection Rate : 0.6953
##           Detection Prevalence : 0.9120
##           Balanced Accuracy : 0.5807
##
##           'Positive' Class : 0
##

```

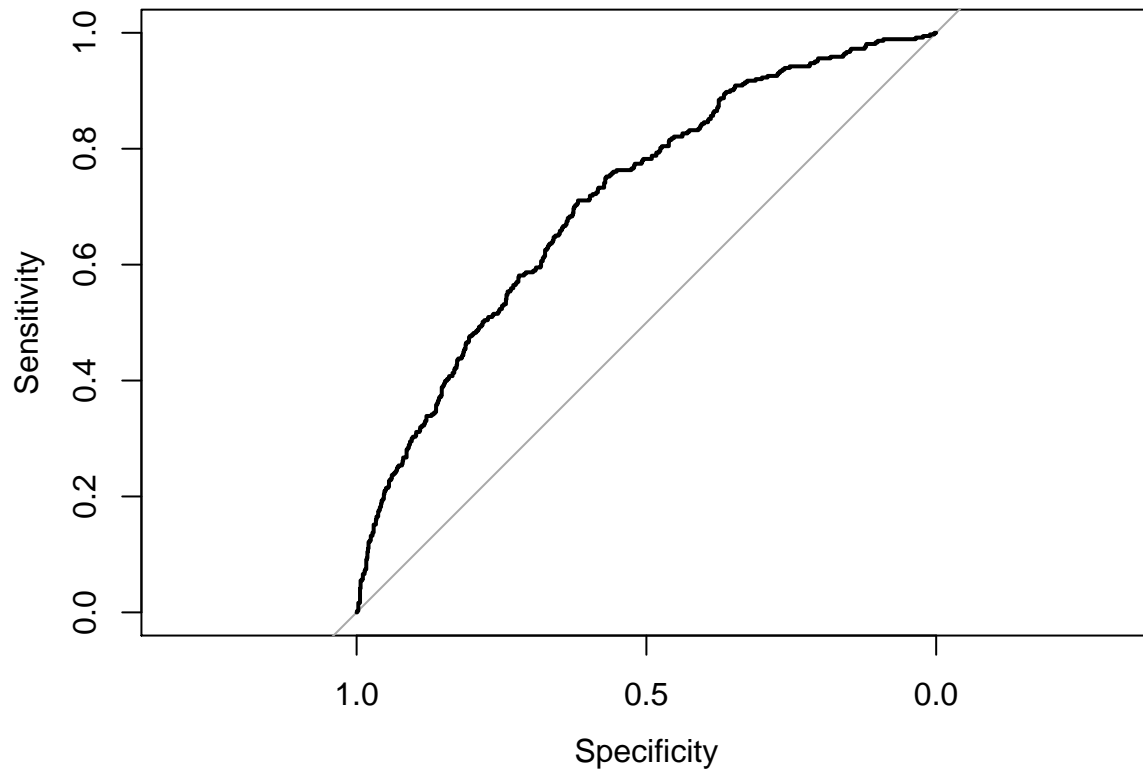


```
## [1] "AUC: 0.716368"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1000 controls (dfPred_raw$class 0) < 375 cases (dfPred_raw$class 1).
## Area under the curve: 0.7164
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7860  -0.7671  -0.5731   0.9067   2.6203
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.644e-01  2.367e-01  -1.963  0.04970 *
## KIDSDRIV     3.061e-01  6.899e-02   4.437 9.12e-06 ***
## AGE          -1.190e-02  4.591e-03  -2.591  0.00956 **
## HOMEKIDS      5.534e-02  3.805e-02   1.454  0.14582
## YOJ          -9.749e-03  8.720e-03  -1.118  0.26357
## INCOME        -1.246e-08  1.048e-06  -0.012  0.99052
## HOME_VAL      -2.702e-06  3.311e-07  -8.160 3.35e-16 ***
## TRAVTIME      8.401e-03  2.147e-03   3.914 9.09e-05 ***
## BLUEBOOK      -6.557e-06  4.641e-06  -1.413  0.15773
```

```

## TIF          -4.388e-02  8.513e-03  -5.155  2.54e-07 ***
## OLDCLAIM     1.908e-06  3.954e-06   0.483  0.62941
## CLM_FREQ     2.940e-01  3.206e-02   9.172  < 2e-16 ***
## MVR_PTS      1.428e-01  1.579e-02   9.047  < 2e-16 ***
## CAR_AGE      -2.114e-02  6.686e-03  -3.162  0.00157 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5852.2  on 5063  degrees of freedom
## Residual deviance: 5240.1  on 5050  degrees of freedom
##    (1348 observations deleted due to missingness)
## AIC: 5268.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  970  287
##           1   51   76
##
##           Accuracy : 0.7558
##           95% CI : (0.7323, 0.7782)
##    No Information Rate : 0.7377
##    P-Value [Acc > NIR] : 0.06636
##
##           Kappa : 0.2017
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9500
##           Specificity : 0.2094
##           Pos Pred Value : 0.7717
##           Neg Pred Value : 0.5984
##           Prevalence : 0.7377
##           Detection Rate : 0.7009
##           Detection Prevalence : 0.9082
##           Balanced Accuracy : 0.5797
##
##           'Positive' Class : 0
##

```



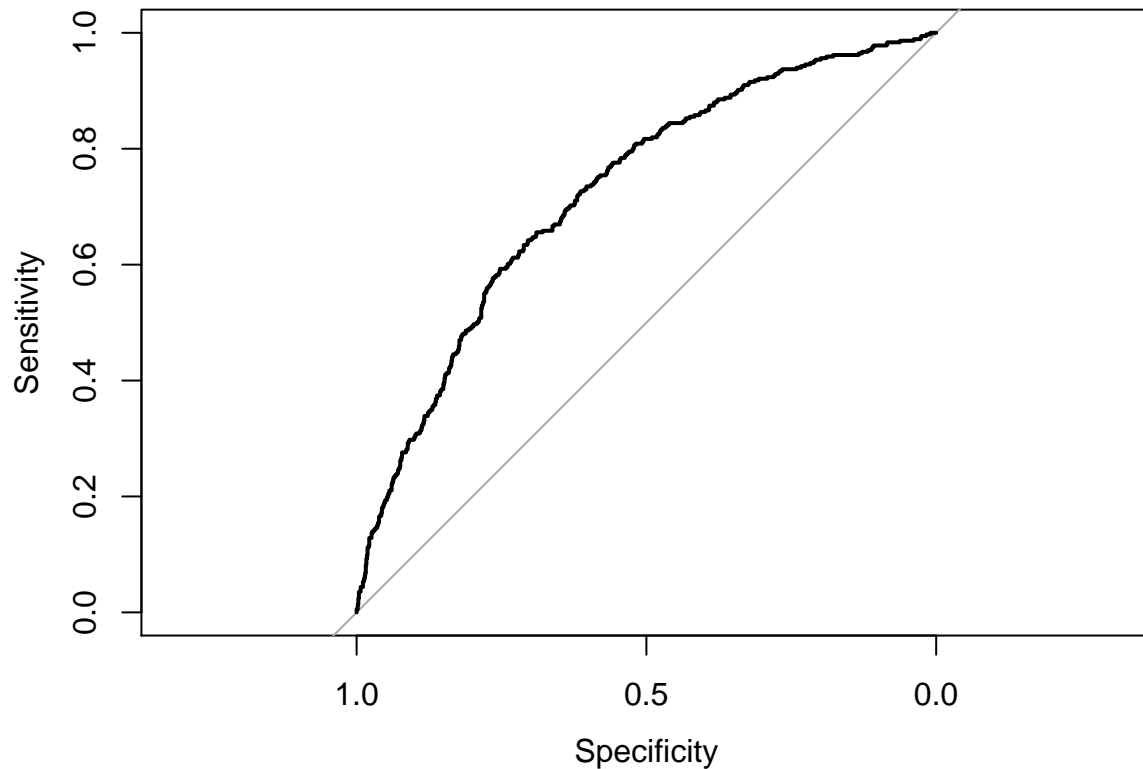
```
## [1] "AUC: 0.712152240956444"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1021 controls (dfPred_raw$class 0) < 363 cases (dfPred_raw$class 1).
## Area under the curve: 0.7122
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9348  -0.7683  -0.5755   0.9079   2.6708
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.909e-01  2.383e-01  -1.221 0.222119
## KIDSDRIV     2.241e-01  7.042e-02   3.183 0.001460 **
## AGE         -1.413e-02  4.592e-03  -3.078 0.002085 **
## HOMEKIDS      6.479e-02  3.760e-02   1.723 0.084811 .
## YOJ          -3.914e-03  8.663e-03  -0.452 0.651415
## INCOME       -5.095e-07  1.038e-06  -0.491 0.623687
## HOME_VAL    -2.437e-06  3.307e-07  -7.370 1.71e-13 ***
## TRAVTIME      6.825e-03  2.137e-03   3.193 0.001407 **
## BLUEBOOK     -1.139e-05  4.640e-06  -2.454 0.014119 *
```

```

## TIF          -3.944e-02  8.519e-03  -4.629 3.67e-06 ***
## OLDCLAIM     6.248e-06  3.992e-06   1.565 0.117526
## CLM_FREQ     2.448e-01  3.231e-02   7.575 3.60e-14 ***
## MVR_PTS      1.499e-01  1.588e-02   9.437 < 2e-16 ***
## CAR_AGE      -2.246e-02  6.758e-03  -3.324 0.000888 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5851.6  on 5072  degrees of freedom
## Residual deviance: 5252.7  on 5059  degrees of freedom
##    (1339 observations deleted due to missingness)
## AIC: 5280.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 959 296
##           1  50  70
##
##           Accuracy : 0.7484
##           95% CI : (0.7245, 0.7711)
##    No Information Rate : 0.7338
##    P-Value [Acc > NIR] : 0.1166
##
##           Kappa : 0.1803
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9504
##           Specificity : 0.1913
##           Pos Pred Value : 0.7641
##           Neg Pred Value : 0.5833
##           Prevalence : 0.7338
##           Detection Rate : 0.6975
##    Detection Prevalence : 0.9127
##           Balanced Accuracy : 0.5709
##
##           'Positive' Class : 0
##

```



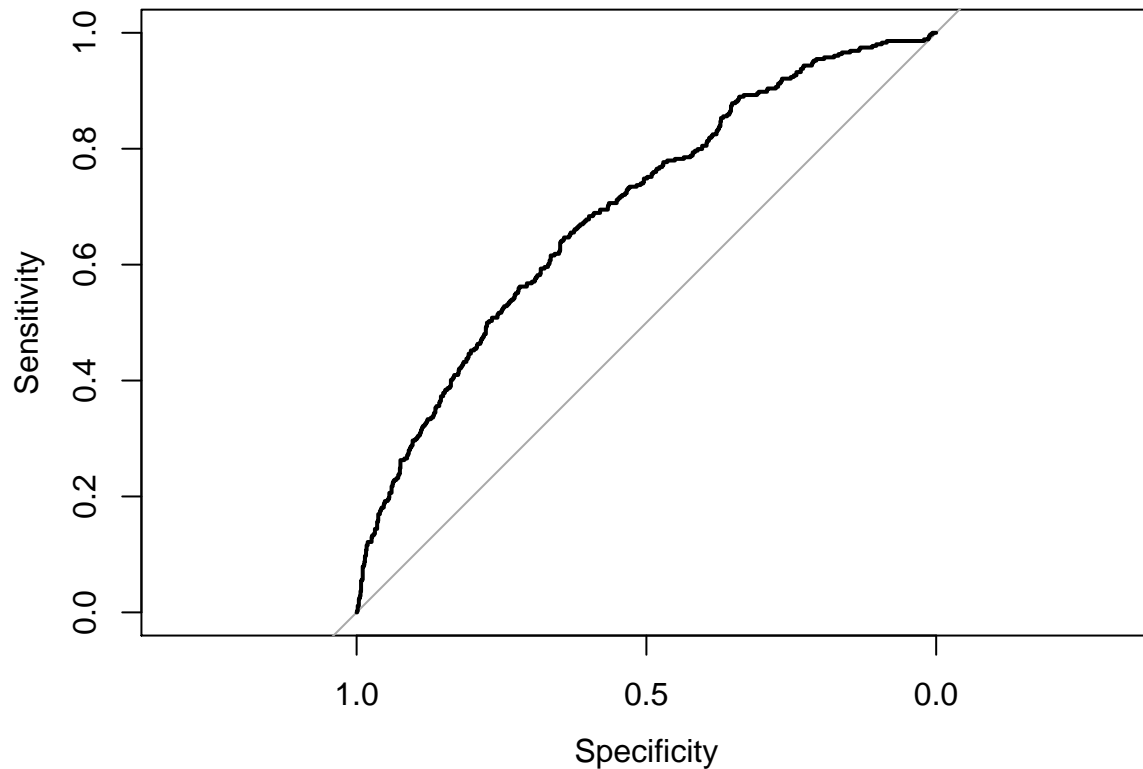


```
## [1] "AUC: 0.724728806858492"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1009 controls (dfPred_raw$class 0) < 366 cases (dfPred_raw$class 1).
## Area under the curve: 0.7247
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7735  -0.7739  -0.5684   0.9226   2.7311
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.329e-01  2.375e-01  -1.402  0.160994
## KIDSDRIV     2.838e-01  6.960e-02   4.078  4.54e-05 ***
## AGE         -1.223e-02  4.572e-03  -2.676  0.007460 **
## HOMEKIDS      5.826e-02  3.804e-02   1.531  0.125674
## YOJ          -2.010e-03  8.771e-03  -0.229  0.818767
## INCOME       -1.375e-07  1.033e-06  -0.133  0.894104
## HOME_VAL    -2.720e-06  3.306e-07  -8.227  < 2e-16 ***
## TRAVTIME      7.667e-03  2.131e-03   3.598  0.000321 ***
## BLUEBOOK    -1.473e-05  4.677e-06  -3.150  0.001631 **
```

```

## TIF          -4.538e-02  8.564e-03  -5.299 1.16e-07 ***
## OLDCLAIM      6.462e-06  3.994e-06   1.618 0.105686
## CLM_FREQ      2.926e-01  3.233e-02   9.052 < 2e-16 ***
## MVR_PTS       1.333e-01  1.596e-02   8.355 < 2e-16 ***
## CAR_AGE       -2.287e-02  6.744e-03  -3.392 0.000695 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5861.2  on 5048  degrees of freedom
## Residual deviance: 5219.8  on 5035  degrees of freedom
##    (1363 observations deleted due to missingness)
## AIC: 5247.8
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 982 278
##           1  63  76
##
##           Accuracy : 0.7563
##           95% CI : (0.7329, 0.7786)
##    No Information Rate : 0.747
##    P-Value [Acc > NIR] : 0.2217
##
##           Kappa : 0.1932
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9397
##           Specificity : 0.2147
##           Pos Pred Value : 0.7794
##           Neg Pred Value : 0.5468
##           Prevalence : 0.7470
##           Detection Rate : 0.7019
##           Detection Prevalence : 0.9006
##           Balanced Accuracy : 0.5772
##
##           'Positive' Class : 0
##

```

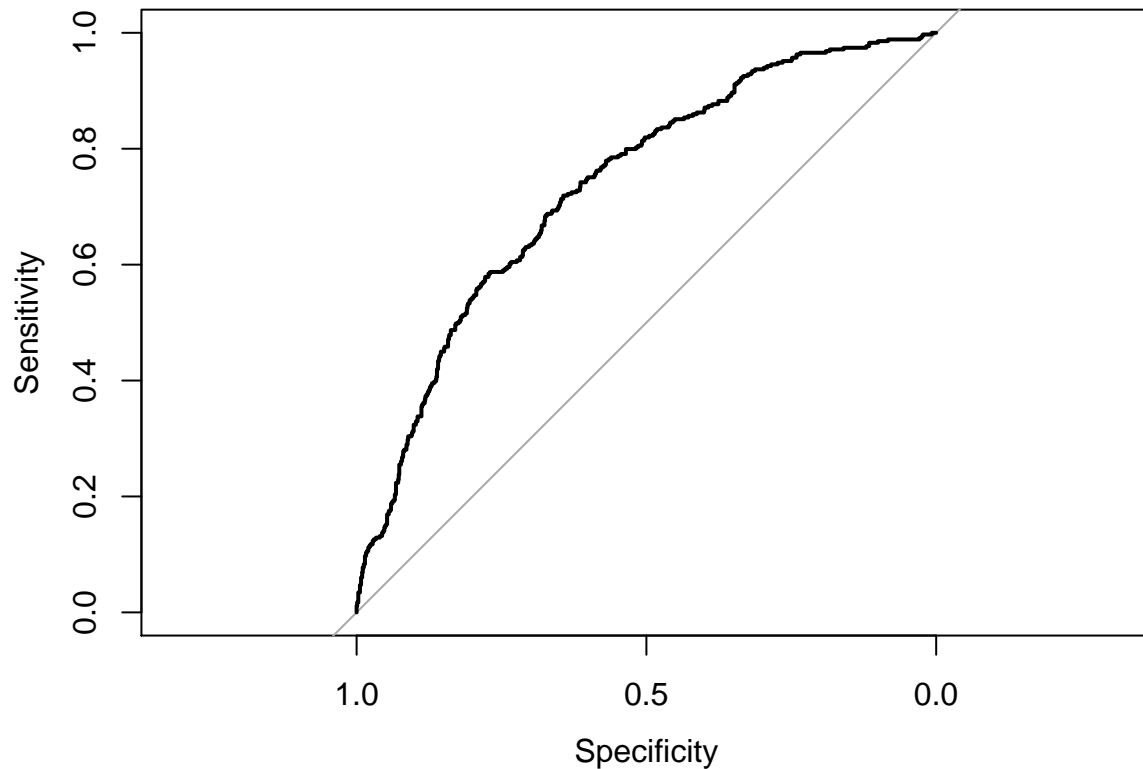


```
## [1] "AUC: 0.693552834320006"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1045 controls (dfPred_raw$class 0) < 354 cases (dfPred_raw$class 1).
## Area under the curve: 0.6936
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0069  -0.7791  -0.5826   0.9434   2.6284
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.142e-01  2.364e-01  -2.175  0.029600 *
## KIDSDRIV     2.395e-01  6.977e-02   3.433  0.000597 ***
## AGE         -9.078e-03  4.503e-03  -2.016  0.043796 *
## HOMEKIDS      7.184e-02  3.742e-02   1.920  0.054865 .
## YOJ         -3.562e-03  8.709e-03  -0.409  0.682516
## INCOME      -8.252e-08  1.033e-06  -0.080  0.936321
## HOME_VAL    -2.537e-06  3.266e-07  -7.767  8.02e-15 ***
## TRAVTIME     8.099e-03  2.092e-03   3.871  0.000108 ***
## BLUEBOOK    -1.251e-05  4.616e-06  -2.710  0.006730 **
```

```

## TIF          -4.573e-02  8.578e-03  -5.331  9.78e-08 ***
## OLDCLAIM      8.309e-06  3.966e-06   2.095  0.036165 *
## CLM_FREQ      2.461e-01  3.222e-02   7.636  2.24e-14 ***
## MVR_PTS       1.392e-01  1.572e-02   8.853  < 2e-16 ***
## CAR_AGE       -2.126e-02  6.682e-03  -3.182  0.001465 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5873.8  on 5052  degrees of freedom
## Residual deviance: 5283.6  on 5039  degrees of freedom
##    (1359 observations deleted due to missingness)
## AIC: 5311.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 988 289
##           1  58  60
##
##           Accuracy : 0.7513
##           95% CI : (0.7277, 0.7737)
##    No Information Rate : 0.7498
##    P-Value [Acc > NIR] : 0.4651
##
##           Kappa : 0.1494
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9446
##           Specificity : 0.1719
##           Pos Pred Value : 0.7737
##           Neg Pred Value : 0.5085
##           Prevalence : 0.7498
##           Detection Rate : 0.7082
##           Detection Prevalence : 0.9154
##           Balanced Accuracy : 0.5582
##
##           'Positive' Class : 0
##

```

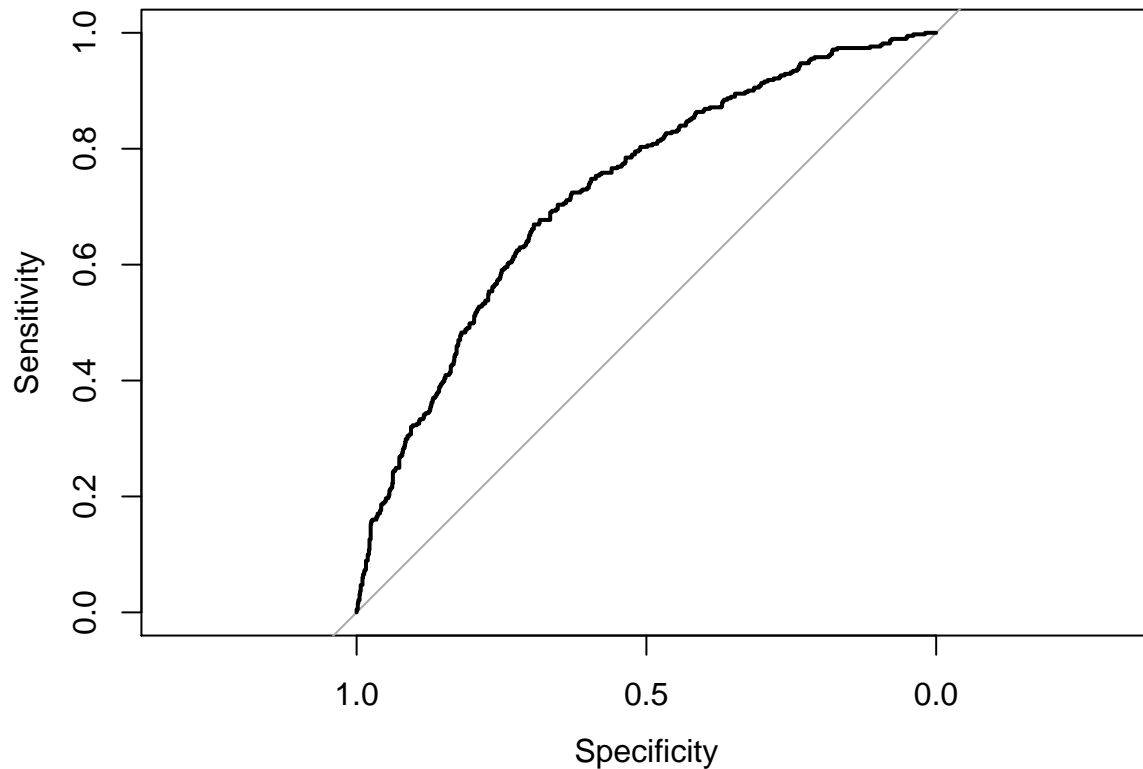


```
## [1] "AUC: 0.735559670624072"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1046 controls (dfPred_raw$class 0) < 349 cases (dfPred_raw$class 1).
## Area under the curve: 0.7356
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9571  -0.7616  -0.5742   0.8954   2.6473
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.374e-01  2.398e-01  -1.824  0.068132 .
## KIDSDRIV     2.104e-01  7.125e-02   2.953  0.003149 **
## AGE          -1.130e-02  4.567e-03  -2.474  0.013345 *
## HOMEKIDS      6.286e-02  3.820e-02   1.645  0.099888 .
## YOJ           -2.883e-03  8.746e-03  -0.330  0.741713
## INCOME        -6.162e-07  1.054e-06  -0.585  0.558877
## HOME_VAL      -2.149e-06  3.334e-07  -6.447  1.14e-10 ***
## TRAVTIME       7.463e-03  2.132e-03   3.500  0.000465 ***
## BLUEBOOK      -1.249e-05  4.691e-06  -2.663  0.007740 **
```

```

## TIF          -4.440e-02  8.628e-03  -5.146  2.66e-07 ***
## OLDCLAIM      8.396e-06  3.947e-06   2.127  0.033384 *
## CLM_FREQ      2.828e-01  3.243e-02   8.718  < 2e-16 ***
## MVR_PTS       1.340e-01  1.606e-02   8.345  < 2e-16 ***
## CAR_AGE       -2.594e-02  6.805e-03  -3.812  0.000138 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5822.4  on 5075  degrees of freedom
## Residual deviance: 5235.7  on 5062  degrees of freedom
##    (1336 observations deleted due to missingness)
## AIC: 5263.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 939 306
##           1  52  75
##
##           Accuracy : 0.7391
##           95% CI : (0.715, 0.7621)
##    No Information Rate : 0.7223
##    P-Value [Acc > NIR] : 0.08688
##
##           Kappa : 0.1816
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9475
##           Specificity : 0.1969
##           Pos Pred Value : 0.7542
##           Neg Pred Value : 0.5906
##           Prevalence : 0.7223
##           Detection Rate : 0.6844
##           Detection Prevalence : 0.9074
##           Balanced Accuracy : 0.5722
##
##           'Positive' Class : 0
##

```



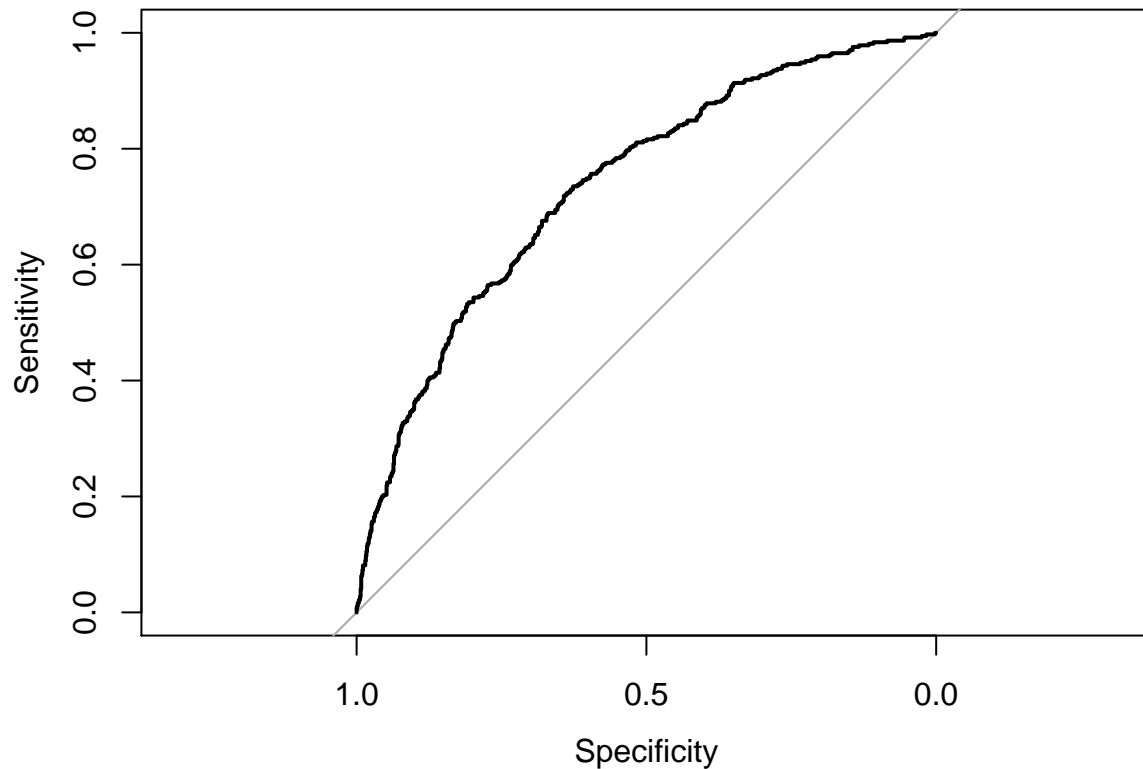
```
## [1] "AUC: 0.726530904121344"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 991 controls (dfPred_raw$class 0) < 381 cases (dfPred_raw$class 1).
## Area under the curve: 0.7265
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0567  -0.7703  -0.5800   0.9137   2.6149
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.764e-01  2.372e-01  -2.009 0.044584 *
## KIDSDRIV     2.791e-01  6.860e-02   4.069 4.72e-05 ***
## AGE         -8.847e-03  4.558e-03  -1.941 0.052294 .
## HOMEKIDS     5.639e-02  3.793e-02   1.487 0.137060
## YOJ         -1.252e-02  8.777e-03  -1.427 0.153645
## INCOME       6.707e-09  1.052e-06   0.006 0.994914
## HOME_VAL    -2.483e-06  3.288e-07  -7.553 4.24e-14 ***
## TRAVTIME     8.786e-03  2.119e-03   4.146 3.38e-05 ***
## BLUEBOOK    -9.075e-06  4.629e-06  -1.961 0.049935 *
```

```

## TIF          -5.216e-02  8.597e-03  -6.067  1.30e-09 ***
## OLDCLAIM     4.802e-06  3.921e-06   1.225  0.220680
## CLM_FREQ     2.592e-01  3.215e-02   8.063  7.46e-16 ***
## MVR_PTS      1.369e-01  1.579e-02   8.667  < 2e-16 ***
## CAR_AGE      -2.431e-02  6.682e-03  -3.638  0.000275 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5842.1  on 5070  degrees of freedom
## Residual deviance: 5269.6  on 5057  degrees of freedom
##    (1341 observations deleted due to missingness)
## AIC: 5297.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 963 297
##           1  44  73
##
##           Accuracy : 0.7524
##           95% CI : (0.7287, 0.775)
##    No Information Rate : 0.7313
##    P-Value [Acc > NIR] : 0.04071
##
##           Kappa : 0.196
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9563
##           Specificity : 0.1973
##           Pos Pred Value : 0.7643
##           Neg Pred Value : 0.6239
##           Prevalence : 0.7313
##           Detection Rate : 0.6993
##           Detection Prevalence : 0.9150
##           Balanced Accuracy : 0.5768
##
##           'Positive' Class : 0
##

```



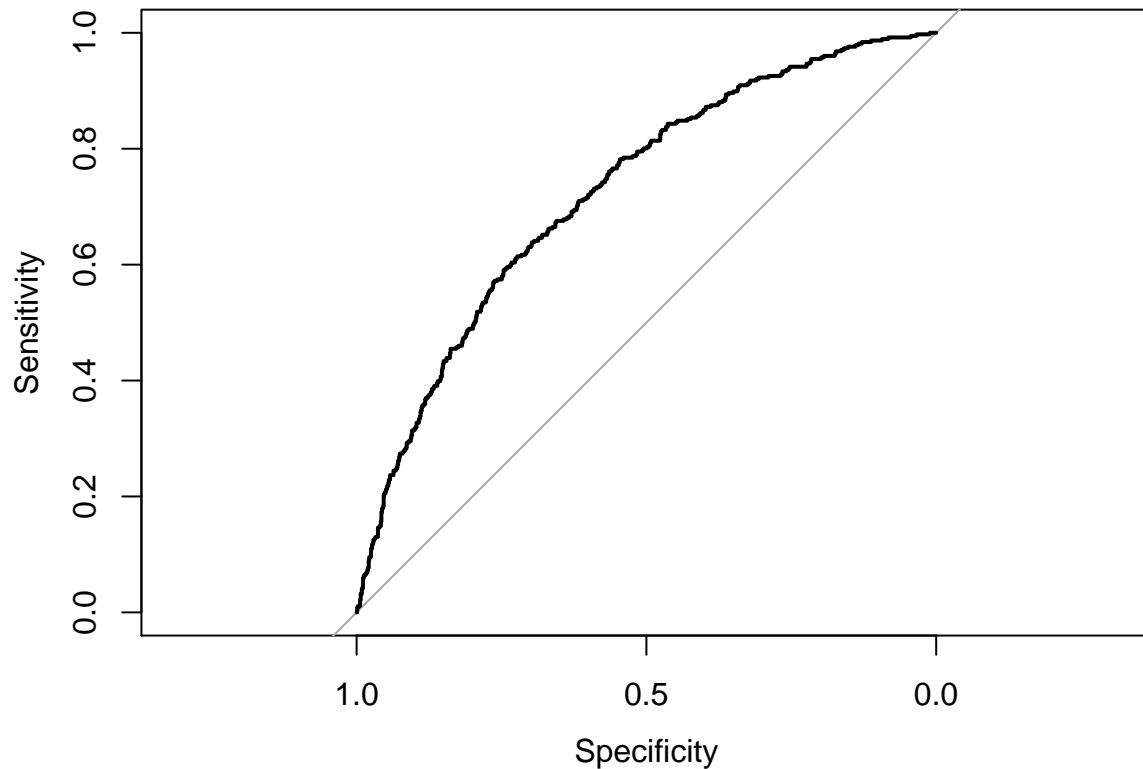


```
## [1] "AUC: 0.736756219973698"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1007 controls (dfPred_raw$class 0) < 370 cases (dfPred_raw$class 1).
## Area under the curve: 0.7368
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9482  -0.7676  -0.5739   0.8862   2.5881
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.031e-01  2.387e-01  -1.688 0.091335 .
## KIDSDRIV     2.237e-01  6.974e-02   3.207 0.001341 **
## AGE          -1.167e-02  4.596e-03  -2.540 0.011084 *
## HOMEKIDS      7.882e-02  3.775e-02   2.088 0.036806 *
## YOJ          -1.259e-02  8.673e-03  -1.451 0.146757
## INCOME        8.034e-07  1.056e-06   0.761 0.446750
## HOME_VAL     -2.597e-06  3.297e-07  -7.877 3.36e-15 ***
## TRAVTIME      7.124e-03  2.123e-03   3.356 0.000792 ***
## BLUEBOOK     -1.020e-05  4.654e-06  -2.193 0.028327 *
```

```

## TIF          -4.136e-02  8.530e-03  -4.848  1.25e-06 ***
## OLDCLAIM     2.918e-06  3.976e-06   0.734  0.463068
## CLM_FREQ     2.600e-01  3.230e-02   8.048  8.44e-16 ***
## MVR_PTS      1.522e-01  1.581e-02   9.626  < 2e-16 ***
## CAR_AGE      -2.427e-02  6.721e-03  -3.611  0.000305 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5826.7  on 5065  degrees of freedom
## Residual deviance: 5231.9  on 5052  degrees of freedom
##    (1346 observations deleted due to missingness)
## AIC: 5259.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 962 311
##           1  44  65
##
##           Accuracy : 0.7431
##           95% CI : (0.7192, 0.766)
##    No Information Rate : 0.7279
##    P-Value [Acc > NIR] : 0.1072
##
##           Kappa : 0.1661
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9563
##           Specificity : 0.1729
##           Pos Pred Value : 0.7557
##           Neg Pred Value : 0.5963
##           Prevalence : 0.7279
##           Detection Rate : 0.6961
##           Detection Prevalence : 0.9211
##           Balanced Accuracy : 0.5646
##
##           'Positive' Class : 0
##

```

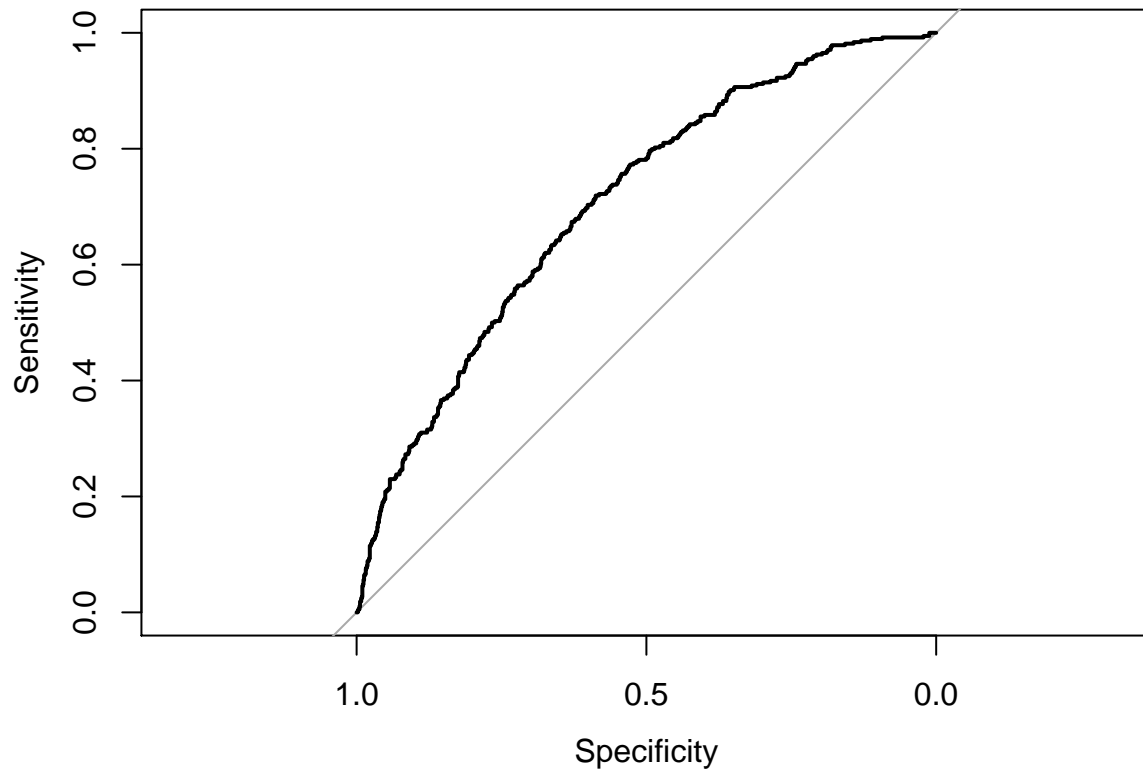


```
## [1] "AUC: 0.725355314918997"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1006 controls (dfPred_raw$class 0) < 376 cases (dfPred_raw$class 1).
## Area under the curve: 0.7254
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7573  -0.7622  -0.5687   0.8928   2.6942
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.026e-01  2.400e-01  -2.094 0.036268 *
## KIDSDRIV     3.071e-01  6.930e-02   4.431 9.37e-06 ***
## AGE          -1.153e-02  4.605e-03  -2.503 0.012297 *
## HOMEKIDS      5.509e-02  3.800e-02   1.450 0.147182
## YOJ           1.542e-03  8.801e-03   0.175 0.860884
## INCOME        -5.857e-07  1.039e-06  -0.564 0.572739
## HOME_VAL      -2.453e-06  3.329e-07  -7.368 1.74e-13 ***
## TRAVTIME       7.696e-03  2.152e-03   3.576 0.000349 ***
## BLUEBOOK      -1.273e-05  4.702e-06  -2.706 0.006805 **
```

```

## TIF          -4.149e-02  8.572e-03  -4.841  1.29e-06 ***
## OLDCLAIM      7.490e-06  3.982e-06   1.881  0.059962 .
## CLM_FREQ      3.016e-01  3.232e-02   9.330  < 2e-16 ***
## MVR_PTS       1.315e-01  1.595e-02   8.245  < 2e-16 ***
## CAR_AGE       -2.197e-02  6.773e-03  -3.243  0.001182 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5829.0  on 5062  degrees of freedom
## Residual deviance: 5207.9  on 5049  degrees of freedom
##    (1349 observations deleted due to missingness)
## AIC: 5235.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  958  295
##           1   53   79
##
##           Accuracy : 0.7487
##           95% CI : (0.725, 0.7714)
##    No Information Rate : 0.73
##    P-Value [Acc > NIR] : 0.06057
##
##           Kappa : 0.1995
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9476
##           Specificity : 0.2112
##           Pos Pred Value : 0.7646
##           Neg Pred Value : 0.5985
##           Prevalence : 0.7300
##           Detection Rate : 0.6917
##           Detection Prevalence : 0.9047
##           Balanced Accuracy : 0.5794
##
##           'Positive' Class : 0
##

```

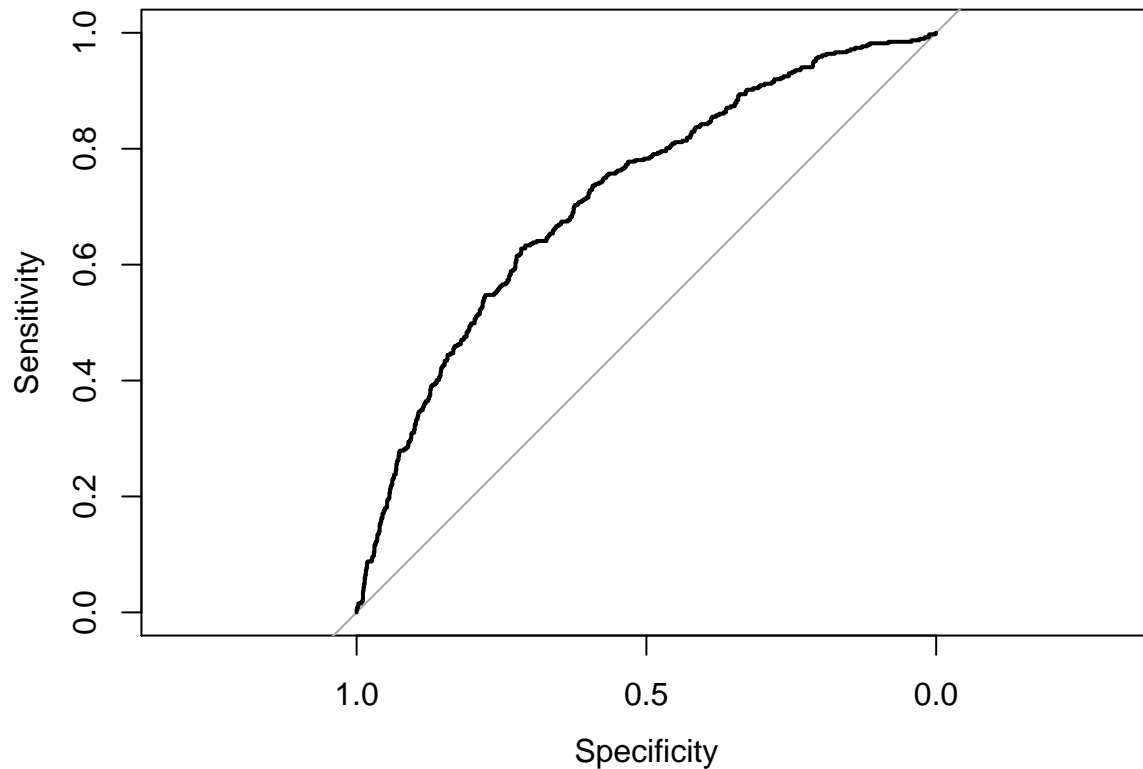


```
## [1] "AUC: 0.706681053862063"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1011 controls (dfPred_raw$class 0) < 374 cases (dfPred_raw$class 1).
## Area under the curve: 0.7067
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0538  -0.7616  -0.5685   0.8658   2.6350
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.463e-01  2.403e-01  -2.689 0.007157 **
## KIDSDRIV     2.805e-01  7.006e-02   4.004 6.23e-05 ***
## AGE          -8.205e-03  4.597e-03  -1.785 0.074279 .
## HOMEKIDS      6.876e-02  3.859e-02   1.782 0.074778 .
## YOJ          -1.963e-03  8.767e-03  -0.224 0.822825
## INCOME       -2.920e-07  1.064e-06  -0.274 0.783864
## HOME_VAL     -2.563e-06  3.340e-07  -7.676 1.65e-14 ***
## TRAVTIME      7.114e-03  2.138e-03   3.327 0.000878 ***
## BLUEBOOK     -9.215e-06  4.718e-06  -1.953 0.050768 .
```

```

## TIF          -4.023e-02  8.542e-03  -4.709  2.49e-06 ***
## OLDCLAIM      6.133e-06  4.077e-06   1.504  0.132518
## CLM_FREQ      2.786e-01  3.245e-02   8.583  < 2e-16 ***
## MVR_PTS       1.503e-01  1.603e-02   9.381  < 2e-16 ***
## CAR_AGE       -2.636e-02  6.807e-03  -3.872  0.000108 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5796.6  on 5053  degrees of freedom
## Residual deviance: 5190.0  on 5040  degrees of freedom
##    (1359 observations deleted due to missingness)
## AIC: 5218
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 950 310
##           1  57  77
##
##           Accuracy : 0.7367
##           95% CI : (0.7128, 0.7597)
##    No Information Rate : 0.7224
##    P-Value [Acc > NIR] : 0.1214
##
##           Kappa : 0.1782
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9434
##           Specificity : 0.1990
##           Pos Pred Value : 0.7540
##           Neg Pred Value : 0.5746
##           Prevalence : 0.7224
##           Detection Rate : 0.6815
##           Detection Prevalence : 0.9039
##           Balanced Accuracy : 0.5712
##
##           'Positive' Class : 0
##

```



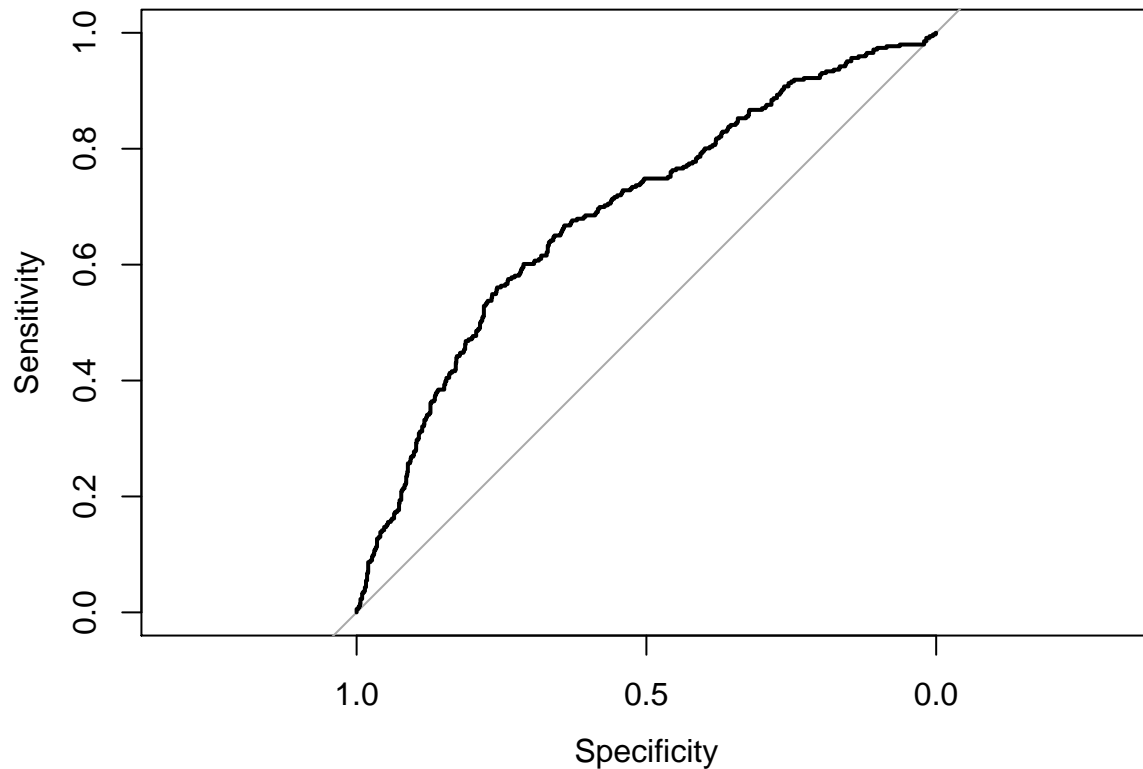
```
## [1] "AUC: 0.71690415155924"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1007 controls (dfPred_raw$class 0) < 387 cases (dfPred_raw$class 1).
## Area under the curve: 0.7169
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0908  -0.7764  -0.5652   0.8998   2.7612
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.649e-01  2.373e-01  -2.381 0.017274 *
## KIDSDRIV     2.433e-01  7.195e-02   3.382 0.000721 ***
## AGE        -7.093e-03  4.540e-03  -1.562 0.118220
## HOMEKIDS     9.978e-02  3.827e-02   2.607 0.009131 **
## YOJ        -2.249e-03  8.733e-03  -0.258 0.796761
## INCOME      -2.976e-07  1.052e-06  -0.283 0.777250
## HOME_VAL    -2.790e-06  3.317e-07  -8.411 < 2e-16 ***
## TRAVTIME     7.880e-03  2.122e-03   3.714 0.000204 ***
## BLUEBOOK    -1.598e-05  4.720e-06  -3.387 0.000708 ***
```

```

## TIF          -4.246e-02  8.570e-03  -4.954  7.28e-07 ***
## OLDCLAIM      7.839e-06  4.004e-06   1.958  0.050249 .
## CLM_FREQ      2.770e-01  3.255e-02   8.512  < 2e-16 ***
## MVR_PTS       1.465e-01  1.606e-02   9.125  < 2e-16 ***
## CAR_AGE       -2.417e-02  6.812e-03  -3.548  0.000388 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5877.9  on 5049  degrees of freedom
## Residual deviance: 5214.7  on 5036  degrees of freedom
##    (1363 observations deleted due to missingness)
## AIC: 5242.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  971  278
##           1   81   68
##
##           Accuracy : 0.7432
##           95% CI : (0.7195, 0.7659)
##    No Information Rate : 0.7525
##    P-Value [Acc > NIR] : 0.799
##
##           Kappa : 0.1478
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9230
##           Specificity : 0.1965
##           Pos Pred Value : 0.7774
##           Neg Pred Value : 0.4564
##           Prevalence : 0.7525
##           Detection Rate : 0.6946
##           Detection Prevalence : 0.8934
##           Balanced Accuracy : 0.5598
##
##           'Positive' Class : 0
##

```



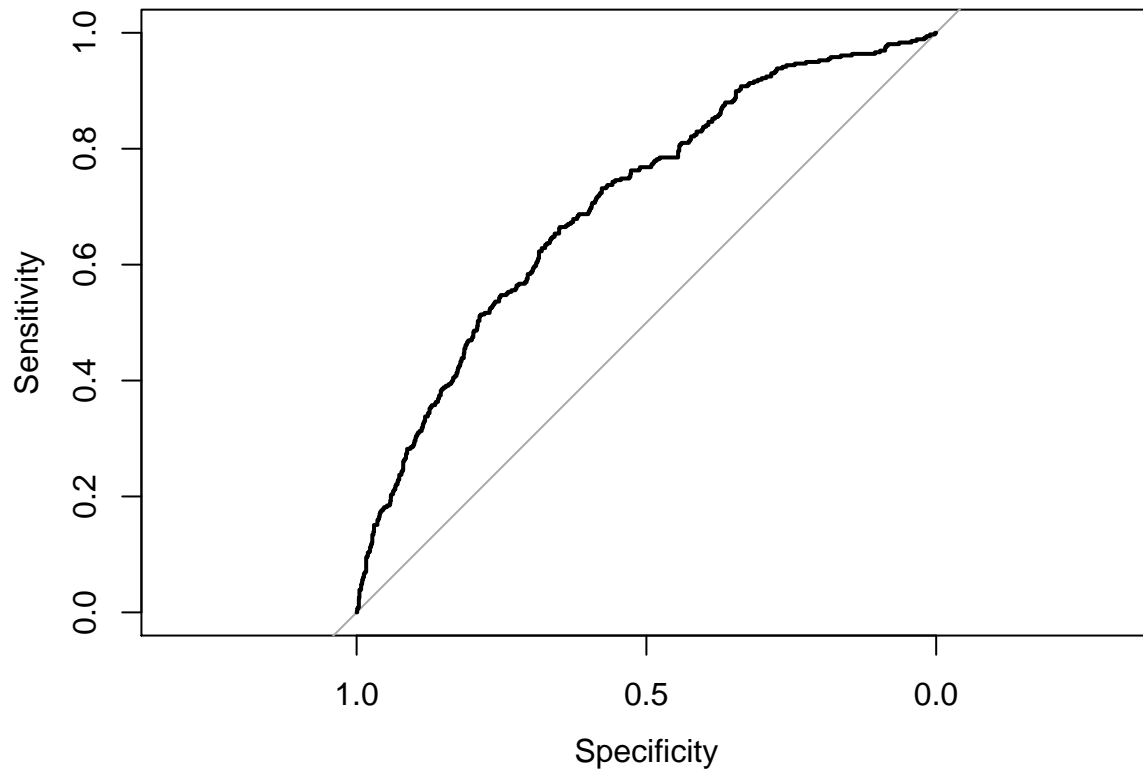


```
## [1] "AUC: 0.688754148442823"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1052 controls (dfPred_raw$class 0) < 346 cases (dfPred_raw$class 1).
## Area under the curve: 0.6888
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8430  -0.7687  -0.5722   0.9143   2.6412
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.158e-01  2.375e-01  -2.593  0.00951 **
## KIDSDRIV     2.168e-01  7.062e-02   3.070  0.00214 **
## AGE          -8.670e-03  4.561e-03  -1.901  0.05731 .
## HOMEKIDS      8.963e-02  3.780e-02   2.371  0.01773 *
## YOJ          -6.170e-03  8.706e-03  -0.709  0.47852
## INCOME       -2.144e-07  1.057e-06  -0.203  0.83927
## HOME_VAL     -2.692e-06  3.331e-07  -8.080 6.47e-16 ***
## TRAVTIME      8.913e-03  2.135e-03   4.174 2.99e-05 ***
## BLUEBOOK     -9.212e-06  4.647e-06  -1.983  0.04742 *
```

```

## TIF          -3.884e-02  8.627e-03  -4.501  6.75e-06 ***
## OLDCLAIM     5.925e-06  3.984e-06   1.487  0.13695
## CLM_FREQ     2.850e-01  3.220e-02   8.851  < 2e-16 ***
## MVR_PTS      1.475e-01  1.596e-02   9.239  < 2e-16 ***
## CAR_AGE      -2.596e-02  6.818e-03  -3.807  0.00014 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5858.1  on 5056  degrees of freedom
## Residual deviance: 5229.1  on 5043  degrees of freedom
##    (1356 observations deleted due to missingness)
## AIC: 5257.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  971 286
##           1   62  72
##
##           Accuracy : 0.7498
##           95% CI : (0.7262, 0.7724)
##    No Information Rate : 0.7426
##    P-Value [Acc > NIR] : 0.2812
##
##           Kappa : 0.1774
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9400
##           Specificity : 0.2011
##           Pos Pred Value : 0.7725
##           Neg Pred Value : 0.5373
##           Prevalence : 0.7426
##           Detection Rate : 0.6981
##           Detection Prevalence : 0.9037
##           Balanced Accuracy : 0.5705
##
##           'Positive' Class : 0
##

```

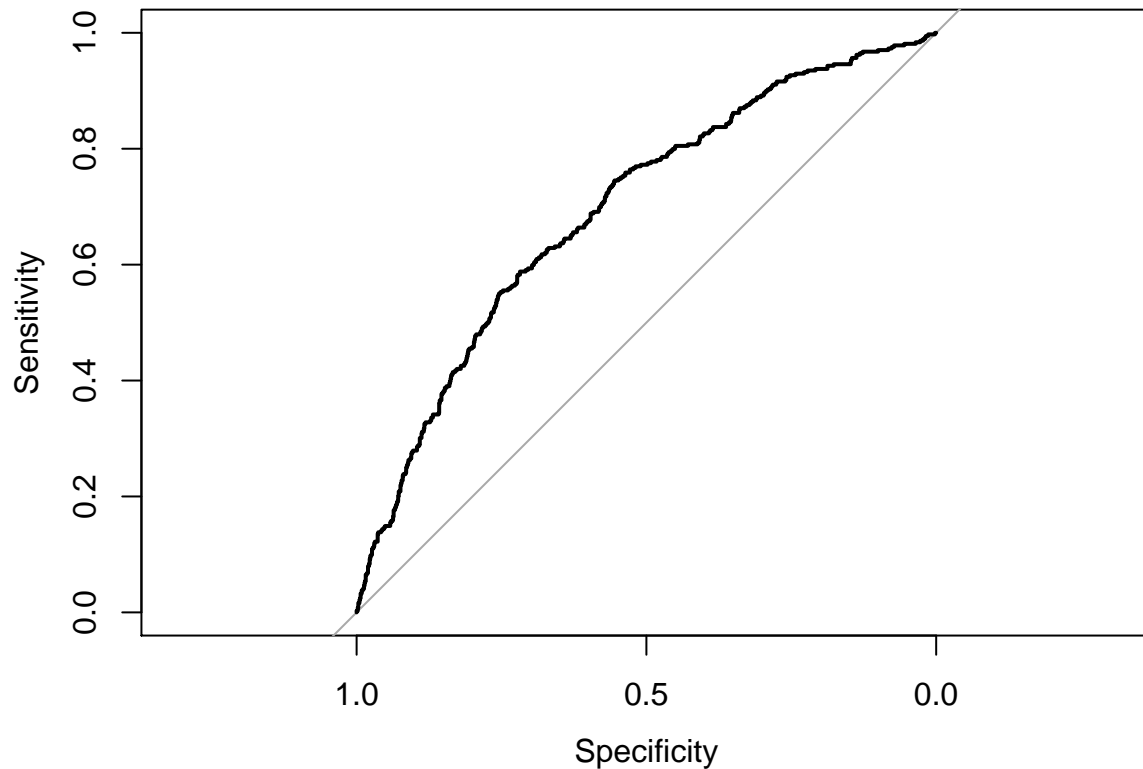


```
## [1] "AUC: 0.705822386388833"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1033 controls (dfPred_raw$class 0) < 358 cases (dfPred_raw$class 1).
## Area under the curve: 0.7058
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9817  -0.7633  -0.5649   0.8774   2.6753
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.769e-01  2.391e-01  -2.413 0.015839 *
## KIDSDRIV     1.840e-01  7.137e-02   2.577 0.009953 **
## AGE          -8.755e-03  4.585e-03  -1.909 0.056204 .
## HOMEKIDS      9.067e-02  3.804e-02   2.384 0.017136 *
## YOJ          -5.259e-03  8.722e-03  -0.603 0.546550
## INCOME        3.292e-07  1.048e-06   0.314 0.753537
## HOME_VAL     -2.672e-06  3.320e-07  -8.047 8.45e-16 ***
## TRAVTIME      7.955e-03  2.144e-03   3.711 0.000206 ***
## BLUEBOOK     -1.307e-05  4.699e-06  -2.781 0.005411 **
```

```

## TIF          -3.878e-02  8.588e-03  -4.515  6.32e-06 ***
## OLDCLAIM      5.704e-06  4.067e-06   1.403  0.160696
## CLM_FREQ      2.606e-01  3.225e-02   8.081  6.43e-16 ***
## MVR_PTS       1.666e-01  1.596e-02  10.444  < 2e-16 ***
## CAR_AGE       -2.639e-02  6.860e-03  -3.847  0.000119 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5838.1  on 5060  degrees of freedom
## Residual deviance: 5192.8  on 5047  degrees of freedom
##    (1352 observations deleted due to missingness)
## AIC: 5220.8
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 949 302
##           1  69  67
##
##           Accuracy : 0.7325
##           95% CI : (0.7084, 0.7557)
##    No Information Rate : 0.734
##    P-Value [Acc > NIR] : 0.5622
##
##           Kappa : 0.1425
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9322
##           Specificity : 0.1816
##           Pos Pred Value : 0.7586
##           Neg Pred Value : 0.4926
##           Prevalence : 0.7340
##           Detection Rate : 0.6842
##           Detection Prevalence : 0.9019
##           Balanced Accuracy : 0.5569
##
##           'Positive' Class : 0
##

```

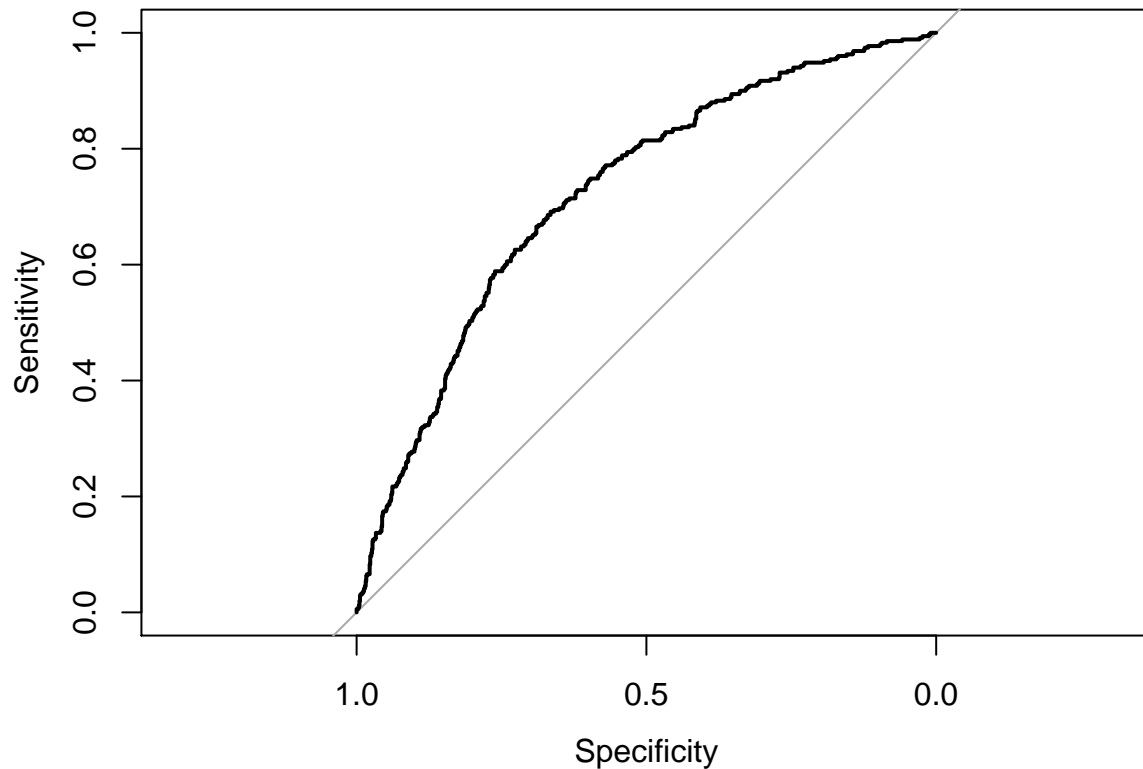


```
## [1] "AUC: 0.694331304806172"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1018 controls (dfPred_raw$class 0) < 369 cases (dfPred_raw$class 1).
## Area under the curve: 0.6943
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0002  -0.7772  -0.5752   0.9200   2.6776
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.973e-01  2.381e-01  -1.669 0.095197 .
## KIDSDRIV     2.469e-01  7.072e-02   3.492 0.000480 ***
## AGE         -1.262e-02  4.583e-03  -2.754 0.005893 **
## HOMEKIDS      8.916e-02  3.728e-02   2.391 0.016784 *
## YOJ          -4.611e-03  8.635e-03  -0.534 0.593300
## INCOME       -7.950e-07  1.041e-06  -0.764 0.444847
## HOME_VAL     -2.483e-06  3.305e-07  -7.513 5.79e-14 ***
## TRAVTIME      7.144e-03  2.127e-03   3.358 0.000785 ***
## BLUEBOOK     -1.089e-05  4.637e-06  -2.348 0.018855 *
```

```

## TIF          -3.737e-02  8.516e-03  -4.388  1.14e-05 ***
## OLDCLAIM     5.969e-06  3.944e-06   1.513  0.130207
## CLM_FREQ     2.640e-01  3.231e-02   8.170  3.08e-16 ***
## MVR_PTS      1.408e-01  1.589e-02   8.866  < 2e-16 ***
## CAR_AGE      -1.878e-02  6.725e-03  -2.792  0.005232 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5886.1  on 5075  degrees of freedom
## Residual deviance: 5274.2  on 5062  degrees of freedom
##    (1336 observations deleted due to missingness)
## AIC: 5302.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 962 283
##           1  60  67
##
##           Accuracy : 0.75
##           95% CI : (0.7262, 0.7727)
##    No Information Rate : 0.7449
##    P-Value [Acc > NIR] : 0.3452
##
##           Kappa : 0.1679
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9413
##           Specificity : 0.1914
##           Pos Pred Value : 0.7727
##           Neg Pred Value : 0.5276
##           Prevalence : 0.7449
##           Detection Rate : 0.7012
##           Detection Prevalence : 0.9074
##           Balanced Accuracy : 0.5664
##
##           'Positive' Class : 0
##

```



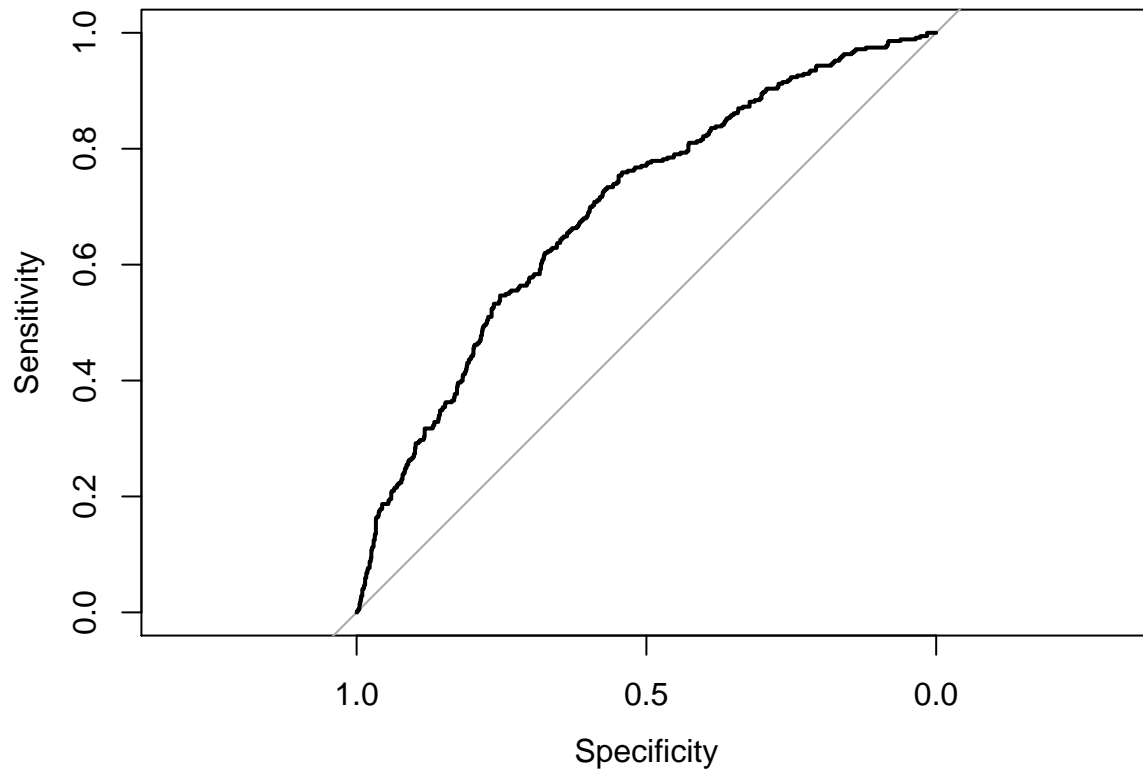
```
## [1] "AUC: 0.723058428850992"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1022 controls (dfPred_raw$class 0) < 350 cases (dfPred_raw$class 1).
## Area under the curve: 0.7231
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8402  -0.7682  -0.5659   0.8900   2.6949
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.473e-01  2.374e-01  -2.305  0.02118 *
## KIDSDRIV     2.801e-01  7.153e-02   3.916 8.99e-05 ***
## AGE          -1.084e-02  4.584e-03  -2.365  0.01803 *
## HOMEKIDS      7.056e-02  3.799e-02   1.857  0.06331 .
## YOJ          -1.347e-03  8.745e-03  -0.154  0.87762
## INCOME        -1.711e-07  1.041e-06  -0.164  0.86939
## HOME_VAL      -2.710e-06  3.320e-07  -8.162 3.28e-16 ***
## TRAVTIME      9.016e-03  2.163e-03   4.168 3.08e-05 ***
## BLUEBOOK     -1.263e-05  4.702e-06  -2.687  0.00722 **
```

```

## TIF          -3.867e-02  8.526e-03  -4.536  5.74e-06 ***
## OLDCLAIM      4.524e-06  3.936e-06   1.149  0.25041
## CLM_FREQ      2.939e-01  3.232e-02   9.094  < 2e-16 ***
## MVR_PTS       1.482e-01  1.599e-02   9.267  < 2e-16 ***
## CAR_AGE       -2.141e-02  6.763e-03  -3.165  0.00155 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5878.1  on 5072  degrees of freedom
## Residual deviance: 5228.4  on 5059  degrees of freedom
##    (1339 observations deleted due to missingness)
## AIC: 5256.4
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  947  274
##           1   75   79
##
##           Accuracy : 0.7462
##           95% CI : (0.7223, 0.769)
##    No Information Rate : 0.7433
##    P-Value [Acc > NIR] : 0.4163
##
##           Kappa : 0.1844
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9266
##           Specificity : 0.2238
##           Pos Pred Value : 0.7756
##           Neg Pred Value : 0.5130
##           Prevalence : 0.7433
##           Detection Rate : 0.6887
##           Detection Prevalence : 0.8880
##           Balanced Accuracy : 0.5752
##
##           'Positive' Class : 0
##

```



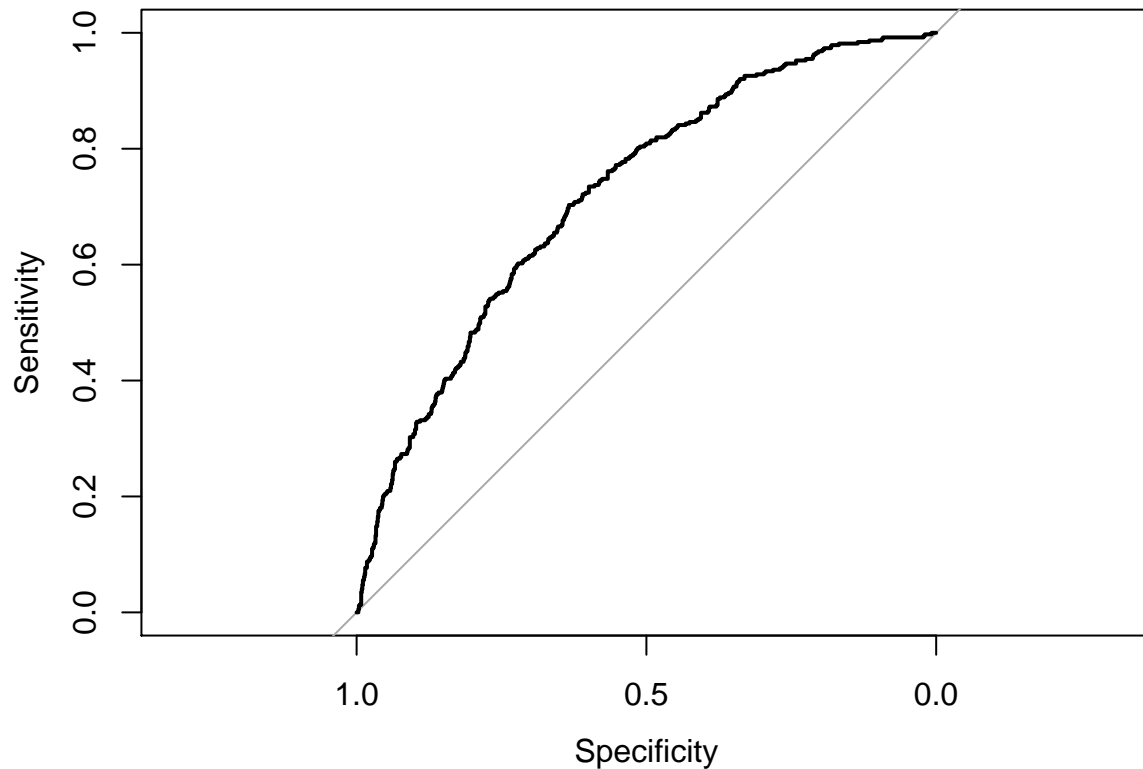


```
## [1] "AUC: 0.694344810763764"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1022 controls (dfPred_raw$class 0) < 353 cases (dfPred_raw$class 1).
## Area under the curve: 0.6943
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7240  -0.7625  -0.5720   0.8955   2.6099
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.957e-01  2.388e-01  -2.495 0.012610 *
## KIDSDRIV     3.350e-01  6.942e-02   4.826 1.40e-06 ***
## AGE         -1.173e-02  4.596e-03  -2.553 0.010672 *
## HOMEKIDS     4.551e-02  3.822e-02   1.191 0.233779
## YOJ          5.287e-04  8.748e-03   0.060 0.951805
## INCOME      -4.410e-07  1.049e-06  -0.420 0.674217
## HOME_VAL    -2.568e-06  3.328e-07  -7.718 1.19e-14 ***
## TRAVTIME     7.480e-03  2.148e-03   3.483 0.000496 ***
## BLUEBOOK    -6.037e-06  4.672e-06  -1.292 0.196267
```

```

## TIF          -3.776e-02  8.476e-03  -4.455  8.39e-06 ***
## OLDCLAIM     4.370e-06  4.006e-06   1.091  0.275286
## CLM_FREQ     2.979e-01  3.230e-02   9.224  < 2e-16 ***
## MVR_PTS      1.402e-01  1.589e-02   8.820  < 2e-16 ***
## CAR_AGE      -2.102e-02  6.700e-03  -3.138  0.001703 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5819.8  on 5057  degrees of freedom
## Residual deviance: 5222.7  on 5044  degrees of freedom
##    (1354 observations deleted due to missingness)
## AIC: 5250.7
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  959  299
##           1   54   78
##
##           Accuracy : 0.746
##           95% CI : (0.7223, 0.7687)
##    No Information Rate : 0.7288
##    P-Value [Acc > NIR] : 0.07742
##
##           Kappa : 0.193
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9467
##           Specificity : 0.2069
##           Pos Pred Value : 0.7623
##           Neg Pred Value : 0.5909
##           Prevalence : 0.7288
##           Detection Rate : 0.6899
##           Detection Prevalence : 0.9050
##           Balanced Accuracy : 0.5768
##
##           'Positive' Class : 0
##

```

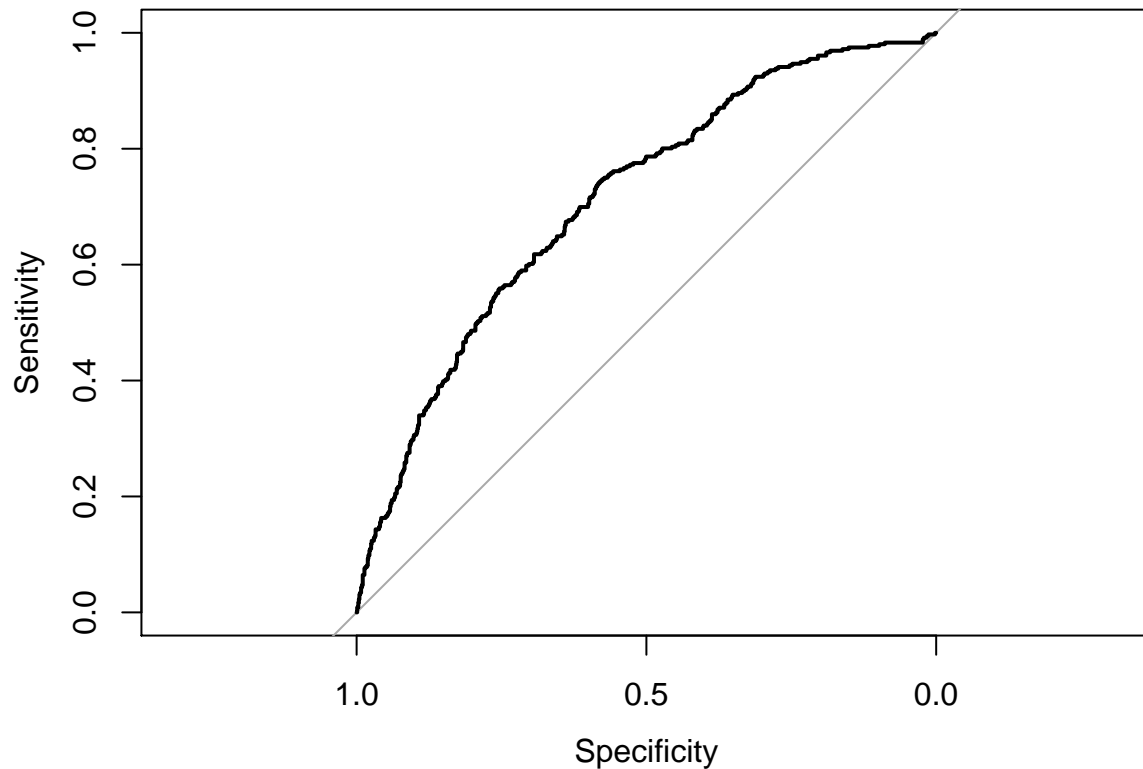


```
## [1] "AUC: 0.721875564609676"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1013 controls (dfPred_raw$class 0) < 377 cases (dfPred_raw$class 1).
## Area under the curve: 0.7219
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8207  -0.7734  -0.5738   0.9151   2.6425
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -7.719e-01  2.367e-01  -3.261 0.001111 **
## KIDSDRIV     2.850e-01  6.969e-02   4.089 4.33e-05 ***
## AGE         -5.880e-03  4.540e-03  -1.295 0.195273
## HOMEKIDS      7.178e-02  3.766e-02   1.906 0.056657 .
## YOJ          -1.174e-03  8.787e-03  -0.134 0.893669
## INCOME       -2.696e-07  1.048e-06  -0.257 0.796941
## HOME_VAL     -2.789e-06  3.308e-07  -8.431 < 2e-16 ***
## TRAVTIME      9.972e-03  2.137e-03   4.667 3.05e-06 ***
## BLUEBOOK     -1.022e-05  4.648e-06  -2.199 0.027872 *
```

```

## TIF          -4.068e-02  8.590e-03  -4.735  2.19e-06 ***
## OLDCLAIM      7.586e-06  3.981e-06   1.905  0.056742 .
## CLM_FREQ      2.528e-01  3.223e-02   7.844  4.37e-15 ***
## MVR_PTS       1.510e-01  1.581e-02   9.548  < 2e-16 ***
## CAR_AGE       -2.466e-02  6.772e-03  -3.641  0.000271 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5859.7  on 5052  degrees of freedom
## Residual deviance: 5239.3  on 5039  degrees of freedom
##    (1360 observations deleted due to missingness)
## AIC: 5267.3
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  973 287
##           1   66  69
##
##           Accuracy : 0.747
##           95% CI : (0.7233, 0.7696)
##    No Information Rate : 0.7448
##    P-Value [Acc > NIR] : 0.4409
##
##           Kappa : 0.1637
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9365
##           Specificity : 0.1938
##           Pos Pred Value : 0.7722
##           Neg Pred Value : 0.5111
##           Prevalence : 0.7448
##           Detection Rate : 0.6975
##           Detection Prevalence : 0.9032
##           Balanced Accuracy : 0.5651
##
##           'Positive' Class : 0
##

```

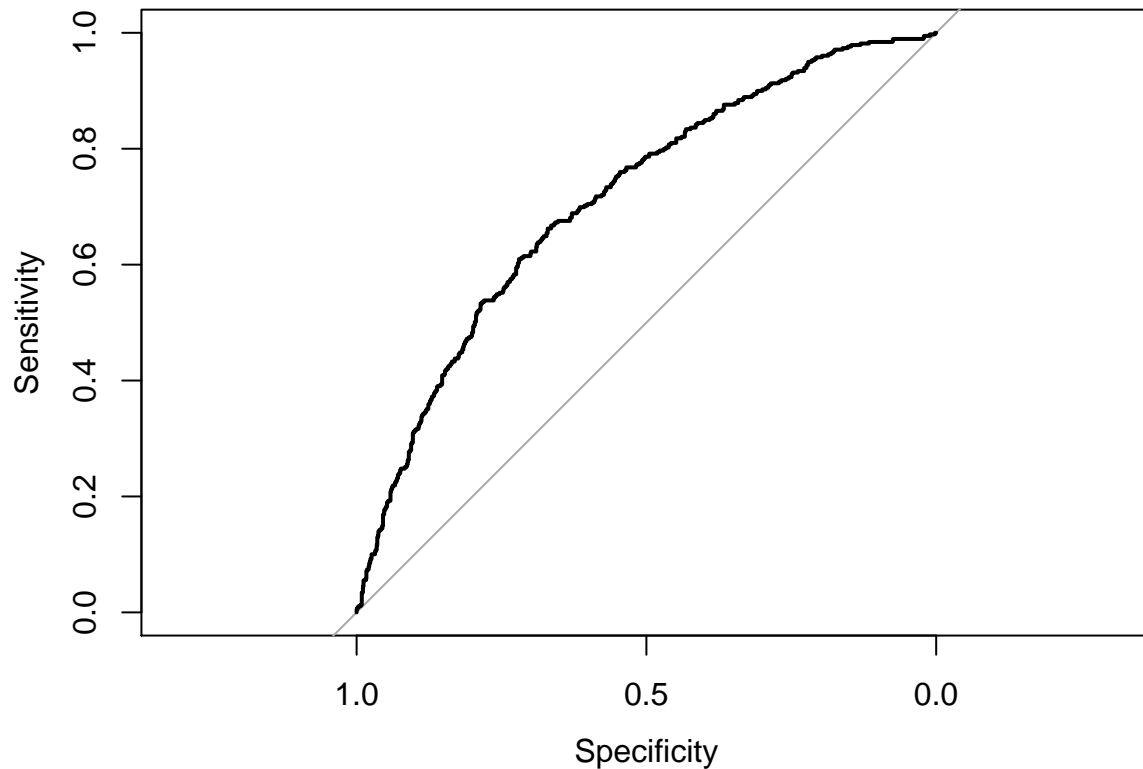


```
## [1] "AUC: 0.712009711152686"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1039 controls (dfPred_raw$class 0) < 356 cases (dfPred_raw$class 1).
## Area under the curve: 0.712
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0264  -0.7645  -0.5675   0.8837   2.7201
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.321e-01  2.388e-01  -2.647 0.008112 **
## KIDSDRIV     2.534e-01  7.070e-02   3.584 0.000339 ***
## AGE         -6.851e-03  4.561e-03  -1.502 0.133067
## HOMEKIDS     8.352e-02  3.816e-02   2.188 0.028634 *
## YOJ         -1.533e-03  8.755e-03  -0.175 0.861047
## INCOME      -3.687e-07  1.065e-06  -0.346 0.729154
## HOME_VAL    -2.480e-06  3.334e-07  -7.439 1.01e-13 ***
## TRAVTIME     7.497e-03  2.132e-03   3.516 0.000437 ***
## BLUEBOOK    -1.540e-05  4.735e-06  -3.252 0.001145 **
```

```

## TIF          -4.001e-02  8.593e-03  -4.656  3.22e-06 ***
## OLDCLAIM      9.841e-06  3.968e-06   2.480  0.013145 *
## CLM_FREQ      2.509e-01  3.263e-02   7.688  1.49e-14 ***
## MVR_PTS       1.479e-01  1.591e-02   9.298  < 2e-16 ***
## CAR_AGE       -2.590e-02  6.843e-03  -3.785  0.000154 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5816.2  on 5058  degrees of freedom
## Residual deviance: 5200.3  on 5045  degrees of freedom
##    (1354 observations deleted due to missingness)
## AIC: 5228.3
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  959  310
##           1   51   69
##
##           Accuracy : 0.7401
##           95% CI : (0.7162, 0.763)
##    No Information Rate : 0.7271
##    P-Value [Acc > NIR] : 0.1458
##
##           Kappa : 0.1673
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9495
##           Specificity : 0.1821
##           Pos Pred Value : 0.7557
##           Neg Pred Value : 0.5750
##           Prevalence : 0.7271
##           Detection Rate : 0.6904
##           Detection Prevalence : 0.9136
##           Balanced Accuracy : 0.5658
##
##           'Positive' Class : 0
##

```



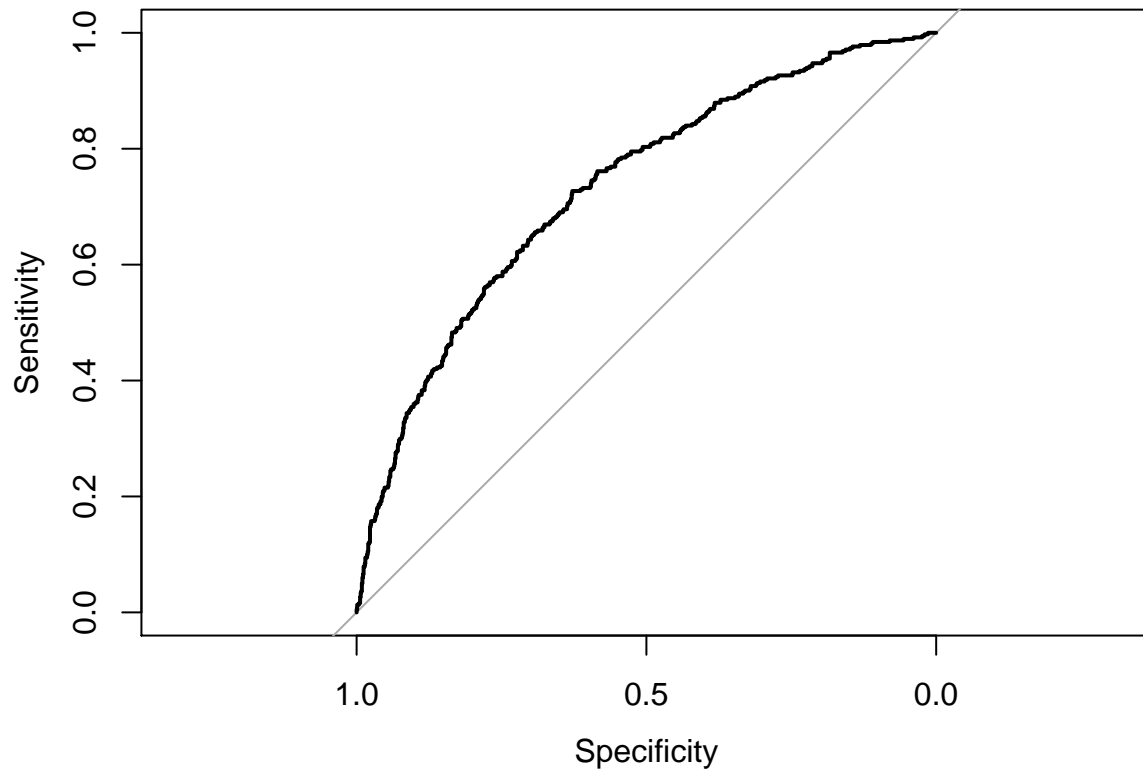
```
## [1] "AUC: 0.712440241385616"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1010 controls (dfPred_raw$class 0) < 379 cases (dfPred_raw$class 1).
## Area under the curve: 0.7124
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0320  -0.7651  -0.5779   0.8920   2.5995
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.687e-01  2.386e-01  -1.964 0.049526 *
## KIDSDRIV     2.617e-01  6.966e-02   3.756 0.000173 ***
## AGE          -9.646e-03  4.578e-03  -2.107 0.035110 *
## HOMEKIDS      4.667e-02  3.823e-02   1.221 0.222143
## YOJ          -7.100e-03  8.812e-03  -0.806 0.420431
## INCOME        3.362e-07  1.051e-06   0.320 0.749039
## HOME_VAL     -2.522e-06  3.297e-07  -7.651 1.99e-14 ***
## TRAVTIME      8.593e-03  2.132e-03   4.030 5.59e-05 ***
## BLUEBOOK     -1.074e-05  4.660e-06  -2.304 0.021202 *
```

```

## TIF          -4.919e-02  8.577e-03  -5.735  9.76e-09 ***
## OLDCLAIM     4.963e-06  3.989e-06   1.244  0.213413
## CLM_FREQ     2.516e-01  3.236e-02   7.778  7.39e-15 ***
## MVR_PTS      1.473e-01  1.596e-02   9.230  < 2e-16 ***
## CAR_AGE      -2.787e-02  6.730e-03  -4.142  3.45e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5818.7  on 5069  degrees of freedom
## Residual deviance: 5241.9  on 5056  degrees of freedom
##    (1342 observations deleted due to missingness)
## AIC: 5269.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  944 299
##           1   53  82
##
##           Accuracy : 0.7446
##           95% CI : (0.7207, 0.7674)
##    No Information Rate : 0.7235
##    P-Value [Acc > NIR] : 0.0422
##
##           Kappa : 0.2024
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9468
##           Specificity : 0.2152
##           Pos Pred Value : 0.7595
##           Neg Pred Value : 0.6074
##           Prevalence : 0.7235
##           Detection Rate : 0.6851
##           Detection Prevalence : 0.9020
##           Balanced Accuracy : 0.5810
##
##           'Positive' Class : 0
##

```



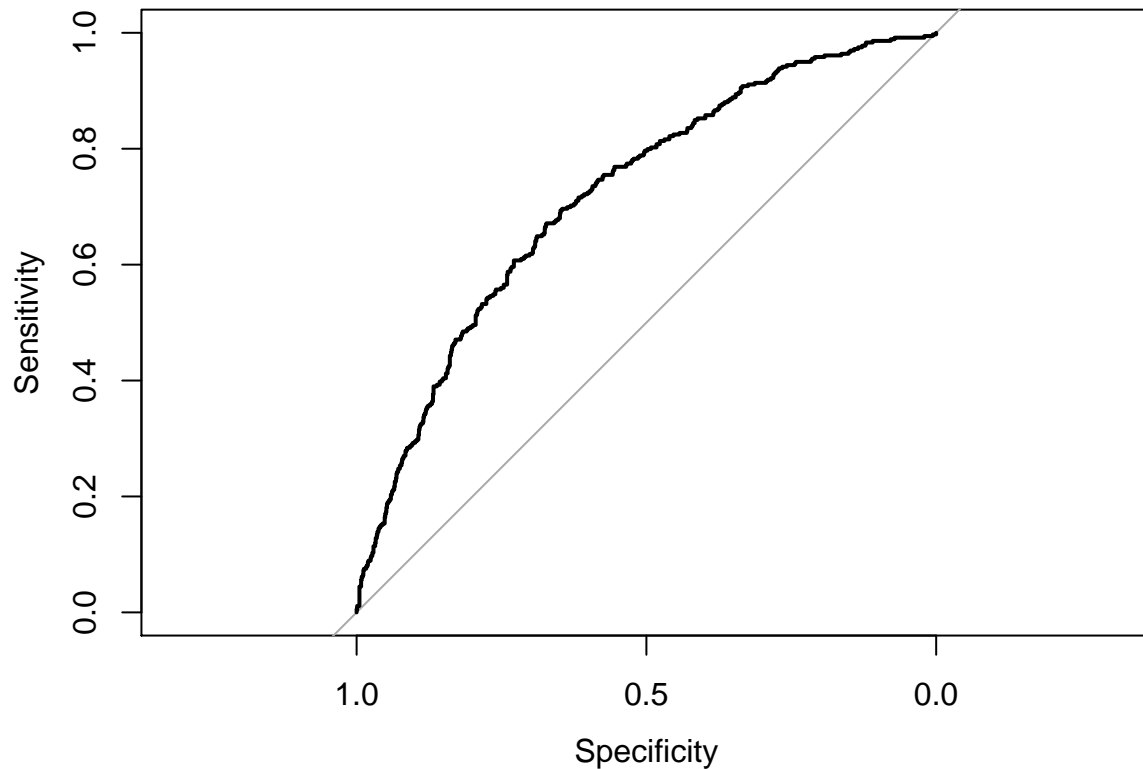


```
## [1] "AUC: 0.73134626978047"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 997 controls (dfPred_raw$class 0) < 381 cases (dfPred_raw$class 1).
## Area under the curve: 0.7313
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0448  -0.7719  -0.5774   0.9226   2.6796
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.321e-01  2.366e-01  -1.826  0.067887 .
## KIDSDRIV     2.916e-01  7.005e-02   4.163  3.14e-05 ***
## AGE          -1.043e-02  4.522e-03  -2.307  0.021048 *
## HOMEKIDS      5.823e-02  3.826e-02   1.522  0.128025
## YOJ          -1.451e-03  8.727e-03  -0.166  0.867919
## INCOME        -9.438e-07  1.042e-06  -0.905  0.365220
## HOME_VAL      -2.346e-06  3.306e-07  -7.095  1.29e-12 ***
## TRAVTIME       7.082e-03  2.113e-03   3.351  0.000805 ***
## BLUEBOOK      -1.389e-05  4.693e-06  -2.959  0.003088 **
```

```

## TIF          -4.728e-02  8.557e-03  -5.526  3.28e-08 ***
## OLDCLAIM     9.275e-06  3.924e-06   2.363  0.018108 *
## CLM_FREQ     2.800e-01  3.238e-02   8.649  < 2e-16 ***
## MVR_PTS      1.267e-01  1.580e-02   8.018  1.07e-15 ***
## CAR_AGE      -1.960e-02  6.717e-03  -2.917  0.003531 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5856.1  on 5056  degrees of freedom
## Residual deviance: 5250.6  on 5043  degrees of freedom
## (1355 observations deleted due to missingness)
## AIC: 5278.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  977  291
##           1   55   68
##
##           Accuracy : 0.7513
##           95% CI : (0.7277, 0.7738)
##    No Information Rate : 0.7419
##    P-Value [Acc > NIR] : 0.2225
##
##           Kappa : 0.1733
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9467
##           Specificity : 0.1894
##           Pos Pred Value : 0.7705
##           Neg Pred Value : 0.5528
##           Prevalence : 0.7419
##           Detection Rate : 0.7024
##           Detection Prevalence : 0.9116
##           Balanced Accuracy : 0.5681
##
##           'Positive' Class : 0
##

```

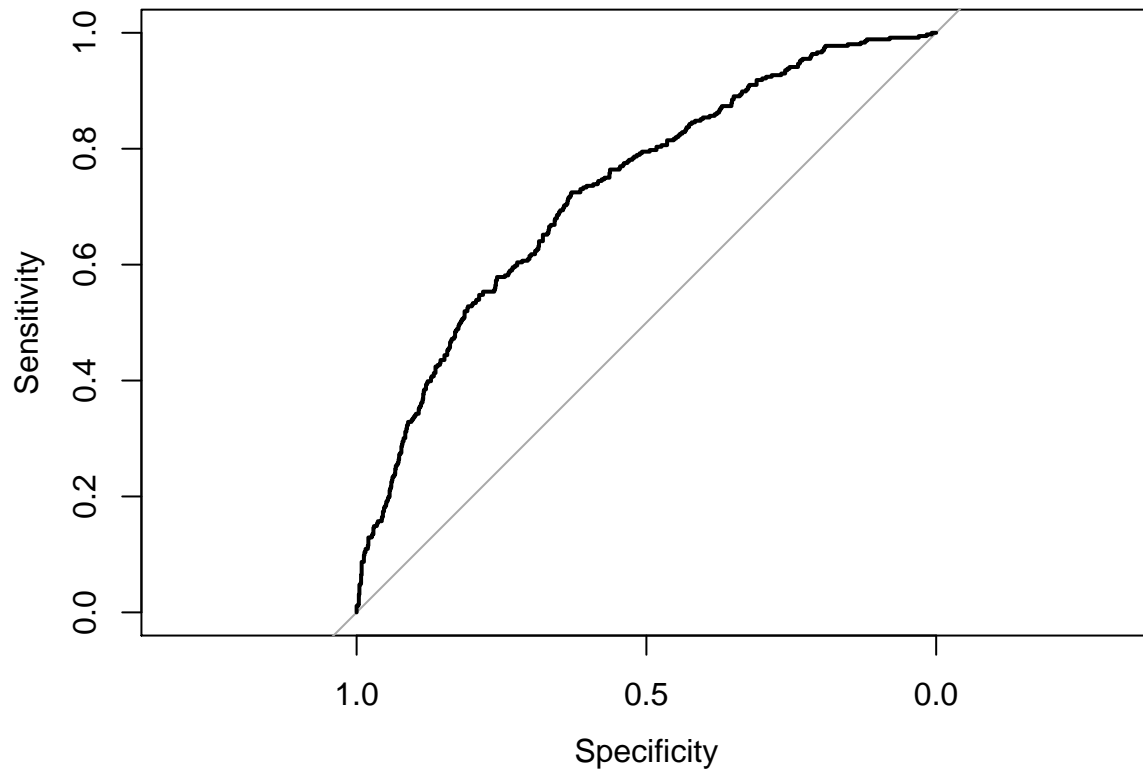


```
## [1] "AUC: 0.720498369717778"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1032 controls (dfPred_raw$class 0) < 359 cases (dfPred_raw$class 1).
## Area under the curve: 0.7205
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0231  -0.7776  -0.5818   0.9336   2.6923
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.305e-01  2.354e-01  -1.829  0.067458 .
## KIDSDRIV     2.586e-01  6.994e-02   3.698  0.000218 ***
## AGE          -7.885e-03  4.497e-03  -1.753  0.079540 .
## HOMEKIDS      6.580e-02  3.771e-02   1.745  0.081013 .
## YOJ          -5.173e-03  8.761e-03  -0.590  0.554916
## INCOME       -2.038e-08  1.047e-06  -0.019  0.984473
## HOME_VAL     -2.542e-06  3.274e-07  -7.764  8.21e-15 ***
## TRAVTIME      8.230e-03  2.094e-03   3.930  8.50e-05 ***
## BLUEBOOK     -1.678e-05  4.652e-06  -3.607  0.000310 ***
```

```

## TIF          -5.006e-02  8.605e-03  -5.818  5.97e-09 ***
## OLDCLAIM     9.837e-06  3.885e-06   2.532  0.011351 *
## CLM_FREQ     2.255e-01  3.260e-02   6.916  4.64e-12 ***
## MVR_PTS      1.341e-01  1.578e-02   8.496  < 2e-16 ***
## CAR_AGE      -2.481e-02  6.704e-03  -3.701  0.000215 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5861.5  on 5055  degrees of freedom
## Residual deviance: 5274.5  on 5042  degrees of freedom
##    (1356 observations deleted due to missingness)
## AIC: 5302.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 983 288
##           1  53  68
##
##           Accuracy : 0.755
##           95% CI : (0.7316, 0.7774)
##    No Information Rate : 0.7443
##    P-Value [Acc > NIR] : 0.1868
##
##           Kappa : 0.1785
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9488
##           Specificity : 0.1910
##           Pos Pred Value : 0.7734
##           Neg Pred Value : 0.5620
##           Prevalence : 0.7443
##           Detection Rate : 0.7062
##           Detection Prevalence : 0.9131
##           Balanced Accuracy : 0.5699
##
##           'Positive' Class : 0
##

```

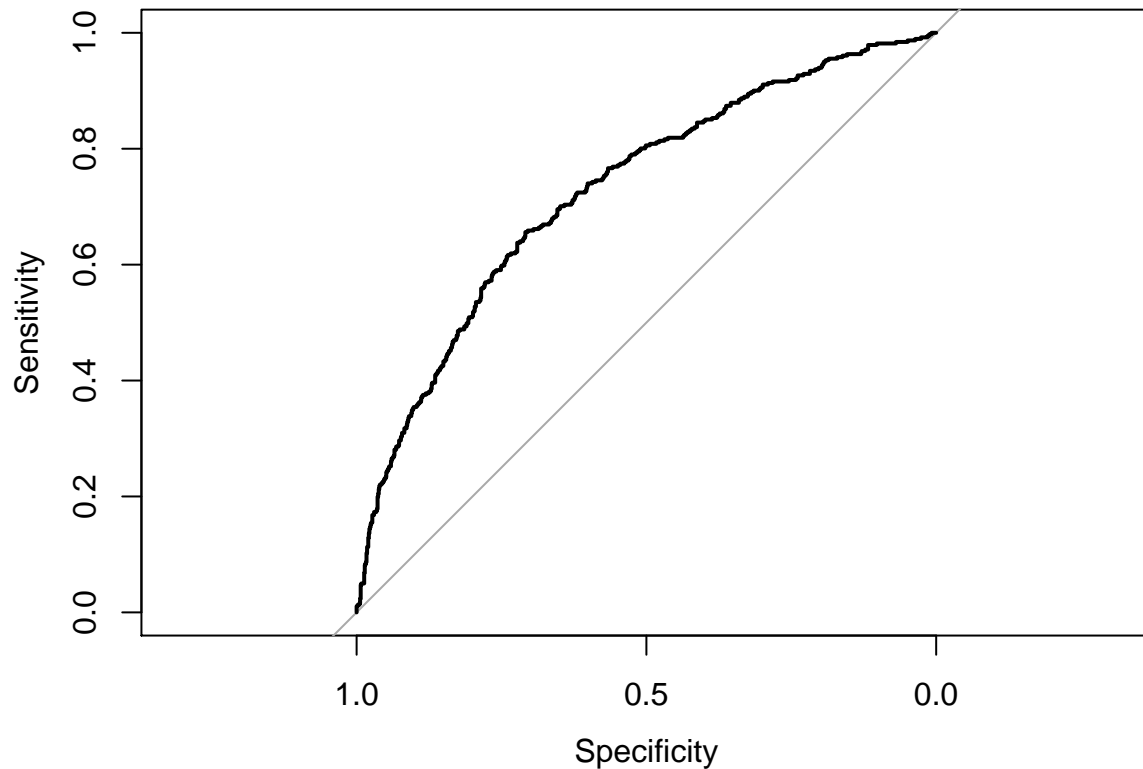


```
## [1] "AUC: 0.728032406403193"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1036 controls (dfPred_raw$class 0) < 356 cases (dfPred_raw$class 1).
## Area under the curve: 0.728
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9780  -0.7662  -0.5768   0.8881   2.6502
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.172e-01  2.389e-01  -0.909  0.363213
## KIDSDRIV     2.695e-01  6.986e-02   3.858  0.000114 ***
## AGE         -1.455e-02  4.622e-03  -3.147  0.001650 **
## HOMEKIDS      4.367e-02  3.828e-02   1.141  0.254040
## YOJ          -5.671e-03  8.720e-03  -0.650  0.515465
## INCOME       -3.842e-07  1.052e-06  -0.365  0.714976
## HOME_VAL     -2.501e-06  3.320e-07  -7.533  4.95e-14 ***
## TRAVTIME      6.649e-03  2.146e-03   3.098  0.001947 **
## BLUEBOOK     -9.679e-06  4.657e-06  -2.078  0.037680 *
```

```

## TIF          -4.377e-02  8.489e-03  -5.156 2.52e-07 ***
## OLDCLAIM     4.227e-06  4.018e-06   1.052 0.292858
## CLM_FREQ     2.522e-01  3.248e-02   7.764 8.25e-15 ***
## MVR_PTS      1.466e-01  1.605e-02   9.137 < 2e-16 ***
## CAR_AGE      -2.625e-02  6.743e-03  -3.894 9.88e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5819.3  on 5070  degrees of freedom
## Residual deviance: 5235.3  on 5057  degrees of freedom
##    (1341 observations deleted due to missingness)
## AIC: 5263.3
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  951 296
##           1   45  85
##
##           Accuracy : 0.7524
##           95% CI : (0.7287, 0.775)
##    No Information Rate : 0.7233
##    P-Value [Acc > NIR] : 0.008159
##
##           Kappa : 0.2233
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.9548
##           Specificity : 0.2231
##           Pos Pred Value : 0.7626
##           Neg Pred Value : 0.6538
##           Prevalence : 0.7233
##           Detection Rate : 0.6906
##           Detection Prevalence : 0.9056
##           Balanced Accuracy : 0.5890
##
##           'Positive' Class : 0
##

```



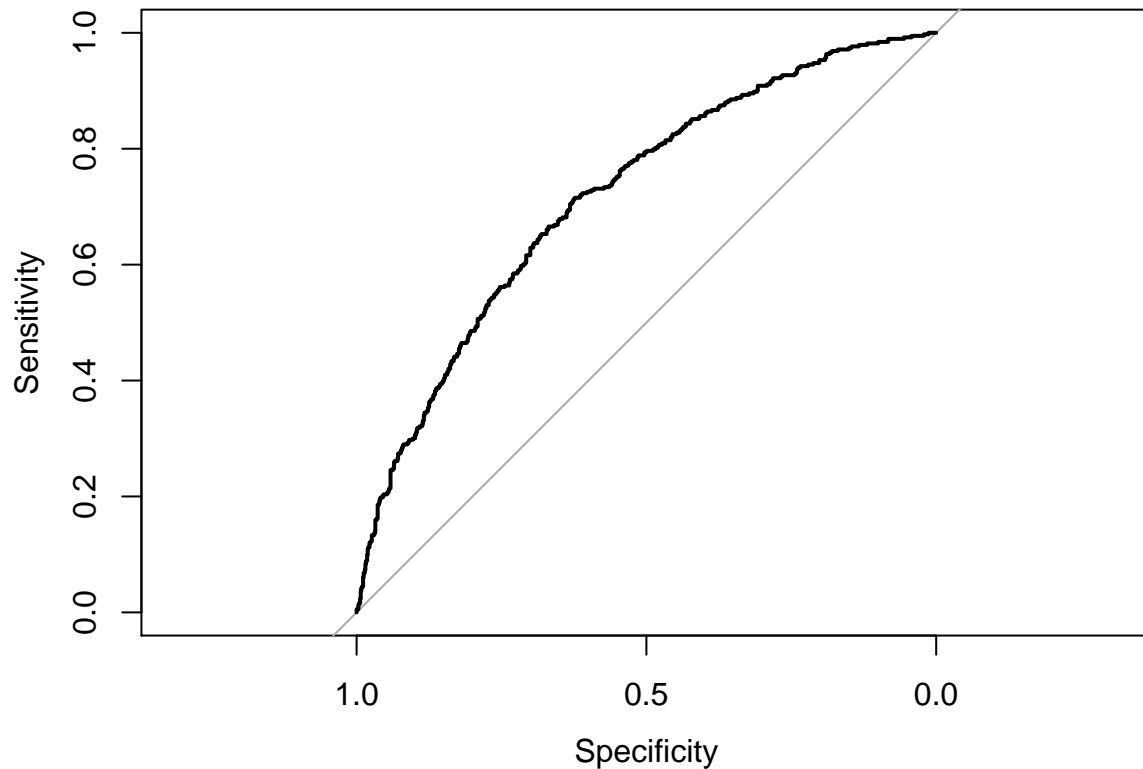
```
## [1] "AUC: 0.727553257650023"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 996 controls (dfPred_raw$class 0) < 381 cases (dfPred_raw$class 1).
## Area under the curve: 0.7276
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9702  -0.7623  -0.5699   0.8971   2.6743
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.427e-01  2.416e-01  -1.418  0.156103
## KIDSDRIV     2.165e-01  6.994e-02   3.095  0.001967 **
## AGE          -1.160e-02  4.596e-03  -2.524  0.011604 *
## HOMEKIDS      7.117e-02  3.801e-02   1.872  0.061149 .
## YOJ          -6.043e-03  8.776e-03  -0.689  0.491040
## INCOME       -5.576e-07  1.052e-06  -0.530  0.596214
## HOME_VAL     -2.129e-06  3.334e-07  -6.385  1.71e-10 ***
## TRAVTIME      7.144e-03  2.125e-03   3.361  0.000776 ***
## BLUEBOOK     -1.415e-05  4.690e-06  -3.017  0.002556 **
```

```

## TIF          -4.827e-02  8.694e-03  -5.552 2.83e-08 ***
## OLDCLAIM     8.655e-06  3.958e-06   2.187 0.028779 *
## CLM_FREQ     2.986e-01  3.240e-02   9.217 < 2e-16 ***
## MVR_PTS      1.213e-01  1.613e-02   7.524 5.31e-14 ***
## CAR_AGE      -2.674e-02  6.819e-03  -3.921 8.82e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5817.0  on 5073  degrees of freedom
## Residual deviance: 5222.3  on 5060  degrees of freedom
##    (1338 observations deleted due to missingness)
## AIC: 5250.3
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 936 302
##           1  55  81
##
##           Accuracy : 0.7402
##           95% CI : (0.7161, 0.7632)
##    No Information Rate : 0.7213
##    P-Value [Acc > NIR] : 0.06171
##
##           Kappa : 0.1945
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9445
##           Specificity : 0.2115
##           Pos Pred Value : 0.7561
##           Neg Pred Value : 0.5956
##           Prevalence : 0.7213
##           Detection Rate : 0.6812
##    Detection Prevalence : 0.9010
##           Balanced Accuracy : 0.5780
##
##           'Positive' Class : 0
##

```



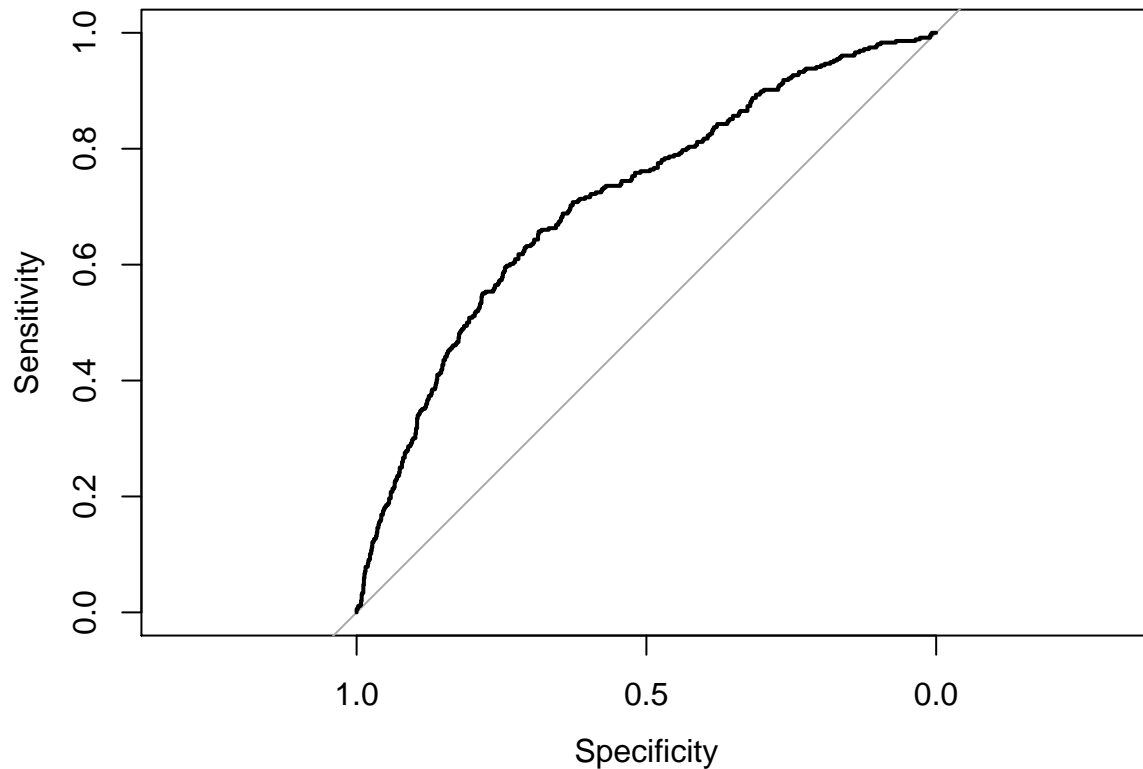


```
## [1] "AUC: 0.718705951474498"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 991 controls (dfPred_raw$class 0) < 383 cases (dfPred_raw$class 1).
## Area under the curve: 0.7187
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0668  -0.7740  -0.5703   0.9065   2.6694
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.934e-01  2.378e-01  -1.654 0.098083 .
## KIDSDRIV     2.776e-01  7.090e-02   3.915 9.03e-05 ***
## AGE          -1.100e-02  4.567e-03  -2.409 0.016008 *
## HOMEKIDS      7.313e-02  3.837e-02   1.906 0.056641 .
## YOJ          -6.024e-03  8.719e-03  -0.691 0.489649
## INCOME        3.701e-08  1.050e-06   0.035 0.971882
## HOME_VAL     -2.701e-06  3.301e-07  -8.182 2.79e-16 ***
## TRAVTIME      7.204e-03  2.118e-03   3.402 0.000670 ***
## BLUEBOOK     -1.201e-05  4.687e-06  -2.562 0.010393 *
```

```

## TIF          -4.518e-02  8.503e-03  -5.314 1.08e-07 ***
## OLDCLAIM     4.087e-06  3.989e-06   1.025 0.305577
## CLM_FREQ     2.917e-01  3.249e-02   8.979 < 2e-16 ***
## MVR_PTS      1.376e-01  1.605e-02   8.573 < 2e-16 ***
## CAR_AGE      -2.282e-02  6.705e-03  -3.403 0.000666 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5862.2  on 5056  degrees of freedom
## Residual deviance: 5239.1  on 5043  degrees of freedom
## (1355 observations deleted due to missingness)
## AIC: 5267.1
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 964 275
##           1  71  81
##
##           Accuracy : 0.7513
##           95% CI : (0.7277, 0.7738)
##    No Information Rate : 0.7441
##    P-Value [Acc > NIR] : 0.2808
##
##           Kappa : 0.1957
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9314
##           Specificity : 0.2275
##           Pos Pred Value : 0.7780
##           Neg Pred Value : 0.5329
##           Prevalence : 0.7441
##           Detection Rate : 0.6930
##           Detection Prevalence : 0.8907
##           Balanced Accuracy : 0.5795
##
##           'Positive' Class : 0
##

```

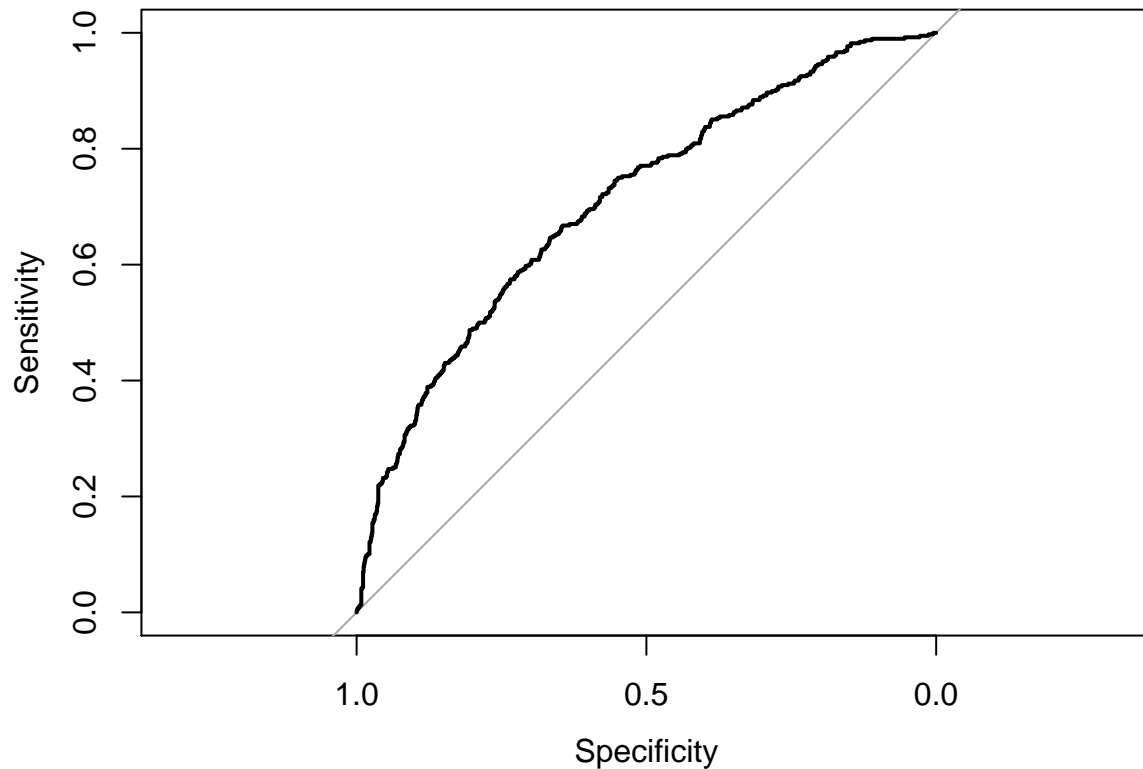


```
## [1] "AUC: 0.710495033382185"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1035 controls (dfPred_raw$class 0) < 356 cases (dfPred_raw$class 1).
## Area under the curve: 0.7105
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9888  -0.7613  -0.5679   0.8684   2.6711
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.298e-01  2.404e-01  -0.956  0.339150
## KIDSDRIV     2.452e-01  6.917e-02   3.545  0.000393 ***
## AGE         -1.183e-02  4.633e-03  -2.555  0.010633 *
## HOMEKIDS      6.541e-02  3.823e-02   1.711  0.087115 .
## YOJ          -1.366e-02  8.786e-03  -1.555  0.119952
## INCOME        7.909e-07  1.059e-06   0.747  0.455175
## HOME_VAL     -2.560e-06  3.313e-07  -7.726  1.11e-14 ***
## TRAVTIME      7.096e-03  2.134e-03   3.325  0.000886 ***
## BLUEBOOK     -1.499e-05  4.700e-06  -3.189  0.001428 **
```

```

## TIF          -4.988e-02  8.614e-03  -5.791  6.98e-09 ***
## OLDCLAIM     3.943e-06  3.979e-06   0.991  0.321611
## CLM_FREQ     2.712e-01  3.251e-02   8.344  < 2e-16 ***
## MVR_PTS      1.405e-01  1.604e-02   8.762  < 2e-16 ***
## CAR_AGE      -2.924e-02  6.780e-03  -4.313  1.61e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5803.5  on 5068  degrees of freedom
## Residual deviance: 5197.6  on 5055  degrees of freedom
## (1343 observations deleted due to missingness)
## AIC: 5225.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  948  302
##           1   43   86
##
##           Accuracy : 0.7498
##           95% CI : (0.7261, 0.7725)
##    No Information Rate : 0.7186
##    P-Value [Acc > NIR] : 0.005074
##
##           Kappa : 0.2237
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.9566
##           Specificity : 0.2216
##           Pos Pred Value : 0.7584
##           Neg Pred Value : 0.6667
##           Prevalence : 0.7186
##           Detection Rate : 0.6875
##           Detection Prevalence : 0.9065
##           Balanced Accuracy : 0.5891
##
##           'Positive' Class : 0
##

```

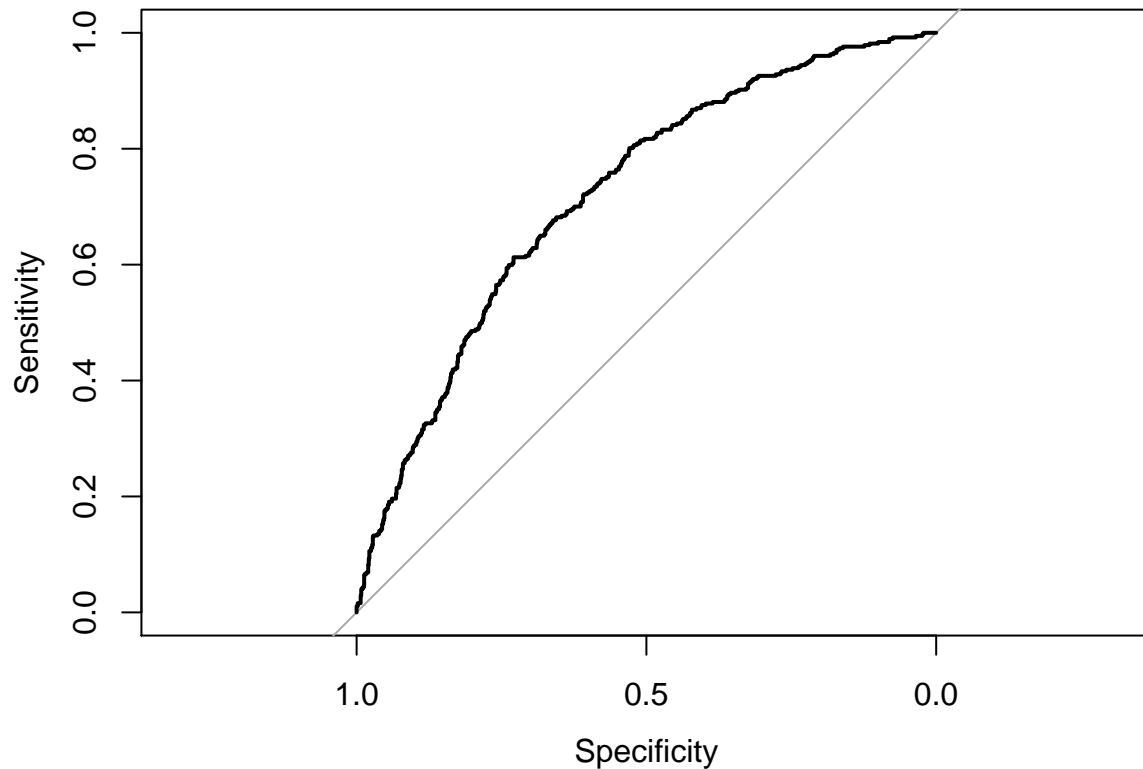


```
## [1] "AUC: 0.709184204229821"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 991 controls (dfPred_raw$class 0) < 388 cases (dfPred_raw$class 1).
## Area under the curve: 0.7092
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9690  -0.7600  -0.5707   0.8798   2.6468
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.427e-01  2.398e-01  -2.263 0.023629 *
## KIDSDRIV     2.339e-01  7.083e-02   3.302 0.000961 ***
## AGE        -1.133e-02  4.577e-03  -2.476 0.013292 *
## HOMEKIDS     6.667e-02  3.797e-02   1.756 0.079156 .
## YOJ          3.858e-04  8.722e-03   0.044 0.964722
## INCOME      -2.953e-07  1.040e-06  -0.284 0.776367
## HOME_VAL    -2.317e-06  3.316e-07  -6.987 2.80e-12 ***
## TRAVTIME     7.131e-03  2.140e-03   3.333 0.000860 ***
## BLUEBOOK    -1.222e-05  4.699e-06  -2.601 0.009292 **
```

```

## TIF          -3.865e-02  8.547e-03  -4.522  6.13e-06 ***
## OLDCLAIM     7.403e-06  4.003e-06   1.850  0.064384 .
## CLM_FREQ     2.745e-01  3.229e-02   8.500  < 2e-16 ***
## MVR_PTS      1.483e-01  1.587e-02   9.343  < 2e-16 ***
## CAR_AGE      -2.147e-02  6.778e-03  -3.167  0.001539 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5827.1  on 5069  degrees of freedom
## Residual deviance: 5222.4  on 5056  degrees of freedom
## (1342 observations deleted due to missingness)
## AIC: 5250.4
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  948  309
##           1   53   68
##
##           Accuracy : 0.7373
##           95% CI : (0.7132, 0.7604)
##    No Information Rate : 0.7264
##    P-Value [Acc > NIR] : 0.1907
##
##           Kappa : 0.1616
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9471
##           Specificity : 0.1804
##           Pos Pred Value : 0.7542
##           Neg Pred Value : 0.5620
##           Prevalence : 0.7264
##           Detection Rate : 0.6880
##           Detection Prevalence : 0.9122
##           Balanced Accuracy : 0.5637
##
##           'Positive' Class : 0
##

```



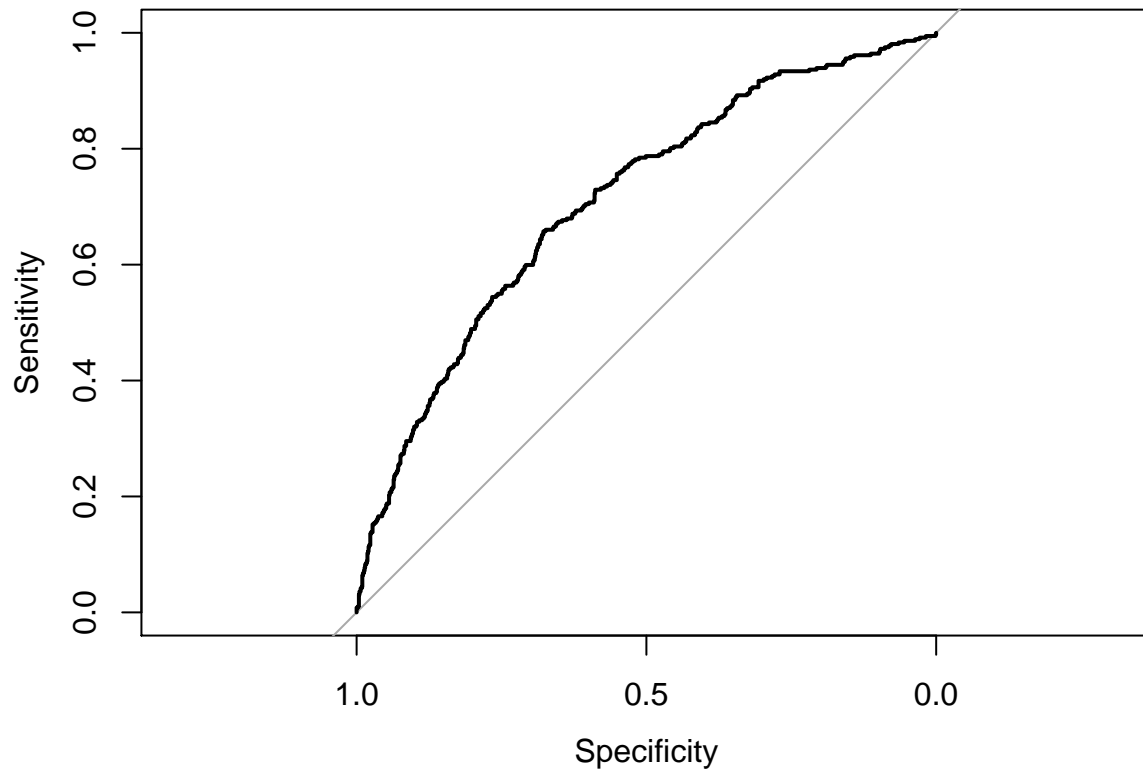
```
## [1] "AUC: 0.720091579508025"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1001 controls (dfPred_raw$class 0) < 377 cases (dfPred_raw$class 1).
## Area under the curve: 0.7201
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9978  -0.7674  -0.5719   0.9117   2.5122
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.851e-01  2.375e-01  -2.042  0.0411 *
## KIDSDRIV     1.796e-01  7.054e-02   2.547  0.0109 *
## AGE         -7.919e-03  4.554e-03  -1.739  0.0820 .
## HOMEKIDS     9.728e-02  3.796e-02   2.563  0.0104 *
## YOJ         -1.365e-02  8.711e-03  -1.567  0.1171
## INCOME      -2.694e-07  1.063e-06  -0.253  0.8000
## HOME_VAL    -2.479e-06  3.326e-07  -7.454 9.05e-14 ***
## TRAVTIME     8.499e-03  2.121e-03   4.007 6.15e-05 ***
## BLUEBOOK    -1.186e-05  4.659e-06  -2.546  0.0109 *
```

```

## TIF          -4.780e-02  8.696e-03  -5.496  3.88e-08 ***
## OLDCLAIM     6.550e-06  3.937e-06   1.664   0.0961 .
## CLM_FREQ     2.763e-01  3.217e-02   8.588  < 2e-16 ***
## MVR_PTS      1.422e-01  1.594e-02   8.923  < 2e-16 ***
## CAR_AGE      -2.681e-02  6.834e-03  -3.923  8.75e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5856.7  on 5067  degrees of freedom
## Residual deviance: 5237.0  on 5054  degrees of freedom
##    (1345 observations deleted due to missingness)
## AIC: 5265
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 965 294
##           1  53  68
##
##           Accuracy : 0.7486
##           95% CI : (0.7248, 0.7713)
##    No Information Rate : 0.7377
##    P-Value [Acc > NIR] : 0.1877
##
##           Kappa : 0.1729
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9479
##           Specificity : 0.1878
##           Pos Pred Value : 0.7665
##           Neg Pred Value : 0.5620
##           Prevalence : 0.7377
##           Detection Rate : 0.6993
##           Detection Prevalence : 0.9123
##           Balanced Accuracy : 0.5679
##
##           'Positive' Class : 0
##

```



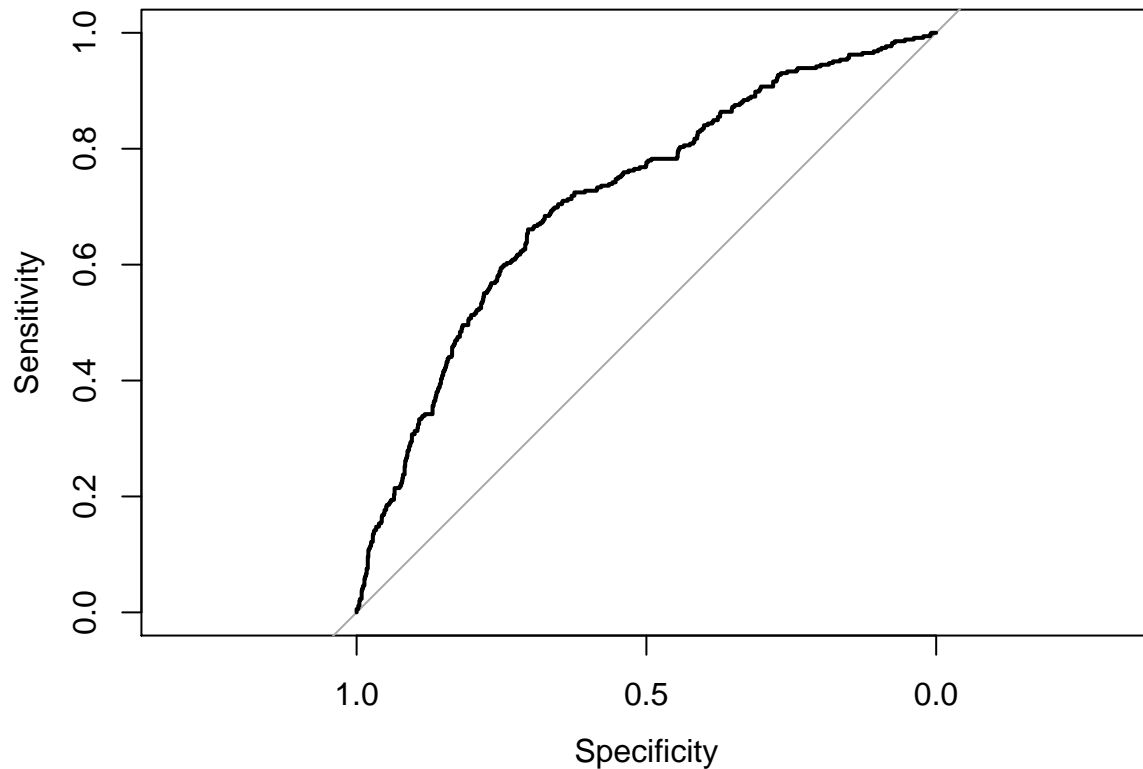


```
## [1] "AUC: 0.710598725699834"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1018 controls (dfPred_raw$class 0) < 362 cases (dfPred_raw$class 1).
## Area under the curve: 0.7106
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0091  -0.7755  -0.5752   0.9317   2.6938
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.729e-01  2.372e-01  -1.572  0.115986
## KIDSDRIV     2.055e-01  7.180e-02   2.862  0.004211 **
## AGE         -1.039e-02  4.526e-03  -2.296  0.021678 *
## HOMEKIDS      9.235e-02  3.792e-02   2.435  0.014876 *
## YOJ          -7.324e-03  8.688e-03  -0.843  0.399213
## INCOME       -4.271e-07  1.050e-06  -0.407  0.684204
## HOME_VAL    -2.476e-06  3.304e-07  -7.493  6.75e-14 ***
## TRAVTIME      7.668e-03  2.103e-03   3.647  0.000265 ***
## BLUEBOOK    -1.410e-05  4.662e-06  -3.025  0.002485 **
```

```

## TIF          -4.615e-02  8.616e-03  -5.356 8.53e-08 ***
## OLDCLAIM     7.117e-06  3.910e-06   1.820 0.068726 .
## CLM_FREQ     2.934e-01  3.246e-02   9.038 < 2e-16 ***
## MVR_PTS      1.266e-01  1.607e-02   7.879 3.30e-15 ***
## CAR_AGE      -2.314e-02  6.745e-03  -3.431 0.000602 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5889.3  on 5064  degrees of freedom
## Residual deviance: 5268.5  on 5051  degrees of freedom
## (1347 observations deleted due to missingness)
## AIC: 5296.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 964 271
##           1  74  74
##
##           Accuracy : 0.7505
##           95% CI : (0.7269, 0.7732)
##    No Information Rate : 0.7505
##    P-Value [Acc > NIR] : 0.5145
##
##           Kappa : 0.1769
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9287
##           Specificity : 0.2145
##           Pos Pred Value : 0.7806
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7505
##           Detection Rate : 0.6970
##           Detection Prevalence : 0.8930
##           Balanced Accuracy : 0.5716
##
##           'Positive' Class : 0
##

```

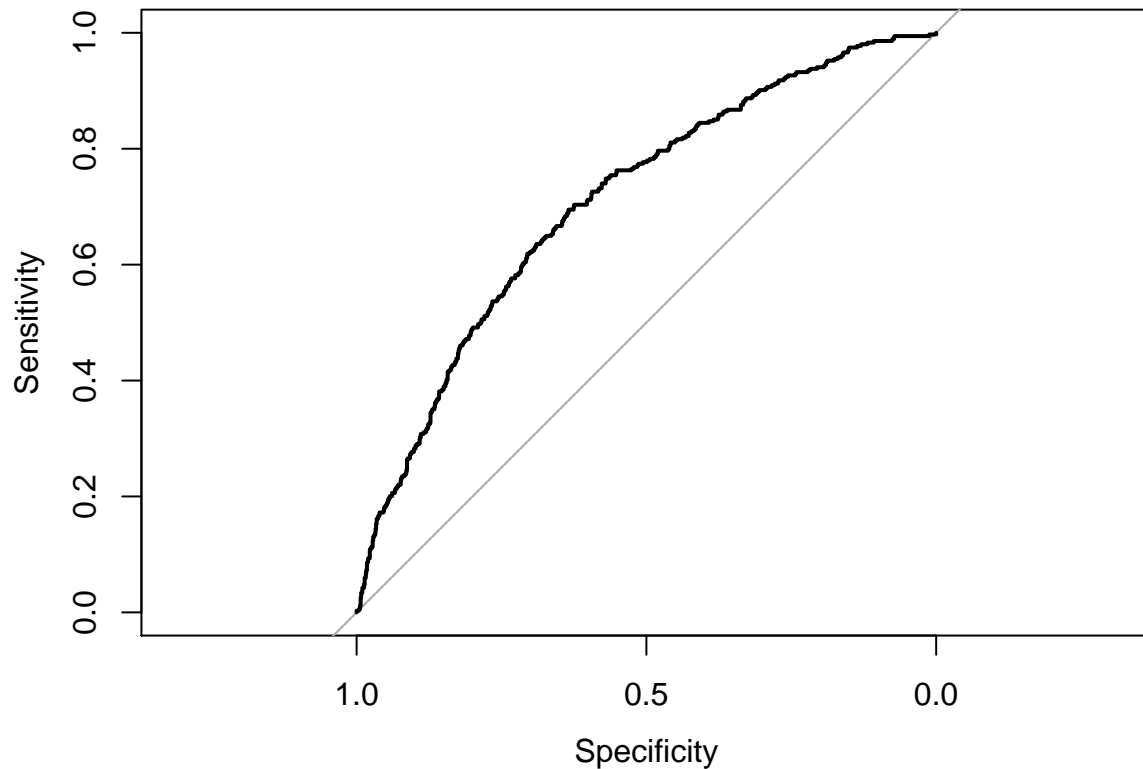


```
## [1] "AUC: 0.715073580743347"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1038 controls (dfPred_raw$class 0) < 345 cases (dfPred_raw$class 1).
## Area under the curve: 0.7151
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0487  -0.7693  -0.5705   0.9058   2.6716
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.243e-01  2.365e-01  -1.794 0.072867 .
## KIDSDRIV     2.455e-01  7.080e-02   3.468 0.000525 ***
## AGE          -8.803e-03  4.556e-03  -1.932 0.053368 .
## HOMEKIDS      9.365e-02  3.775e-02   2.481 0.013110 *
## YOJ          -1.267e-02  8.710e-03  -1.455 0.145731
## INCOME        -1.079e-07  1.056e-06  -0.102 0.918597
## HOME_VAL      -2.499e-06  3.304e-07  -7.563 3.94e-14 ***
## TRAVTIME      8.443e-03  2.130e-03   3.964 7.38e-05 ***
## BLUEBOOK     -1.646e-05  4.698e-06  -3.503 0.000461 ***
```

```

## TIF          -4.727e-02  8.601e-03  -5.496 3.88e-08 ***
## OLDCLAIM      6.076e-06  3.821e-06   1.590 0.111769
## CLM_FREQ      2.700e-01  3.240e-02   8.333 < 2e-16 ***
## MVR_PTS       1.370e-01  1.583e-02   8.657 < 2e-16 ***
## CAR_AGE       -2.139e-02  6.746e-03  -3.171 0.001520 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5881.7  on 5081  degrees of freedom
## Residual deviance: 5254.6  on 5068  degrees of freedom
##    (1330 observations deleted due to missingness)
## AIC: 5282.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 957 285
##           1  55  69
##
##           Accuracy : 0.7511
##           95% CI : (0.7273, 0.7738)
##    No Information Rate : 0.7408
##    P-Value [Acc > NIR] : 0.2027
##
##           Kappa : 0.1782
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9457
##           Specificity : 0.1949
##           Pos Pred Value : 0.7705
##           Neg Pred Value : 0.5565
##           Prevalence : 0.7408
##           Detection Rate : 0.7006
##           Detection Prevalence : 0.9092
##           Balanced Accuracy : 0.5703
##
##           'Positive' Class : 0
##

```

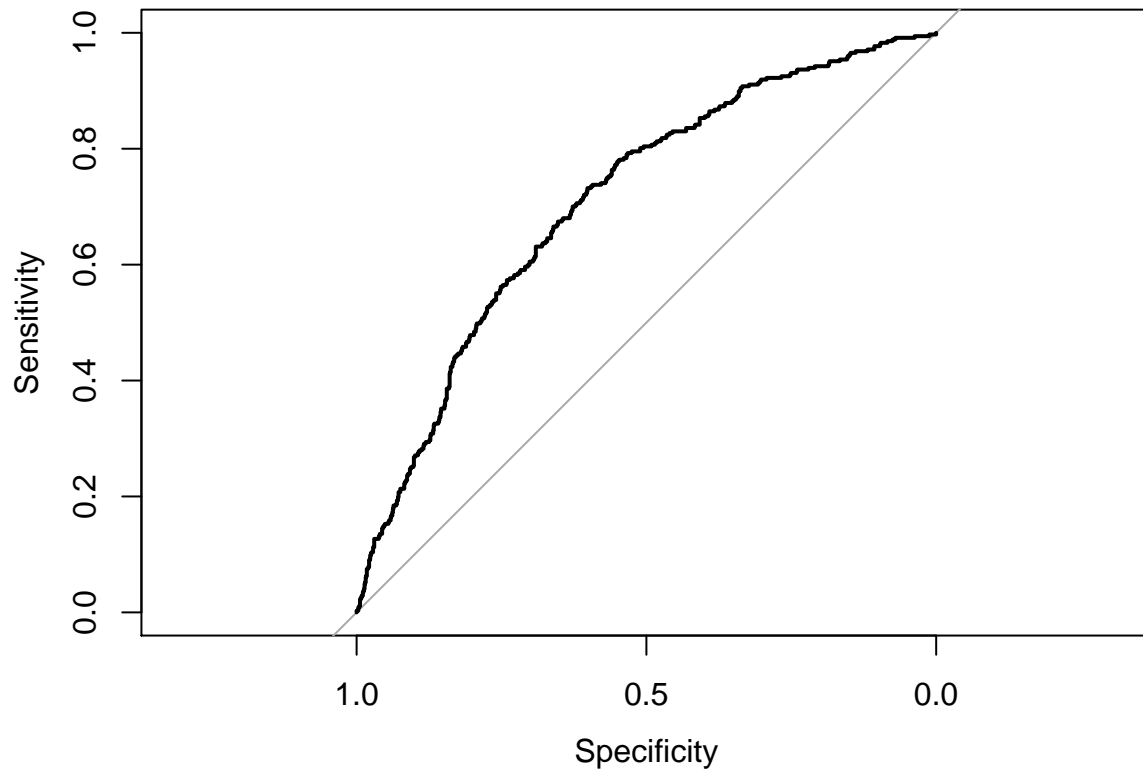


```
## [1] "AUC: 0.708986512136844"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1012 controls (dfPred_raw$class 0) < 354 cases (dfPred_raw$class 1).
## Area under the curve: 0.709
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0298  -0.7684  -0.5703   0.9082   2.6270
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.050e-01  2.375e-01  -2.126  0.03351 *
## KIDSDRIV     2.250e-01  7.071e-02   3.182  0.00146 **
## AGE          -1.007e-02  4.564e-03  -2.206  0.02737 *
## HOMEKIDS      9.935e-02  3.750e-02   2.649  0.00807 **
## YOJ           -1.074e-02  8.664e-03  -1.239  0.21517
## INCOME        -1.684e-07  1.041e-06  -0.162  0.87146
## HOME_VAL      -2.501e-06  3.296e-07  -7.588 3.25e-14 ***
## TRAVTIME       8.303e-03  2.128e-03   3.901 9.58e-05 ***
## BLUEBOOK      -1.217e-05  4.663e-06  -2.610  0.00906 **
```

```

## TIF          -4.304e-02  8.581e-03  -5.016  5.27e-07 ***
## OLDCLAIM      4.506e-06  3.896e-06   1.157  0.24743
## CLM_FREQ      2.890e-01  3.206e-02   9.016  < 2e-16 ***
## MVR_PTS       1.424e-01  1.578e-02   9.026  < 2e-16 ***
## CAR_AGE       -1.794e-02  6.729e-03  -2.666  0.00767 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5894.0  on 5078  degrees of freedom
## Residual deviance: 5260.6  on 5065  degrees of freedom
##    (1333 observations deleted due to missingness)
## AIC: 5288.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 960 289
##           1  62  58
##
##           Accuracy : 0.7436
##           95% CI : (0.7196, 0.7666)
##    No Information Rate : 0.7465
##    P-Value [Acc > NIR] : 0.6119
##
##           Kappa : 0.1358
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9393
##           Specificity : 0.1671
##           Pos Pred Value : 0.7686
##           Neg Pred Value : 0.4833
##           Prevalence : 0.7465
##           Detection Rate : 0.7012
##    Detection Prevalence : 0.9123
##           Balanced Accuracy : 0.5532
##
##           'Positive' Class : 0
##

```



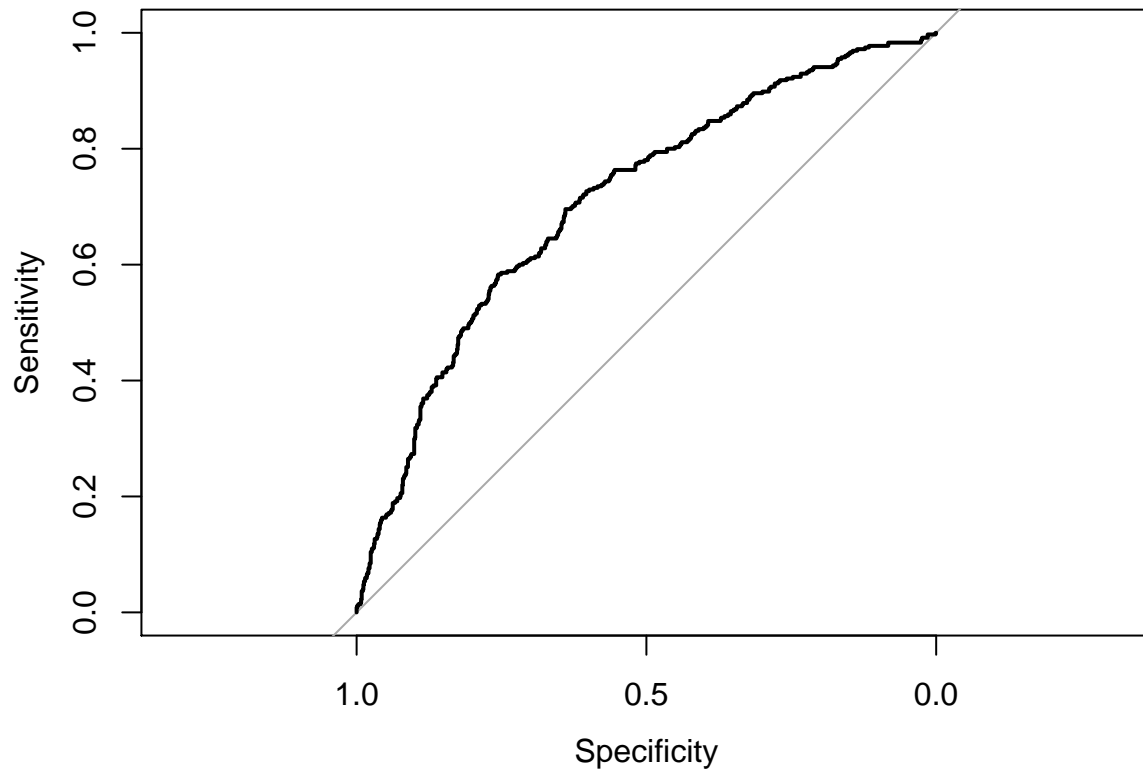
```
## [1] "AUC: 0.709740182836389"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1022 controls (dfPred_raw$class 0) < 347 cases (dfPred_raw$class 1).
## Area under the curve: 0.7097
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1189  -0.7731  -0.5722   0.8993   2.6865
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -7.393e-01  2.381e-01  -3.105 0.001901 **
## KIDSDRIV     2.504e-01  7.094e-02   3.530 0.000416 ***
## AGE          -4.298e-03  4.543e-03  -0.946 0.344136
## HOMEKIDS      9.671e-02  3.778e-02   2.560 0.010461 *
## YOJ          -4.413e-03  8.791e-03  -0.502 0.615681
## INCOME        -2.823e-07  1.055e-06  -0.268 0.789021
## HOME_VAL      -2.656e-06  3.308e-07  -8.029 9.84e-16 ***
## TRAVTIME       9.616e-03  2.126e-03   4.523 6.10e-06 ***
## BLUEBOOK      -1.413e-05  4.688e-06  -3.014 0.002575 **
```

```

## TIF          -4.451e-02  8.633e-03  -5.156 2.52e-07 ***
## OLDCLAIM      8.385e-06  3.954e-06   2.121 0.033939 *
## CLM_FREQ      2.549e-01  3.242e-02   7.863 3.75e-15 ***
## MVR_PTS       1.474e-01  1.597e-02   9.230 < 2e-16 ***
## CAR_AGE       -2.547e-02  6.805e-03  -3.743 0.000181 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5869.8  on 5065  degrees of freedom
## Residual deviance: 5241.0  on 5052  degrees of freedom
## (1347 observations deleted due to missingness)
## AIC: 5269
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 955 285
##           1  72  70
##
##           Accuracy : 0.7417
##           95% CI : (0.7177, 0.7646)
##    No Information Rate : 0.7431
##    P-Value [Acc > NIR] : 0.5631
##
##           Kappa : 0.1581
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9299
##           Specificity : 0.1972
##           Pos Pred Value : 0.7702
##           Neg Pred Value : 0.4930
##           Prevalence : 0.7431
##           Detection Rate : 0.6910
##           Detection Prevalence : 0.8973
##           Balanced Accuracy : 0.5635
##
##           'Positive' Class : 0
##

```



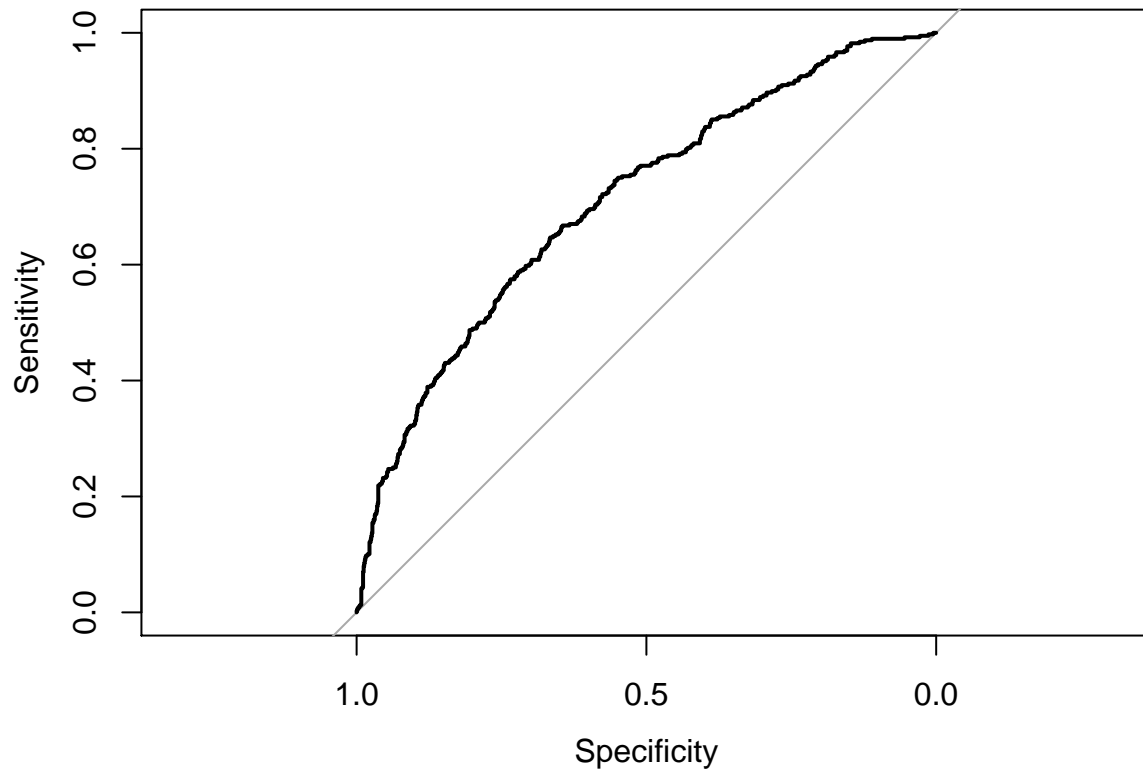


```
## [1] "AUC: 0.710147153613012"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1027 controls (dfPred_raw$class 0) < 355 cases (dfPred_raw$class 1).
## Area under the curve: 0.7101
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9888  -0.7613  -0.5679   0.8684   2.6711
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.298e-01  2.404e-01  -0.956  0.339150
## KIDSDRIV     2.452e-01  6.917e-02   3.545  0.000393 ***
## AGE         -1.183e-02  4.633e-03  -2.555  0.010633 *
## HOMEKIDS      6.541e-02  3.823e-02   1.711  0.087115 .
## YOJ          -1.366e-02  8.786e-03  -1.555  0.119952
## INCOME        7.909e-07  1.059e-06   0.747  0.455175
## HOME_VAL     -2.560e-06  3.313e-07  -7.726  1.11e-14 ***
## TRAVTIME      7.096e-03  2.134e-03   3.325  0.000886 ***
## BLUEBOOK     -1.499e-05  4.700e-06  -3.189  0.001428 **
```

```

## TIF          -4.988e-02  8.614e-03  -5.791  6.98e-09 ***
## OLDCLAIM     3.943e-06  3.979e-06   0.991  0.321611
## CLM_FREQ     2.712e-01  3.251e-02   8.344  < 2e-16 ***
## MVRPTS       1.405e-01  1.604e-02   8.762  < 2e-16 ***
## CARAGE       -2.924e-02  6.780e-03  -4.313  1.61e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5803.5  on 5068  degrees of freedom
## Residual deviance: 5197.6  on 5055  degrees of freedom
## (1343 observations deleted due to missingness)
## AIC: 5225.6
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0  948 302
##           1   43  86
##
##           Accuracy : 0.7498
##           95% CI : (0.7261, 0.7725)
##    No Information Rate : 0.7186
##    P-Value [Acc > NIR] : 0.005074
##
##           Kappa : 0.2237
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.9566
##           Specificity : 0.2216
##           Pos Pred Value : 0.7584
##           Neg Pred Value : 0.6667
##           Prevalence : 0.7186
##           Detection Rate : 0.6875
##           Detection Prevalence : 0.9065
##           Balanced Accuracy : 0.5891
##
##           'Positive' Class : 0
##

```

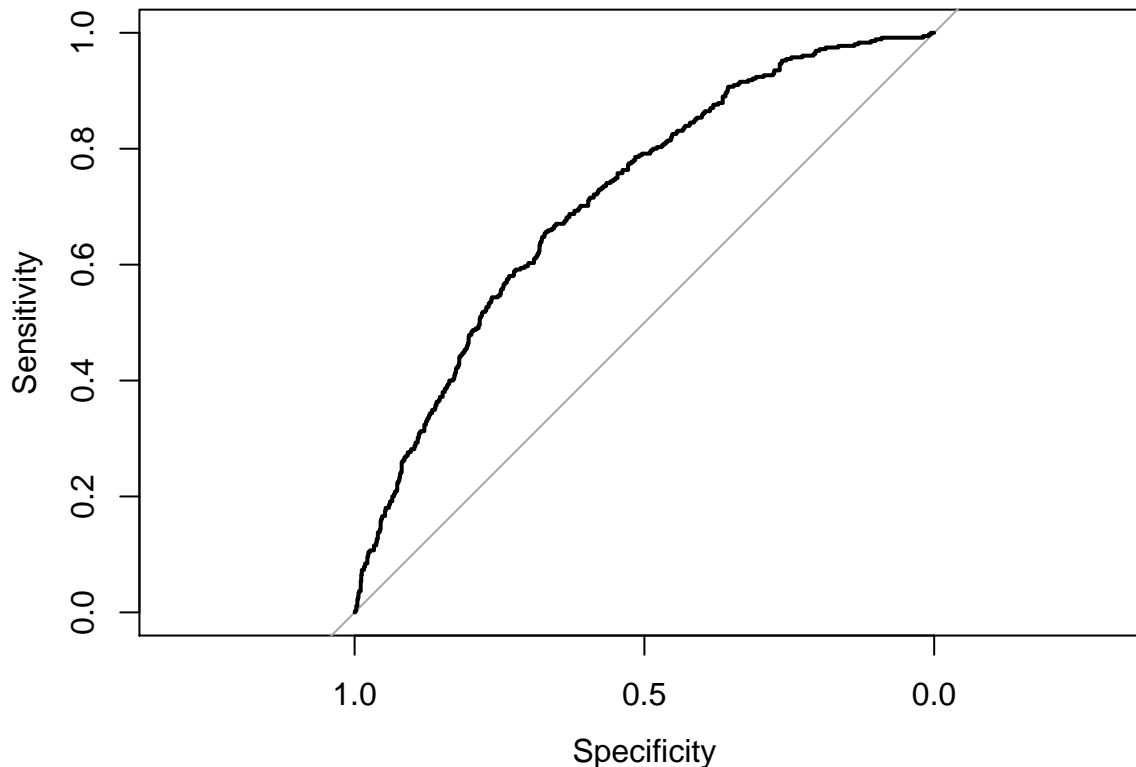


```
## [1] "AUC: 0.709184204229821"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 991 controls (dfPred_raw$class 0) < 388 cases (dfPred_raw$class 1).
## Area under the curve: 0.7092
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7546  -0.7691  -0.5747   0.9263   2.6858
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.626e-01  2.366e-01  -2.378  0.017403 *
## KIDSDRIV     3.273e-01  7.023e-02   4.660  3.16e-06 ***
## AGE          -1.112e-02  4.530e-03  -2.454  0.014137 *
## HOMEKIDS      5.125e-02  3.809e-02   1.345  0.178466
## YOJ           5.965e-03  8.734e-03   0.683  0.494617
## INCOME        -8.953e-07  1.036e-06  -0.864  0.387730
## HOME_VAL      -2.562e-06  3.313e-07  -7.732  1.06e-14 ***
## TRAVTIME       7.360e-03  2.127e-03   3.461  0.000539 ***
## BLUEBOOK      -1.108e-05  4.680e-06  -2.368  0.017876 *
```

```

## TIF          -3.871e-02  8.491e-03  -4.559 5.14e-06 ***
## OLDCLAIM      8.604e-06  3.974e-06   2.165 0.030379 *
## CLM_FREQ      2.885e-01  3.242e-02   8.900 < 2e-16 ***
## MVR_PTS       1.323e-01  1.583e-02   8.358 < 2e-16 ***
## CAR_AGE       -1.864e-02  6.702e-03  -2.781 0.005426 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5857.4  on 5045  degrees of freedom
## Residual deviance: 5240.2  on 5032  degrees of freedom
##    (1366 observations deleted due to missingness)
## AIC: 5268.2
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 979 287
##           1  68  68
##
##           Accuracy : 0.7468
##           95% CI : (0.7232, 0.7694)
##    No Information Rate : 0.7468
##    P-Value [Acc > NIR] : 0.5143
##
##           Kappa : 0.159
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9351
##           Specificity : 0.1915
##           Pos Pred Value : 0.7733
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7468
##           Detection Rate : 0.6983
##           Detection Prevalence : 0.9030
##           Balanced Accuracy : 0.5633
##
##           'Positive' Class : 0
##

```

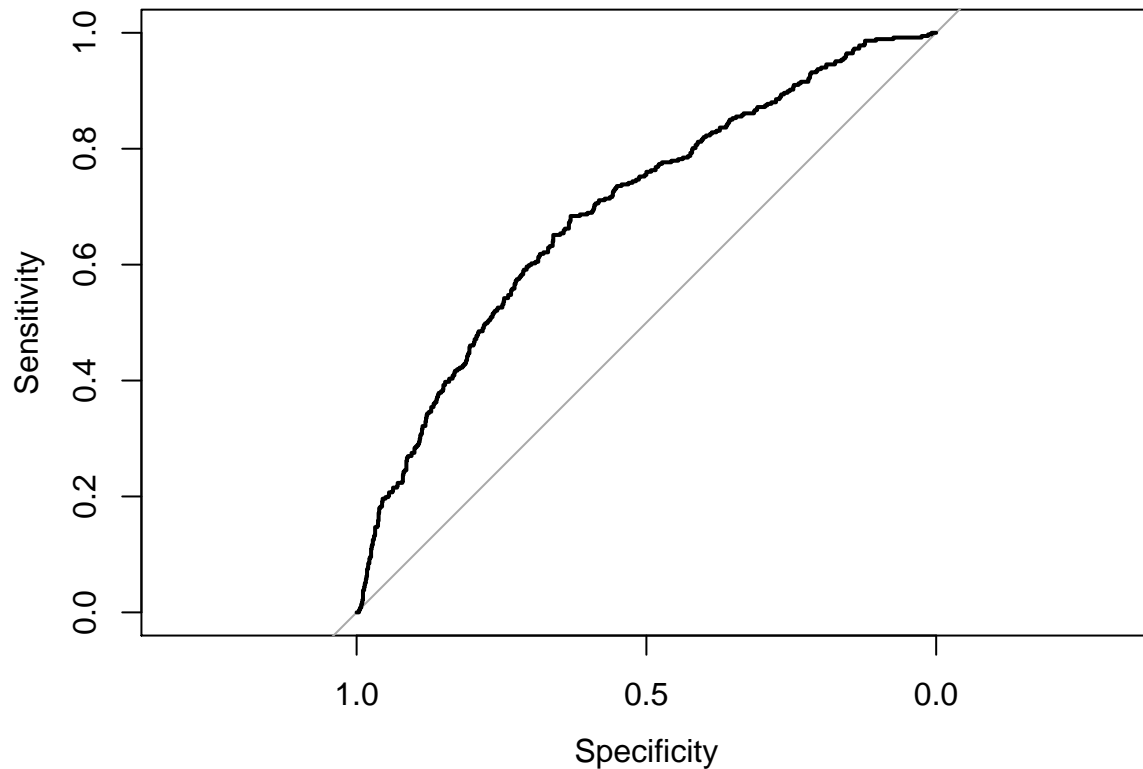


```
## [1] "AUC: 0.713480500961836"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1047 controls (dfPred_raw$class 0) < 355 cases (dfPred_raw$class 1).
## Area under the curve: 0.7135
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0316  -0.7672  -0.5625   0.8859   2.7621
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.273e-01  2.397e-01  -1.365 0.172126
## KIDSDRIV     2.313e-01  7.063e-02   3.276 0.001054 **
## AGE         -9.699e-03  4.604e-03  -2.107 0.035122 *
## HOMEKIDS      9.258e-02  3.789e-02   2.444 0.014540 *
## YOJ         -1.057e-02  8.778e-03  -1.204 0.228719
## INCOME        2.580e-07  1.053e-06   0.245 0.806447
## HOME_VAL     -2.525e-06  3.314e-07  -7.619 2.56e-14 ***
## TRAVTIME      7.827e-03  2.136e-03   3.664 0.000248 ***
## BLUEBOOK     -1.970e-05  4.732e-06  -4.162 3.15e-05 ***
```

```

## TIF          -4.807e-02  8.647e-03  -5.559 2.71e-08 ***
## OLDCLAIM      6.530e-06  3.896e-06   1.676 0.093702 .
## CLM_FREQ      2.790e-01  3.258e-02   8.565 < 2e-16 ***
## MVR_PTS       1.349e-01  1.606e-02   8.402 < 2e-16 ***
## CAR_AGE       -2.568e-02  6.808e-03  -3.772 0.000162 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5853.2  on 5078  degrees of freedom
## Residual deviance: 5213.0  on 5065  degrees of freedom
##    (1333 observations deleted due to missingness)
## AIC: 5241
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 942 291
##           1  60  76
##
##           Accuracy : 0.7436
##           95% CI : (0.7196, 0.7666)
##    No Information Rate : 0.7319
##    P-Value [Acc > NIR] : 0.1723
##
##           Kappa : 0.1839
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9401
##           Specificity : 0.2071
##           Pos Pred Value : 0.7640
##           Neg Pred Value : 0.5588
##           Prevalence : 0.7319
##           Detection Rate : 0.6881
##           Detection Prevalence : 0.9007
##           Balanced Accuracy : 0.5736
##
##           'Positive' Class : 0
##

```



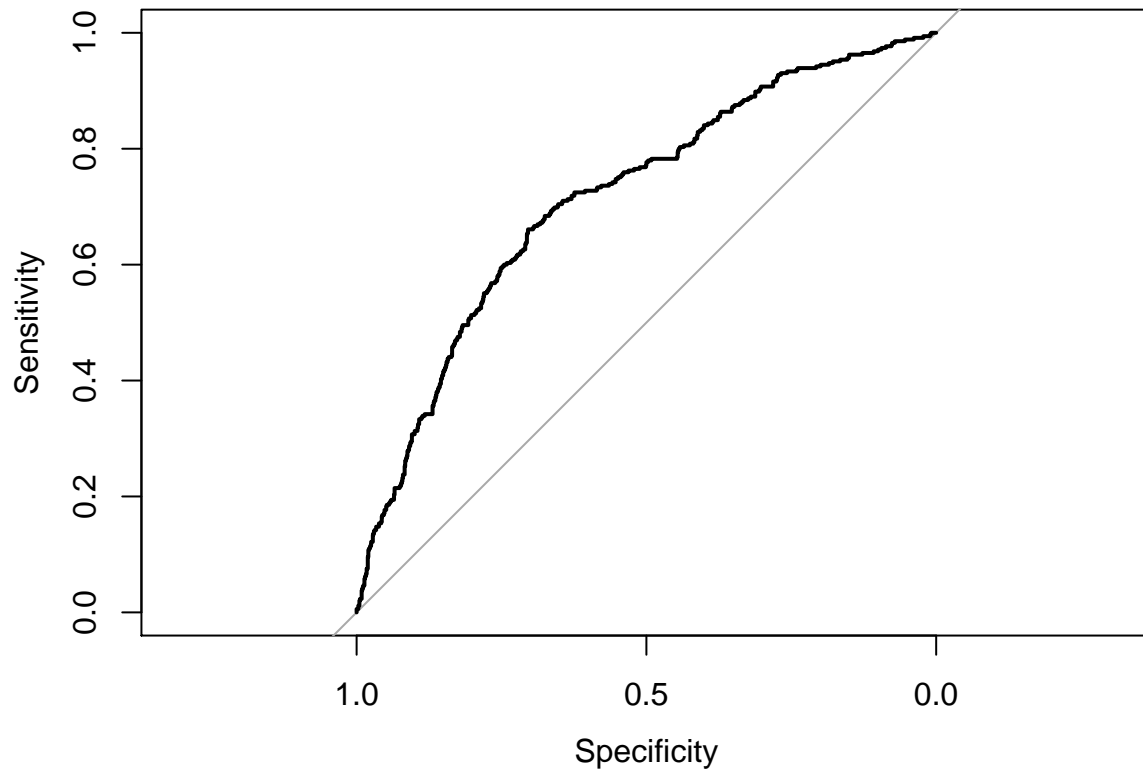
```
## [1] "AUC: 0.694583584873849"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1002 controls (dfPred_raw$class 0) < 367 cases (dfPred_raw$class 1).
## Area under the curve: 0.6946
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0091  -0.7755  -0.5752   0.9317   2.6938
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.729e-01  2.372e-01  -1.572  0.115986
## KIDSDRIV     2.055e-01  7.180e-02   2.862  0.004211 **
## AGE          -1.039e-02  4.526e-03  -2.296  0.021678 *
## HOMEKIDS      9.235e-02  3.792e-02   2.435  0.014876 *
## YOJ          -7.324e-03  8.688e-03  -0.843  0.399213
## INCOME        -4.271e-07  1.050e-06  -0.407  0.684204
## HOME_VAL      -2.476e-06  3.304e-07  -7.493  6.75e-14 ***
## TRAVTIME       7.668e-03  2.103e-03   3.647  0.000265 ***
## BLUEBOOK      -1.410e-05  4.662e-06  -3.025  0.002485 **
```

```

## TIF          -4.615e-02  8.616e-03  -5.356 8.53e-08 ***
## OLDCLAIM     7.117e-06  3.910e-06   1.820 0.068726 .
## CLM_FREQ     2.934e-01  3.246e-02   9.038 < 2e-16 ***
## MVR_PTS      1.266e-01  1.607e-02   7.879 3.30e-15 ***
## CAR_AGE      -2.314e-02  6.745e-03  -3.431 0.000602 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5889.3  on 5064  degrees of freedom
## Residual deviance: 5268.5  on 5051  degrees of freedom
##    (1347 observations deleted due to missingness)
## AIC: 5296.5
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 964 271
##           1  74  74
##
##           Accuracy : 0.7505
##           95% CI : (0.7269, 0.7732)
##    No Information Rate : 0.7505
##    P-Value [Acc > NIR] : 0.5145
##
##           Kappa : 0.1769
##
## Mcnemar's Test P-Value : <2e-16
##
##           Sensitivity : 0.9287
##           Specificity : 0.2145
##           Pos Pred Value : 0.7806
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7505
##           Detection Rate : 0.6970
##           Detection Prevalence : 0.8930
##           Balanced Accuracy : 0.5716
##
##           'Positive' Class : 0
##

```



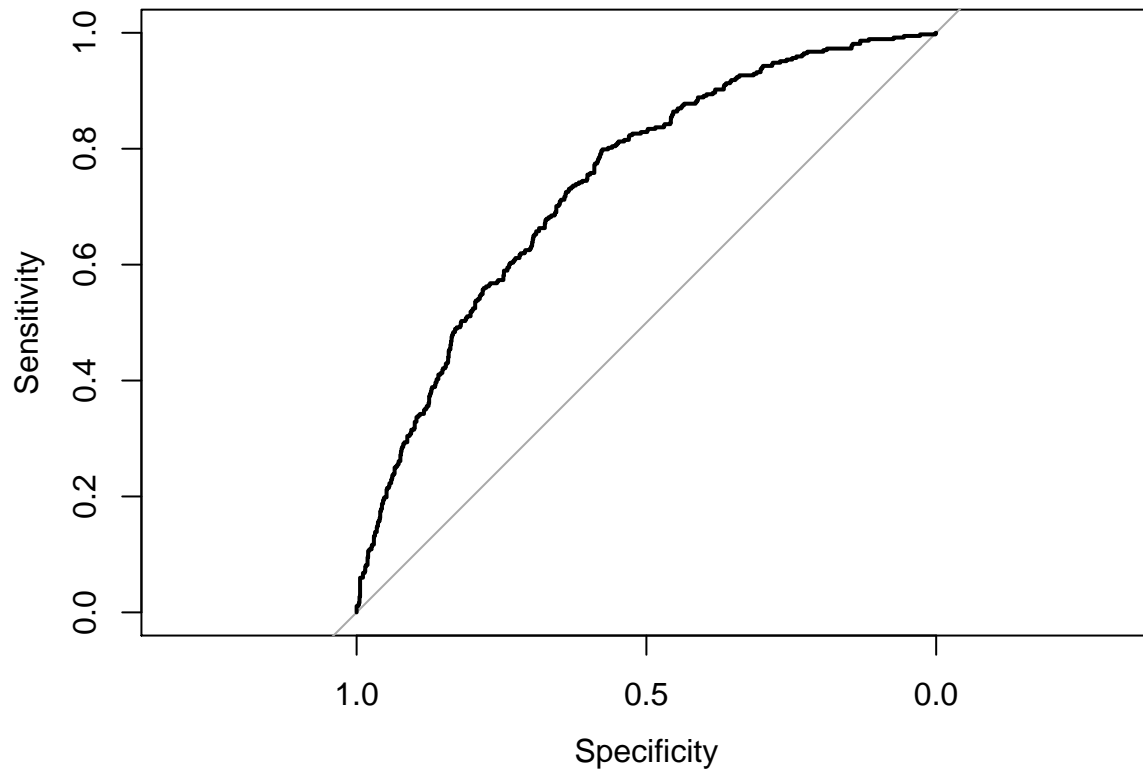


```
## [1] "AUC: 0.715073580743347"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1038 controls (dfPred_raw$class 0) < 345 cases (dfPred_raw$class 1).
## Area under the curve: 0.7151
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0695  -0.7696  -0.5815   0.9258   2.6470
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.029e-01  2.375e-01  -2.539  0.01112 *
## KIDSDRIV     2.959e-01  6.910e-02   4.282 1.85e-05 ***
## AGE          -7.679e-03  4.526e-03  -1.697  0.08978 .
## HOMEKIDS      5.599e-02  3.778e-02   1.482  0.13839
## YOJ          -3.311e-03  8.789e-03  -0.377  0.70637
## INCOME        -9.286e-07  1.045e-06  -0.888  0.37445
## HOME_VAL      -2.220e-06  3.298e-07  -6.731 1.68e-11 ***
## TRAVTIME      8.742e-03  2.119e-03   4.125 3.70e-05 ***
## BLUEBOOK      -1.209e-05  4.659e-06  -2.595  0.00946 **
```

```

## TIF          -4.948e-02  8.621e-03  -5.739  9.52e-09 ***
## OLDCLAIM     9.826e-06  3.878e-06   2.534  0.01129 *
## CLM_FREQ     2.583e-01  3.226e-02   8.006  1.18e-15 ***
## MVR_PTS      1.274e-01  1.573e-02   8.096  5.68e-16 ***
## CAR_AGE      -2.104e-02  6.710e-03  -3.135  0.00172 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5847.5  on 5072  degrees of freedom
## Residual deviance: 5274.9  on 5059  degrees of freedom
##    (1339 observations deleted due to missingness)
## AIC: 5302.9
##
## Number of Fisher Scoring iterations: 4
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 963 299
##           1  44  69
##
##           Accuracy : 0.7505
##           95% CI : (0.7268, 0.7732)
##    No Information Rate : 0.7324
##    P-Value [Acc > NIR] : 0.06701
##
##           Kappa : 0.1843
##
## Mcnemar's Test P-Value : < 2e-16
##
##           Sensitivity : 0.9563
##           Specificity : 0.1875
##           Pos Pred Value : 0.7631
##           Neg Pred Value : 0.6106
##           Prevalence : 0.7324
##           Detection Rate : 0.7004
##           Detection Prevalence : 0.9178
##           Balanced Accuracy : 0.5719
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.740423017140883"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 1007 controls (dfPred_raw$class 0) < 368 cases (dfPred_raw$class 1).
## Area under the curve: 0.7404
## [1] "Accuracy: 0.748154624686216"
## [1] "AIC: 5265.33743871079"
## [1] "AUC: 0.714855781543545"
```

The base model has an accuracy of .748, an AIC of 5270 and an AUC of .716.

**Create Model 2 - a model with all of the transformed and added variables.**

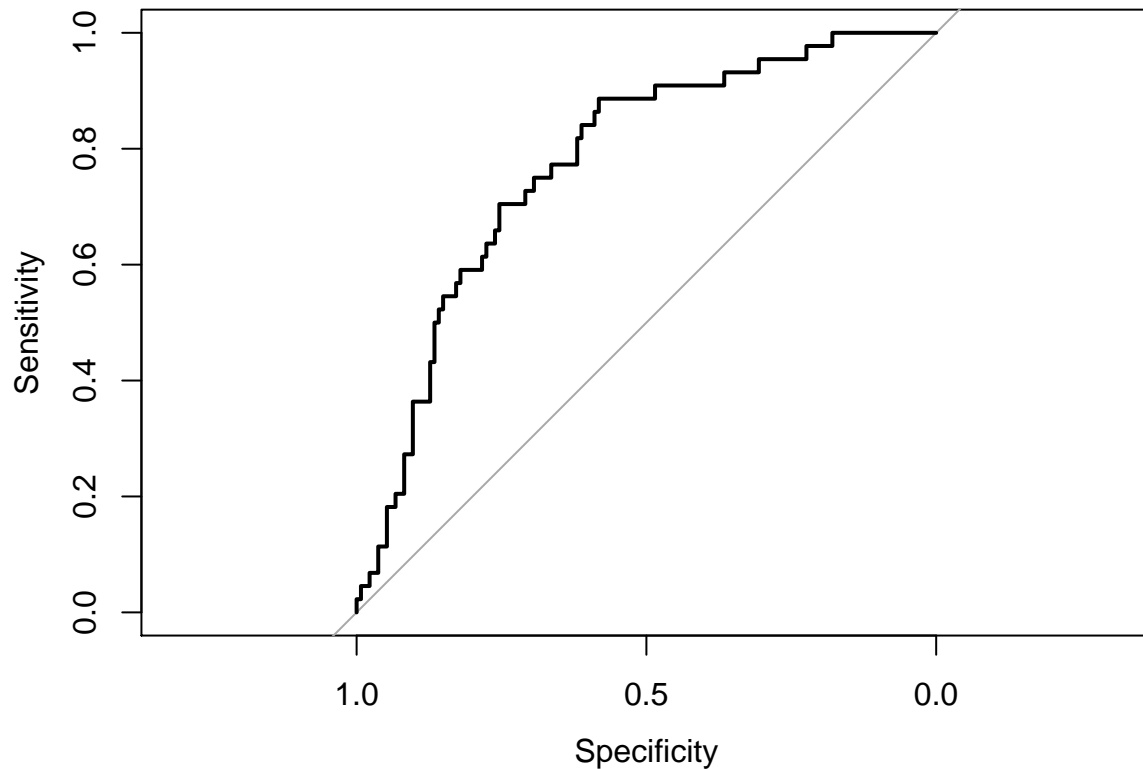
```
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2910  -0.7093  -0.3641   0.5468   2.9284
##
## Coefficients:
##                                     Estimate Std. Error z value Pr(>|z|)
```

## (Intercept)	1.912e+01	1.418e+01	1.348	0.177526	
## KIDSDRIV	-1.059e-01	1.939e+00	-0.055	0.956447	
## AGE	-5.026e-02	1.078e-01	-0.466	0.641187	
## HOMEKIDS	-5.584e-01	7.184e-01	-0.777	0.436971	
## YOJ	-2.478e-01	1.302e-01	-1.903	0.057092	.
## INCOME	-1.149e-05	1.311e-05	-0.876	0.381057	
## HOME_VAL	6.718e-06	8.657e-06	0.776	0.437735	
## TRAVTIME	1.682e-02	2.139e-02	0.787	0.431543	
## BLUEBOOK	4.467e-05	3.894e-05	1.147	0.251279	
## TIF	-7.228e-02	9.206e-02	-0.785	0.432389	
## OLDCLAIM	-2.034e-05	2.220e-05	-0.916	0.359542	
## CLM_FREQ	8.582e-01	9.691e-01	0.886	0.375860	
## MVR_PTS	-3.075e-02	1.646e-01	-0.187	0.851853	
## CAR_AGE	5.256e-04	7.361e-02	0.007	0.994303	
## PARENT1_Yes	9.141e-01	4.776e-01	1.914	0.055616	.
## MSTATUS_Yes	-4.166e-01	3.246e-01	-1.283	0.199367	
## SEX_z_F	-7.086e-01	4.555e-01	-1.556	0.119747	
## EDUCATION_.High.School	-1.733e-01	7.731e-01	-0.224	0.822601	
## EDUCATION_Bachelors	-1.935e-01	6.601e-01	-0.293	0.769427	
## EDUCATION_Masters	7.664e-01	5.566e-01	1.377	0.168559	
## EDUCATION_z_High.School	9.230e-02	7.114e-01	0.130	0.896765	
## JOB_	-1.010e+00	6.872e-01	-1.469	0.141845	
## JOB_Clerical	-4.370e-01	4.851e-01	-0.901	0.367634	
## JOB_Doctor	-1.230e-01	9.176e-01	-0.134	0.893339	
## JOB_Home.Maker	-7.938e-01	6.765e-01	-1.173	0.240654	
## JOB_Lawyer	-1.089e+00	6.577e-01	-1.656	0.097652	.
## JOB_Manager	-1.003e+00	4.697e-01	-2.135	0.032752	*
## JOB_Student	-1.371e+00	7.431e-01	-1.844	0.065125	.
## JOB_z_Blue.Collar	-4.552e-01	4.427e-01	-1.028	0.303900	
## CAR_USE_Commercial	8.744e-01	3.560e-01	2.457	0.014027	*
## CAR_TYPE_Panel.Truck	-2.586e-01	6.233e-01	-0.415	0.678231	
## CAR_TYPE_Pickup	8.474e-01	4.038e-01	2.099	0.035854	*
## CAR_TYPE_Sports.Car	1.179e+00	5.316e-01	2.217	0.026624	*
## CAR_TYPE_Van	1.475e-01	4.684e-01	0.315	0.752896	
## CAR_TYPE_z_SUV	1.594e+00	4.512e-01	3.532	0.000412	***
## RED_CAR_no	-4.933e-02	3.403e-01	-0.145	0.884725	
## REVOKED_Yes	9.829e-01	3.985e-01	2.466	0.013653	*
## URBANICITY_z_Highly.Rural..Rural	-2.446e+00	4.315e-01	-5.667	1.45e-08	***
## YOJ_NA	-3.299e-01	4.613e-01	-0.715	0.474525	
## INCOME_NA	3.327e-01	5.745e-01	0.579	0.562591	
## CAR_AGE_NA	-1.989e-01	4.583e-01	-0.434	0.664317	
## HOME_VAL_NA	-3.553e-01	2.850e-01	-1.246	0.212597	
## ageSquared	6.623e-04	1.158e-03	0.572	0.567499	
## yojSquared	1.237e-02	6.614e-03	1.870	0.061421	.
## income_log	8.257e-02	3.045e-01	0.271	0.786267	
## homeval_log	-1.537e+00	1.336e+00	-1.150	0.250007	
## travtime_log	1.952e-01	6.066e-01	0.322	0.747665	
## bluebook_log	-2.063e-01	4.729e-01	-0.436	0.662587	
## carage_log	-3.534e-01	4.744e-01	-0.745	0.456336	
## oldclaim_log	1.555e-01	1.677e-01	0.927	0.353804	
## clm_freq_log	-2.327e+00	3.016e+00	-0.772	0.440367	
## mvr_pts_log	1.785e-01	4.870e-01	0.367	0.713897	
## tif_log	1.312e-02	5.197e-01	0.025	0.979856	
## kidsdriv_log	1.094e+00	2.436e+00	0.449	0.653272	

```

## homekids_log          1.028e+00  1.521e+00  0.676 0.498997
## inter                 5.632e-03  3.409e-02  0.165 0.868799
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 741.72  on 637  degrees of freedom
## Residual deviance: 548.52  on 582  degrees of freedom
##    (1 observation deleted due to missingness)
## AIC: 660.52
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 117  28
##           1  17  16
##
##           Accuracy : 0.7472
##           95% CI : (0.6767, 0.8092)
##    No Information Rate : 0.7528
##    P-Value [Acc > NIR] : 0.608
##
##           Kappa : 0.2585
##
## Mcnemar's Test P-Value : 0.136
##
##           Sensitivity : 0.8731
##           Specificity : 0.3636
##           Pos Pred Value : 0.8069
##           Neg Pred Value : 0.4848
##           Prevalence : 0.7528
##           Detection Rate : 0.6573
##    Detection Prevalence : 0.8146
##           Balanced Accuracy : 0.6184
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.774762550881954"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7748
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3571  -0.6316  -0.3463   0.4966   3.1034
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.157e+01  1.481e+01   1.456  0.145411
## KIDSDRIV      -1.312e-01  1.907e+00  -0.069  0.945173
## AGE           -5.203e-02  1.171e-01  -0.444  0.656895
## HOMEKIDS      -1.289e+00  7.074e-01  -1.822  0.068437 .
## YOJ           -2.251e-01  1.413e-01  -1.593  0.111206
## INCOME        -2.823e-05  1.352e-05  -2.088  0.036803 *
## HOME_VAL       1.599e-05  8.927e-06   1.791  0.073303 .
## TRAVTIME       1.466e-02  2.166e-02   0.677  0.498681
## BLUEBOOK       7.606e-05  3.893e-05   1.954  0.050728 .
```

```

## TIF -6.094e-02 9.565e-02 -0.637 0.524042
## OLDCLAIM -6.523e-06 2.190e-05 -0.298 0.765839
## CLM_FREQ 6.789e-01 9.508e-01 0.714 0.475236
## MVR_PTS 6.396e-02 1.713e-01 0.373 0.708897
## CAR_AGE 3.940e-02 7.461e-02 0.528 0.597447
## PARENT1_Yes 2.505e-01 4.831e-01 0.519 0.604090
## MSTATUS_Yes -7.653e-01 3.364e-01 -2.275 0.022914 *
## SEX_z_F -7.589e-01 4.749e-01 -1.598 0.110041
## EDUCATION_.High.School 4.734e-01 8.279e-01 0.572 0.567442
## EDUCATION_Bachelors 2.281e-01 7.218e-01 0.316 0.752037
## EDUCATION_Masters 7.951e-01 6.200e-01 1.282 0.199683
## EDUCATION_z_High.School 6.394e-01 7.670e-01 0.834 0.404463
## JOB_ -8.104e-01 7.486e-01 -1.083 0.278958
## JOB_Clerical -2.315e-01 4.778e-01 -0.484 0.628068
## JOB_Doctor 2.971e-02 1.059e+00 0.028 0.977621
## JOB_Home.Maker -2.440e-01 6.908e-01 -0.353 0.723878
## JOB_Lawyer -7.696e-01 6.587e-01 -1.168 0.242636
## JOB_Manager -1.027e+00 4.838e-01 -2.123 0.033745 *
## JOB_Student -4.821e-01 7.293e-01 -0.661 0.508628
## JOB_z_Blue.Collar -1.445e-01 4.545e-01 -0.318 0.750510
## CAR_USE_Commercial 4.188e-01 3.462e-01 1.210 0.226469
## CAR_TYPE_Panel.Truck 1.963e-01 6.203e-01 0.316 0.751717
## CAR_TYPE_Pickup 1.152e+00 4.115e-01 2.799 0.005118 **
## CAR_TYPE_Sports.Car 2.103e+00 5.458e-01 3.854 0.000116 ***
## CAR_TYPE_Van 3.372e-01 4.671e-01 0.722 0.470369
## CAR_TYPE_z_SUV 1.870e+00 4.906e-01 3.812 0.000138 ***
## RED_CAR_no -1.182e-01 3.306e-01 -0.358 0.720712
## REVOKED_Yes 6.125e-01 4.116e-01 1.488 0.136697
## URBANICITY_z_Highly.Rural..Rural -2.302e+00 4.086e-01 -5.634 1.76e-08 ***
## YOJ_NA -6.597e-01 4.414e-01 -1.495 0.135008
## INCOME_NA 7.084e-01 6.410e-01 1.105 0.269132
## CAR_AGE_NA 5.177e-02 5.015e-01 0.103 0.917778
## HOME_VAL_NA -2.039e-01 2.936e-01 -0.694 0.487373
## ageSquared 2.704e-04 1.278e-03 0.211 0.832506
## yojSquared 1.058e-02 7.120e-03 1.486 0.137245
## income_log 2.047e-01 3.009e-01 0.680 0.496323
## homeval_log -1.662e+00 1.397e+00 -1.189 0.234329
## travtime_log 1.693e-01 6.128e-01 0.276 0.782352
## bluebook_log -5.537e-01 4.719e-01 -1.173 0.240671
## carage_log -5.262e-01 4.751e-01 -1.107 0.268096
## oldclaim_log 1.272e-01 1.633e-01 0.779 0.436247
## clm_freq_log -1.701e+00 2.959e+00 -0.575 0.565273
## mvr_pts_log -7.991e-02 5.073e-01 -0.158 0.874840
## tif_log -8.911e-02 5.369e-01 -0.166 0.868185
## kidsdriv_log 6.662e-01 2.228e+00 0.299 0.764961
## homekids_log 2.587e+00 1.502e+00 1.722 0.085066 .
## inter 7.588e-03 3.573e-02 0.212 0.831820
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 746.22 on 641 degrees of freedom
## Residual deviance: 537.32 on 586 degrees of freedom

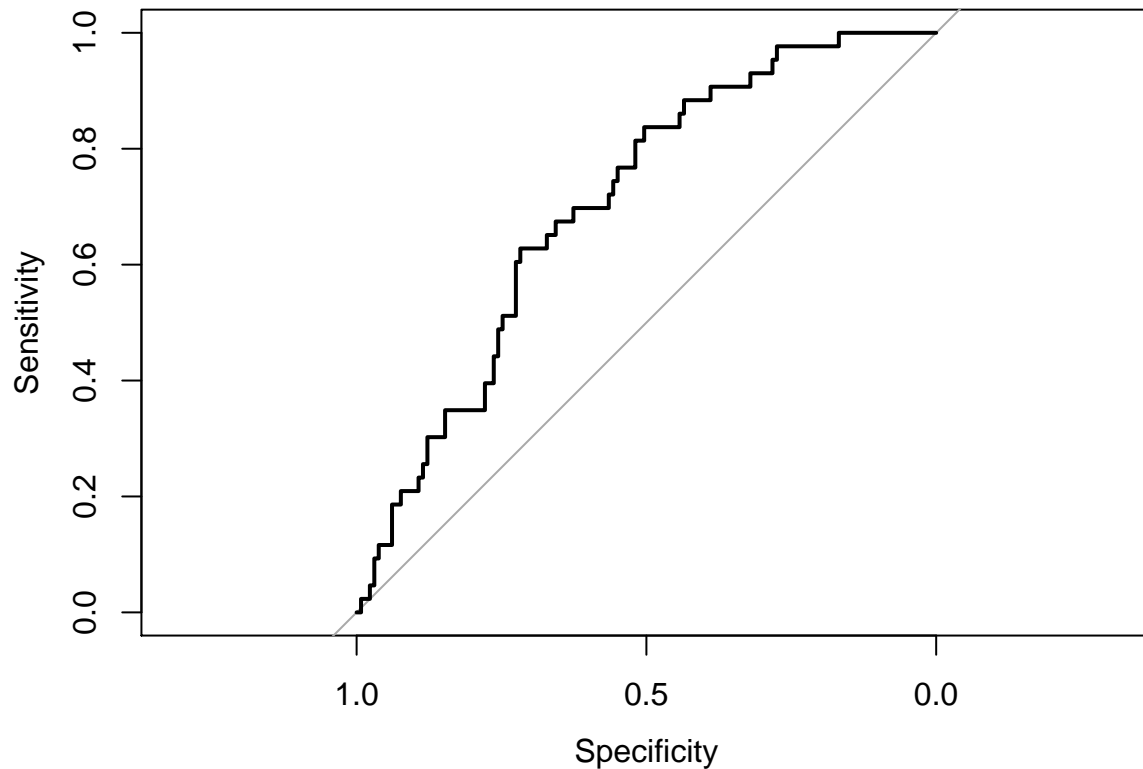
```

```

## (1 observation deleted due to missingness)
## AIC: 649.32
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 110  28
##           1  21  15
##
##           Accuracy : 0.7184
##           95% CI : (0.6453, 0.7838)
##           No Information Rate : 0.7529
##           P-Value [Acc > NIR] : 0.8726
##
##           Kappa : 0.1994
##
## Mcnemar's Test P-Value : 0.3914
##
##           Sensitivity : 0.8397
##           Specificity : 0.3488
##           Pos Pred Value : 0.7971
##           Neg Pred Value : 0.4167
##           Prevalence : 0.7529
##           Detection Rate : 0.6322
##           Detection Prevalence : 0.7931
##           Balanced Accuracy : 0.5943
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.704242854606781"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7042
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0820  -0.6559  -0.3291   0.5140   3.0585
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.197e+01  1.505e+01   1.460  0.14438
## KIDSDRIV      -1.585e+00  2.156e+00  -0.735  0.46225
## AGE           -1.473e-01  1.110e-01  -1.327  0.18466
## HOMEKIDS      -6.371e-01  7.456e-01  -0.855  0.39278
## YOJ           -2.828e-01  1.377e-01  -2.054  0.03998 *
## INCOME        -2.284e-05  1.348e-05  -1.694  0.09019 .
## HOME_VAL       1.411e-05  9.076e-06   1.555  0.12006
## TRAVTIME       4.278e-02  2.144e-02   1.995  0.04599 *
## BLUEBOOK       1.744e-05  4.118e-05   0.423  0.67202
```

```

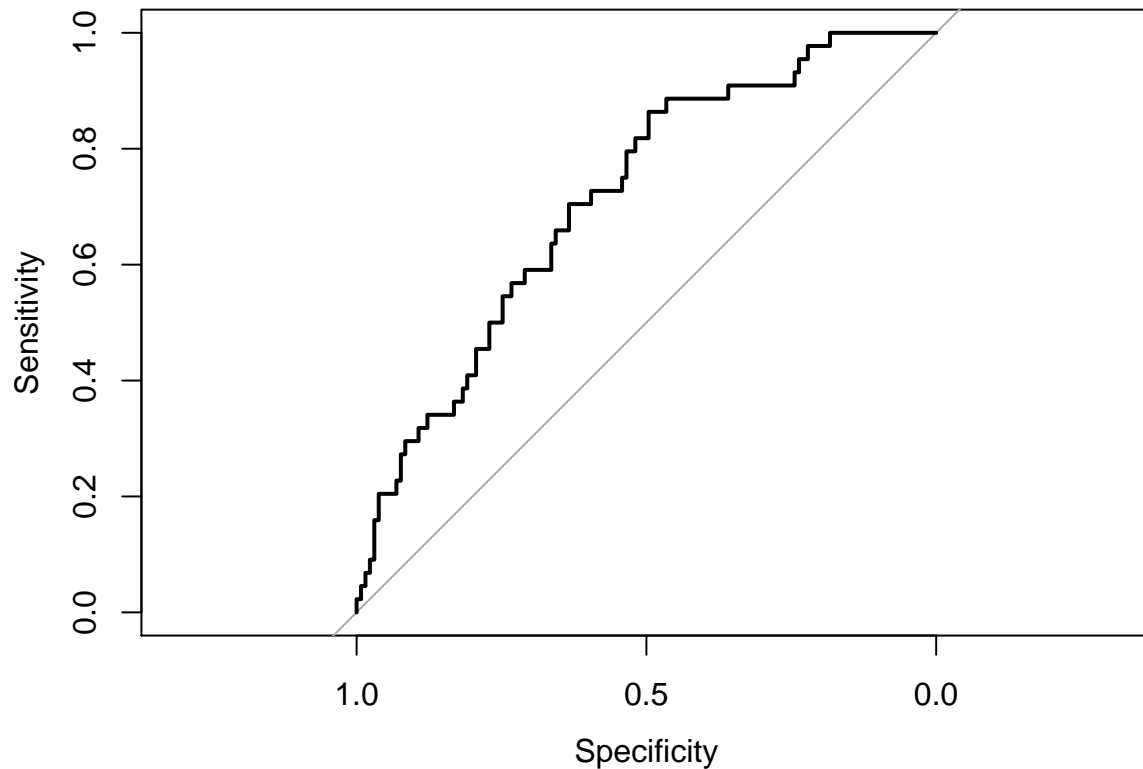
## TIF -1.215e-01 1.036e-01 -1.173 0.24094
## OLDCLAIM -7.815e-06 2.296e-05 -0.340 0.73359
## CLM_FREQ 5.709e-01 9.800e-01 0.583 0.56020
## MVR_PTS 6.940e-02 1.741e-01 0.399 0.69021
## CAR_AGE -1.288e-02 7.308e-02 -0.176 0.86004
## PARENT1_Yes 4.281e-01 4.742e-01 0.903 0.36658
## MSTATUS_Yes -6.165e-01 3.409e-01 -1.809 0.07052 .
## SEX_z_F -7.042e-01 4.617e-01 -1.525 0.12719
## EDUCATION_.High.School -1.122e-02 7.828e-01 -0.014 0.98857
## EDUCATION_Bachelors 1.625e-02 6.708e-01 0.024 0.98068
## EDUCATION_Masters 8.962e-01 5.760e-01 1.556 0.11975
## EDUCATION_z_High.School 3.826e-01 7.192e-01 0.532 0.59471
## JOB_ -1.072e+00 7.163e-01 -1.496 0.13470
## JOB_Clerical -1.502e-01 4.947e-01 -0.304 0.76134
## JOB_Doctor 4.637e-01 9.341e-01 0.496 0.61958
## JOB_Home.Maker -6.768e-01 7.031e-01 -0.963 0.33576
## JOB_Lawyer -9.879e-01 6.971e-01 -1.417 0.15647
## JOB_Manager -9.279e-01 4.974e-01 -1.866 0.06209 .
## JOB_Student -1.785e+00 7.626e-01 -2.341 0.01923 *
## JOB_z_Blue.Collar -1.840e-01 4.460e-01 -0.413 0.67987
## CAR_USE_Commercial 8.161e-01 3.502e-01 2.331 0.01977 *
## CAR_TYPE_Panel.Truck 3.356e-01 6.235e-01 0.538 0.59036
## CAR_TYPE_Pickup 1.299e+00 4.088e-01 3.177 0.00149 **
## CAR_TYPE_Sports.Car 1.549e+00 5.437e-01 2.850 0.00438 **
## CAR_TYPE_Van 5.603e-01 4.838e-01 1.158 0.24680
## CAR_TYPE_z_SUV 1.992e+00 4.866e-01 4.092 4.27e-05 ***
## RED_CAR_no -1.469e-01 3.423e-01 -0.429 0.66775
## REVOKED_Yes 9.875e-01 4.172e-01 2.367 0.01793 *
## URBANICITY_z_Highly.Rural..Rural -2.395e+00 4.120e-01 -5.813 6.14e-09 ***
## YOJ_NA 9.978e-02 4.732e-01 0.211 0.83302
## INCOME_NA 7.462e-01 5.951e-01 1.254 0.20991
## CAR_AGE_NA -3.110e-01 4.647e-01 -0.669 0.50328
## HOME_VAL_NA -2.324e-01 2.829e-01 -0.822 0.41135
## ageSquared 1.527e-03 1.188e-03 1.285 0.19890
## yojSquared 1.327e-02 6.930e-03 1.915 0.05552 .
## income_log -9.708e-02 2.959e-01 -0.328 0.74289
## homeval_log -1.546e+00 1.395e+00 -1.108 0.26773
## travtime_log -5.756e-01 5.963e-01 -0.965 0.33444
## bluebook_log -4.165e-03 5.071e-01 -0.008 0.99345
## carage_log -2.242e-01 4.689e-01 -0.478 0.63252
## oldclaim_log 1.254e-01 1.654e-01 0.758 0.44834
## clm_freq_log -1.552e+00 3.024e+00 -0.513 0.60773
## mvr_pts_log -1.750e-01 5.136e-01 -0.341 0.73336
## tif_log 2.172e-01 5.549e-01 0.391 0.69544
## kidsdriv_log -1.434e+00 2.617e+00 -0.548 0.58367
## homekids_log 1.467e+00 1.589e+00 0.923 0.35583
## inter 7.526e-02 4.775e-02 1.576 0.11500
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 743.59 on 640 degrees of freedom
## Residual deviance: 531.70 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 643.7
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  29
##           1  22  15
##
##           Accuracy : 0.7086
##           95% CI : (0.6352, 0.7747)
##           No Information Rate : 0.7486
##           P-Value [Acc > NIR] : 0.9028
##
##           Kappa : 0.1826
##
## Mcnemar's Test P-Value : 0.4008
##
##           Sensitivity : 0.8321
##           Specificity : 0.3409
##           Pos Pred Value : 0.7899
##           Neg Pred Value : 0.4054
##           Prevalence : 0.7486
##           Detection Rate : 0.6229
##           Detection Prevalence : 0.7886
##           Balanced Accuracy : 0.5865
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.714954892435808"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.715
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0598  -0.6625  -0.3730   0.5628   2.9228
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.522e+01  1.405e+01   1.083  0.27862
## KIDSDRIV      -2.324e+00  1.931e+00  -1.204  0.22872
## AGE           -1.479e-01  1.030e-01  -1.436  0.15104
## HOMEKIDS      -8.131e-01  6.956e-01  -1.169  0.24244
## YOJ           -2.727e-01  1.296e-01  -2.104  0.03541 *
## INCOME        -1.648e-06  1.370e-05  -0.120  0.90422
## HOME_VAL       2.979e-07  8.952e-06   0.033  0.97345
## TRAVTIME       2.097e-02  2.242e-02   0.935  0.34972
## BLUEBOOK       8.036e-05  3.740e-05   2.149  0.03167 *
```

```

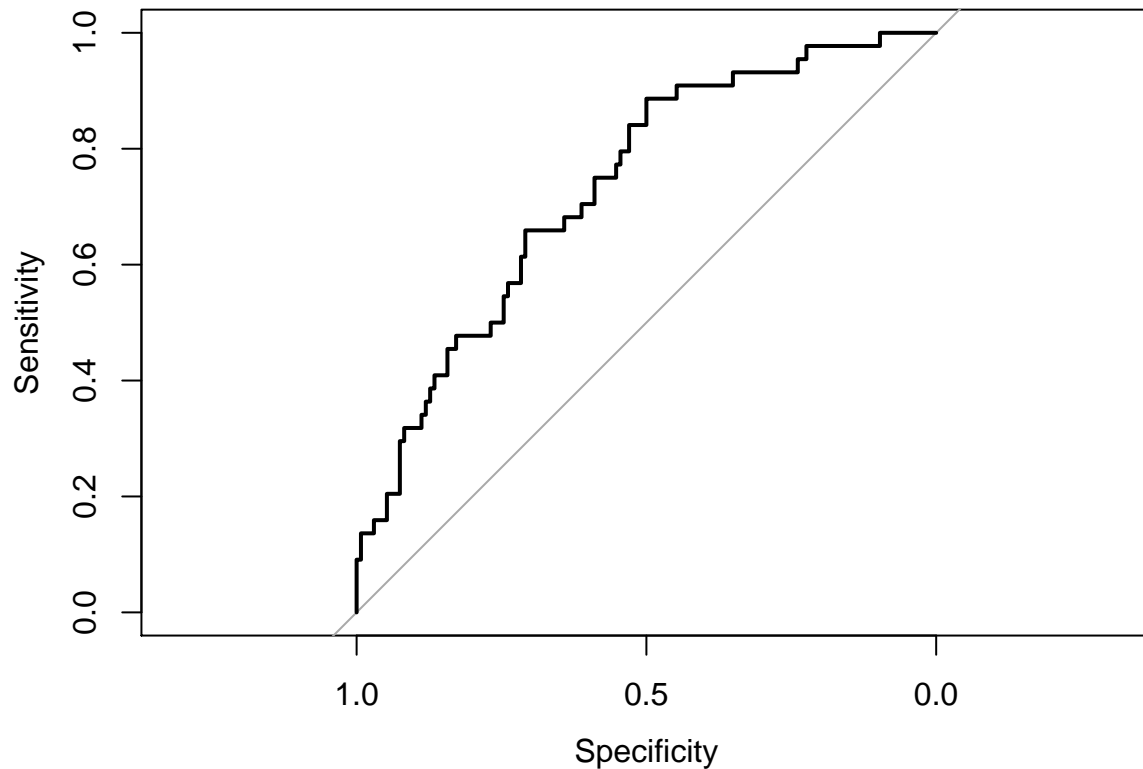
## TIF -1.156e-02 8.821e-02 -0.131 0.89574
## OLDCLAIM 3.884e-06 2.208e-05 0.176 0.86041
## CLM_FREQ -1.210e-01 1.025e+00 -0.118 0.90606
## MVR_PTS 7.056e-02 1.665e-01 0.424 0.67176
## CAR_AGE -3.149e-02 7.412e-02 -0.425 0.67089
## PARENT1_Yes 3.656e-01 4.744e-01 0.771 0.44090
## MSTATUS_Yes -6.135e-01 3.269e-01 -1.877 0.06056
## SEX_z_F -6.380e-01 4.677e-01 -1.364 0.17253
## EDUCATION_.High.School -4.252e-01 8.050e-01 -0.528 0.59740
## EDUCATION_Bachelors -3.383e-01 6.710e-01 -0.504 0.61416
## EDUCATION_Masters 5.653e-01 5.771e-01 0.979 0.32737
## EDUCATION_z_High.School 2.731e-01 7.273e-01 0.376 0.70726
## JOB_ -1.108e+00 6.839e-01 -1.619 0.10535
## JOB_Clerical -5.393e-01 4.921e-01 -1.096 0.27310
## JOB_Doctor 2.192e-01 9.208e-01 0.238 0.81184
## JOB_Home.Maker -6.092e-01 6.606e-01 -0.922 0.35647
## JOB_Lawyer -9.173e-01 6.526e-01 -1.406 0.15985
## JOB_Manager -1.073e+00 4.715e-01 -2.275 0.02288 *
## JOB_Student -1.118e+00 7.194e-01 -1.555 0.12003
## JOB_z_Blue.Collar -1.935e-01 4.686e-01 -0.413 0.67961
## CAR_USE_Commercial 5.503e-01 3.518e-01 1.565 0.11770
## CAR_TYPE_Panel.Truck -1.941e-01 6.256e-01 -0.310 0.75638
## CAR_TYPE_Pickup 1.073e+00 4.001e-01 2.681 0.00734 **
## CAR_TYPE_Sports.Car 1.654e+00 5.395e-01 3.066 0.00217 **
## CAR_TYPE_Van 1.035e-01 4.591e-01 0.225 0.82170
## CAR_TYPE_z_SUV 1.874e+00 4.684e-01 4.001 6.31e-05 ***
## RED_CAR_no -5.099e-01 3.507e-01 -1.454 0.14588
## REVOKED_Yes 2.722e-01 4.244e-01 0.641 0.52131
## URBANICITY_z_Highly.Rural..Rural -2.559e+00 4.315e-01 -5.931 3.00e-09 ***
## YOJ_NA -6.382e-01 4.133e-01 -1.544 0.12260
## INCOME_NA 3.330e-01 5.288e-01 0.630 0.52886
## CAR_AGE_NA -1.763e-01 4.592e-01 -0.384 0.70098
## HOME_VAL_NA -1.294e-01 2.857e-01 -0.453 0.65051
## ageSquared 1.534e-03 1.125e-03 1.364 0.17257
## yojSquared 1.331e-02 6.663e-03 1.997 0.04577 *
## income_log -9.799e-02 2.909e-01 -0.337 0.73621
## homeval_log -5.633e-01 1.340e+00 -0.420 0.67421
## travtime_log 2.137e-01 6.278e-01 0.340 0.73353
## bluebook_log -4.825e-01 4.499e-01 -1.073 0.28345
## carage_log -1.626e-01 4.690e-01 -0.347 0.72876
## oldclaim_log 5.083e-02 1.723e-01 0.295 0.76799
## clm_freq_log 4.006e-02 3.151e+00 0.013 0.98986
## mvr_pts_log -1.793e-01 4.908e-01 -0.365 0.71482
## tif_log -2.481e-01 5.064e-01 -0.490 0.62419
## kidsdriv_log 1.285e+00 2.208e+00 0.582 0.56044
## homekids_log 1.763e+00 1.502e+00 1.174 0.24048
## inter 5.080e-02 3.452e-02 1.472 0.14112
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 741.72 on 637 degrees of freedom
## Residual deviance: 546.82 on 582 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 658.82
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  27
##           1  18  17
##
##           Accuracy : 0.7472
##           95% CI : (0.6767, 0.8092)
##           No Information Rate : 0.7528
##           P-Value [Acc > NIR] : 0.608
##
##           Kappa : 0.2706
##
## Mcnemar's Test P-Value : 0.233
##
##           Sensitivity : 0.8657
##           Specificity : 0.3864
##           Pos Pred Value : 0.8112
##           Neg Pred Value : 0.4857
##           Prevalence : 0.7528
##           Detection Rate : 0.6517
##           Detection Prevalence : 0.8034
##           Balanced Accuracy : 0.6260
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.733887381275441"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7339
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9778  -0.6424  -0.3061   0.4908   2.5832
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.164e+01  1.491e+01   1.451  0.14664
## KIDSDRIV      -4.265e-01  1.748e+00  -0.244  0.80727
## AGE           -2.756e-02  1.182e-01  -0.233  0.81569
## HOMEKIDS      -1.022e+00  6.777e-01  -1.508  0.13155
## YOJ           -7.442e-02  1.380e-01  -0.539  0.58971
## INCOME        -1.901e-05  1.363e-05  -1.395  0.16295
## HOME_VAL       1.350e-05  9.039e-06   1.494  0.13528
## TRAVTIME       4.135e-02  2.301e-02   1.797  0.07233
## BLUEBOOK       5.608e-05  3.953e-05   1.419  0.15596
```

```

## TIF -4.306e-02 9.452e-02 -0.456 0.64869
## OLDCLAIM 3.492e-06 2.077e-05 0.168 0.86651
## CLM_FREQ 6.847e-01 9.343e-01 0.733 0.46365
## MVR_PTS 5.766e-02 1.786e-01 0.323 0.74680
## CAR_AGE -9.587e-02 7.201e-02 -1.331 0.18306
## PARENT1_Yes 3.998e-01 4.860e-01 0.823 0.41062
## MSTATUS_Yes -5.686e-01 3.461e-01 -1.643 0.10045
## SEX_z_F -7.718e-01 4.821e-01 -1.601 0.10942
## EDUCATION_.High.School -5.924e-02 7.893e-01 -0.075 0.94018
## EDUCATION_Bachelors -2.230e-01 6.629e-01 -0.336 0.73654
## EDUCATION_Masters 7.546e-01 5.510e-01 1.370 0.17083
## EDUCATION_z_High.School 3.038e-01 7.169e-01 0.424 0.67172
## JOB_ -5.810e-01 6.827e-01 -0.851 0.39476
## JOB_Clerical 9.017e-03 4.947e-01 0.018 0.98546
## JOB_Doctor -2.264e-01 9.871e-01 -0.229 0.81859
## JOB_Home.Maker -5.095e-01 7.378e-01 -0.691 0.48981
## JOB_Lawyer -2.389e-01 6.698e-01 -0.357 0.72139
## JOB_Manager -7.689e-01 5.067e-01 -1.517 0.12918
## JOB_Student -7.291e-02 7.802e-01 -0.093 0.92555
## JOB_z_Blue.Collar -1.353e-01 4.639e-01 -0.292 0.77057
## CAR_USE_Commercial 5.935e-01 3.503e-01 1.694 0.09017 .
## CAR_TYPE_Panel.Truck 7.375e-01 6.576e-01 1.122 0.26207
## CAR_TYPE_Pickup 1.399e+00 4.400e-01 3.181 0.00147 **
## CAR_TYPE_Sports.Car 2.552e+00 5.664e-01 4.505 6.64e-06 ***
## CAR_TYPE_Van 1.066e+00 4.881e-01 2.184 0.02899 *
## CAR_TYPE_z_SUV 2.093e+00 5.155e-01 4.061 4.88e-05 ***
## RED_CAR_no -8.316e-02 3.396e-01 -0.245 0.80656
## REVOKED_Yes 2.124e-01 4.195e-01 0.506 0.61268
## URBANICITY_z_Highly.Rural..Rural -2.502e+00 4.591e-01 -5.449 5.07e-08 ***
## YOJ_NA -2.476e-01 4.308e-01 -0.575 0.56545
## INCOME_NA 6.236e-02 6.063e-01 0.103 0.91808
## CAR_AGE_NA 9.898e-02 4.998e-01 0.198 0.84301
## HOME_VAL_NA -2.452e-01 2.951e-01 -0.831 0.40606
## ageSquared 7.474e-05 1.295e-03 0.058 0.95399
## yojSquared 4.122e-03 7.158e-03 0.576 0.56474
## income_log -8.625e-02 2.965e-01 -0.291 0.77111
## homeval_log -1.426e+00 1.407e+00 -1.013 0.31098
## travtime_log -6.012e-01 6.390e-01 -0.941 0.34677
## bluebook_log -5.692e-01 4.786e-01 -1.189 0.23429
## carage_log 2.791e-01 4.732e-01 0.590 0.55538
## oldclaim_log 1.749e-01 1.590e-01 1.100 0.27123
## clm_freq_log -1.944e+00 2.915e+00 -0.667 0.50481
## mvr_pts_log -9.188e-02 5.229e-01 -0.176 0.86051
## tif_log -1.391e-01 5.334e-01 -0.261 0.79424
## kidsdriv_log -1.403e-01 2.271e+00 -0.062 0.95072
## homekids_log 2.329e+00 1.470e+00 1.584 0.11318
## inter 2.459e-02 3.067e-02 0.802 0.42267
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 730.57 on 639 degrees of freedom
## Residual deviance: 522.07 on 584 degrees of freedom

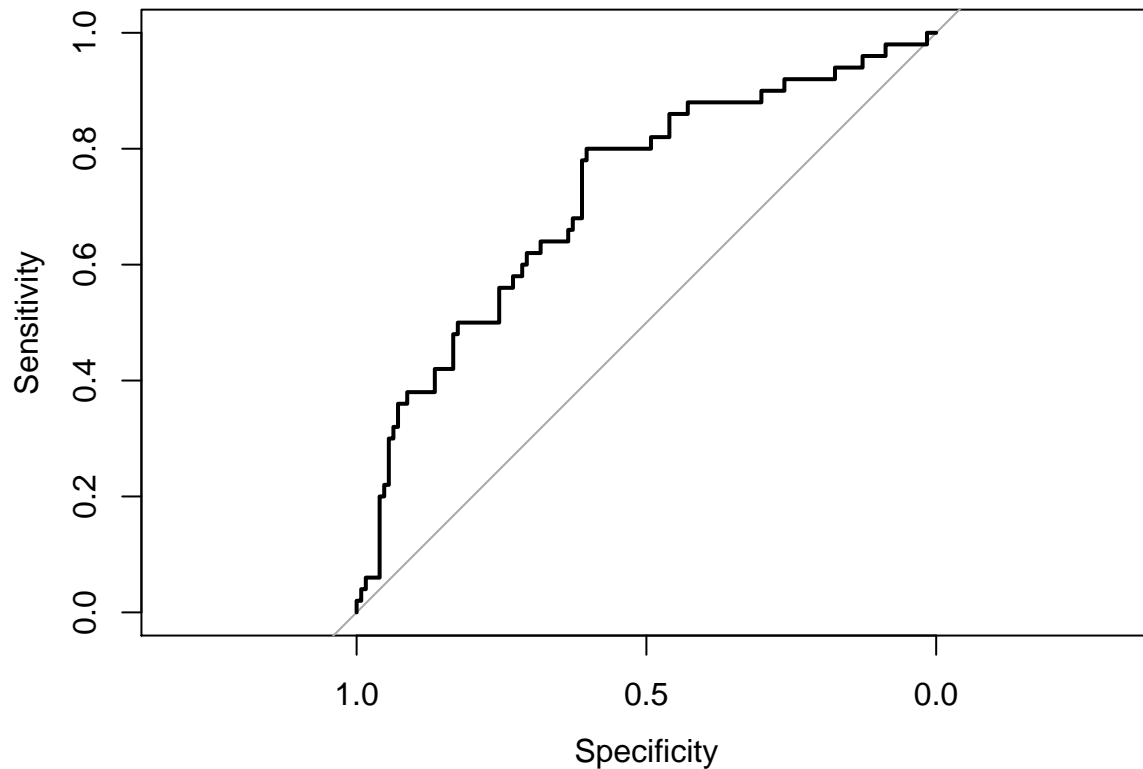
```



```

## AIC: 634.07
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 113  31
##           1  13  19
##
##           Accuracy : 0.75
##           95% CI : (0.6793, 0.8121)
##           No Information Rate : 0.7159
##           P-Value [Acc > NIR] : 0.17945
##
##           Kappa : 0.3105
##
## Mcnemar's Test P-Value : 0.01038
##
##           Sensitivity : 0.8968
##           Specificity : 0.3800
##           Pos Pred Value : 0.7847
##           Neg Pred Value : 0.5938
##           Prevalence : 0.7159
##           Detection Rate : 0.6420
##           Detection Prevalence : 0.8182
##           Balanced Accuracy : 0.6384
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.72031746031746"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 126 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7203
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0279  -0.6700  -0.3421   0.3471   3.2173
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.797e+01  1.442e+01   1.940 0.052346 .
## KIDSDRIV     -1.786e+00  1.704e+00  -1.048 0.294684
## AGE          -1.799e-01  1.118e-01  -1.609 0.107590
## HOMEKIDS     -1.207e+00  6.844e-01  -1.763 0.077929 .
## YOJ          -1.724e-01  1.356e-01  -1.272 0.203450
## INCOME       -1.067e-05  1.371e-05  -0.778 0.436464
## HOME_VAL      9.134e-06  8.930e-06   1.023 0.306365
## TRAVTIME      3.045e-02  2.258e-02   1.349 0.177378
## BLUEBOOK      7.405e-05  3.799e-05   1.949 0.051312 .
```

```

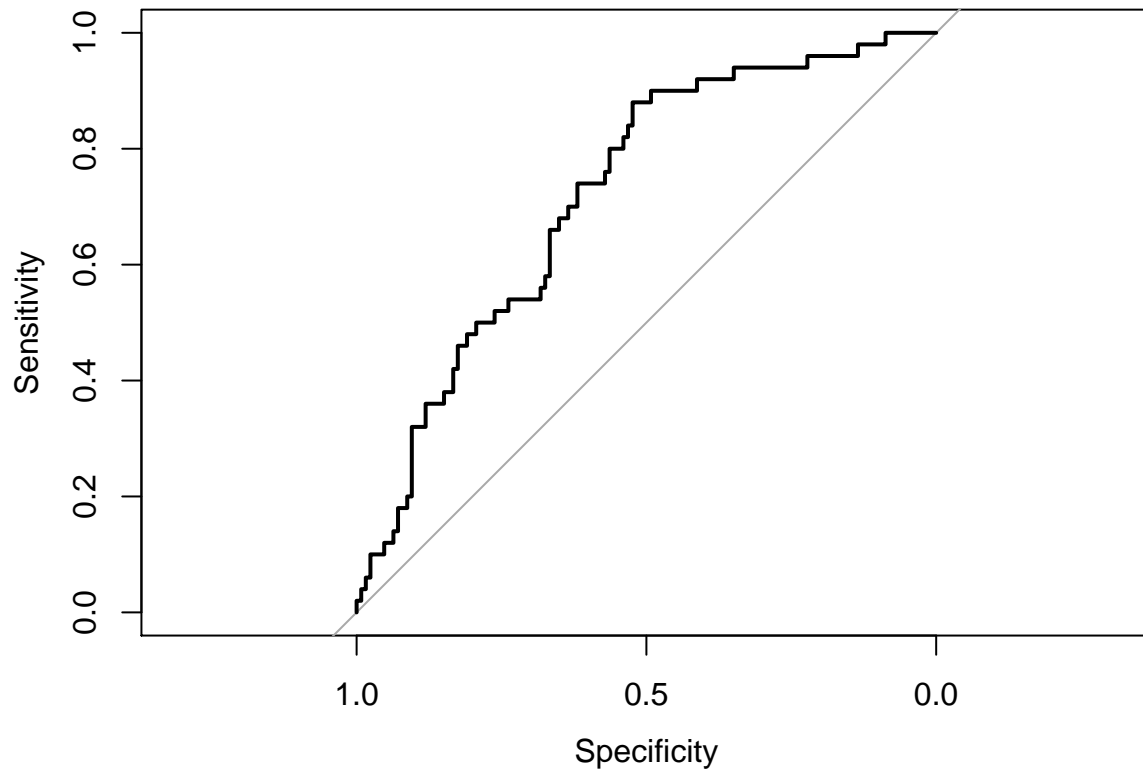
## TIF -5.512e-02 9.517e-02 -0.579 0.562452
## OLDCLAIM 5.738e-06 2.178e-05 0.263 0.792169
## CLM_FREQ 4.155e-01 1.009e+00 0.412 0.680540
## MVRPTS 6.332e-02 1.812e-01 0.349 0.726733
## CAR_AGE -3.895e-02 7.447e-02 -0.523 0.600924
## PARENT1_Yes 6.353e-01 4.822e-01 1.317 0.187690
## MSTATUS_Yes -5.874e-01 3.374e-01 -1.741 0.081731 .
## SEX_z_F -1.225e+00 4.896e-01 -2.502 0.012357 *
## EDUCATION_.High.School 1.155e-02 8.116e-01 0.014 0.988645
## EDUCATION_Bachelors -9.879e-02 6.842e-01 -0.144 0.885202
## EDUCATION_Masters 7.238e-01 5.821e-01 1.243 0.213727
## EDUCATION_z_High.School 3.627e-01 7.420e-01 0.489 0.624941
## JOB_ -7.099e-01 7.466e-01 -0.951 0.341688
## JOB_Clerical -9.178e-02 5.087e-01 -0.180 0.856823
## JOB_Doctor 3.059e-01 9.653e-01 0.317 0.751292
## JOB_Home.Maker -8.148e-01 7.405e-01 -1.100 0.271179
## JOB_Lawyer -3.023e-01 6.880e-01 -0.439 0.660424
## JOB_Manager -9.275e-01 5.128e-01 -1.809 0.070478 .
## JOB_Student -8.191e-01 7.334e-01 -1.117 0.264083
## JOB_z_Blue.Collar -8.446e-02 4.591e-01 -0.184 0.854034
## CAR_USE_Commercial 7.180e-01 3.595e-01 1.997 0.045778 *
## CAR_TYPE_Panel.Truck -2.249e-01 6.313e-01 -0.356 0.721683
## CAR_TYPE_Pickup 1.125e+00 4.186e-01 2.687 0.007201 **
## CAR_TYPE_Sports.Car 2.299e+00 5.690e-01 4.041 5.33e-05 ***
## CAR_TYPE_Van 5.410e-01 4.810e-01 1.125 0.260648
## CAR_TYPE_z_SUV 1.955e+00 5.089e-01 3.841 0.000122 ***
## RED_CAR_no -1.509e-01 3.398e-01 -0.444 0.656930
## REVOKED_Yes 3.858e-01 4.319e-01 0.893 0.371726
## URBANICITY_z_Highly.Rural..Rural -2.625e+00 4.355e-01 -6.028 1.66e-09 ***
## YOJ_NA -7.803e-01 4.319e-01 -1.806 0.070854 .
## INCOME_NA 2.879e-01 5.701e-01 0.505 0.613506
## CAR_AGE_NA -1.842e-01 4.849e-01 -0.380 0.704023
## HOME_VAL_NA -2.953e-01 2.843e-01 -1.039 0.298925
## ageSquared 1.831e-03 1.205e-03 1.520 0.128589
## yojSquared 7.583e-03 6.981e-03 1.086 0.277358
## income_log -4.130e-02 3.115e-01 -0.133 0.894506
## homeval_log -1.592e+00 1.373e+00 -1.159 0.246257
## travtime_log -2.210e-01 6.295e-01 -0.351 0.725522
## bluebook_log -6.420e-01 4.527e-01 -1.418 0.156184
## carage_log -7.047e-03 4.777e-01 -0.015 0.988229
## oldclaim_log 8.094e-02 1.672e-01 0.484 0.628402
## clm_freq_log -1.351e+00 3.108e+00 -0.435 0.663886
## mvr_pts_log -3.510e-03 5.238e-01 -0.007 0.994654
## tif_log 1.007e-01 5.370e-01 0.187 0.851308
## kidsdriv_log 1.418e+00 2.146e+00 0.661 0.508828
## homekids_log 2.587e+00 1.493e+00 1.732 0.083245 .
## inter 4.067e-02 3.140e-02 1.295 0.195152
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 730.57 on 639 degrees of freedom
## Residual deviance: 522.14 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 634.14
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  36
##           1  12  14
##
##           Accuracy : 0.7273
##           95% CI : (0.6552, 0.7916)
##           No Information Rate : 0.7159
##           P-Value [Acc > NIR] : 0.4054532
##
##           Kappa : 0.216
##
## Mcnemar's Test P-Value : 0.0009009
##
##           Sensitivity : 0.9048
##           Specificity : 0.2800
##           Pos Pred Value : 0.7600
##           Neg Pred Value : 0.5385
##           Prevalence : 0.7159
##           Detection Rate : 0.6477
##           Detection Prevalence : 0.8523
##           Balanced Accuracy : 0.5924
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC:  0.722222222222222"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 126 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7222
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0990  -0.6739  -0.3548   0.5027   3.0441
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.738e+01  1.512e+01   1.149  0.250402
## KIDSDRIV       1.134e+00  1.763e+00   0.643  0.520151
## AGE          -1.879e-01  1.011e-01  -1.859  0.063050 .
## HOMEKIDS     -1.140e+00  6.891e-01  -1.655  0.098003 .
## YOJ          -1.587e-01  1.260e-01  -1.259  0.208040
## INCOME       -2.454e-05  1.308e-05  -1.875  0.060747 .
## HOME_VAL      1.357e-05  8.928e-06   1.520  0.128606
## TRAVTIME      3.254e-02  2.164e-02   1.504  0.132663
## BLUEBOOK      6.543e-05  4.196e-05   1.559  0.118910
```

```

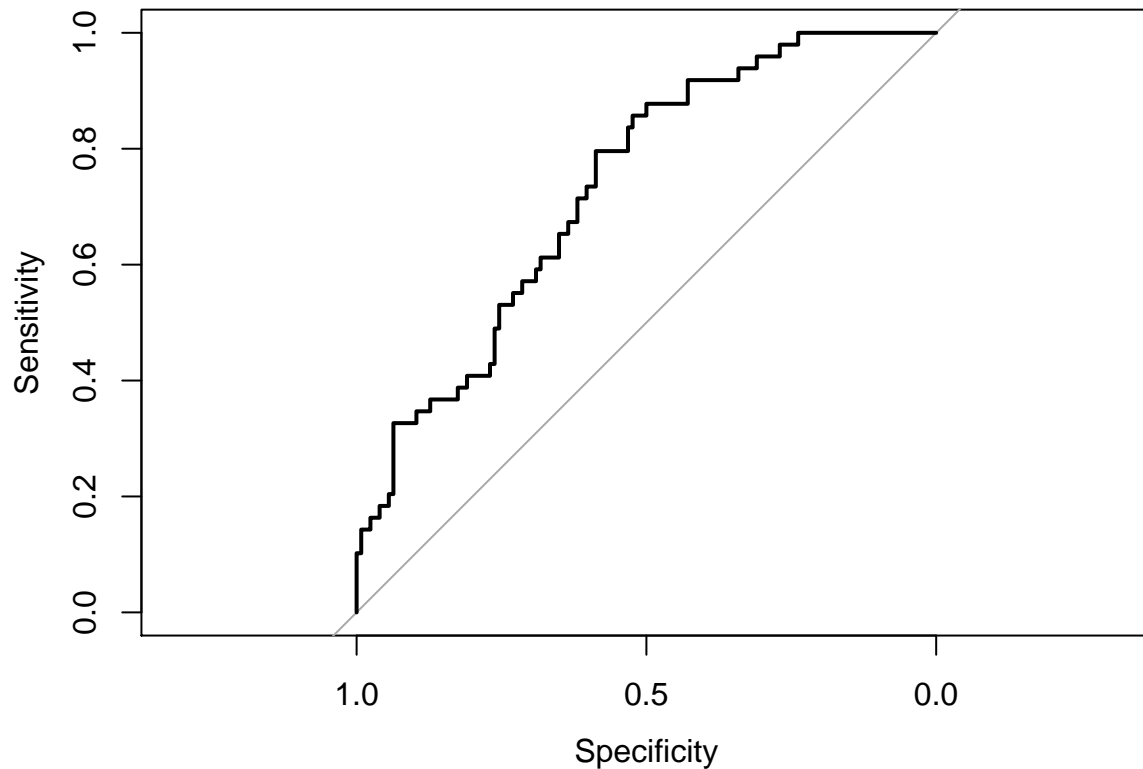
## TIF -4.312e-02 8.860e-02 -0.487 0.626503
## OLDCLAIM 4.441e-05 2.206e-05 2.013 0.044108 *
## CLM_FREQ -1.525e+00 1.081e+00 -1.411 0.158112
## MVRPTS 1.258e-01 1.731e-01 0.726 0.467624
## CAR_AGE -3.857e-02 7.166e-02 -0.538 0.590426
## PARENT1_Yes 7.386e-01 4.682e-01 1.578 0.114670
## MSTATUS_Yes -7.052e-01 3.343e-01 -2.110 0.034894 *
## SEX_z_F -1.215e+00 4.931e-01 -2.464 0.013758 *
## EDUCATION_.High.School 3.007e-01 8.520e-01 0.353 0.724170
## EDUCATION_Bachelors -2.016e-02 7.336e-01 -0.027 0.978083
## EDUCATION_Masters 1.004e+00 6.360e-01 1.578 0.114472
## EDUCATION_z_High.School 4.926e-01 7.817e-01 0.630 0.528554
## JOB_ -7.146e-01 7.129e-01 -1.002 0.316186
## JOB_Clerical -4.225e-02 4.982e-01 -0.085 0.932412
## JOB_Doctor -9.575e-02 1.024e+00 -0.094 0.925499
## JOB_Home.Maker -5.388e-02 6.619e-01 -0.081 0.935130
## JOB_Lawyer -2.475e-01 6.878e-01 -0.360 0.718920
## JOB_Manager -5.504e-01 4.843e-01 -1.137 0.255734
## JOB_Student -5.191e-01 7.173e-01 -0.724 0.469273
## JOB_z_Blue.Collar 1.821e-01 4.653e-01 0.391 0.695589
## CAR_USE_Commercial 2.577e-01 3.457e-01 0.746 0.455882
## CAR_TYPE_Panel.Truck 1.831e-01 6.214e-01 0.295 0.768284
## CAR_TYPE_Pickup 1.510e+00 4.151e-01 3.637 0.000275 ***
## CAR_TYPE_Sports.Car 2.285e+00 5.603e-01 4.079 4.53e-05 ***
## CAR_TYPE_Van 8.888e-01 4.831e-01 1.840 0.065795 .
## CAR_TYPE_z_SUV 2.386e+00 5.180e-01 4.607 4.08e-06 ***
## RED_CAR_no -1.140e-01 3.292e-01 -0.346 0.729126
## REVOKED_Yes -3.707e-01 4.407e-01 -0.841 0.400235
## URBANICITY_z_Highly.Rural..Rural -2.401e+00 4.112e-01 -5.840 5.24e-09 ***
## YOJ_NA -2.369e-01 4.443e-01 -0.533 0.593812
## INCOME_NA 1.210e-01 5.373e-01 0.225 0.821814
## CAR_AGE_NA 1.531e-01 5.083e-01 0.301 0.763274
## HOME_VAL_NA -1.154e-01 2.899e-01 -0.398 0.690570
## ageSquared 2.058e-03 1.098e-03 1.873 0.061001 .
## yojSquared 8.383e-03 6.553e-03 1.279 0.200797
## income_log 6.267e-03 2.819e-01 0.022 0.982262
## homeval_log -1.073e+00 1.409e+00 -0.762 0.446185
## travtime_log -4.056e-01 6.107e-01 -0.664 0.506603
## bluebook_log -3.199e-01 5.140e-01 -0.622 0.533738
## carage_log -6.593e-02 4.591e-01 -0.144 0.885807
## oldclaim_log -1.860e-01 1.723e-01 -1.080 0.280352
## clm_freq_log 4.509e+00 3.289e+00 1.371 0.170329
## mvr_pts_log -3.132e-01 5.065e-01 -0.618 0.536365
## tif_log 5.566e-02 5.075e-01 0.110 0.912669
## kidsdriv_log -1.238e+00 2.308e+00 -0.536 0.591876
## homekids_log 2.425e+00 1.468e+00 1.651 0.098712 .
## inter 5.932e-03 3.085e-02 0.192 0.847514
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.27 on 640 degrees of freedom
## Residual deviance: 542.69 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 654.69
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 107  31
##           1  19  18
##
##           Accuracy : 0.7143
##           95% CI : (0.6412, 0.7799)
##           No Information Rate : 0.72
##           P-Value [Acc > NIR] : 0.6041
##
##           Kappa : 0.2341
##
## Mcnemar's Test P-Value : 0.1198
##
##           Sensitivity : 0.8492
##           Specificity : 0.3673
##           Pos Pred Value : 0.7754
##           Neg Pred Value : 0.4865
##           Prevalence : 0.7200
##           Detection Rate : 0.6114
##           Detection Prevalence : 0.7886
##           Balanced Accuracy : 0.6083
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.732912212504049"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 126 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.7329
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3584  -0.6626  -0.3493   0.4886   2.9724
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.187e+01  1.480e+01   1.478 0.139477
## KIDSDRIV      -7.244e-02  1.881e+00  -0.039 0.969284
## AGE           -1.476e-01  1.045e-01  -1.412 0.157969
## HOMEKIDS      -1.324e+00  7.080e-01  -1.870 0.061541 .
## YOJ           -2.549e-01  1.369e-01  -1.862 0.062548 .
## INCOME        -2.726e-05  1.330e-05  -2.050 0.040364 *
## HOME_VAL       1.582e-05  8.932e-06   1.771 0.076589 .
## TRAVTIME       2.395e-02  2.167e-02   1.106 0.268938
## BLUEBOOK       7.969e-05  3.951e-05   2.017 0.043705 *
```



```

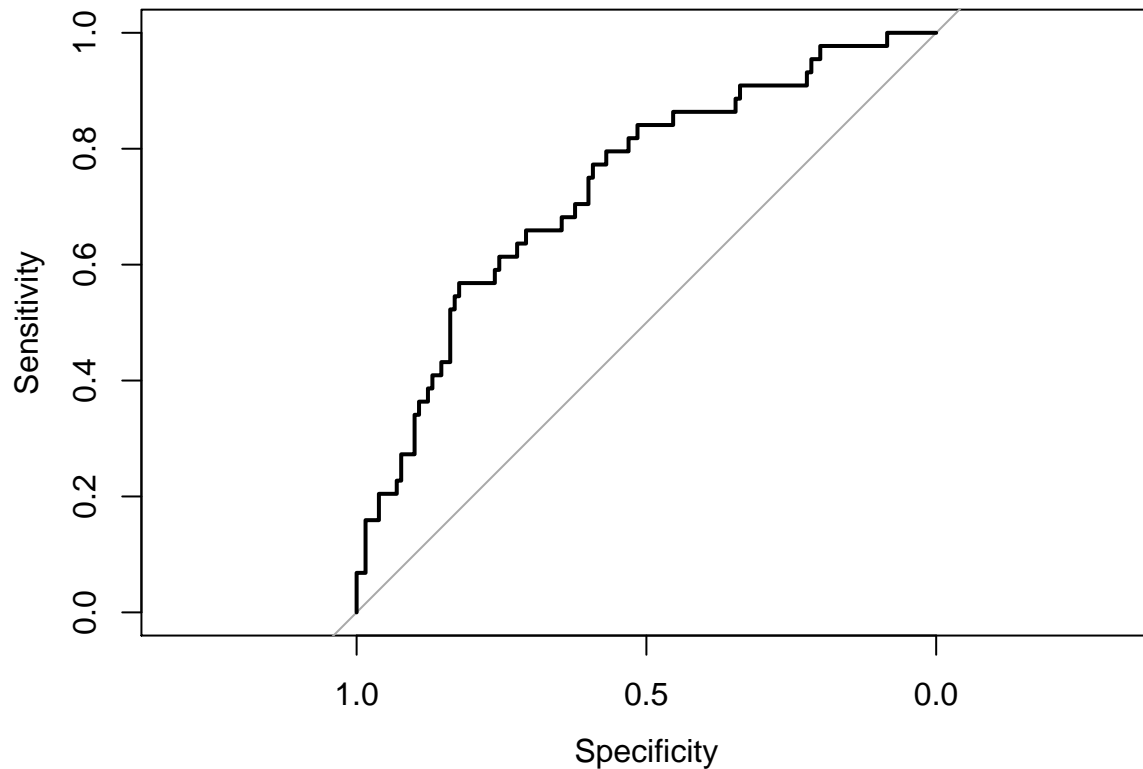
## TIF -1.244e-02 9.027e-02 -0.138 0.890359
## OLDCLAIM -1.209e-05 2.157e-05 -0.561 0.575001
## CLM_FREQ 8.840e-02 9.553e-01 0.093 0.926270
## MVRPTS 5.033e-02 1.666e-01 0.302 0.762586
## CAR_AGE -6.602e-03 6.871e-02 -0.096 0.923452
## PARENT1_Yes 7.156e-01 4.796e-01 1.492 0.135695
## MSTATUS_Yes -6.028e-01 3.297e-01 -1.828 0.067476 .
## SEX_z_F -1.054e+00 4.711e-01 -2.238 0.025217 *
## EDUCATION_.High.School 1.001e+00 8.515e-01 1.176 0.239551
## EDUCATION_Bachelors 9.588e-01 7.432e-01 1.290 0.197048
## EDUCATION_Masters 1.293e+00 6.321e-01 2.045 0.040850 *
## EDUCATION_z_High.School 8.653e-01 7.989e-01 1.083 0.278784
## JOB_ -3.373e-01 7.077e-01 -0.477 0.633647
## JOB_Clerical -2.524e-01 4.818e-01 -0.524 0.600388
## JOB_Doctor 8.564e-01 9.460e-01 0.905 0.365361
## JOB_Home.Maker -7.877e-01 6.853e-01 -1.150 0.250346
## JOB_Lawyer -4.910e-01 6.748e-01 -0.728 0.466871
## JOB_Manager -7.419e-01 4.876e-01 -1.521 0.128136
## JOB_Student -8.039e-01 7.288e-01 -1.103 0.270040
## JOB_z_Blue.Collar -3.382e-01 4.554e-01 -0.743 0.457707
## CAR_USE_Commercial 6.177e-01 3.578e-01 1.727 0.084254 .
## CAR_TYPE_Panel.Truck -4.493e-02 6.056e-01 -0.074 0.940850
## CAR_TYPE_Pickup 7.909e-01 4.124e-01 1.918 0.055155 .
## CAR_TYPE_Sports.Car 1.933e+00 5.384e-01 3.590 0.000331 ***
## CAR_TYPE_Van 5.577e-01 4.757e-01 1.172 0.241008
## CAR_TYPE_z_SUV 1.865e+00 4.880e-01 3.821 0.000133 ***
## RED_CAR_no 7.485e-03 3.312e-01 0.023 0.981969
## REVOKED_Yes 6.236e-01 4.114e-01 1.516 0.129606
## URBANICITY_z_Highly.Rural..Rural -2.474e+00 4.249e-01 -5.823 5.78e-09 ***
## YOJ_NA -5.863e-01 4.483e-01 -1.308 0.190927
## INCOME_NA 4.617e-01 6.061e-01 0.762 0.446220
## CAR_AGE_NA -3.217e-01 5.328e-01 -0.604 0.546039
## HOME_VAL_NA -1.144e-01 2.871e-01 -0.399 0.690226
## ageSquared 1.583e-03 1.123e-03 1.409 0.158797
## yojSquared 1.160e-02 6.942e-03 1.671 0.094769 .
## income_log -1.167e-01 3.191e-01 -0.366 0.714701
## homeval_log -1.258e+00 1.418e+00 -0.887 0.374999
## travtime_log -1.984e-01 6.089e-01 -0.326 0.744570
## bluebook_log -4.936e-01 4.542e-01 -1.087 0.277157
## carage_log -3.917e-01 4.540e-01 -0.863 0.388241
## oldclaim_log 5.574e-02 1.601e-01 0.348 0.727664
## clm_freq_log -1.467e-01 2.962e+00 -0.050 0.960512
## mvr_pts_log 1.589e-02 4.930e-01 0.032 0.974287
## tif_log -8.062e-02 5.144e-01 -0.157 0.875471
## kidsdriv_log -5.828e-01 2.368e+00 -0.246 0.805601
## homekids_log 2.846e+00 1.521e+00 1.871 0.061307 .
## inter 2.459e-02 3.396e-02 0.724 0.469069
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 744.21 on 641 degrees of freedom
## Residual deviance: 546.87 on 586 degrees of freedom

```

```

## AIC: 658.87
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 113  27
##           1  17  17
##
##           Accuracy : 0.7471
##           95% CI : (0.6758, 0.8099)
##           No Information Rate : 0.7471
##           P-Value [Acc > NIR] : 0.5404
##
##           Kappa : 0.2764
##
## Mcnemar's Test P-Value : 0.1748
##
##           Sensitivity : 0.8692
##           Specificity : 0.3864
##           Pos Pred Value : 0.8071
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7471
##           Detection Rate : 0.6494
##           Detection Prevalence : 0.8046
##           Balanced Accuracy : 0.6278
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.734615384615385"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 130 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7346
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2261  -0.6803  -0.3594   0.4859   3.2436
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.125e+01  1.468e+01   1.448  0.1476
## KIDSDRIV      -9.071e-01  1.741e+00  -0.521  0.6024
## AGE           -7.031e-02  1.149e-01  -0.612  0.5404
## HOMEKIDS      -9.759e-01  6.876e-01  -1.419  0.1559
## YOJ           -8.548e-02  1.306e-01  -0.654  0.5129
## INCOME        -7.305e-06  1.273e-05  -0.574  0.5662
## HOME_VAL       7.478e-06  8.704e-06   0.859  0.3903
## TRAVTIME       4.036e-02  2.114e-02   1.909  0.0562
## BLUEBOOK       6.164e-05  3.798e-05   1.623  0.1046
```

```

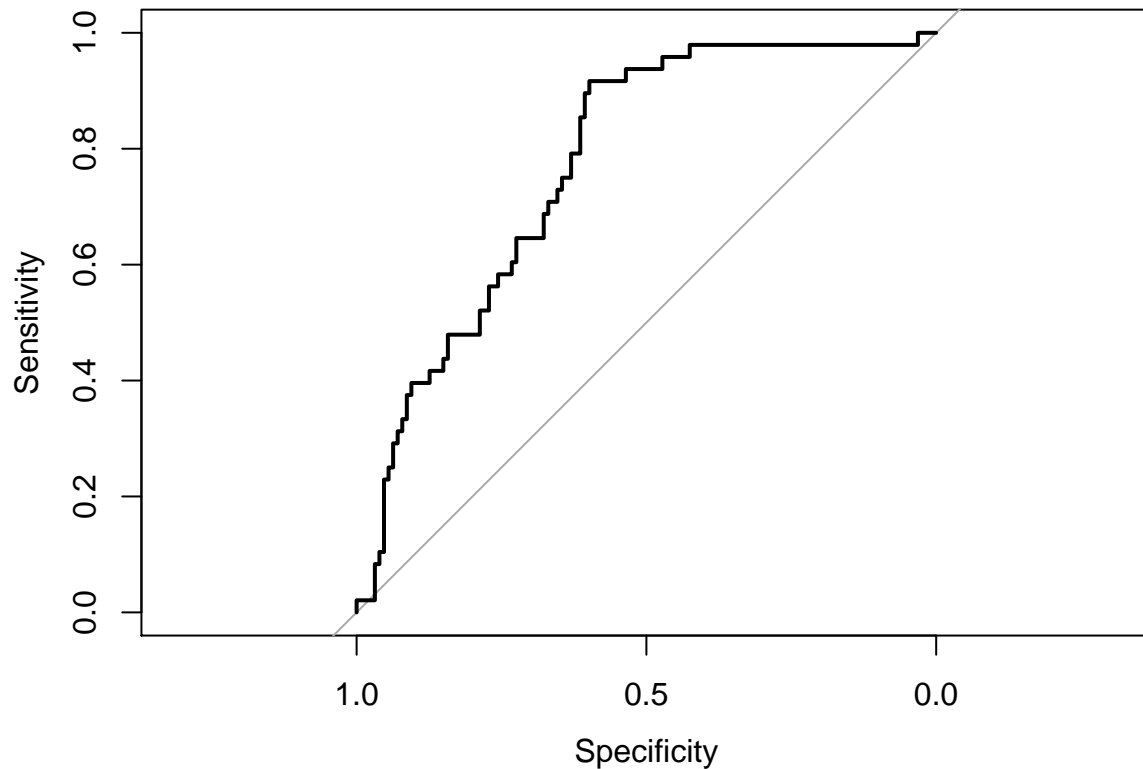
## TIF -4.795e-02 8.952e-02 -0.536 0.5922
## OLDCLAIM 6.867e-06 2.148e-05 0.320 0.7491
## CLM_FREQ 5.488e-01 9.495e-01 0.578 0.5633
## MVR_PTS 5.981e-02 1.744e-01 0.343 0.7317
## CAR_AGE -6.875e-02 6.903e-02 -0.996 0.3193
## PARENT1_Yes 5.763e-01 4.800e-01 1.201 0.2299
## MSTATUS_Yes -7.087e-01 3.295e-01 -2.151 0.0315 *
## SEX_z_F -1.126e+00 4.719e-01 -2.387 0.0170 *
## EDUCATION_.High.School 9.480e-02 7.850e-01 0.121 0.9039
## EDUCATION_Bachelors -1.553e-01 6.415e-01 -0.242 0.8087
## EDUCATION_Masters 6.741e-01 5.406e-01 1.247 0.2124
## EDUCATION_z_High.School 5.558e-01 6.930e-01 0.802 0.4226
## JOB_ -5.093e-01 6.760e-01 -0.753 0.4512
## JOB_Clerical 8.068e-02 4.978e-01 0.162 0.8712
## JOB_Doctor -3.468e-02 8.799e-01 -0.039 0.9686
## JOB_Home.Maker 8.370e-02 6.958e-01 0.120 0.9042
## JOB_Lawyer -5.261e-01 6.643e-01 -0.792 0.4283
## JOB_Manager -6.982e-01 5.008e-01 -1.394 0.1633
## JOB_Student -4.199e-01 7.364e-01 -0.570 0.5685
## JOB_z_Blue.Collar 1.193e-01 4.714e-01 0.253 0.8003
## CAR_USE_Commercial 4.796e-01 3.564e-01 1.346 0.1784
## CAR_TYPE_Panel.Truck -1.364e-01 6.122e-01 -0.223 0.8236
## CAR_TYPE_Pickup 1.028e+00 3.982e-01 2.583 0.0098 **
## CAR_TYPE_Sports.Car 2.249e+00 5.594e-01 4.021 5.80e-05 ***
## CAR_TYPE_Van 5.903e-01 4.519e-01 1.306 0.1914
## CAR_TYPE_z_SUV 1.997e+00 5.070e-01 3.939 8.18e-05 ***
## RED_CAR_no -1.083e-01 3.181e-01 -0.340 0.7336
## REVOKED_Yes 1.574e-01 4.306e-01 0.366 0.7146
## URBANICITY_z_Highly.Rural..Rural -2.545e+00 4.435e-01 -5.739 9.54e-09 ***
## YOJ_NA -4.125e-01 4.014e-01 -1.028 0.3041
## INCOME_NA 1.487e-01 5.094e-01 0.292 0.7704
## CAR_AGE_NA -1.673e-01 4.691e-01 -0.357 0.7213
## HOME_VAL_NA -1.620e-01 2.801e-01 -0.578 0.5631
## ageSquared 6.919e-04 1.264e-03 0.547 0.5842
## yojSquared 4.888e-03 6.761e-03 0.723 0.4697
## income_log 3.249e-02 2.795e-01 0.116 0.9075
## homeval_log -1.423e+00 1.402e+00 -1.015 0.3101
## travtime_log -5.329e-01 5.992e-01 -0.889 0.3738
## bluebook_log -5.511e-01 4.585e-01 -1.202 0.2294
## carage_log 3.778e-01 4.615e-01 0.819 0.4131
## oldclaim_log 1.303e-01 1.565e-01 0.833 0.4051
## clm_freq_log -1.622e+00 2.916e+00 -0.556 0.5779
## mvr_pts_log -8.777e-03 5.096e-01 -0.017 0.9863
## tif_log 4.695e-02 5.190e-01 0.090 0.9279
## kidsdriv_log 1.008e+00 2.172e+00 0.464 0.6427
## homekids_log 2.000e+00 1.483e+00 1.349 0.1774
## inter 2.355e-02 3.322e-02 0.709 0.4783
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 735.37 on 640 degrees of freedom
## Residual deviance: 541.52 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 653.52
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 113  29
##           1  14  19
##
##           Accuracy : 0.7543
##           95% CI : (0.6836, 0.8161)
##           No Information Rate : 0.7257
##           P-Value [Acc > NIR] : 0.22453
##
##           Kappa : 0.3163
##
## Mcnemar's Test P-Value : 0.03276
##
##           Sensitivity : 0.8898
##           Specificity : 0.3958
##           Pos Pred Value : 0.7958
##           Neg Pred Value : 0.5758
##           Prevalence : 0.7257
##           Detection Rate : 0.6457
##           Detection Prevalence : 0.8114
##           Balanced Accuracy : 0.6428
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.773950131233596"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 127 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.774
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5210  -0.6595  -0.3506   0.4717   3.0584
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.360e+01  1.499e+01   0.907  0.36418
## KIDSDRIV      -4.322e+00  2.107e+00  -2.051  0.04023 *
## AGE           -1.282e-01  1.099e-01  -1.167  0.24331
## HOMEKIDS       2.434e-01  6.739e-01   0.361  0.71793
## YOJ            -1.386e-01  1.347e-01  -1.029  0.30370
## INCOME         -1.847e-05  1.360e-05  -1.358  0.17440
## HOME_VAL       9.772e-06  9.209e-06   1.061  0.28862
## TRAVTIME       2.180e-02  2.226e-02   0.979  0.32744
## BLUEBOOK       5.571e-05  3.843e-05   1.450  0.14715
```

```

## TIF -8.698e-02 1.014e-01 -0.857 0.39119
## OLDCLAIM -7.399e-06 2.270e-05 -0.326 0.74446
## CLM_FREQ 1.129e+00 9.519e-01 1.186 0.23556
## MVR_PTS 7.791e-02 1.830e-01 0.426 0.67024
## CAR_AGE 7.325e-03 7.201e-02 0.102 0.91898
## PARENT1_Yes 7.038e-01 4.731e-01 1.488 0.13684
## MSTATUS_Yes -9.547e-01 3.431e-01 -2.783 0.00539 **
## SEX_z_F -7.446e-01 4.779e-01 -1.558 0.11921
## EDUCATION_.High.School -9.035e-01 8.022e-01 -1.126 0.26004
## EDUCATION_Bachelors -8.410e-01 6.694e-01 -1.256 0.20897
## EDUCATION_Masters 6.488e-01 5.807e-01 1.117 0.26381
## EDUCATION_z_High.School -4.459e-02 7.194e-01 -0.062 0.95057
## JOB_ -1.820e+00 7.205e-01 -2.526 0.01152 *
## JOB_Clerical -8.259e-02 4.814e-01 -0.172 0.86377
## JOB_Doctor -1.364e+00 1.006e+00 -1.357 0.17480
## JOB_Home.Maker 2.425e-01 6.777e-01 0.358 0.72048
## JOB_Lawyer -1.721e+00 6.726e-01 -2.558 0.01052 *
## JOB_Manager -1.153e+00 5.083e-01 -2.268 0.02333 *
## JOB_Student -3.105e-01 7.222e-01 -0.430 0.66725
## JOB_z_Blue.Collar 2.121e-02 4.396e-01 0.048 0.96152
## CAR_USE_Commercial 6.494e-01 3.513e-01 1.849 0.06450 .
## CAR_TYPE_Panel.Truck 2.357e-01 6.502e-01 0.363 0.71697
## CAR_TYPE_Pickup 1.162e+00 4.233e-01 2.745 0.00604 **
## CAR_TYPE_Sports.Car 2.286e+00 5.455e-01 4.190 2.79e-05 ***
## CAR_TYPE_Van 7.669e-01 4.646e-01 1.651 0.09878 .
## CAR_TYPE_z_SUV 2.108e+00 4.855e-01 4.342 1.41e-05 ***
## RED_CAR_no -3.294e-01 3.526e-01 -0.934 0.35029
## REVOKED_Yes 6.253e-01 4.159e-01 1.504 0.13271
## URBANICITY_z_Highly.Rural..Rural -2.262e+00 3.963e-01 -5.708 1.14e-08 ***
## YOJ_NA -2.641e-01 4.276e-01 -0.618 0.53683
## INCOME_NA 2.981e-01 5.527e-01 0.539 0.58965
## CAR_AGE_NA -6.218e-02 4.579e-01 -0.136 0.89199
## HOME_VAL_NA 1.098e-01 2.951e-01 0.372 0.70988
## ageSquared 1.412e-03 1.191e-03 1.186 0.23564
## yojSquared 7.800e-03 6.923e-03 1.127 0.25988
## income_log 2.139e-01 2.954e-01 0.724 0.46891
## homeval_log -9.505e-01 1.440e+00 -0.660 0.50922
## travtime_log 3.689e-01 6.385e-01 0.578 0.56340
## bluebook_log -5.611e-01 4.700e-01 -1.194 0.23254
## carage_log -3.028e-01 4.698e-01 -0.645 0.51915
## oldclaim_log 2.189e-01 1.667e-01 1.313 0.18921
## clm_freq_log -3.385e+00 2.975e+00 -1.138 0.25521
## mvr_pts_log -1.134e-01 5.275e-01 -0.215 0.82974
## tif_log 2.241e-01 5.602e-01 0.400 0.68916
## kidsdriv_log 3.659e+00 2.196e+00 1.667 0.09557 .
## homekids_log -4.385e-01 1.474e+00 -0.298 0.76600
## inter 7.238e-02 4.154e-02 1.743 0.08141 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 740.94 on 639 degrees of freedom
## Residual deviance: 527.30 on 584 degrees of freedom

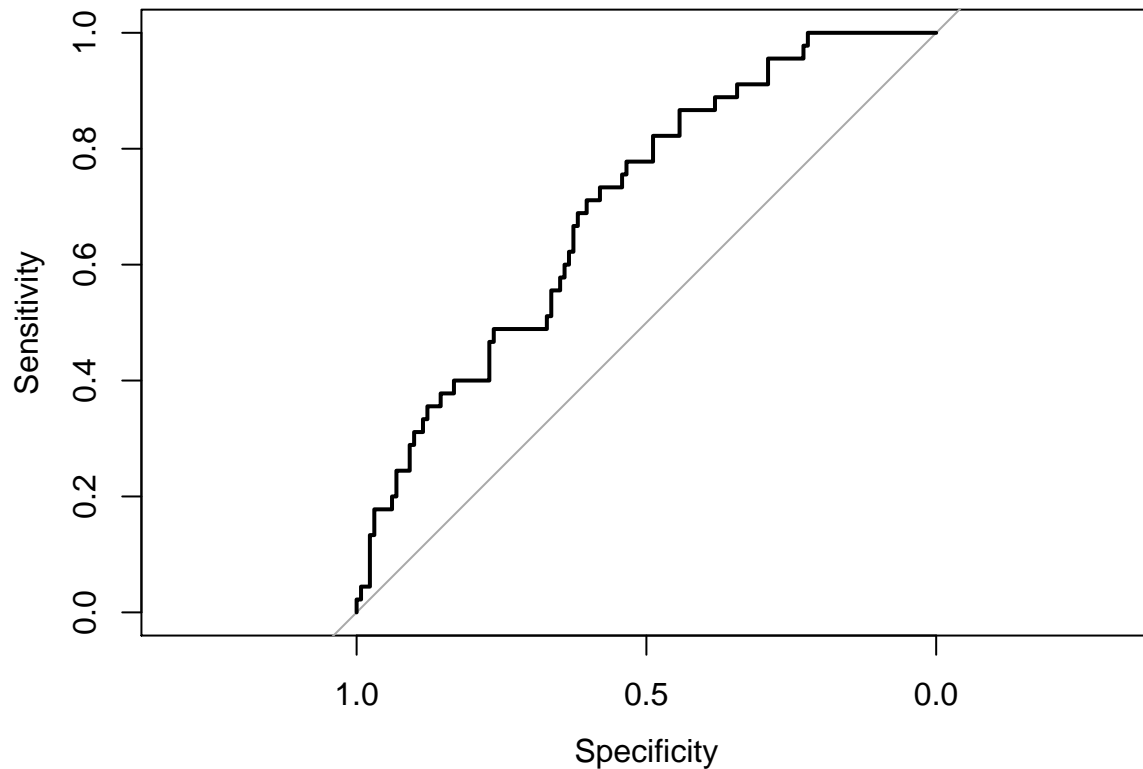
```

```

## (1 observation deleted due to missingness)
## AIC: 639.3
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  28
##           1  22  17
##
##           Accuracy : 0.7159
##           95% CI : (0.6432, 0.7812)
##           No Information Rate : 0.7443
##           P-Value [Acc > NIR] : 0.8294
##
##           Kappa : 0.2194
##
## Mcnemar's Test P-Value : 0.4795
##
##           Sensitivity : 0.8321
##           Specificity : 0.3778
##           Pos Pred Value : 0.7956
##           Neg Pred Value : 0.4359
##           Prevalence : 0.7443
##           Detection Rate : 0.6193
##           Detection Prevalence : 0.7784
##           Balanced Accuracy : 0.6049
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.701272264631043"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7013
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1939  -0.6475  -0.3616   0.4887   3.0086
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.253e+00  1.520e+01   0.214  0.83053
## KIDSDRIV      -1.179e+00  1.840e+00  -0.641  0.52183
## AGE           -1.846e-01  1.061e-01  -1.740  0.08193
## HOMEKIDS      -7.561e-01  6.705e-01  -1.128  0.25943
## YOJ           -1.840e-01  1.291e-01  -1.425  0.15408
## INCOME        -1.287e-05  1.402e-05  -0.918  0.35839
## HOME_VAL       3.716e-06  9.232e-06   0.403  0.68729
## TRAVTIME       2.177e-02  2.259e-02   0.964  0.33529
## BLUEBOOK       2.427e-08  4.227e-05   0.001  0.99954
```

```

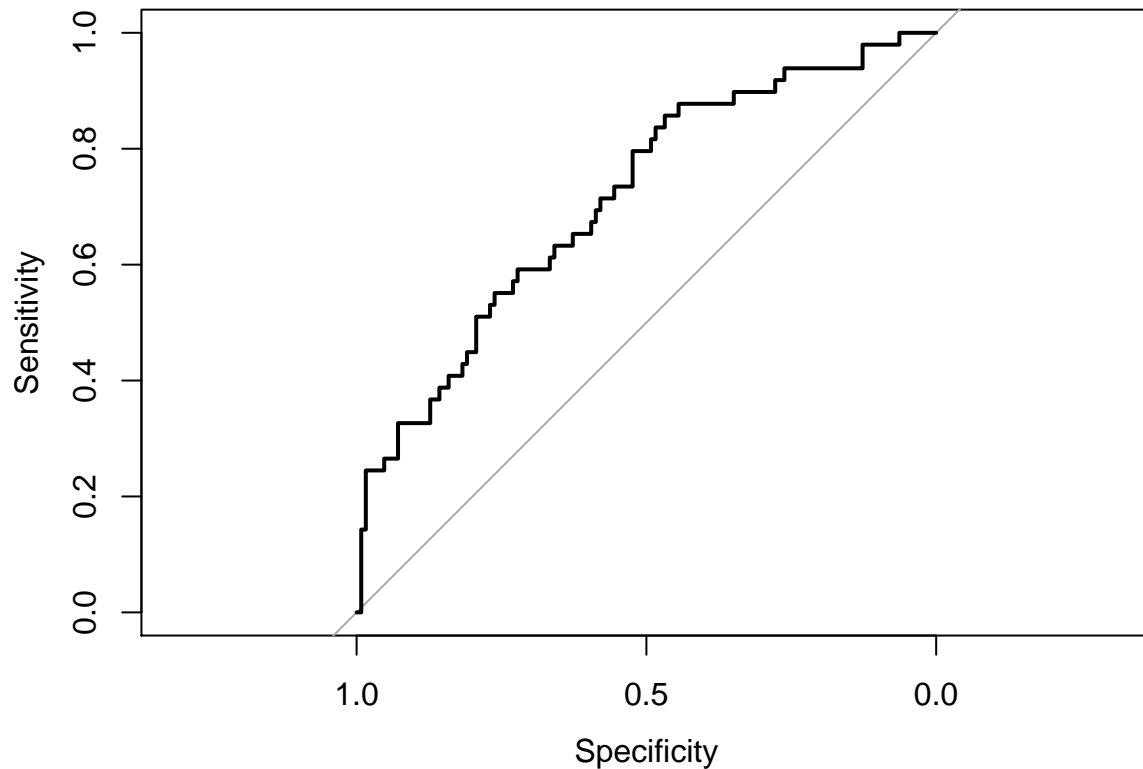
## TIF -1.034e-02 9.129e-02 -0.113 0.90984
## OLDCLAIM -1.107e-05 2.281e-05 -0.485 0.62755
## CLM_FREQ -1.554e-01 1.070e+00 -0.145 0.88455
## MVR_PTS 4.226e-02 1.671e-01 0.253 0.80029
## CAR_AGE -9.576e-02 7.430e-02 -1.289 0.19749
## PARENT1_Yes 5.382e-01 4.899e-01 1.099 0.27195
## MSTATUS_Yes -3.737e-01 3.380e-01 -1.106 0.26879
## SEX_z_F -9.208e-01 4.666e-01 -1.973 0.04847 *
## EDUCATION_.High.School -6.419e-01 8.339e-01 -0.770 0.44141
## EDUCATION_Bachelors -9.251e-01 7.036e-01 -1.315 0.18857
## EDUCATION_Masters 7.755e-01 6.185e-01 1.254 0.20987
## EDUCATION_z_High.School -1.911e-01 7.552e-01 -0.253 0.80023
## JOB_ -1.599e+00 7.011e-01 -2.281 0.02257 *
## JOB_Clerical -4.763e-01 4.945e-01 -0.963 0.33550
## JOB_Doctor -5.292e-01 9.251e-01 -0.572 0.56728
## JOB_Home.Maker -1.705e-01 6.399e-01 -0.266 0.78991
## JOB_Lawyer -1.549e+00 6.758e-01 -2.292 0.02192 *
## JOB_Manager -9.496e-01 4.969e-01 -1.911 0.05599 .
## JOB_Student -6.697e-01 7.421e-01 -0.902 0.36682
## JOB_z_Blue.Collar -2.526e-01 4.586e-01 -0.551 0.58173
## CAR_USE_Commercial 9.165e-01 3.567e-01 2.570 0.01018 *
## CAR_TYPE_Panel.Truck -5.010e-01 6.481e-01 -0.773 0.43947
## CAR_TYPE_Pickup 9.683e-01 4.047e-01 2.392 0.01674 *
## CAR_TYPE_Sports.Car 1.457e+00 5.461e-01 2.668 0.00763 **
## CAR_TYPE_Van 1.661e-01 4.871e-01 0.341 0.73311
## CAR_TYPE_z_SUV 1.968e+00 4.657e-01 4.227 2.37e-05 ***
## RED_CAR_no -1.071e-01 3.534e-01 -0.303 0.76176
## REVOKED_Yes 7.683e-01 4.177e-01 1.840 0.06583 .
## URBANICITY_z_Highly.Rural..Rural -2.516e+00 4.484e-01 -5.611 2.01e-08 ***
## YOJ_NA 1.064e-01 4.420e-01 0.241 0.80980
## INCOME_NA 1.136e-01 5.373e-01 0.211 0.83253
## CAR_AGE_NA -4.152e-01 4.989e-01 -0.832 0.40520
## HOME_VAL_NA -1.499e-02 2.954e-01 -0.051 0.95953
## ageSquared 1.971e-03 1.157e-03 1.703 0.08854 .
## yojSquared 8.584e-03 6.658e-03 1.289 0.19732
## income_log 5.720e-02 3.095e-01 0.185 0.85336
## homeval_log -1.890e-01 1.409e+00 -0.134 0.89330
## travtime_log -1.694e-02 6.368e-01 -0.027 0.97878
## bluebook_log 2.494e-01 5.178e-01 0.482 0.63002
## carage_log 2.741e-01 4.728e-01 0.580 0.56213
## oldclaim_log 3.940e-02 1.759e-01 0.224 0.82282
## clm_freq_log 4.331e-01 3.281e+00 0.132 0.89496
## mvr_pts_log -2.816e-02 4.983e-01 -0.057 0.95493
## tif_log -7.242e-02 5.179e-01 -0.140 0.88878
## kidsdriv_log -9.177e-01 2.326e+00 -0.395 0.69312
## homekids_log 1.594e+00 1.463e+00 1.089 0.27613
## inter 5.371e-02 3.423e-02 1.569 0.11665
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.27 on 640 degrees of freedom
## Residual deviance: 535.10 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 647.1
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 110  31
##           1   16  18
##
##           Accuracy : 0.7314
##           95% CI : (0.6593, 0.7955)
##           No Information Rate : 0.72
##           P-Value [Acc > NIR] : 0.40486
##
##           Kappa : 0.2652
##
## Mcnemar's Test P-Value : 0.04114
##
##           Sensitivity : 0.8730
##           Specificity : 0.3673
##           Pos Pred Value : 0.7801
##           Neg Pred Value : 0.5294
##           Prevalence : 0.7200
##           Detection Rate : 0.6286
##           Detection Prevalence : 0.8057
##           Balanced Accuracy : 0.6202
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.71379980563654"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 126 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.7138
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0370  -0.6977  -0.3725   0.5286   2.4453
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.577e+01  1.466e+01   1.076  0.28208
## KIDSDRIV       4.743e-01  1.875e+00   0.253  0.80030
## AGE          -6.290e-02  1.025e-01  -0.614  0.53934
## HOMEKIDS      -4.172e-01  6.631e-01  -0.629  0.52927
## YOJ          -2.264e-01  1.298e-01  -1.744  0.08110 .
## INCOME       -2.517e-05  1.322e-05  -1.905  0.05680 .
## HOME_VAL      1.291e-05  8.738e-06   1.478  0.13952
## TRAVTIME      1.357e-02  2.158e-02   0.629  0.52954
## BLUEBOOK      6.546e-06  4.447e-05   0.147  0.88298
```

```

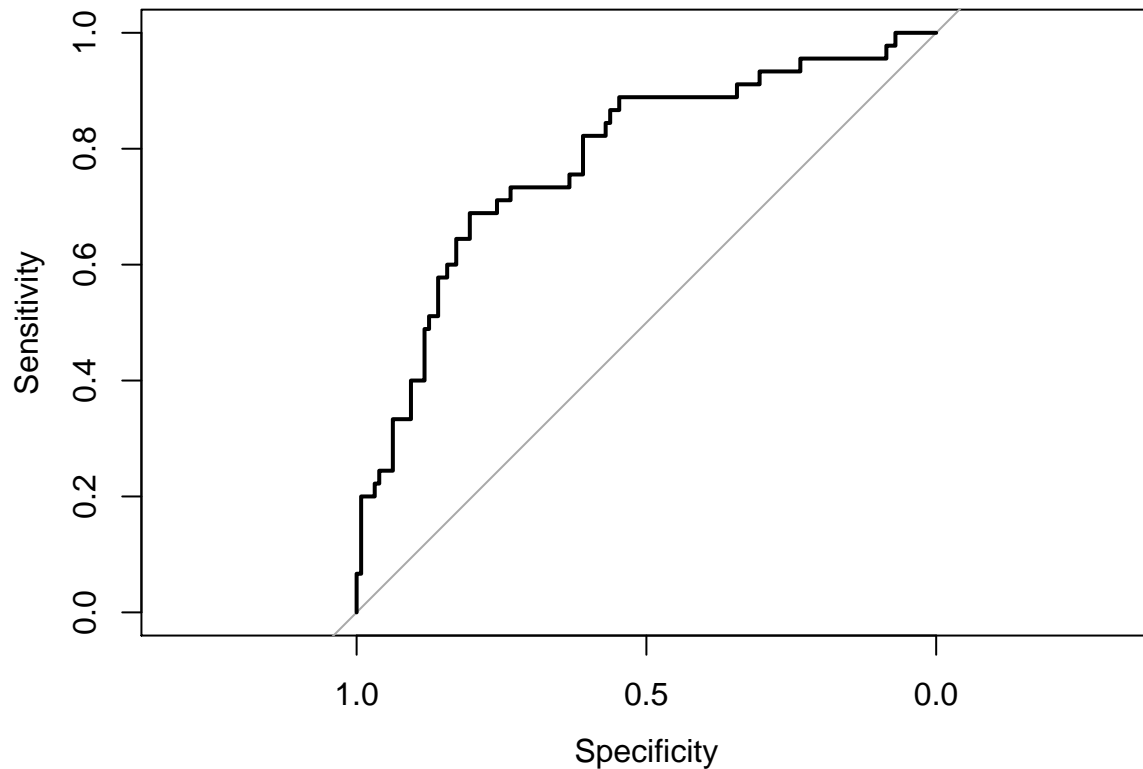
## TIF -5.880e-03 8.837e-02 -0.067 0.94695
## OLDCLAIM 8.176e-06 2.200e-05 0.372 0.71014
## CLM_FREQ -4.240e-01 9.818e-01 -0.432 0.66588
## MVR_PTS -1.079e-01 1.689e-01 -0.639 0.52281
## CAR_AGE 2.856e-03 7.138e-02 0.040 0.96809
## PARENT1_Yes 4.922e-01 4.710e-01 1.045 0.29594
## MSTATUS_Yes -4.907e-01 3.297e-01 -1.489 0.13662
## SEX_z_F -6.757e-01 4.702e-01 -1.437 0.15071
## EDUCATION_.High.School 1.621e-01 8.234e-01 0.197 0.84397
## EDUCATION_Bachelors 6.439e-02 7.169e-01 0.090 0.92843
## EDUCATION_Masters 8.924e-01 6.171e-01 1.446 0.14815
## EDUCATION_z_High.School 3.505e-01 7.634e-01 0.459 0.64615
## JOB_ -1.153e+00 6.887e-01 -1.674 0.09414 .
## JOB_Clerical -3.796e-01 4.795e-01 -0.792 0.42853
## JOB_Doctor -1.561e-01 1.006e+00 -0.155 0.87673
## JOB_Home.Maker -5.585e-01 6.699e-01 -0.834 0.40450
## JOB_Lawyer -9.528e-01 6.505e-01 -1.465 0.14302
## JOB_Manager -8.408e-01 4.718e-01 -1.782 0.07473 .
## JOB_Student -9.360e-01 7.080e-01 -1.322 0.18614
## JOB_z_Blue.Collar -5.396e-01 4.430e-01 -1.218 0.22325
## CAR_USE_Commercial 6.259e-01 3.316e-01 1.888 0.05908 .
## CAR_TYPE_Panel.Truck 5.610e-02 6.216e-01 0.090 0.92809
## CAR_TYPE_Pickup 8.705e-01 3.951e-01 2.203 0.02757 *
## CAR_TYPE_Sports.Car 1.450e+00 5.327e-01 2.722 0.00649 **
## CAR_TYPE_Van 2.108e-01 4.796e-01 0.439 0.66030
## CAR_TYPE_z_SUV 1.398e+00 4.676e-01 2.990 0.00279 **
## RED_CAR_no -1.044e-01 3.280e-01 -0.318 0.75030
## REVOKED_Yes 6.843e-01 4.147e-01 1.650 0.09890 .
## URBANICITY_z_Highly.Rural..Rural -2.407e+00 4.594e-01 -5.240 1.61e-07 ***
## YOJ_NA -2.589e-01 4.378e-01 -0.592 0.55418
## INCOME_NA 3.464e-01 5.510e-01 0.629 0.52958
## CAR_AGE_NA -4.296e-01 4.836e-01 -0.888 0.37440
## HOME_VAL_NA -1.986e-01 2.870e-01 -0.692 0.48908
## ageSquared 6.319e-04 1.119e-03 0.565 0.57222
## yojSquared 1.023e-02 6.677e-03 1.533 0.12529
## income_log 1.496e-01 2.741e-01 0.546 0.58516
## homeval_log -1.510e+00 1.355e+00 -1.114 0.26519
## travtime_log 1.618e-01 6.091e-01 0.266 0.79053
## bluebook_log 1.417e-01 5.447e-01 0.260 0.79477
## carage_log -3.464e-01 4.571e-01 -0.758 0.44852
## oldclaim_log -8.357e-02 1.653e-01 -0.505 0.61325
## clm_freq_log 1.857e+00 3.032e+00 0.613 0.54012
## mvr_pts_log 5.097e-01 4.945e-01 1.031 0.30265
## tif_log -3.265e-01 5.007e-01 -0.652 0.51432
## kidsdriv_log -1.239e+00 2.320e+00 -0.534 0.59339
## homekids_log 9.521e-01 1.435e+00 0.663 0.50710
## inter 1.706e-02 3.187e-02 0.535 0.59245
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 742.79 on 642 degrees of freedom
## Residual deviance: 558.02 on 587 degrees of freedom

```

```

## AIC: 670.02
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  30
##           1  10  15
##
##           Accuracy : 0.7688
##           95% CI : (0.6987, 0.8294)
##           No Information Rate : 0.7399
##           P-Value [Acc > NIR] : 0.219369
##
##           Kappa : 0.2982
##
## Mcnemar's Test P-Value : 0.002663
##
##           Sensitivity : 0.9219
##           Specificity : 0.3333
##           Pos Pred Value : 0.7973
##           Neg Pred Value : 0.6000
##           Prevalence : 0.7399
##           Detection Rate : 0.6821
##           Detection Prevalence : 0.8555
##           Balanced Accuracy : 0.6276
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.780381944444444"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 128 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7804
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1759  -0.6499  -0.3234   0.4342   2.6840
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.711e+01  1.519e+01   1.127 0.259888
## KIDSDRIV      -4.319e-01  1.671e+00  -0.258 0.796062
## AGE           -2.262e-01  1.074e-01  -2.106 0.035224 *
## HOMEKIDS      -2.525e-01  6.766e-01  -0.373 0.708998
## YOJ           -2.415e-01  1.375e-01  -1.756 0.079011 .
## INCOME        -1.490e-05  1.371e-05  -1.087 0.277167
## HOME_VAL       7.787e-06  9.283e-06   0.839 0.401565
## TRAVTIME       3.570e-02  2.297e-02   1.555 0.120061
## BLUEBOOK       1.661e-05  3.984e-05   0.417 0.676836
```

```

## TIF -3.038e-02 8.934e-02 -0.340 0.733848
## OLDCLAIM 2.408e-05 2.340e-05 1.029 0.303336
## CLM_FREQ -3.428e-01 1.063e+00 -0.322 0.747187
## MVR_PTS -1.069e-01 1.828e-01 -0.585 0.558567
## CAR_AGE -5.838e-02 7.355e-02 -0.794 0.427382
## PARENT1_Yes 1.791e-01 4.794e-01 0.374 0.708692
## MSTATUS_Yes -7.169e-01 3.442e-01 -2.083 0.037262 *
## SEX_z_F -1.037e+00 4.808e-01 -2.156 0.031066 *
## EDUCATION_.High.School 2.216e-01 8.464e-01 0.262 0.793446
## EDUCATION_Bachelors -2.602e-01 7.426e-01 -0.350 0.726003
## EDUCATION_Masters 6.056e-01 6.377e-01 0.950 0.342225
## EDUCATION_z_High.School 2.269e-01 7.915e-01 0.287 0.774402
## JOB_ -1.346e+00 7.124e-01 -1.890 0.058804 .
## JOB_Clerical -1.566e-01 5.077e-01 -0.308 0.757760
## JOB_Doctor 1.314e-01 1.002e+00 0.131 0.895675
## JOB_Home.Maker -7.856e-01 7.417e-01 -1.059 0.289544
## JOB_Lawyer -7.041e-01 7.039e-01 -1.000 0.317161
## JOB_Manager -9.365e-01 5.135e-01 -1.824 0.068212 .
## JOB_Student -1.590e+00 7.718e-01 -2.060 0.039401 *
## JOB_z_Blue.Collar -3.309e-01 4.503e-01 -0.735 0.462460
## CAR_USE_Commercial 9.831e-01 3.531e-01 2.784 0.005371 **
## CAR_TYPE_Panel.Truck -5.593e-02 6.390e-01 -0.088 0.930259
## CAR_TYPE_Pickup 1.189e+00 4.048e-01 2.938 0.003304 **
## CAR_TYPE_Sports.Car 1.684e+00 5.595e-01 3.009 0.002621 **
## CAR_TYPE_Van 6.595e-01 4.802e-01 1.373 0.169650
## CAR_TYPE_z_SUV 1.820e+00 4.922e-01 3.697 0.000218 ***
## RED_CAR_no -1.295e-01 3.350e-01 -0.386 0.699163
## REVOKED_Yes 5.662e-01 4.264e-01 1.328 0.184242
## URBANICITY_z_Highly.Rural..Rural -2.546e+00 4.513e-01 -5.640 1.7e-08 ***
## YOJ_NA -5.031e-01 4.725e-01 -1.065 0.287014
## INCOME_NA 2.264e-01 5.539e-01 0.409 0.682720
## CAR_AGE_NA 7.826e-02 5.124e-01 0.153 0.878598
## HOME_VAL_NA -2.439e-01 2.842e-01 -0.858 0.390839
## ageSquared 2.405e-03 1.163e-03 2.067 0.038721 *
## yojSquared 1.195e-02 7.014e-03 1.704 0.088459 .
## income_log -5.338e-02 2.878e-01 -0.185 0.852865
## homeval_log -1.049e+00 1.423e+00 -0.737 0.461299
## travtime_log -2.182e-01 6.490e-01 -0.336 0.736674
## bluebook_log 2.559e-02 4.739e-01 0.054 0.956942
## carage_log 1.650e-01 4.743e-01 0.348 0.727851
## oldclaim_log -8.778e-02 1.761e-01 -0.498 0.618238
## clm_freq_log 1.610e+00 3.264e+00 0.493 0.621864
## mvr_pts_log 3.948e-01 5.302e-01 0.745 0.456565
## tif_log -1.175e-01 5.106e-01 -0.230 0.817962
## kidsdriv_log 5.558e-01 2.152e+00 0.258 0.796157
## homekids_log 8.673e-01 1.463e+00 0.593 0.553310
## inter 1.979e-02 3.084e-02 0.642 0.521133
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 740.94 on 639 degrees of freedom
## Residual deviance: 521.64 on 584 degrees of freedom

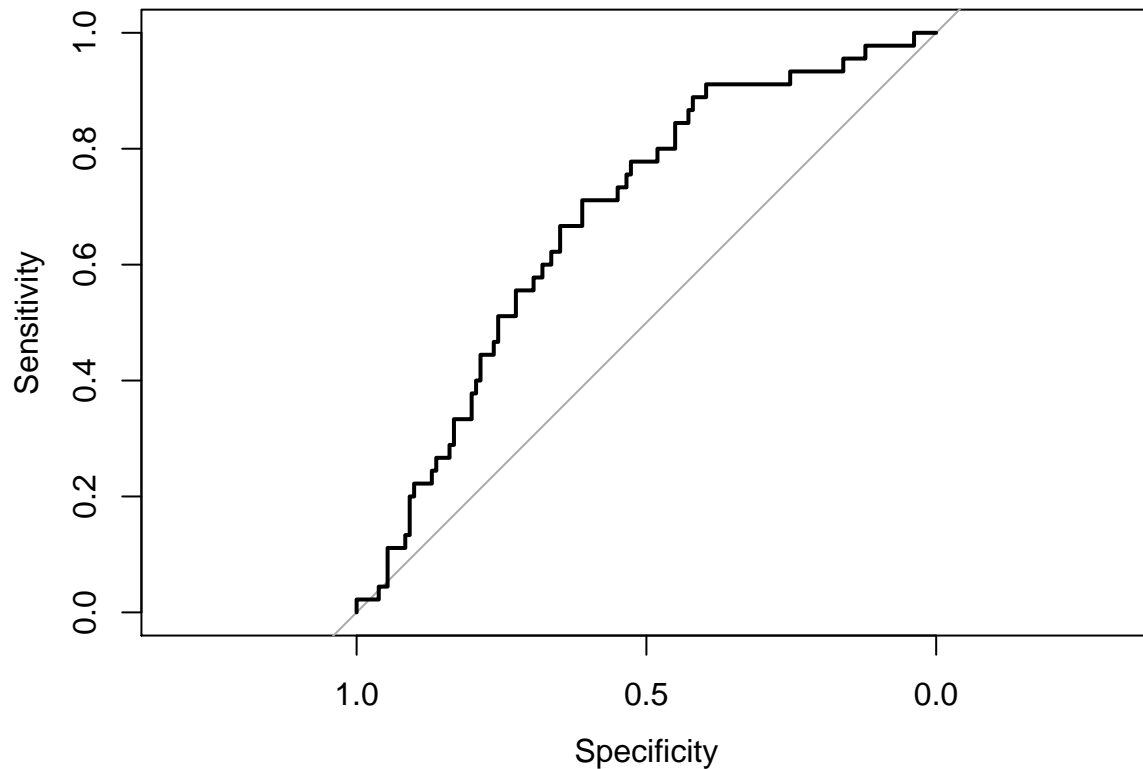
```



```

## AIC: 633.64
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  35
##           1   15  10
##
##           Accuracy : 0.7159
##           95% CI : (0.6432, 0.7812)
##           No Information Rate : 0.7443
##           P-Value [Acc > NIR] : 0.82943
##
##           Kappa : 0.1261
##
## Mcnemar's Test P-Value : 0.00721
##
##           Sensitivity : 0.8855
##           Specificity : 0.2222
##           Pos Pred Value : 0.7682
##           Neg Pred Value : 0.4000
##           Prevalence : 0.7443
##           Detection Rate : 0.6591
##           Detection Prevalence : 0.8580
##           Balanced Accuracy : 0.5539
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.680916030534351"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.6809
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0209  -0.6705  -0.3620   0.5177   2.9058
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.120e+01  1.567e+01   1.353 0.176019
## KIDSDRIV      -9.536e-01  2.076e+00  -0.459 0.645927
## AGE          -1.525e-01  1.017e-01  -1.499 0.133876
## HOMEKIDS      -8.631e-01  7.269e-01  -1.187 0.235107
## YOJ           -1.771e-01  1.264e-01  -1.401 0.161311
## INCOME        -2.888e-05  1.355e-05  -2.132 0.032995 *
## HOME_VAL       1.556e-05  9.282e-06   1.677 0.093579 .
## TRAVTIME       4.882e-02  2.077e-02   2.351 0.018741 *
## BLUEBOOK       4.054e-05  4.489e-05   0.903 0.366472
```

```

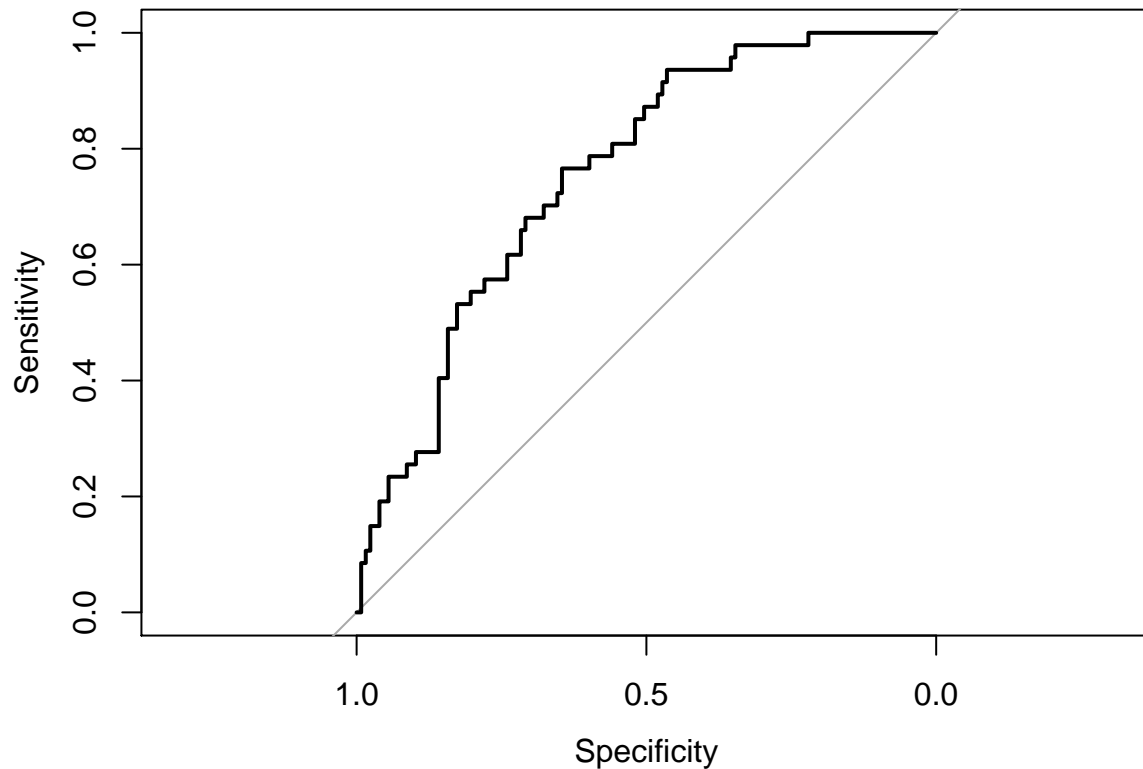
## TIF -1.567e-02 9.145e-02 -0.171 0.863975
## OLDCLAIM 5.100e-07 2.204e-05 0.023 0.981542
## CLM_FREQ 4.451e-02 9.895e-01 0.045 0.964120
## MVR_PTS -9.973e-02 1.697e-01 -0.588 0.556725
## CAR_AGE 1.991e-02 7.151e-02 0.278 0.780688
## PARENT1_Yes 1.028e+00 4.779e-01 2.152 0.031421 *
## MSTATUS_Yes -4.093e-01 3.280e-01 -1.248 0.212047
## SEX_z_F -7.419e-01 4.826e-01 -1.537 0.124187
## EDUCATION_.High.School 1.340e-02 8.592e-01 0.016 0.987562
## EDUCATION_Bachelors 2.026e-02 7.413e-01 0.027 0.978195
## EDUCATION_Masters 9.028e-01 6.519e-01 1.385 0.166077
## EDUCATION_z_High.School 1.635e-01 7.866e-01 0.208 0.835364
## JOB_ -1.383e+00 7.049e-01 -1.962 0.049749 *
## JOB_Clerical -2.860e-01 4.656e-01 -0.614 0.539021
## JOB_Doctor -3.003e-01 9.767e-01 -0.307 0.758475
## JOB_Home.Maker -1.903e-01 6.247e-01 -0.305 0.760678
## JOB_Lawyer -9.713e-01 6.535e-01 -1.486 0.137204
## JOB_Manager -1.161e+00 4.988e-01 -2.328 0.019898 *
## JOB_Student -9.146e-01 7.062e-01 -1.295 0.195291
## JOB_z_Blue.Collar -1.739e-01 4.444e-01 -0.391 0.695598
## CAR_USE_Commercial 3.816e-01 3.424e-01 1.115 0.265000
## CAR_TYPE_Panel.Truck 4.236e-01 6.235e-01 0.680 0.496821
## CAR_TYPE_Pickup 1.393e+00 3.947e-01 3.529 0.000417 ***
## CAR_TYPE_Sports.Car 1.798e+00 5.419e-01 3.318 0.000907 ***
## CAR_TYPE_Van 4.411e-01 4.882e-01 0.904 0.366176
## CAR_TYPE_z_SUV 1.893e+00 4.838e-01 3.913 9.10e-05 ***
## RED_CAR_no -1.088e-01 3.368e-01 -0.323 0.746568
## REVOKED_Yes 5.668e-01 4.175e-01 1.358 0.174584
## URBANICITY_z_Highly.Rural..Rural -2.065e+00 3.928e-01 -5.257 1.46e-07 ***
## YOJ_NA -4.568e-02 4.510e-01 -0.101 0.919316
## INCOME_NA 2.598e-01 5.348e-01 0.486 0.627049
## CAR_AGE_NA 1.034e-01 4.968e-01 0.208 0.835091
## HOME_VAL_NA 1.821e-02 2.890e-01 0.063 0.949756
## ageSquared 1.658e-03 1.106e-03 1.500 0.133648
## yojSquared 8.274e-03 6.518e-03 1.269 0.204314
## income_log 2.147e-01 2.800e-01 0.767 0.443197
## homeval_log -1.623e+00 1.428e+00 -1.136 0.255755
## travtime_log -7.866e-01 5.884e-01 -1.337 0.181320
## bluebook_log -1.929e-01 5.471e-01 -0.353 0.724338
## carage_log -5.345e-01 4.594e-01 -1.163 0.244688
## oldclaim_log 6.254e-02 1.626e-01 0.385 0.700479
## clm_freq_log 1.095e-01 3.028e+00 0.036 0.971150
## mvr_pts_log 4.597e-01 5.002e-01 0.919 0.358083
## tif_log -9.138e-02 5.168e-01 -0.177 0.859653
## kidsdriv_log 2.519e-01 2.438e+00 0.103 0.917691
## homekids_log 1.516e+00 1.527e+00 0.993 0.320650
## inter 3.592e-02 3.463e-02 1.037 0.299537
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 738.05 on 641 degrees of freedom
## Residual deviance: 544.19 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 656.19
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  33
##           1  18  14
##
##           Accuracy : 0.7069
##           95% CI : (0.6333, 0.7733)
##           No Information Rate : 0.7299
##           P-Value [Acc > NIR] : 0.78042
##
##           Kappa : 0.1736
##
## Mcnemar's Test P-Value : 0.04995
##
##           Sensitivity : 0.8583
##           Specificity : 0.2979
##           Pos Pred Value : 0.7676
##           Neg Pred Value : 0.4375
##           Prevalence : 0.7299
##           Detection Rate : 0.6264
##           Detection Prevalence : 0.8161
##           Balanced Accuracy : 0.5781
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.756743173060814"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 127 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.7567
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2792  -0.6716  -0.3536   0.6673   2.5532
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.431e+01  1.460e+01   0.980 0.326959
## KIDSDRIV      -1.686e+00  2.010e+00  -0.839 0.401430
## AGE           -6.866e-02  1.061e-01  -0.647 0.517739
## HOMEKIDS      -3.081e-01  6.592e-01  -0.467 0.640283
## YOJ           -1.370e-01  1.309e-01  -1.047 0.295310
## INCOME        -1.767e-05  1.333e-05  -1.326 0.184864
## HOME_VAL       8.188e-06  9.040e-06   0.906 0.365062
## TRAVTIME       3.016e-02  2.115e-02   1.426 0.153810
## BLUEBOOK       8.183e-05  3.918e-05   2.088 0.036762 *
```

```

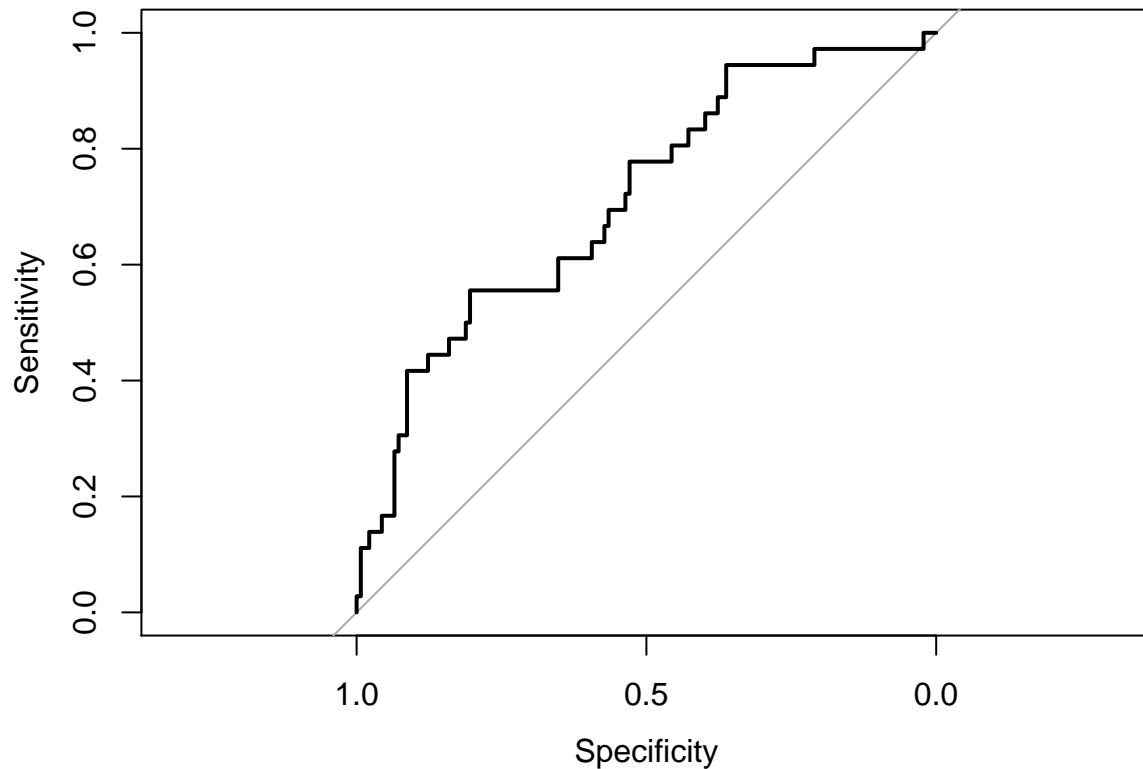
## TIF -6.311e-02 9.105e-02 -0.693 0.488258
## OLDCLAIM -1.458e-05 2.189e-05 -0.666 0.505237
## CLM_FREQ 2.078e-01 9.472e-01 0.219 0.826337
## MVRPTS -1.329e-01 1.769e-01 -0.751 0.452406
## CAR_AGE 2.804e-02 7.441e-02 0.377 0.706302
## PARENT1_Yes 5.298e-01 4.640e-01 1.142 0.253454
## MSTATUS_Yes -8.216e-01 3.268e-01 -2.514 0.011924 *
## SEX_z_F -4.751e-01 4.695e-01 -1.012 0.311542
## EDUCATION_.High.School -1.584e-01 7.679e-01 -0.206 0.836595
## EDUCATION_Bachelors -2.886e-01 6.510e-01 -0.443 0.657584
## EDUCATION_Masters 5.120e-01 5.650e-01 0.906 0.364844
## EDUCATION_z_High.School 1.634e-01 6.982e-01 0.234 0.814973
## JOB_ -1.856e+00 7.117e-01 -2.608 0.009095 **
## JOB_Clerical -2.590e-01 4.571e-01 -0.567 0.570962
## JOB_Doctor -5.745e-01 8.974e-01 -0.640 0.522061
## JOB_Home.Maker -1.767e-01 6.497e-01 -0.272 0.785624
## JOB_Lawyer -7.386e-01 6.266e-01 -1.179 0.238539
## JOB_Manager -1.018e+00 4.674e-01 -2.177 0.029463 *
## JOB_Student -3.343e-01 7.099e-01 -0.471 0.637684
## JOB_z_Blue.Collar -3.043e-01 4.332e-01 -0.703 0.482310
## CAR_USE_Commercial 6.351e-01 3.404e-01 1.866 0.062042 .
## CAR_TYPE_Panel.Truck 3.256e-01 6.310e-01 0.516 0.605827
## CAR_TYPE_Pickup 1.493e+00 4.076e-01 3.663 0.000249 ***
## CAR_TYPE_Sports.Car 1.926e+00 5.309e-01 3.627 0.000287 ***
## CAR_TYPE_Van 6.043e-01 4.627e-01 1.306 0.191563
## CAR_TYPE_z_SUV 1.874e+00 4.778e-01 3.922 8.77e-05 ***
## RED_CAR_no -3.935e-01 3.379e-01 -1.165 0.244127
## REVOKED_Yes 5.491e-01 4.086e-01 1.344 0.179041
## URBANICITY_z_Highly.Rural..Rural -2.183e+00 4.089e-01 -5.340 9.28e-08 ***
## YOJ_NA -3.422e-01 4.589e-01 -0.746 0.455797
## INCOME_NA -5.392e-02 5.381e-01 -0.100 0.920173
## CAR_AGE_NA -4.439e-01 4.539e-01 -0.978 0.328010
## HOME_VAL_NA 8.773e-02 2.849e-01 0.308 0.758098
## ageSquared 7.732e-04 1.138e-03 0.680 0.496739
## yojSquared 7.054e-03 6.792e-03 1.039 0.299026
## income_log 1.140e-01 2.818e-01 0.405 0.685733
## homeval_log -7.160e-01 1.377e+00 -0.520 0.603125
## travtime_log -2.258e-02 6.046e-01 -0.037 0.970207
## bluebook_log -8.100e-01 4.752e-01 -1.705 0.088251 .
## carage_log -3.777e-01 4.779e-01 -0.790 0.429333
## oldclaim_log 7.945e-02 1.611e-01 0.493 0.621858
## clm_freq_log -2.319e-01 2.931e+00 -0.079 0.936935
## mvr_pts_log 5.781e-01 5.070e-01 1.140 0.254176
## tif_log 2.515e-02 5.101e-01 0.049 0.960679
## kidsdriv_log 2.422e+00 2.234e+00 1.085 0.278135
## homekids_log 6.032e-01 1.416e+00 0.426 0.670049
## inter 2.230e-02 3.697e-02 0.603 0.546310
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 759.91 on 641 degrees of freedom
## Residual deviance: 556.16 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 668.16
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 121  20
##           1  17  16
##
##           Accuracy : 0.7874
##           95% CI : (0.719, 0.8456)
##           No Information Rate : 0.7931
##           P-Value [Acc > NIR] : 0.6169
##
##           Kappa : 0.3315
##
## Mcnemar's Test P-Value : 0.7423
##
##           Sensitivity : 0.8768
##           Specificity : 0.4444
##           Pos Pred Value : 0.8582
##           Neg Pred Value : 0.4848
##           Prevalence : 0.7931
##           Detection Rate : 0.6954
##           Detection Prevalence : 0.8103
##           Balanced Accuracy : 0.6606
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.711553945249597"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 36 cases (dfPred_raw$class 1).
## Area under the curve: 0.7116
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2862  -0.6706  -0.3421   0.5351   3.1247
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.852e+01  1.500e+01   2.569  0.01021 *
## KIDSDRIV      -7.657e-01  1.662e+00  -0.461  0.64490
## AGE           -7.186e-02  1.090e-01  -0.659  0.50958
## HOMEKIDS      -8.470e-01  6.955e-01  -1.218  0.22325
## YOJ           -1.329e-01  1.336e-01  -0.995  0.31974
## INCOME        -3.015e-05  1.386e-05  -2.176  0.02955 *
## HOME_VAL       2.129e-05  9.145e-06   2.328  0.01990 *
## TRAVTIME       2.327e-02  2.123e-02   1.096  0.27288
## BLUEBOOK       7.029e-05  3.853e-05   1.824  0.06813 .
```



```

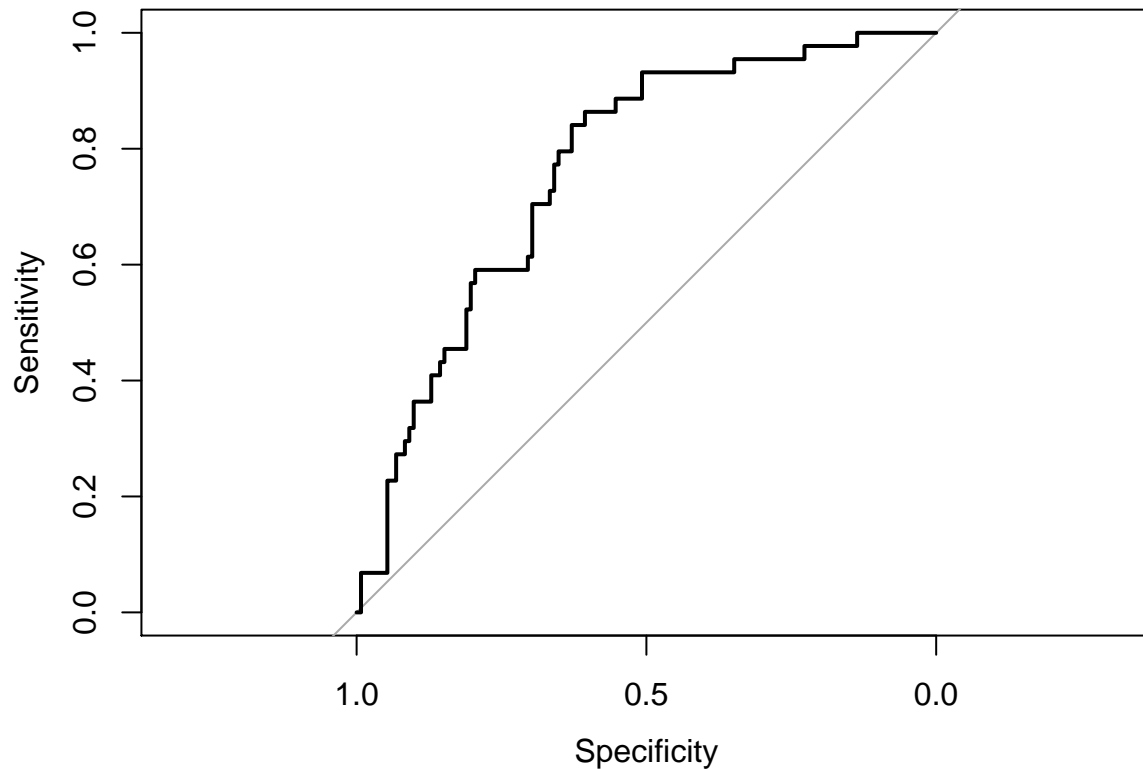
## TIF -8.527e-02 9.265e-02 -0.920 0.35737
## OLDCLAIM 9.323e-07 2.076e-05 0.045 0.96418
## CLM_FREQ 5.977e-01 9.354e-01 0.639 0.52286
## MVR_PTS 4.977e-03 1.740e-01 0.029 0.97719
## CAR_AGE -8.967e-02 7.044e-02 -1.273 0.20301
## PARENT1_Yes 8.301e-01 4.741e-01 1.751 0.07997
## MSTATUS_Yes -5.456e-01 3.376e-01 -1.616 0.10603
## SEX_z_F -6.085e-01 4.737e-01 -1.284 0.19897
## EDUCATION_High.School 3.947e-01 7.519e-01 0.525 0.59964
## EDUCATION_Bachelors -2.643e-01 6.267e-01 -0.422 0.67325
## EDUCATION_Masters 6.183e-01 5.220e-01 1.184 0.23624
## EDUCATION_z_High.School 5.120e-01 6.808e-01 0.752 0.45199
## JOB_ -4.378e-01 6.800e-01 -0.644 0.51970
## JOB_Clerical -4.949e-01 4.843e-01 -1.022 0.30677
## JOB_Doctor -4.068e-02 9.017e-01 -0.045 0.96401
## JOB_Home.Maker -4.477e-01 6.832e-01 -0.655 0.51227
## JOB_Lawyer -4.133e-02 6.662e-01 -0.062 0.95053
## JOB_Manager -5.093e-01 4.868e-01 -1.046 0.29548
## JOB_Student -1.008e+00 7.296e-01 -1.382 0.16708
## JOB_z_Blue.Collar -2.741e-01 4.567e-01 -0.600 0.54850
## CAR_USE_Commercial 5.494e-01 3.473e-01 1.582 0.11367
## CAR_TYPE_Panel.Truck -9.056e-02 6.362e-01 -0.142 0.88681
## CAR_TYPE_Pickup 1.142e+00 4.037e-01 2.829 0.00468 **
## CAR_TYPE_Sports.Car 1.610e+00 5.471e-01 2.944 0.00324 **
## CAR_TYPE_Van 4.003e-01 4.725e-01 0.847 0.39683
## CAR_TYPE_z_SUV 1.926e+00 4.775e-01 4.032 5.52e-05 ***
## RED_CAR_no -3.901e-01 3.419e-01 -1.141 0.25389
## REVOKED_Yes 6.395e-01 3.923e-01 1.630 0.10303
## URBANICITY_z_Highly.Rural..Rural -2.440e+00 4.397e-01 -5.550 2.86e-08 ***
## YOJ_NA -4.215e-01 4.184e-01 -1.007 0.31370
## INCOME_NA 3.956e-02 5.596e-01 0.071 0.94364
## CAR_AGE_NA -1.763e-01 4.939e-01 -0.357 0.72109
## HOME_VAL_NA -7.412e-02 2.891e-01 -0.256 0.79763
## ageSquared 7.667e-04 1.184e-03 0.648 0.51730
## yojSquared 7.124e-03 6.911e-03 1.031 0.30265
## income_log 3.386e-01 2.898e-01 1.168 0.24263
## homeval_log -3.263e+00 1.389e+00 -2.349 0.01882 *
## travtime_log -8.074e-02 6.014e-01 -0.134 0.89321
## bluebook_log -5.889e-01 4.751e-01 -1.239 0.21520
## carage_log 3.463e-01 4.561e-01 0.759 0.44765
## oldclaim_log 9.026e-02 1.567e-01 0.576 0.56448
## clm_freq_log -1.310e+00 2.873e+00 -0.456 0.64849
## mvr_pts_log 1.216e-01 5.132e-01 0.237 0.81269
## tif_log 1.703e-01 5.213e-01 0.327 0.74393
## kidsdriv_log 1.202e+00 2.194e+00 0.548 0.58367
## homekids_log 1.760e+00 1.481e+00 1.188 0.23476
## inter 1.633e-02 3.070e-02 0.532 0.59471
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 742.96 on 639 degrees of freedom
## Residual deviance: 541.85 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 653.85
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 117  28
##           1  15  16
##
##           Accuracy : 0.7557
##           95% CI : (0.6853, 0.8172)
##           No Information Rate : 0.75
##           P-Value [Acc > NIR] : 0.47109
##
##           Kappa : 0.2773
##
## Mcnemar's Test P-Value : 0.06725
##
##           Sensitivity : 0.8864
##           Specificity : 0.3636
##           Pos Pred Value : 0.8069
##           Neg Pred Value : 0.5161
##           Prevalence : 0.7500
##           Detection Rate : 0.6648
##           Detection Prevalence : 0.8239
##           Balanced Accuracy : 0.6250
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.764807162534435"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 132 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7648
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2098  -0.7096  -0.3965   0.6288   2.9160
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.220e+01  1.367e+01   0.892 0.372222
## KIDSDRIV      -1.206e+00  1.875e+00  -0.643 0.519957
## AGE           -7.035e-02  1.032e-01  -0.682 0.495448
## HOMEKIDS      -1.003e+00  6.807e-01  -1.474 0.140503
## YOJ           -1.919e-01  1.256e-01  -1.528 0.126550
## INCOME        -1.175e-05  1.276e-05  -0.921 0.357220
## HOME_VAL       5.862e-06  8.354e-06   0.702 0.482854
## TRAVTIME       3.878e-02  2.022e-02   1.918 0.055087
## BLUEBOOK       4.749e-05  3.790e-05   1.253 0.210139
```

```

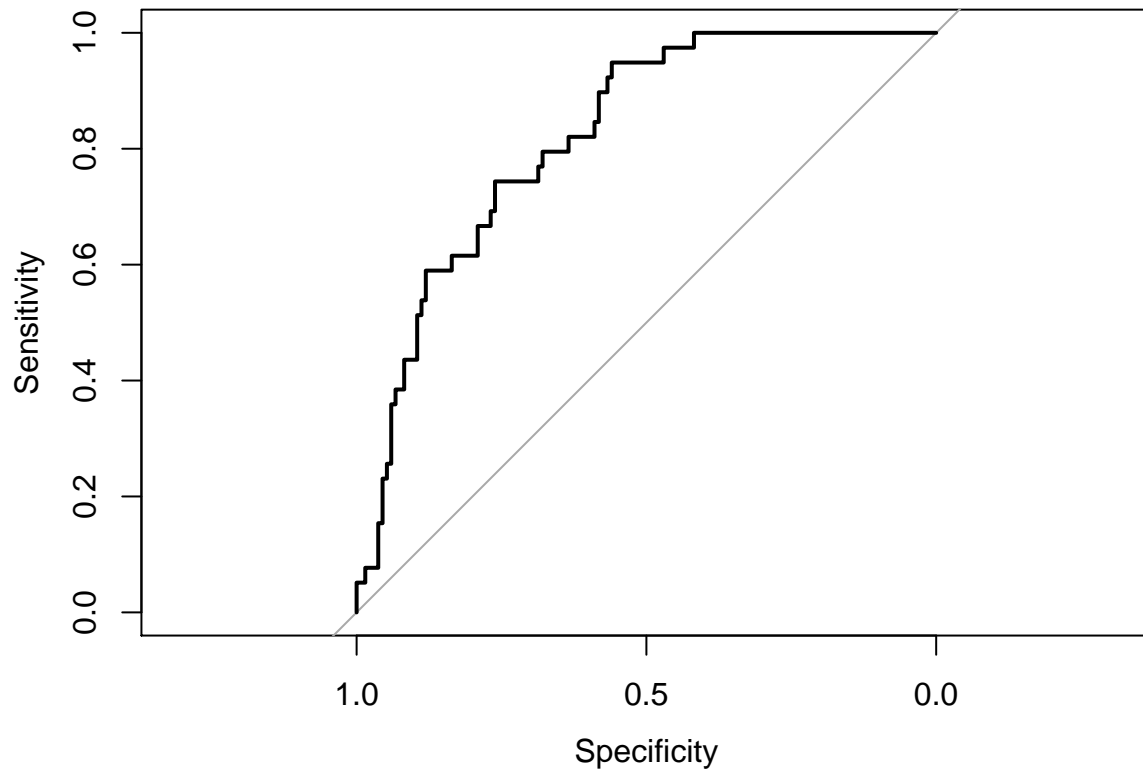
## TIF -1.400e-01 9.910e-02 -1.413 0.157602
## OLDCLAIM 5.759e-06 2.093e-05 0.275 0.783219
## CLM_FREQ -2.989e-01 9.614e-01 -0.311 0.755910
## MVR_PTS 1.299e-01 1.635e-01 0.795 0.426862
## CAR_AGE -1.411e-02 6.663e-02 -0.212 0.832288
## PARENT1_Yes 7.004e-01 4.539e-01 1.543 0.122790
## MSTATUS_Yes -7.195e-01 3.225e-01 -2.231 0.025696 *
## SEX_z_F -6.795e-01 4.563e-01 -1.489 0.136435
## EDUCATION_.High.School 2.709e-01 7.393e-01 0.366 0.714017
## EDUCATION_Bachelors -4.699e-02 5.987e-01 -0.078 0.937441
## EDUCATION_Masters 7.264e-01 5.131e-01 1.416 0.156900
## EDUCATION_z_High.School 4.333e-01 6.563e-01 0.660 0.509104
## JOB_ -6.506e-01 6.738e-01 -0.966 0.334275
## JOB_Clerical -4.089e-02 4.791e-01 -0.085 0.931987
## JOB_Doctor -2.919e-01 8.708e-01 -0.335 0.737491
## JOB_Home.Maker -2.168e-01 6.641e-01 -0.326 0.744086
## JOB_Lawyer -6.184e-01 6.087e-01 -1.016 0.309721
## JOB_Manager -5.200e-01 4.523e-01 -1.150 0.250320
## JOB_Student -5.111e-01 7.094e-01 -0.720 0.471237
## JOB_z_Blue.Collar -2.289e-01 4.421e-01 -0.518 0.604633
## CAR_USE_Commercial 5.321e-01 3.422e-01 1.555 0.119940
## CAR_TYPE_Panel.Truck -2.960e-01 6.130e-01 -0.483 0.629225
## CAR_TYPE_Pickup 1.127e+00 3.970e-01 2.839 0.004525 **
## CAR_TYPE_Sports.Car 1.768e+00 5.136e-01 3.441 0.000579 ***
## CAR_TYPE_Van 3.237e-01 4.373e-01 0.740 0.459130
## CAR_TYPE_z_SUV 1.595e+00 4.639e-01 3.438 0.000586 ***
## RED_CAR_no -2.647e-01 3.298e-01 -0.803 0.422171
## REVOKED_Yes 1.956e-01 4.138e-01 0.473 0.636467
## URBANICITY_z_Highly.Rural..Rural -2.179e+00 3.972e-01 -5.486 4.12e-08 ***
## YOJ_NA -3.630e-01 4.074e-01 -0.891 0.372942
## INCOME_NA -1.560e-01 4.964e-01 -0.314 0.753372
## CAR_AGE_NA -3.708e-01 4.814e-01 -0.770 0.441162
## HOME_VAL_NA -4.522e-02 2.787e-01 -0.162 0.871125
## ageSquared 7.998e-04 1.115e-03 0.717 0.473318
## yojSquared 1.058e-02 6.512e-03 1.625 0.104106
## income_log 8.244e-02 2.974e-01 0.277 0.781653
## homeval_log -7.923e-01 1.297e+00 -0.611 0.541136
## travtime_log -4.659e-01 5.562e-01 -0.838 0.402251
## bluebook_log -2.672e-01 4.747e-01 -0.563 0.573523
## carage_log -1.107e-01 4.409e-01 -0.251 0.801839
## oldclaim_log 2.891e-02 1.599e-01 0.181 0.856485
## clm_freq_log 7.074e-01 2.956e+00 0.239 0.810842
## mvr_pts_log -2.627e-01 4.829e-01 -0.544 0.586409
## tif_log 4.633e-01 5.404e-01 0.857 0.391223
## kidsdriv_log 1.118e+00 2.097e+00 0.533 0.593884
## homekids_log 1.910e+00 1.458e+00 1.310 0.190127
## inter 3.112e-02 3.842e-02 0.810 0.417900
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 754.78 on 642 degrees of freedom
## Residual deviance: 574.23 on 587 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 686.23
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 121  22
##           1  13  17
##
##           Accuracy : 0.7977
##           95% CI : (0.73, 0.8549)
##           No Information Rate : 0.7746
##           P-Value [Acc > NIR] : 0.2654
##
##           Kappa : 0.3691
##
## Mcnemar's Test P-Value : 0.1763
##
##           Sensitivity : 0.9030
##           Specificity : 0.4359
##           Pos Pred Value : 0.8462
##           Neg Pred Value : 0.5667
##           Prevalence : 0.7746
##           Detection Rate : 0.6994
##           Detection Prevalence : 0.8266
##           Balanced Accuracy : 0.6694
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.821278224263299"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 39 cases (dfPred_raw$class 1).
## Area under the curve: 0.8213
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2251  -0.6672  -0.3617   0.4924   2.9210
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.913e+01  1.488e+01   1.286 0.198523
## KIDSDRIV      -3.668e-01  1.897e+00  -0.193 0.846671
## AGE           -2.053e-01  1.083e-01  -1.895 0.058133 .
## HOMEKIDS      -5.782e-01  6.770e-01  -0.854 0.393093
## YOJ           -2.124e-01  1.301e-01  -1.632 0.102590
## INCOME        -2.421e-05  1.337e-05  -1.811 0.070136 .
## HOME_VAL       1.388e-05  9.167e-06   1.514 0.129914
## TRAVTIME       4.463e-02  2.062e-02   2.165 0.030423 *
## BLUEBOOK       5.090e-05  4.022e-05   1.266 0.205672
```

```

## TIF -8.766e-02 9.644e-02 -0.909 0.363397
## OLDCLAIM 1.875e-05 2.131e-05 0.879 0.379154
## CLM_FREQ 1.606e-01 1.003e+00 0.160 0.872734
## MVR_PTS 2.534e-02 1.735e-01 0.146 0.883912
## CAR_AGE 2.641e-02 7.058e-02 0.374 0.708270
## PARENT1_Yes 4.934e-01 4.612e-01 1.070 0.284777
## MSTATUS_Yes -7.129e-01 3.319e-01 -2.148 0.031740 *
## SEX_z_F -7.192e-01 4.665e-01 -1.542 0.123189
## EDUCATION_.High.School 6.737e-01 7.847e-01 0.859 0.390538
## EDUCATION_Bachelors 3.693e-01 6.613e-01 0.558 0.576536
## EDUCATION_Masters 1.237e+00 5.786e-01 2.137 0.032576 *
## EDUCATION_z_High.School 9.627e-01 7.168e-01 1.343 0.179239
## JOB_ -1.282e+00 7.041e-01 -1.821 0.068572 .
## JOB_Clerical -3.111e-01 5.015e-01 -0.620 0.534992
## JOB_Doctor 4.068e-01 9.084e-01 0.448 0.654307
## JOB_Home.Maker 9.682e-02 6.474e-01 0.150 0.881116
## JOB_Lawyer -7.593e-01 6.565e-01 -1.156 0.247494
## JOB_Manager -1.022e+00 5.142e-01 -1.988 0.046769 *
## JOB_Student -1.159e+00 7.318e-01 -1.584 0.113263
## JOB_z_Blue.Collar -3.941e-01 4.618e-01 -0.853 0.393433
## CAR_USE_Commercial 8.753e-01 3.601e-01 2.431 0.015065 *
## CAR_TYPE_Panel.Truck -1.868e-01 6.249e-01 -0.299 0.764995
## CAR_TYPE_Pickup 1.104e+00 4.087e-01 2.702 0.006886 **
## CAR_TYPE_Sports.Car 1.828e+00 5.270e-01 3.469 0.000522 ***
## CAR_TYPE_Van 5.718e-01 4.707e-01 1.215 0.224407
## CAR_TYPE_z_SUV 1.930e+00 4.783e-01 4.036 5.44e-05 ***
## RED_CAR_no -4.008e-01 3.395e-01 -1.181 0.237692
## REVOKED_Yes 1.689e-01 4.380e-01 0.386 0.699711
## URBANICITY_z_Highly.Rural..Rural -2.330e+00 4.049e-01 -5.756 8.62e-09 ***
## YOJ_NA -3.055e-01 4.244e-01 -0.720 0.471621
## INCOME_NA 3.660e-01 5.461e-01 0.670 0.502744
## CAR_AGE_NA -2.398e-01 4.621e-01 -0.519 0.603718
## HOME_VAL_NA -2.582e-01 2.810e-01 -0.919 0.358155
## ageSquared 2.268e-03 1.176e-03 1.928 0.053893 .
## yojSquared 1.220e-02 6.740e-03 1.810 0.070250 .
## income_log 7.783e-02 2.646e-01 0.294 0.768670
## homeval_log -1.285e+00 1.412e+00 -0.911 0.362482
## travtime_log -7.840e-01 5.736e-01 -1.367 0.171725
## bluebook_log -1.215e-01 4.910e-01 -0.247 0.804571
## carage_log -4.209e-01 4.558e-01 -0.923 0.355802
## oldclaim_log 1.344e-02 1.650e-01 0.081 0.935082
## clm_freq_log -2.636e-01 3.060e+00 -0.086 0.931350
## mvr_pts_log 4.202e-02 5.061e-01 0.083 0.933824
## tif_log 2.218e-01 5.297e-01 0.419 0.675456
## kidsdriv_log 2.588e-01 2.272e+00 0.114 0.909306
## homekids_log 1.267e+00 1.470e+00 0.862 0.388893
## inter 2.753e-02 3.663e-02 0.752 0.452328
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 739.51 on 640 degrees of freedom
## Residual deviance: 546.66 on 585 degrees of freedom

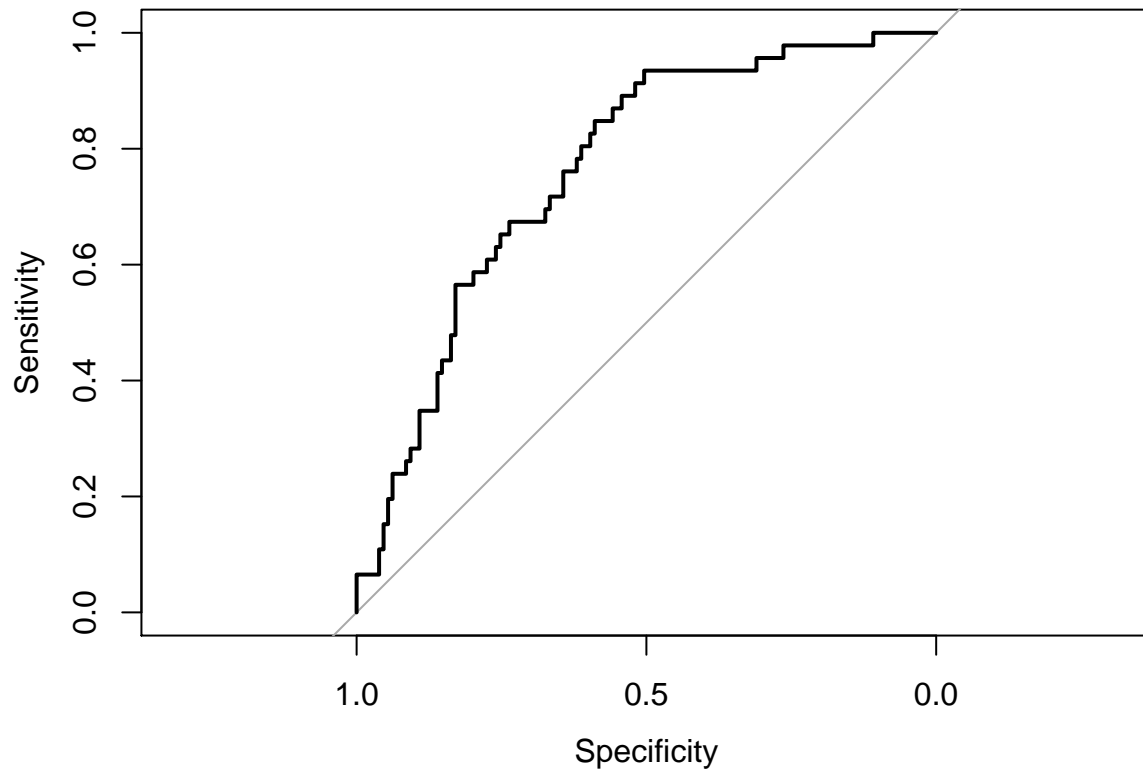
```

```

## (1 observation deleted due to missingness)
## AIC: 658.66
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  30
##           1  15  16
##
##           Accuracy : 0.7429
##           95% CI : (0.6715, 0.8058)
##           No Information Rate : 0.7371
##           P-Value [Acc > NIR] : 0.47119
##
##           Kappa : 0.2587
##
## Mcnemar's Test P-Value : 0.03689
##
##           Sensitivity : 0.8837
##           Specificity : 0.3478
##           Pos Pred Value : 0.7917
##           Neg Pred Value : 0.5161
##           Prevalence : 0.7371
##           Detection Rate : 0.6514
##           Detection Prevalence : 0.8229
##           Balanced Accuracy : 0.6158
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.765082574991574"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 129 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7651
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2919  -0.6540  -0.3698   0.5117   3.1286
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.348e+01  1.495e+01   0.902 0.367164
## KIDSDRIV      -1.132e+00  1.757e+00  -0.644 0.519447
## AGE           -1.413e-02  1.168e-01  -0.121 0.903724
## HOMEKIDS      -1.556e+00  7.102e-01  -2.191 0.028444 *
## YOJ           -1.785e-01  1.308e-01  -1.364 0.172449
## INCOME        -1.801e-05  1.288e-05  -1.398 0.162073
## HOME_VAL       1.086e-05  8.715e-06   1.246 0.212663
## TRAVTIME       2.425e-02  2.074e-02   1.169 0.242390
## BLUEBOOK       3.370e-05  4.138e-05   0.814 0.415437
```

```

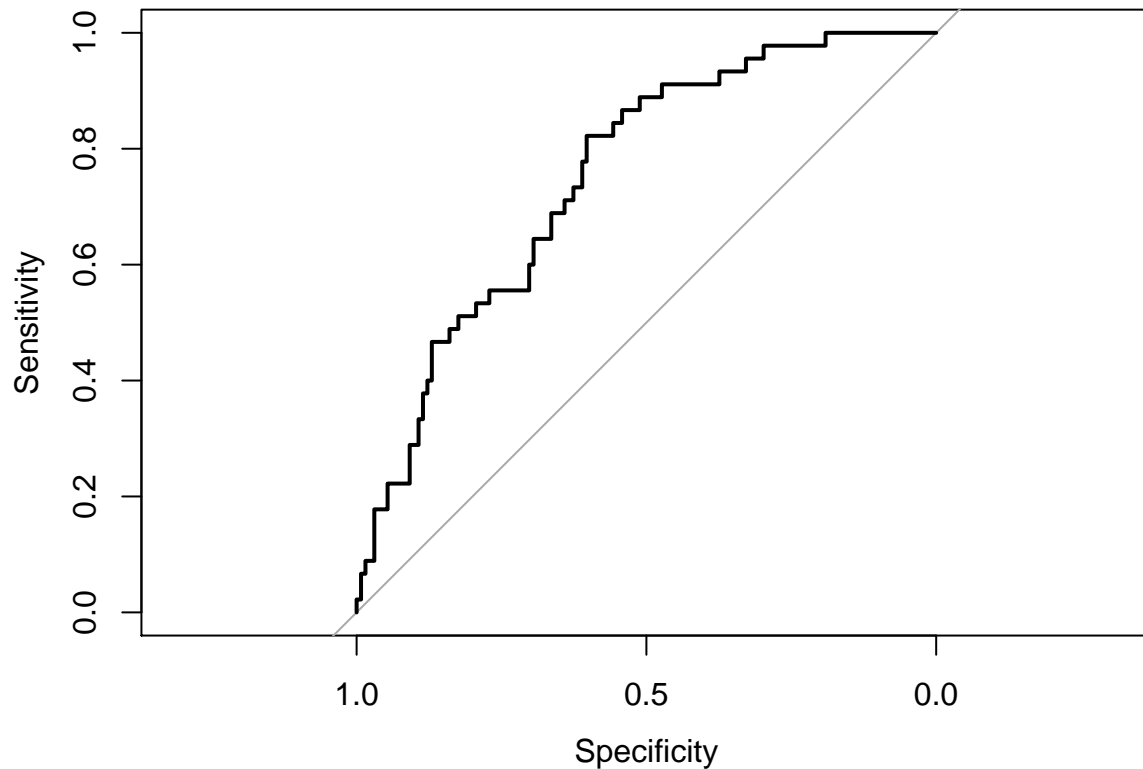
## TIF -5.942e-02 8.983e-02 -0.662 0.508261
## OLDCLAIM 3.531e-06 2.269e-05 0.156 0.876345
## CLM_FREQ 1.984e-01 9.822e-01 0.202 0.839906
## MVRPTS 4.937e-02 1.721e-01 0.287 0.774209
## CAR_AGE -5.056e-02 7.359e-02 -0.687 0.492022
## PARENT1_Yes 7.121e-01 4.849e-01 1.469 0.141929
## MSTATUS_Yes -5.665e-01 3.261e-01 -1.737 0.082409 .
## SEX_z_F -1.120e+00 4.832e-01 -2.318 0.020466 *
## EDUCATION_.High.School 9.859e-01 8.144e-01 1.211 0.226067
## EDUCATION_Bachelors 3.150e-01 6.921e-01 0.455 0.648985
## EDUCATION_Masters 9.201e-01 6.039e-01 1.524 0.127605
## EDUCATION_z_High.School 8.315e-01 7.371e-01 1.128 0.259270
## JOB_ -6.095e-01 6.756e-01 -0.902 0.366919
## JOB_Clerical -5.084e-01 4.888e-01 -1.040 0.298331
## JOB_Doctor 6.253e-01 9.076e-01 0.689 0.490854
## JOB_Home.Maker -4.497e-01 6.794e-01 -0.662 0.508013
## JOB_Lawyer -3.360e-01 6.500e-01 -0.517 0.605186
## JOB_Manager -7.057e-01 4.568e-01 -1.545 0.122349
## JOB_Student -7.392e-01 7.441e-01 -0.993 0.320511
## JOB_z_Blue.Collar -2.782e-01 4.459e-01 -0.624 0.532640
## CAR_USE_Commercial 3.631e-01 3.489e-01 1.041 0.298068
## CAR_TYPE_Panel.Truck -1.441e-01 6.313e-01 -0.228 0.819391
## CAR_TYPE_Pickup 1.247e+00 3.959e-01 3.150 0.001633 **
## CAR_TYPE_Sports.Car 1.999e+00 5.505e-01 3.631 0.000282 ***
## CAR_TYPE_Van 4.090e-01 4.635e-01 0.882 0.377515
## CAR_TYPE_z_SUV 2.332e+00 4.927e-01 4.733 2.21e-06 ***
## RED_CAR_no -1.176e-01 3.326e-01 -0.353 0.723785
## REVOKED_Yes 3.209e-02 4.315e-01 0.074 0.940717
## URBANICITY_z_Highly.Rural..Rural -2.477e+00 4.133e-01 -5.992 2.07e-09 ***
## YOJ_NA -1.919e-01 4.512e-01 -0.425 0.670663
## INCOME_NA 8.594e-02 5.524e-01 0.156 0.876361
## CAR_AGE_NA 1.696e-01 4.969e-01 0.341 0.732831
## HOME_VAL_NA -2.948e-01 2.864e-01 -1.029 0.303369
## ageSquared 1.267e-05 1.282e-03 0.010 0.992110
## yojSquared 9.072e-03 6.735e-03 1.347 0.177936
## income_log 5.120e-02 2.819e-01 0.182 0.855856
## homeval_log -1.411e+00 1.369e+00 -1.030 0.302792
## travtime_log -1.390e-01 5.768e-01 -0.241 0.809585
## bluebook_log 1.054e-01 5.075e-01 0.208 0.835546
## carage_log 1.404e-02 4.784e-01 0.029 0.976584
## oldclaim_log 1.162e-01 1.620e-01 0.718 0.473048
## clm_freq_log -9.762e-01 3.031e+00 -0.322 0.747380
## mvr_pts_log 5.412e-02 5.045e-01 0.107 0.914570
## tif_log 1.677e-01 5.201e-01 0.323 0.747054
## kidsdriv_log 1.567e+00 2.191e+00 0.715 0.474405
## homekids_log 3.006e+00 1.497e+00 2.007 0.044727 *
## inter 2.318e-02 3.367e-02 0.688 0.491168
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 740.94 on 639 degrees of freedom
## Residual deviance: 546.02 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 658.02
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  24
##           1  17  21
##
##           Accuracy : 0.767
##           95% CI : (0.6975, 0.8273)
##           No Information Rate : 0.7443
##           P-Value [Acc > NIR] : 0.2757
##
##           Kappa : 0.355
##
## Mcnemar's Test P-Value : 0.3487
##
##           Sensitivity : 0.8702
##           Specificity : 0.4667
##           Pos Pred Value : 0.8261
##           Neg Pred Value : 0.5526
##           Prevalence : 0.7443
##           Detection Rate : 0.6477
##           Detection Prevalence : 0.7841
##           Balanced Accuracy : 0.6684
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.751823579304495"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7518
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.87208  -0.66535  -0.35758  -0.08857   2.83599
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.074e+01  1.519e+01   0.707  0.4793
## KIDSDRIV      -4.477e-01  1.923e+00  -0.233  0.8159
## AGE           -1.310e-01  1.036e-01  -1.265  0.2060
## HOMEKIDS      -7.860e-01  6.553e-01  -1.199  0.2304
## YOJ           -2.075e-01  1.302e-01  -1.593  0.1112
## INCOME        -2.313e-05  1.351e-05  -1.712  0.0868
## HOME_VAL       1.228e-05  9.042e-06   1.358  0.1746
## TRAVTIME       2.102e-02  2.263e-02   0.929  0.3530
## BLUEBOOK       2.616e-05  4.244e-05   0.616  0.5377
```

```

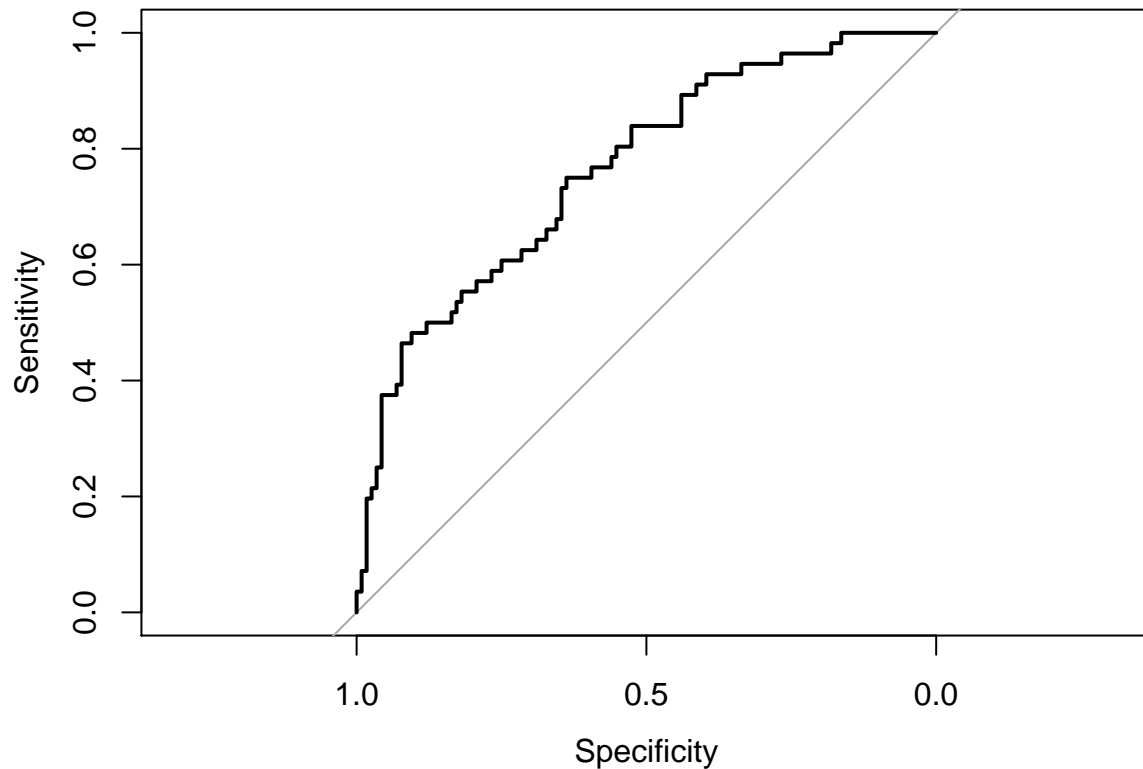
## TIF -3.474e-02 9.560e-02 -0.363 0.7163
## OLDCLAIM 7.899e-06 2.134e-05 0.370 0.7112
## CLM_FREQ -4.900e-01 1.087e+00 -0.451 0.6521
## MVR_PTS 1.813e-01 1.811e-01 1.001 0.3167
## CAR_AGE 3.532e-04 6.989e-02 0.005 0.9960
## PARENT1_Yes 4.817e-01 4.709e-01 1.023 0.3063
## MSTATUS_Yes -6.641e-01 3.418e-01 -1.943 0.0520 .
## SEX_z_F -1.216e+00 4.933e-01 -2.465 0.0137 *
## EDUCATION_.High.School -9.159e-02 8.579e-01 -0.107 0.9150
## EDUCATION_Bachelors -3.077e-01 7.199e-01 -0.427 0.6690
## EDUCATION_Masters 7.051e-01 6.011e-01 1.173 0.2408
## EDUCATION_z_High.School 2.410e-02 7.763e-01 0.031 0.9752
## JOB_ -1.459e+00 7.245e-01 -2.014 0.0440 *
## JOB_Clerical -3.735e-01 4.874e-01 -0.766 0.4434
## JOB_Doctor -4.696e-01 1.007e+00 -0.466 0.6409
## JOB_Home.Maker 1.098e-01 6.596e-01 0.167 0.8677
## JOB_Lawyer -1.170e+00 6.743e-01 -1.735 0.0828 .
## JOB_Manager -1.040e+00 5.184e-01 -2.007 0.0448 *
## JOB_Student -6.840e-01 7.115e-01 -0.961 0.3364
## JOB_z_Blue.Collar -3.296e-01 4.511e-01 -0.731 0.4649
## CAR_USE_Commercial 7.772e-01 3.558e-01 2.185 0.0289 *
## CAR_TYPE_Panel.Truck 1.988e-03 6.254e-01 0.003 0.9975
## CAR_TYPE_Pickup 8.567e-01 4.088e-01 2.095 0.0361 *
## CAR_TYPE_Sports.Car 2.446e+00 5.630e-01 4.344 1.40e-05 ***
## CAR_TYPE_Van 5.231e-01 4.911e-01 1.065 0.2868
## CAR_TYPE_z_SUV 2.063e+00 5.103e-01 4.043 5.27e-05 ***
## RED_CAR_no -8.556e-02 3.290e-01 -0.260 0.7948
## REVOKED_Yes 4.416e-01 4.121e-01 1.071 0.2840
## URBANICITY_z_Highly.Rural..Rural -1.945e+00 3.966e-01 -4.904 9.39e-07 ***
## YOJ_NA -2.704e-01 4.383e-01 -0.617 0.5373
## INCOME_NA 3.930e-01 5.628e-01 0.698 0.4850
## CAR_AGE_NA -1.675e-01 4.964e-01 -0.337 0.7359
## HOME_VAL_NA -3.404e-01 2.847e-01 -1.196 0.2318
## ageSquared 1.295e-03 1.120e-03 1.157 0.2474
## yojSquared 1.214e-02 6.656e-03 1.823 0.0683 .
## income_log -1.331e-02 2.802e-01 -0.048 0.9621
## homeval_log -7.995e-01 1.411e+00 -0.567 0.5709
## travtime_log 6.933e-02 6.460e-01 0.107 0.9145
## bluebook_log 6.906e-03 5.254e-01 0.013 0.9895
## carage_log -2.113e-01 4.591e-01 -0.460 0.6453
## oldclaim_log -4.954e-02 1.709e-01 -0.290 0.7719
## clm_freq_log 1.721e+00 3.279e+00 0.525 0.5998
## mvr_pts_log -4.983e-01 5.241e-01 -0.951 0.3417
## tif_log -7.400e-02 5.300e-01 -0.140 0.8890
## kidsdriv_log -9.209e-01 2.282e+00 -0.404 0.6866
## homekids_log 1.637e+00 1.425e+00 1.149 0.2506
## inter 3.971e-02 3.636e-02 1.092 0.2748
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 719.86 on 643 degrees of freedom
## Residual deviance: 538.65 on 588 degrees of freedom

```

```

## AIC: 650.65
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 110  35
##           1   6  21
##
##           Accuracy : 0.7616
##           95% CI : (0.6908, 0.8232)
##           No Information Rate : 0.6744
##           P-Value [Acc > NIR] : 0.007938
##
##           Kappa : 0.3733
##
## Mcnemar's Test P-Value : 1.226e-05
##
##           Sensitivity : 0.9483
##           Specificity : 0.3750
##           Pos Pred Value : 0.7586
##           Neg Pred Value : 0.7778
##           Prevalence : 0.6744
##           Detection Rate : 0.6395
##           Detection Prevalence : 0.8430
##           Balanced Accuracy : 0.6616
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.764470443349754"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 116 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7645
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1799  -0.6571  -0.3667   0.5120   3.0534
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.368e+01  1.462e+01   1.620 0.105220
## KIDSDRIV      -2.010e+00  2.093e+00  -0.960 0.336839
## AGE           -2.288e-01  1.086e-01  -2.107 0.035132 *
## HOMEKIDS      -7.456e-01  6.905e-01  -1.080 0.280201
## YOJ           -3.007e-01  1.332e-01  -2.258 0.023963 *
## INCOME        -2.237e-05  1.314e-05  -1.703 0.088592 .
## HOME_VAL       1.422e-05  8.811e-06   1.614 0.106619
## TRAVTIME       4.331e-02  2.038e-02   2.125 0.033558 *
## BLUEBOOK       4.508e-05  4.030e-05   1.119 0.263314
```

```

## TIF -7.055e-02 9.427e-02 -0.748 0.454235
## OLDCLAIM -1.119e-05 2.105e-05 -0.531 0.595092
## CLM_FREQ 3.374e-01 9.277e-01 0.364 0.716108
## MVR_PTS 4.766e-02 1.657e-01 0.288 0.773664
## CAR_AGE 1.186e-02 7.170e-02 0.165 0.868659
## PARENT1_Yes 4.546e-01 4.916e-01 0.925 0.355130
## MSTATUS_Yes -6.114e-01 3.289e-01 -1.859 0.063009 .
## SEX_z_F -6.303e-01 4.434e-01 -1.422 0.155134
## EDUCATION_.High.School 3.298e-01 7.910e-01 0.417 0.676761
## EDUCATION_Bachelors 1.274e-01 6.656e-01 0.191 0.848238
## EDUCATION_Masters 1.128e+00 5.987e-01 1.884 0.059593 .
## EDUCATION_z_High.School 2.904e-01 7.221e-01 0.402 0.687619
## JOB_ -1.057e+00 6.819e-01 -1.550 0.121029
## JOB_Clerical 1.598e-02 4.776e-01 0.033 0.973307
## JOB_Doctor -9.225e-02 9.083e-01 -0.102 0.919099
## JOB_Home.Maker -4.328e-01 6.582e-01 -0.658 0.510853
## JOB_Lawyer -9.783e-01 6.636e-01 -1.474 0.140434
## JOB_Manager -8.968e-01 4.938e-01 -1.816 0.069357 .
## JOB_Student -8.026e-01 7.536e-01 -1.065 0.286866
## JOB_z_Blue.Collar -5.691e-02 4.468e-01 -0.127 0.898645
## CAR_USE_Commercial 7.348e-01 3.544e-01 2.073 0.038150 *
## CAR_TYPE_Panel.Truck -1.936e-01 6.187e-01 -0.313 0.754376
## CAR_TYPE_Pickup 7.435e-01 4.037e-01 1.842 0.065499 .
## CAR_TYPE_Sports.Car 1.608e+00 5.099e-01 3.154 0.001611 **
## CAR_TYPE_Van 6.167e-01 4.574e-01 1.348 0.177566
## CAR_TYPE_z_SUV 1.579e+00 4.523e-01 3.490 0.000482 ***
## RED_CAR_no -6.446e-02 3.291e-01 -0.196 0.844720
## REVOKED_Yes 3.696e-01 4.218e-01 0.876 0.380901
## URBANICITY_z_Highly.Rural..Rural -2.304e+00 4.222e-01 -5.458 4.82e-08 ***
## YOJ_NA 1.387e-01 4.437e-01 0.313 0.754640
## INCOME_NA 5.663e-01 5.909e-01 0.958 0.337852
## CAR_AGE_NA -1.727e-02 5.194e-01 -0.033 0.973472
## HOME_VAL_NA -3.636e-02 2.906e-01 -0.125 0.900429
## ageSquared 2.310e-03 1.172e-03 1.971 0.048707 *
## yojSquared 1.504e-02 6.791e-03 2.215 0.026785 *
## income_log 1.957e-02 2.962e-01 0.066 0.947335
## homeval_log -1.409e+00 1.363e+00 -1.034 0.301245
## travtime_log -7.798e-01 5.659e-01 -1.378 0.168178
## bluebook_log -2.733e-01 4.789e-01 -0.571 0.568208
## carage_log -5.134e-01 4.690e-01 -1.095 0.273710
## oldclaim_log 1.428e-01 1.581e-01 0.904 0.366247
## clm_freq_log -1.034e+00 2.893e+00 -0.357 0.720879
## mvr_pts_log 3.531e-02 4.925e-01 0.072 0.942842
## tif_log 1.594e-01 5.300e-01 0.301 0.763641
## kidsdriv_log 2.819e+00 2.515e+00 1.121 0.262405
## homekids_log 1.250e+00 1.490e+00 0.839 0.401604
## inter 2.631e-02 3.618e-02 0.727 0.467111
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 738.90 on 639 degrees of freedom
## Residual deviance: 546.23 on 584 degrees of freedom

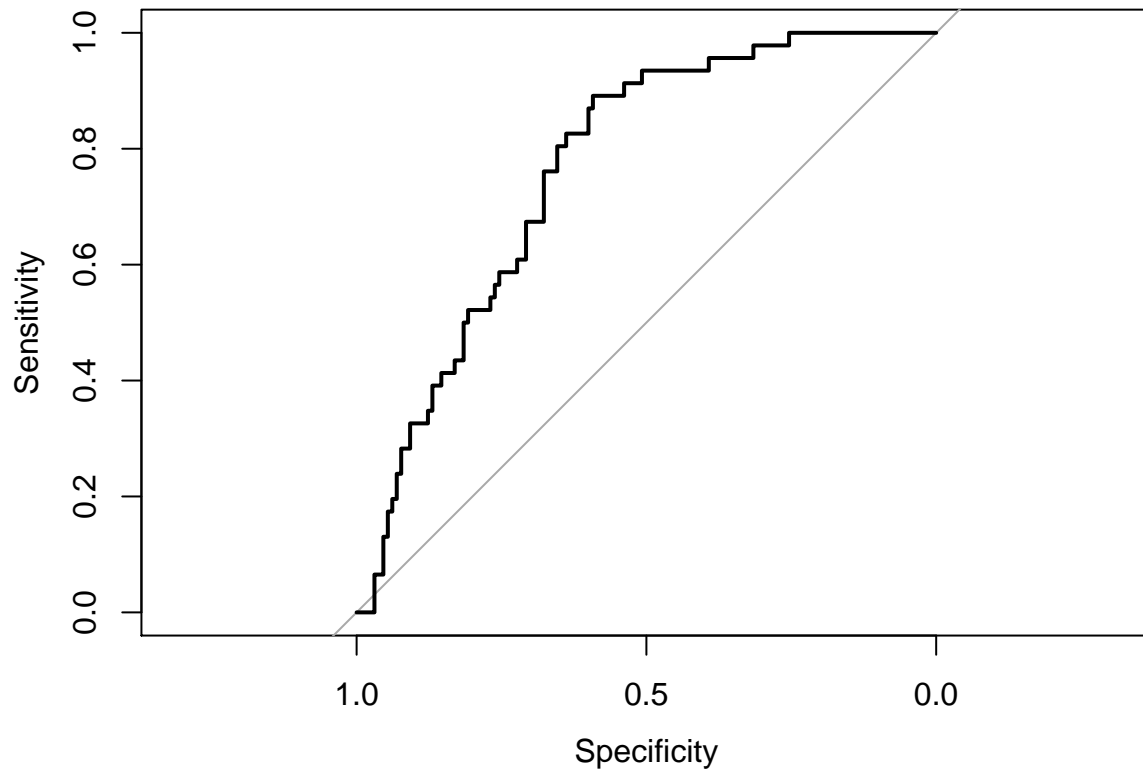
```



```

## (1 observation deleted due to missingness)
## AIC: 658.23
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 112  28
##           1  18  18
##
##           Accuracy : 0.7386
##           95% CI : (0.6672, 0.8019)
##           No Information Rate : 0.7386
##           P-Value [Acc > NIR] : 0.5396
##
##           Kappa : 0.2719
##
## Mcnemar's Test P-Value : 0.1845
##
##           Sensitivity : 0.8615
##           Specificity : 0.3913
##           Pos Pred Value : 0.8000
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7386
##           Detection Rate : 0.6364
##           Detection Prevalence : 0.7955
##           Balanced Accuracy : 0.6264
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.766555183946488"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 130 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7666
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2403  -0.6838  -0.3626   0.5593   2.9657
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.420e+01  1.458e+01   2.346 0.018974 *
## KIDSDRIV      -1.309e+00  1.721e+00  -0.760 0.447046
## AGE           -1.282e-01  1.035e-01  -1.238 0.215543
## HOMEKIDS       -5.039e-01  6.993e-01  -0.721 0.471104
## YOJ            -1.955e-01  1.334e-01  -1.466 0.142740
## INCOME         -2.494e-05  1.323e-05  -1.885 0.059476 .
## HOME_VAL        1.876e-05  8.881e-06   2.113 0.034619 *
## TRAVTIME        3.465e-02  2.115e-02   1.638 0.101421
## BLUEBOOK        6.144e-05  3.948e-05   1.556 0.119598
```

```

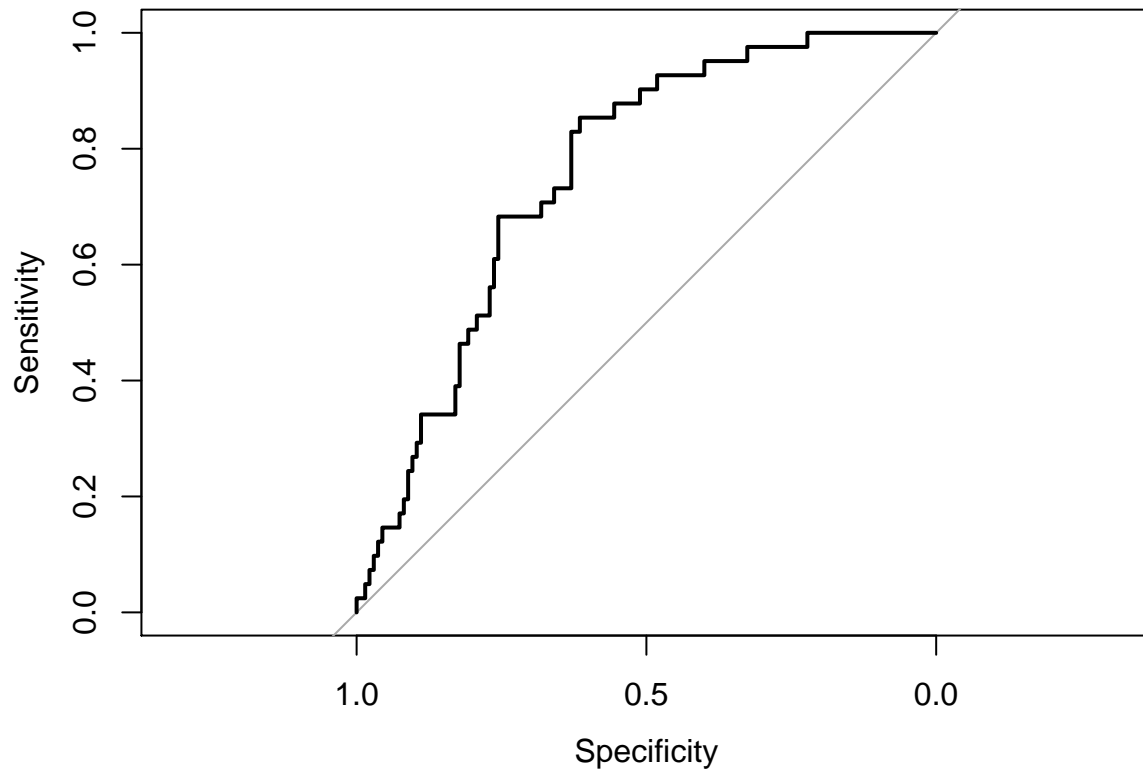
## TIF -4.630e-02 9.335e-02 -0.496 0.619923
## OLDCLAIM -4.515e-07 2.173e-05 -0.021 0.983422
## CLM_FREQ 3.102e-01 8.966e-01 0.346 0.729366
## MVRPTS 2.062e-01 1.693e-01 1.219 0.223016
## CAR_AGE -2.565e-02 6.775e-02 -0.379 0.705039
## PARENT1_Yes 7.967e-01 4.775e-01 1.669 0.095187 .
## MSTATUS_Yes -6.136e-01 3.321e-01 -1.848 0.064634 .
## SEX_z_F -8.119e-01 4.502e-01 -1.803 0.071350 .
## EDUCATION_.High.School 3.734e-01 7.915e-01 0.472 0.637125
## EDUCATION_Bachelors 3.470e-01 6.712e-01 0.517 0.605195
## EDUCATION_Masters 1.129e+00 5.757e-01 1.961 0.049888 *
## EDUCATION_z_High.School 6.282e-01 7.303e-01 0.860 0.389703
## JOB_ -5.273e-01 6.833e-01 -0.772 0.440292
## JOB_Clerical -1.956e-02 4.794e-01 -0.041 0.967453
## JOB_Doctor 4.813e-01 9.120e-01 0.528 0.597727
## JOB_Home.Maker -5.319e-01 6.815e-01 -0.780 0.435144
## JOB_Lawyer -6.783e-01 6.496e-01 -1.044 0.296381
## JOB_Manager -5.772e-01 4.794e-01 -1.204 0.228575
## JOB_Student -7.325e-01 7.278e-01 -1.006 0.314184
## JOB_z_Blue.Collar -1.114e-01 4.501e-01 -0.247 0.804523
## CAR_USE_Commercial 3.175e-01 3.472e-01 0.914 0.360478
## CAR_TYPE_Panel.Truck 3.107e-01 6.102e-01 0.509 0.610666
## CAR_TYPE_Pickup 1.304e+00 4.049e-01 3.220 0.001282 **
## CAR_TYPE_Sports.Car 2.006e+00 5.430e-01 3.694 0.000221 ***
## CAR_TYPE_Van 4.185e-01 4.789e-01 0.874 0.382239
## CAR_TYPE_z_SUV 1.676e+00 4.755e-01 3.525 0.000424 ***
## RED_CAR_no -7.373e-02 3.319e-01 -0.222 0.824191
## REVOKED_Yes 6.418e-01 4.247e-01 1.511 0.130709
## URBANICITY_z_Highly.Rural..Rural -2.370e+00 4.301e-01 -5.511 3.58e-08 ***
## YOJ_NA -2.690e-01 4.136e-01 -0.650 0.515512
## INCOME_NA 8.141e-01 5.999e-01 1.357 0.174776
## CAR_AGE_NA -3.045e-01 4.838e-01 -0.629 0.529132
## HOME_VAL_NA 1.510e-02 2.908e-01 0.052 0.958593
## ageSquared 1.400e-03 1.120e-03 1.250 0.211200
## yojSquared 7.720e-03 6.765e-03 1.141 0.253738
## income_log 2.849e-01 3.125e-01 0.912 0.361926
## homeval_log -2.823e+00 1.420e+00 -1.988 0.046766 *
## travtime_log -3.314e-01 5.950e-01 -0.557 0.577550
## bluebook_log -4.176e-01 4.769e-01 -0.876 0.381177
## carage_log -3.876e-02 4.508e-01 -0.086 0.931475
## oldclaim_log 4.990e-02 1.541e-01 0.324 0.746021
## clm_freq_log -5.445e-01 2.785e+00 -0.195 0.845008
## mvr_pts_log -3.781e-01 5.001e-01 -0.756 0.449562
## tif_log -1.430e-01 5.276e-01 -0.271 0.786352
## kidsdriv_log 2.600e+00 2.190e+00 1.187 0.235164
## homekids_log 1.074e+00 1.499e+00 0.717 0.473631
## inter 1.114e-02 3.027e-02 0.368 0.712816
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 748.95 on 639 degrees of freedom
## Residual deviance: 550.34 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 662.34
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  27
##           1  19  14
##
##           Accuracy : 0.7386
##           95% CI : (0.6672, 0.8019)
##           No Information Rate : 0.767
##           P-Value [Acc > NIR] : 0.8369
##
##           Kappa : 0.2154
##
## Mcnemar's Test P-Value : 0.3020
##
##           Sensitivity : 0.8593
##           Specificity : 0.3415
##           Pos Pred Value : 0.8112
##           Neg Pred Value : 0.4242
##           Prevalence : 0.7670
##           Detection Rate : 0.6591
##           Detection Prevalence : 0.8125
##           Balanced Accuracy : 0.6004
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.759168925022584"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 135 controls (dfPred_raw$class 0) < 41 cases (dfPred_raw$class 1).
## Area under the curve: 0.7592
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0948  -0.6440  -0.3525   0.5454   3.0643
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.171e+01  1.589e+01   0.737 0.461310
## KIDSDRIV      -9.870e-01  1.876e+00  -0.526 0.598711
## AGE           -1.542e-01  1.069e-01  -1.443 0.149134
## HOMEKIDS      -9.965e-01  6.792e-01  -1.467 0.142354
## YOJ           -1.290e-01  1.276e-01  -1.011 0.312078
## INCOME        -1.463e-05  1.339e-05  -1.093 0.274574
## HOME_VAL       6.119e-06  9.410e-06   0.650 0.515495
## TRAVTIME       5.882e-02  2.211e-02   2.661 0.007802 **
## BLUEBOOK       5.654e-05  3.866e-05   1.463 0.143546
```

```

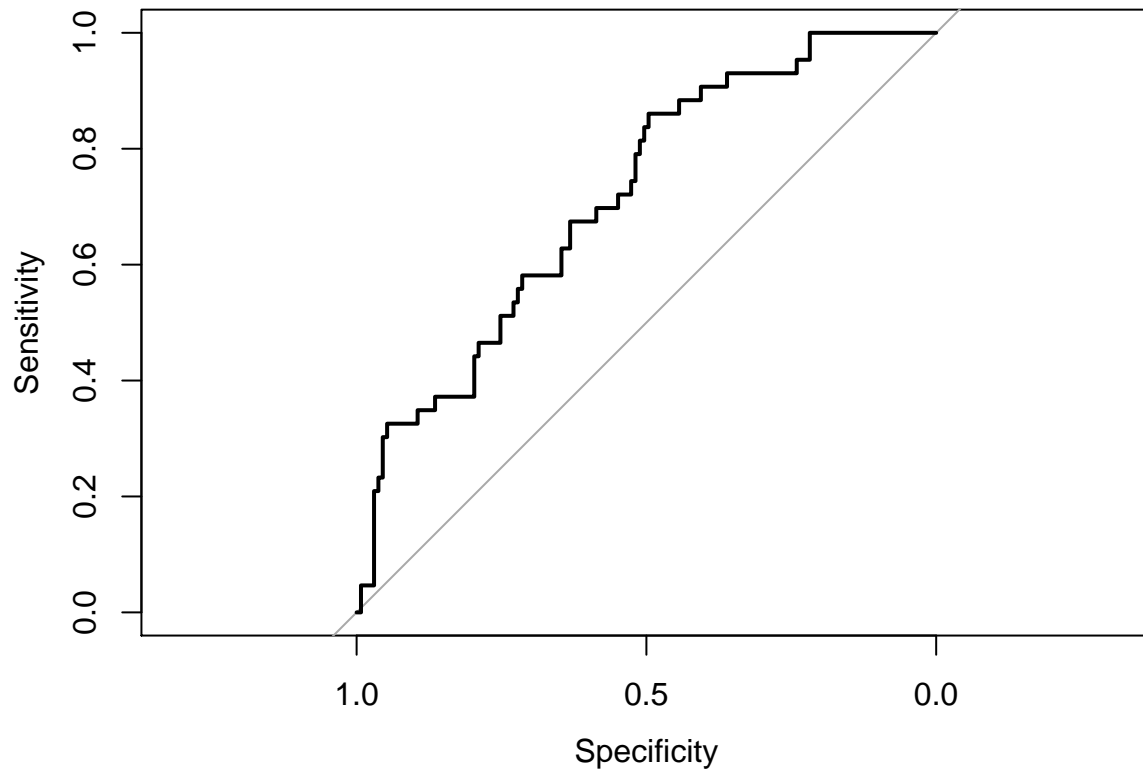
## TIF -2.291e-02 8.927e-02 -0.257 0.797482
## OLDCLAIM -3.490e-06 2.191e-05 -0.159 0.873422
## CLM_FREQ -2.069e-01 9.773e-01 -0.212 0.832309
## MVRPTS 2.856e-02 1.717e-01 0.166 0.867849
## CAR_AGE -6.158e-02 7.393e-02 -0.833 0.404899
## PARENT1_Yes 3.679e-01 4.732e-01 0.777 0.436892
## MSTATUS_Yes -7.192e-01 3.336e-01 -2.155 0.031125 *
## SEX_z_F -5.897e-01 4.729e-01 -1.247 0.212417
## EDUCATION_.High.School -1.301e-02 7.963e-01 -0.016 0.986961
## EDUCATION_Bachelors -6.045e-02 6.729e-01 -0.090 0.928414
## EDUCATION_Masters 6.700e-01 5.733e-01 1.169 0.242549
## EDUCATION_z_High.School 3.430e-01 7.223e-01 0.475 0.634872
## JOB_ -1.232e+00 6.984e-01 -1.764 0.077669 .
## JOB_Clerical -1.507e-01 4.853e-01 -0.311 0.756119
## JOB_Doctor -2.659e-02 8.950e-01 -0.030 0.976303
## JOB_Home.Maker -2.030e-01 6.643e-01 -0.306 0.759981
## JOB_Lawyer -9.609e-02 6.352e-01 -0.151 0.879759
## JOB_Manager -9.112e-01 4.706e-01 -1.936 0.052826 .
## JOB_Student -8.213e-01 7.510e-01 -1.094 0.274099
## JOB_z_Blue.Collar -2.123e-01 4.427e-01 -0.480 0.631503
## CAR_USE_Commercial 6.018e-01 3.527e-01 1.706 0.087953 .
## CAR_TYPE_Panel.Truck 3.404e-01 6.244e-01 0.545 0.585631
## CAR_TYPE_Pickup 1.363e+00 4.137e-01 3.294 0.000987 ***
## CAR_TYPE_Sports.Car 1.948e+00 5.327e-01 3.657 0.000255 ***
## CAR_TYPE_Van 8.236e-01 4.681e-01 1.759 0.078525 .
## CAR_TYPE_z_SUV 2.130e+00 4.854e-01 4.387 1.15e-05 ***
## RED_CAR_no -5.382e-01 3.476e-01 -1.548 0.121535
## REVOKED_Yes 5.535e-01 3.971e-01 1.394 0.163400
## URBANICITY_z_Highly.Rural..Rural -2.384e+00 4.182e-01 -5.701 1.19e-08 ***
## YOJ_NA -1.139e-01 4.504e-01 -0.253 0.800373
## INCOME_NA 2.835e-02 5.249e-01 0.054 0.956930
## CAR_AGE_NA 2.157e-01 5.057e-01 0.426 0.669799
## HOME_VAL_NA -3.376e-01 2.879e-01 -1.173 0.240946
## ageSquared 1.674e-03 1.149e-03 1.457 0.145252
## yojSquared 7.004e-03 6.591e-03 1.063 0.287919
## income_log -5.271e-04 2.938e-01 -0.002 0.998569
## homeval_log -5.250e-01 1.500e+00 -0.350 0.726273
## travtime_log -9.206e-01 6.037e-01 -1.525 0.127256
## bluebook_log -2.879e-01 4.718e-01 -0.610 0.541660
## carage_log 1.256e-01 4.833e-01 0.260 0.794940
## oldclaim_log 5.936e-02 1.624e-01 0.366 0.714733
## clm_freq_log 6.172e-01 3.029e+00 0.204 0.838539
## mvr_pts_log -1.123e-01 5.017e-01 -0.224 0.822914
## tif_log -2.628e-03 5.121e-01 -0.005 0.995905
## kidsdriv_log 1.299e+00 2.232e+00 0.582 0.560767
## homekids_log 2.280e+00 1.457e+00 1.565 0.117612
## inter 1.974e-02 3.526e-02 0.560 0.575614
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 744.97 on 639 degrees of freedom
## Residual deviance: 538.52 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 650.52
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 111  27
##           1  22  16
##
##           Accuracy : 0.7216
##           95% CI : (0.6492, 0.7864)
##           No Information Rate : 0.7557
##           P-Value [Acc > NIR] : 0.8721
##
##           Kappa : 0.2151
##
## Mcnemar's Test P-Value : 0.5677
##
##           Sensitivity : 0.8346
##           Specificity : 0.3721
##           Pos Pred Value : 0.8043
##           Neg Pred Value : 0.4211
##           Prevalence : 0.7557
##           Detection Rate : 0.6307
##           Detection Prevalence : 0.7841
##           Balanced Accuracy : 0.6033
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.71655883895786"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7166
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0747  -0.6558  -0.3248   0.1660   3.1180
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.118e+01  1.501e+01   2.076 0.037852 *
## KIDSDRIV      -1.291e+00  1.901e+00  -0.679 0.496986
## AGE           -1.878e-01  1.088e-01  -1.726 0.084355 .
## HOMEKIDS      -5.162e-01  6.793e-01  -0.760 0.447345
## YOJ           -1.515e-01  1.302e-01  -1.163 0.244682
## INCOME        -2.650e-05  1.388e-05  -1.910 0.056174 .
## HOME_VAL       1.663e-05  9.206e-06   1.806 0.070935 .
## TRAVTIME       3.547e-02  2.189e-02   1.620 0.105146
## BLUEBOOK       4.316e-05  3.955e-05   1.091 0.275116
```



```

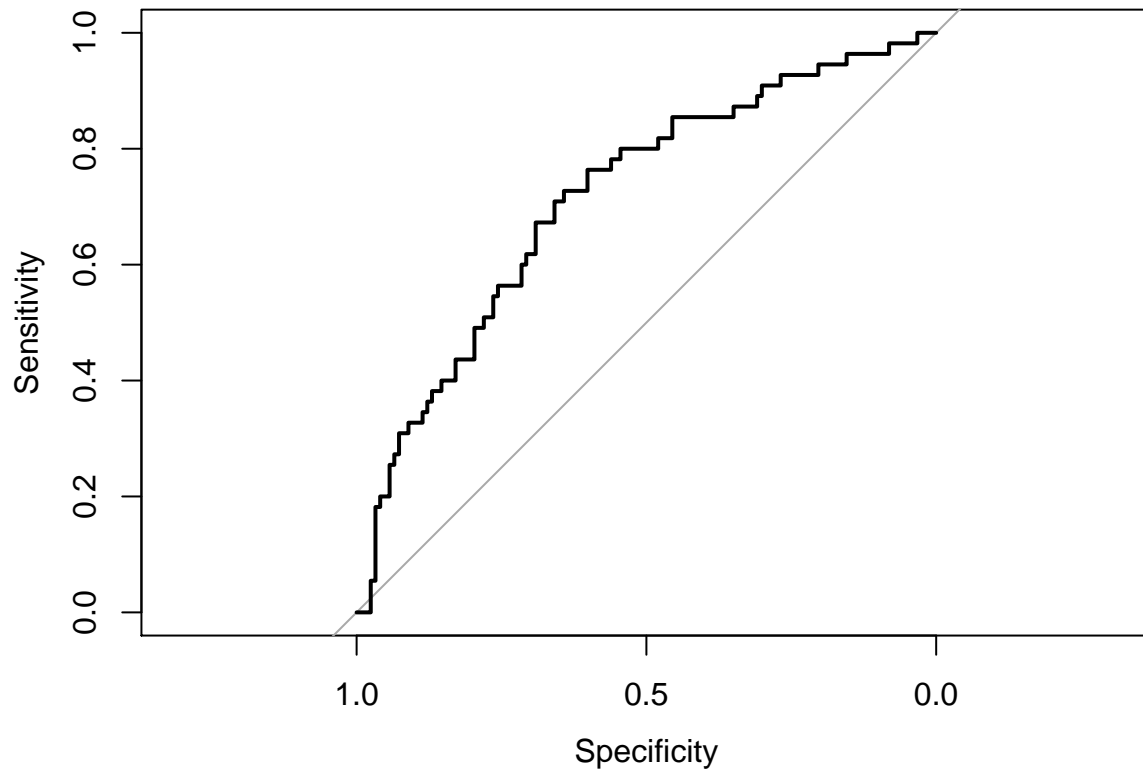
## TIF -4.517e-02 9.391e-02 -0.481 0.630535
## OLDCLAIM -1.343e-05 2.223e-05 -0.604 0.545701
## CLM_FREQ 1.121e+00 1.005e+00 1.115 0.264967
## MVRPTS -1.050e-01 1.749e-01 -0.601 0.548056
## CAR_AGE -5.843e-02 7.372e-02 -0.793 0.428050
## PARENT1_Yes 1.044e+00 4.836e-01 2.159 0.030848 *
## MSTATUS_Yes -3.286e-01 3.475e-01 -0.946 0.344351
## SEX_z_F -7.809e-01 4.835e-01 -1.615 0.106266
## EDUCATION_.High.School -1.655e-01 8.012e-01 -0.207 0.836368
## EDUCATION_Bachelors -1.799e-01 6.814e-01 -0.264 0.791722
## EDUCATION_Masters 1.110e+00 5.954e-01 1.865 0.062201 .
## EDUCATION_z_High.School 3.562e-01 7.339e-01 0.485 0.627447
## JOB_ -6.851e-01 6.956e-01 -0.985 0.324675
## JOB_Clerical -4.463e-01 5.159e-01 -0.865 0.386970
## JOB_Doctor -5.924e-02 9.693e-01 -0.061 0.951264
## JOB_Home.Maker -4.020e-01 6.857e-01 -0.586 0.557686
## JOB_Lawyer -6.704e-01 7.280e-01 -0.921 0.357133
## JOB_Manager -1.023e+00 5.422e-01 -1.887 0.059110 .
## JOB_Student -1.316e+00 7.544e-01 -1.745 0.080991 .
## JOB_z_Blue.Collar -8.721e-02 4.688e-01 -0.186 0.852411
## CAR_USE_Commercial 9.678e-01 3.639e-01 2.660 0.007819 **
## CAR_TYPE_Panel.Truck 1.068e-01 6.440e-01 0.166 0.868331
## CAR_TYPE_Pickup 8.244e-01 4.210e-01 1.958 0.050193 .
## CAR_TYPE_Sports.Car 1.174e+00 5.593e-01 2.099 0.035800 *
## CAR_TYPE_Van 6.686e-01 4.840e-01 1.381 0.167192
## CAR_TYPE_z_SUV 1.866e+00 4.808e-01 3.880 0.000104 ***
## RED_CAR_no -1.874e-01 3.532e-01 -0.531 0.595741
## REVOKED_Yes 8.189e-01 4.180e-01 1.959 0.050127 .
## URBANICITY_z_Highly.Rural..Rural -2.589e+00 4.700e-01 -5.509 3.62e-08 ***
## YOJ_NA -2.198e-01 4.403e-01 -0.499 0.617619
## INCOME_NA 1.823e-01 5.717e-01 0.319 0.749784
## CAR_AGE_NA -3.428e-01 4.945e-01 -0.693 0.488204
## HOME_VAL_NA -1.047e-01 3.027e-01 -0.346 0.729515
## ageSquared 2.235e-03 1.179e-03 1.896 0.057951 .
## yojSquared 7.273e-03 6.709e-03 1.084 0.278326
## income_log 1.004e-01 2.885e-01 0.348 0.727699
## homeval_log -2.125e+00 1.410e+00 -1.507 0.131893
## travtime_log -5.542e-01 6.238e-01 -0.888 0.374276
## bluebook_log -4.906e-01 4.620e-01 -1.062 0.288313
## carage_log -1.244e-01 4.721e-01 -0.263 0.792214
## oldclaim_log 1.862e-01 1.694e-01 1.099 0.271904
## clm_freq_log -2.952e+00 3.110e+00 -0.949 0.342465
## mvr_pts_log 3.790e-01 5.157e-01 0.735 0.462397
## tif_log 1.156e-01 5.367e-01 0.215 0.829459
## kidsdriv_log 1.865e+00 2.327e+00 0.802 0.422749
## homekids_log 1.102e+00 1.481e+00 0.744 0.456684
## inter 1.964e-02 3.578e-02 0.549 0.583120
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 718.64 on 637 degrees of freedom
## Residual deviance: 512.90 on 582 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 624.9
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 106  34
##           1  17  21
##
##           Accuracy : 0.7135
##           95% CI : (0.6411, 0.7786)
##           No Information Rate : 0.691
##           P-Value [Acc > NIR] : 0.28757
##
##           Kappa : 0.2664
##
## Mcnemar's Test P-Value : 0.02506
##
##           Sensitivity : 0.8618
##           Specificity : 0.3818
##           Pos Pred Value : 0.7571
##           Neg Pred Value : 0.5526
##           Prevalence : 0.6910
##           Detection Rate : 0.5955
##           Detection Prevalence : 0.7865
##           Balanced Accuracy : 0.6218
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.715594974131559"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 123 controls (dfPred_raw$class 0) < 55 cases (dfPred_raw$class 1).
## Area under the curve: 0.7156
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2572  -0.6670  -0.3432   0.4391   3.1184
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.335e+01  1.456e+01   1.604  0.10874
## KIDSDRIV      -1.440e+00  2.108e+00  -0.683  0.49447
## AGE           -4.086e-03  1.173e-01  -0.035  0.97221
## HOMEKIDS      -5.766e-01  6.996e-01  -0.824  0.40982
## YOJ           -1.883e-01  1.292e-01  -1.457  0.14519
## INCOME        -2.285e-05  1.366e-05  -1.673  0.09429
## HOME_VAL       1.419e-05  9.082e-06   1.562  0.11831
## TRAVTIME       3.515e-02  2.169e-02   1.621  0.10506
## BLUEBOOK       3.952e-05  3.766e-05   1.049  0.29396
```

```

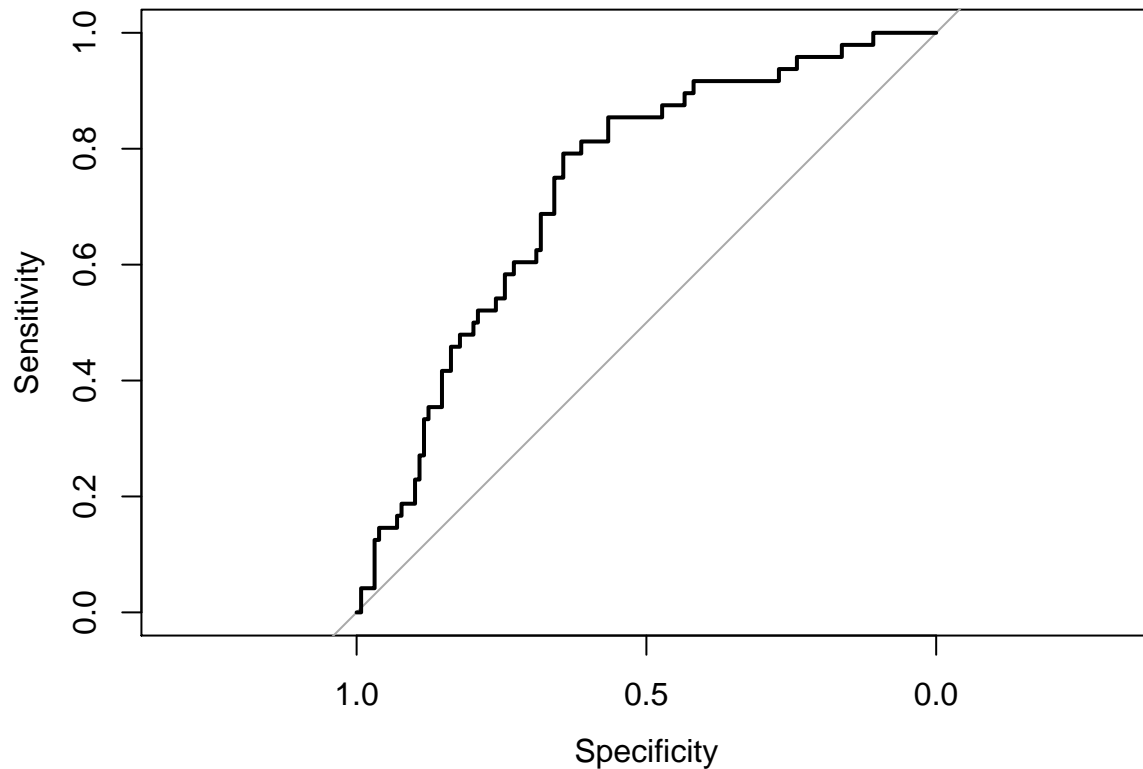
## TIF 3.646e-02 8.825e-02 0.413 0.67947
## OLDCLAIM -1.348e-05 2.266e-05 -0.595 0.55192
## CLM_FREQ 8.045e-01 9.277e-01 0.867 0.38581
## MVR_PTS 8.578e-02 1.731e-01 0.495 0.62027
## CAR_AGE -1.727e-02 7.028e-02 -0.246 0.80593
## PARENT1_Yes 7.925e-01 4.873e-01 1.626 0.10389
## MSTATUS_Yes -5.260e-01 3.297e-01 -1.595 0.11067
## SEX_z_F -7.053e-01 4.692e-01 -1.503 0.13274
## EDUCATION_.High.School -4.140e-01 8.342e-01 -0.496 0.61967
## EDUCATION_Bachelors -3.133e-01 6.910e-01 -0.453 0.65025
## EDUCATION_Masters 4.344e-01 5.668e-01 0.766 0.44339
## EDUCATION_z_High.School 1.595e-01 7.520e-01 0.212 0.83202
## JOB_ -1.309e+00 7.123e-01 -1.838 0.06601
## JOB_Clerical -4.263e-01 4.741e-01 -0.899 0.36853
## JOB_Doctor -5.287e-01 9.750e-01 -0.542 0.58764
## JOB_Home.Maker -8.677e-01 6.963e-01 -1.246 0.21273
## JOB_Lawyer -9.571e-01 6.388e-01 -1.498 0.13404
## JOB_Manager -1.317e+00 5.057e-01 -2.603 0.00924 **
## JOB_Student -9.407e-01 7.506e-01 -1.253 0.21007
## JOB_z_Blue.Collar -5.344e-01 4.447e-01 -1.202 0.22948
## CAR_USE_Commercial 8.169e-01 3.564e-01 2.292 0.02189 *
## CAR_TYPE_Panel.Truck 3.630e-01 6.311e-01 0.575 0.56510
## CAR_TYPE_Pickup 9.368e-01 4.219e-01 2.220 0.02641 *
## CAR_TYPE_Sports.Car 2.314e+00 5.343e-01 4.331 1.48e-05 ***
## CAR_TYPE_Van 6.261e-01 4.603e-01 1.360 0.17375
## CAR_TYPE_z_SUV 1.879e+00 4.673e-01 4.021 5.78e-05 ***
## RED_CAR_no -1.684e-01 3.374e-01 -0.499 0.61772
## REVOKED_Yes 7.208e-01 4.144e-01 1.740 0.08192
## URBANICITY_z_Highly.Rural..Rural -2.473e+00 4.468e-01 -5.535 3.11e-08 ***
## YOJ_NA -6.282e-01 4.383e-01 -1.433 0.15180
## INCOME_NA 2.710e-01 5.294e-01 0.512 0.60875
## CAR_AGE_NA 1.154e-01 4.882e-01 0.236 0.81316
## HOME_VAL_NA -1.872e-01 2.909e-01 -0.643 0.51997
## ageSquared -1.285e-04 1.296e-03 -0.099 0.92097
## yojSquared 9.581e-03 6.674e-03 1.435 0.15115
## income_log 1.611e-01 3.024e-01 0.533 0.59425
## homeval_log -1.948e+00 1.384e+00 -1.407 0.15938
## travtime_log -3.533e-01 6.113e-01 -0.578 0.56334
## bluebook_log -2.628e-01 4.455e-01 -0.590 0.55524
## carage_log -1.052e-01 4.685e-01 -0.225 0.82235
## oldclaim_log 1.172e-01 1.605e-01 0.730 0.46535
## clm_freq_log -1.911e+00 2.883e+00 -0.663 0.50744
## mvr_pts_log -1.473e-01 5.094e-01 -0.289 0.77243
## tif_log -4.243e-01 5.135e-01 -0.826 0.40862
## kidsdriv_log 2.784e+00 2.250e+00 1.237 0.21603
## homekids_log 9.688e-01 1.503e+00 0.645 0.51922
## inter 1.163e-02 3.895e-02 0.299 0.76525
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 734.16 on 638 degrees of freedom
## Residual deviance: 530.93 on 583 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 642.93
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  33
##           1  15  15
##
##           Accuracy : 0.7288
##           95% CI : (0.657, 0.7928)
##           No Information Rate : 0.7288
##           P-Value [Acc > NIR] : 0.53877
##
##           Kappa : 0.2224
##
## Mcnemar's Test P-Value : 0.01414
##
##           Sensitivity : 0.8837
##           Specificity : 0.3125
##           Pos Pred Value : 0.7755
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7288
##           Detection Rate : 0.6441
##           Detection Prevalence : 0.8305
##           Balanced Accuracy : 0.5981
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.734496124031008"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 129 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.7345
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3213  -0.6365  -0.3326   0.4939   2.6255
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.294e+01  1.567e+01   0.826 0.408638
## KIDSDRIV      -2.051e+00  1.878e+00  -1.092 0.274742
## AGE           -5.758e-02  1.184e-01  -0.487 0.626595
## HOMEKIDS      -1.142e+00  7.025e-01  -1.626 0.103885
## YOJ           -9.056e-02  1.450e-01  -0.624 0.532301
## INCOME        -4.694e-06  1.321e-05  -0.355 0.722394
## HOME_VAL       3.001e-06  9.140e-06   0.328 0.742615
## TRAVTIME       2.739e-02  2.330e-02   1.175 0.239893
## BLUEBOOK       7.317e-05  3.903e-05   1.875 0.060816 .
```

```

## TIF -1.249e-01 9.486e-02 -1.317 0.187797
## OLDCLAIM 2.175e-05 2.222e-05 0.979 0.327718
## CLM_FREQ 1.863e-01 9.658e-01 0.193 0.847066
## MVR_PTS -1.568e-02 1.714e-01 -0.091 0.927101
## CAR_AGE -3.263e-02 7.297e-02 -0.447 0.654724
## PARENT1_Yes 4.288e-01 4.824e-01 0.889 0.374035
## MSTATUS_Yes -4.668e-01 3.397e-01 -1.374 0.169385
## SEX_z_F -1.035e+00 4.853e-01 -2.134 0.032866 *
## EDUCATION_.High.School 4.894e-01 8.164e-01 0.599 0.548861
## EDUCATION_Bachelors -1.213e-02 7.062e-01 -0.017 0.986294
## EDUCATION_Masters 5.714e-01 5.894e-01 0.969 0.332302
## EDUCATION_z_High.School 4.805e-01 7.551e-01 0.636 0.524517
## JOB_ -3.313e-02 7.095e-01 -0.047 0.962754
## JOB_Clerical 4.105e-01 4.980e-01 0.824 0.409792
## JOB_Doctor 8.648e-01 9.507e-01 0.910 0.363009
## JOB_Home.Maker -7.347e-02 7.483e-01 -0.098 0.921786
## JOB_Lawyer -2.847e-02 7.148e-01 -0.040 0.968231
## JOB_Manager -4.479e-01 5.080e-01 -0.882 0.378001
## JOB_Student -3.763e-01 7.639e-01 -0.493 0.622266
## JOB_z_Blue.Collar 5.241e-01 4.600e-01 1.139 0.254523
## CAR_USE_Commercial 2.803e-01 3.493e-01 0.803 0.422234
## CAR_TYPE_Panel.Truck -3.786e-02 6.430e-01 -0.059 0.953043
## CAR_TYPE_Pickup 1.407e+00 4.012e-01 3.505 0.000456 ***
## CAR_TYPE_Sports.Car 1.615e+00 5.474e-01 2.951 0.003171 **
## CAR_TYPE_Van 5.907e-01 4.701e-01 1.257 0.208907
## CAR_TYPE_z_SUV 1.982e+00 4.929e-01 4.022 5.77e-05 ***
## RED_CAR_no -1.923e-01 3.384e-01 -0.568 0.569764
## REVOKED_Yes 2.555e-01 4.206e-01 0.607 0.543563
## URBANICITY_z_Highly.Rural..Rural -2.974e+00 4.645e-01 -6.402 1.54e-10 ***
## YOJ_NA -7.264e-01 4.404e-01 -1.649 0.099104 .
## INCOME_NA 2.356e-01 5.787e-01 0.407 0.683934
## CAR_AGE_NA 3.520e-01 5.394e-01 0.653 0.513997
## HOME_VAL_NA -1.941e-01 2.911e-01 -0.667 0.504801
## ageSquared 6.331e-04 1.315e-03 0.482 0.630119
## yojSquared 9.240e-04 7.411e-03 0.125 0.900773
## income_log -7.832e-02 2.873e-01 -0.273 0.785153
## homeval_log -8.130e-01 1.463e+00 -0.556 0.578501
## travtime_log 1.060e-01 6.475e-01 0.164 0.870011
## bluebook_log -4.740e-01 4.488e-01 -1.056 0.290845
## carage_log -1.073e-01 4.758e-01 -0.226 0.821551
## oldclaim_log 4.276e-02 1.652e-01 0.259 0.795808
## clm_freq_log -3.715e-01 3.020e+00 -0.123 0.902101
## mvr_pts_log 9.038e-02 5.090e-01 0.178 0.859059
## tif_log 4.931e-01 5.369e-01 0.918 0.358457
## kidsdriv_log -1.402e-01 2.213e+00 -0.063 0.949500
## homekids_log 2.818e+00 1.492e+00 1.889 0.058916 .
## inter 6.403e-02 3.724e-02 1.720 0.085496 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 750.21 on 641 degrees of freedom
## Residual deviance: 526.32 on 586 degrees of freedom

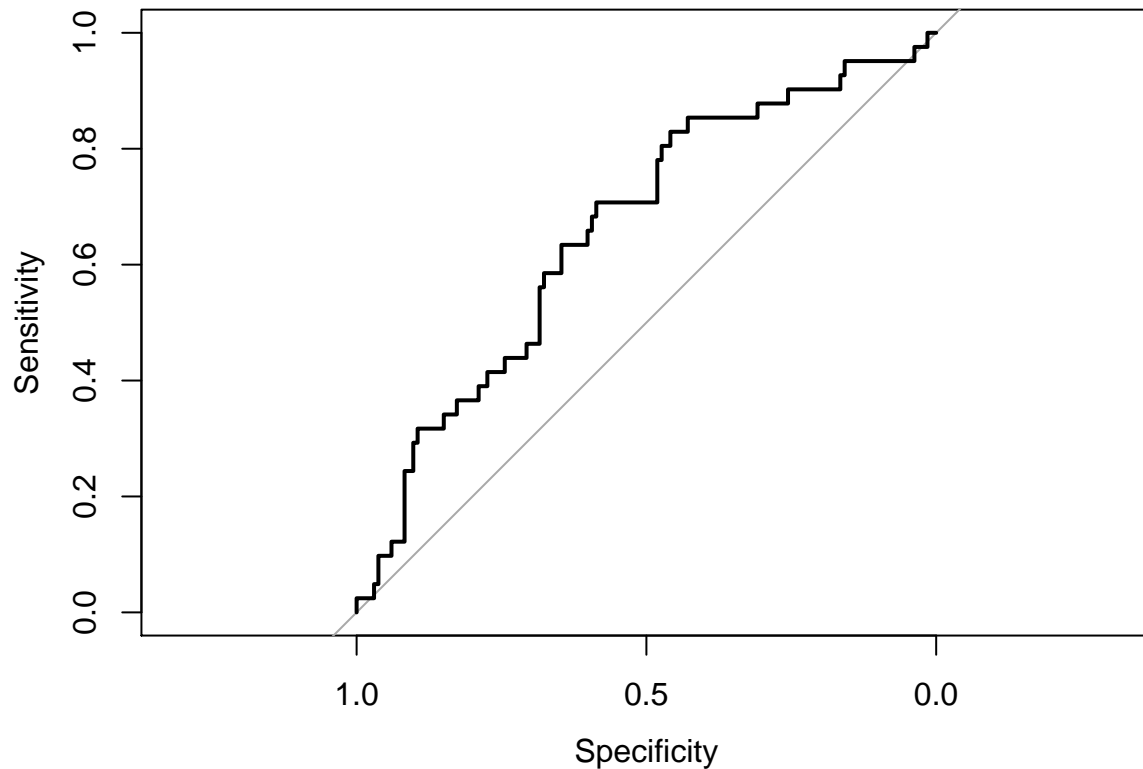
```

```

## (1 observation deleted due to missingness)
## AIC: 638.32
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 110  26
##           1  23  15
##
##           Accuracy : 0.7184
##           95% CI : (0.6453, 0.7838)
##           No Information Rate : 0.7644
##           P-Value [Acc > NIR] : 0.9332
##
##           Kappa : 0.1979
##
## Mcnemar's Test P-Value : 0.7751
##
##           Sensitivity : 0.8271
##           Specificity : 0.3659
##           Pos Pred Value : 0.8088
##           Neg Pred Value : 0.3947
##           Prevalence : 0.7644
##           Detection Rate : 0.6322
##           Detection Prevalence : 0.7816
##           Balanced Accuracy : 0.5965
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.659636897120851"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 41 cases (dfPred_raw$class 1).
## Area under the curve: 0.6596
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0317  -0.6703  -0.3537   0.5694   2.8830
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.713e+01  1.558e+01   1.100  0.271499
## KIDSDRIV       1.385e-01  1.714e+00   0.081  0.935603
## AGE          -1.617e-01  1.122e-01  -1.441  0.149581
## HOMEKIDS      -6.720e-01  6.450e-01  -1.042  0.297428
## YOJ           -1.637e-01  1.385e-01  -1.182  0.237094
## INCOME        -3.292e-05  1.403e-05  -2.347  0.018903 *
## HOME_VAL       1.606e-05  9.343e-06   1.719  0.085567 .
## TRAVTIME       1.176e-02  2.261e-02   0.520  0.602878
## BLUEBOOK       7.174e-05  3.952e-05   1.815  0.069482 .
```

```

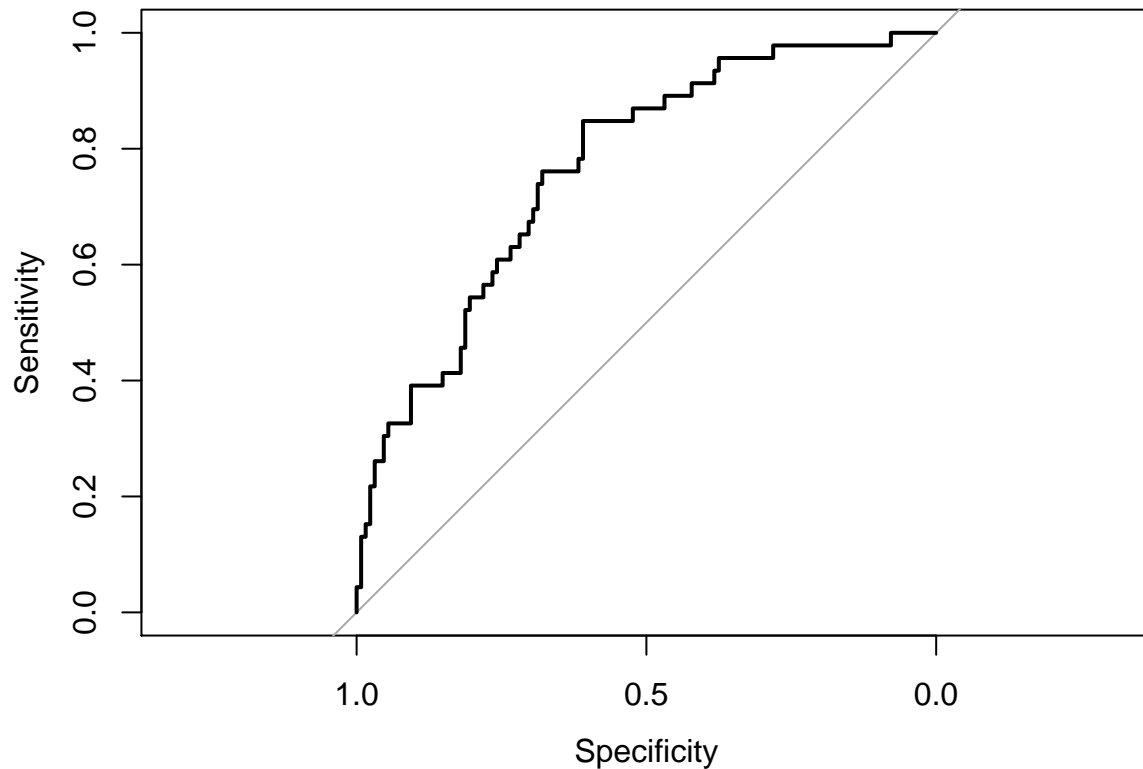
## TIF 1.230e-02 9.000e-02 0.137 0.891327
## OLDCLAIM 1.446e-05 2.222e-05 0.651 0.515210
## CLM_FREQ -4.684e-01 1.000e+00 -0.468 0.639551
## MVRPTS 1.163e-01 1.695e-01 0.686 0.492477
## CAR_AGE -1.021e-01 7.165e-02 -1.426 0.153978
## PARENT1_Yes 3.023e-01 4.643e-01 0.651 0.515068
## MSTATUS_Yes -5.908e-01 3.372e-01 -1.752 0.079732 .
## SEX_z_F -9.602e-01 4.599e-01 -2.088 0.036823 *
## EDUCATION_.High.School 2.311e-01 8.284e-01 0.279 0.780278
## EDUCATION_Bachelors 4.670e-02 7.108e-01 0.066 0.947621
## EDUCATION_Masters 1.103e+00 5.994e-01 1.840 0.065736 .
## EDUCATION_z_High.School 5.987e-01 7.573e-01 0.791 0.429205
## JOB_ -8.766e-01 6.689e-01 -1.310 0.190041
## JOB_Clerical -2.489e-01 4.878e-01 -0.510 0.609886
## JOB_Doctor 8.913e-02 1.015e+00 0.088 0.930024
## JOB_Home.Maker -1.260e-01 6.764e-01 -0.186 0.852203
## JOB_Lawyer -6.073e-01 6.568e-01 -0.925 0.355170
## JOB_Manager -1.042e+00 4.934e-01 -2.111 0.034754 *
## JOB_Student -5.318e-01 7.227e-01 -0.736 0.461855
## JOB_z_Blue.Collar -2.528e-01 4.517e-01 -0.560 0.575605
## CAR_USE_Commercial 3.197e-01 3.435e-01 0.931 0.351932
## CAR_TYPE_Panel.Truck 6.291e-02 6.082e-01 0.103 0.917625
## CAR_TYPE_Pickup 1.189e+00 4.087e-01 2.909 0.003622 **
## CAR_TYPE_Sports.Car 2.058e+00 5.490e-01 3.749 0.000177 ***
## CAR_TYPE_Van 6.239e-01 4.616e-01 1.352 0.176455
## CAR_TYPE_z_SUV 2.007e+00 4.806e-01 4.177 2.95e-05 ***
## RED_CAR_no -3.275e-01 3.286e-01 -0.997 0.318989
## REVOKED_Yes 4.098e-01 3.874e-01 1.058 0.290153
## URBANICITY_z_Highly.Rural..Rural -2.304e+00 4.192e-01 -5.498 3.85e-08 ***
## YOJ_NA -4.139e-01 4.181e-01 -0.990 0.322128
## INCOME_NA 2.685e-01 5.989e-01 0.448 0.653893
## CAR_AGE_NA -7.189e-02 5.077e-01 -0.142 0.887392
## HOME_VAL_NA -3.438e-01 2.862e-01 -1.201 0.229654
## ageSquared 1.623e-03 1.230e-03 1.320 0.186835
## yojSquared 8.533e-03 6.910e-03 1.235 0.216860
## income_log 3.184e-01 3.305e-01 0.963 0.335423
## homeval_log -1.380e+00 1.496e+00 -0.922 0.356270
## travtime_log 2.561e-01 6.463e-01 0.396 0.691913
## bluebook_log -3.683e-01 4.842e-01 -0.761 0.446863
## carage_log 2.845e-01 4.598e-01 0.619 0.536128
## oldclaim_log -8.700e-02 1.685e-01 -0.516 0.605528
## clm_freq_log 1.806e+00 3.098e+00 0.583 0.559957
## mvr_pts_log -2.191e-01 4.993e-01 -0.439 0.660718
## tif_log -3.733e-01 5.138e-01 -0.726 0.467574
## kidsdriv_log -1.426e-01 2.183e+00 -0.065 0.947922
## homekids_log 1.776e+00 1.401e+00 1.268 0.204938
## inter 1.193e-02 3.010e-02 0.396 0.691897
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 740.12 on 641 degrees of freedom
## Residual deviance: 548.24 on 586 degrees of freedom

```

```

## AIC: 660.24
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  28
##           1  12  18
##
##           Accuracy : 0.7701
##           95% CI : (0.7004, 0.8304)
##           No Information Rate : 0.7356
##           P-Value [Acc > NIR] : 0.17254
##
##           Kappa : 0.3349
##
## Mcnemar's Test P-Value : 0.01771
##
##           Sensitivity : 0.9062
##           Specificity : 0.3913
##           Pos Pred Value : 0.8056
##           Neg Pred Value : 0.6000
##           Prevalence : 0.7356
##           Detection Rate : 0.6667
##           Detection Prevalence : 0.8276
##           Balanced Accuracy : 0.6488
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.767663043478261"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 128 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7677
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1094  -0.6868  -0.3741   0.5657   2.7506
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.476e+01  1.524e+01   0.968  0.33295
## KIDSDRIV      -1.764e+00  1.762e+00  -1.001  0.31691
## AGE           -1.801e-01  1.021e-01  -1.765  0.07760 .
## HOMEKIDS      -3.126e-01  6.544e-01  -0.478  0.63285
## YOJ           -2.200e-01  1.378e-01  -1.596  0.11055
## INCOME        -2.702e-05  1.371e-05  -1.971  0.04870 *
## HOME_VAL       1.253e-05  9.098e-06   1.377  0.16837
## TRAVTIME       2.192e-02  2.213e-02   0.991  0.32190
## BLUEBOOK       2.128e-05  3.954e-05   0.538  0.59034
```

```

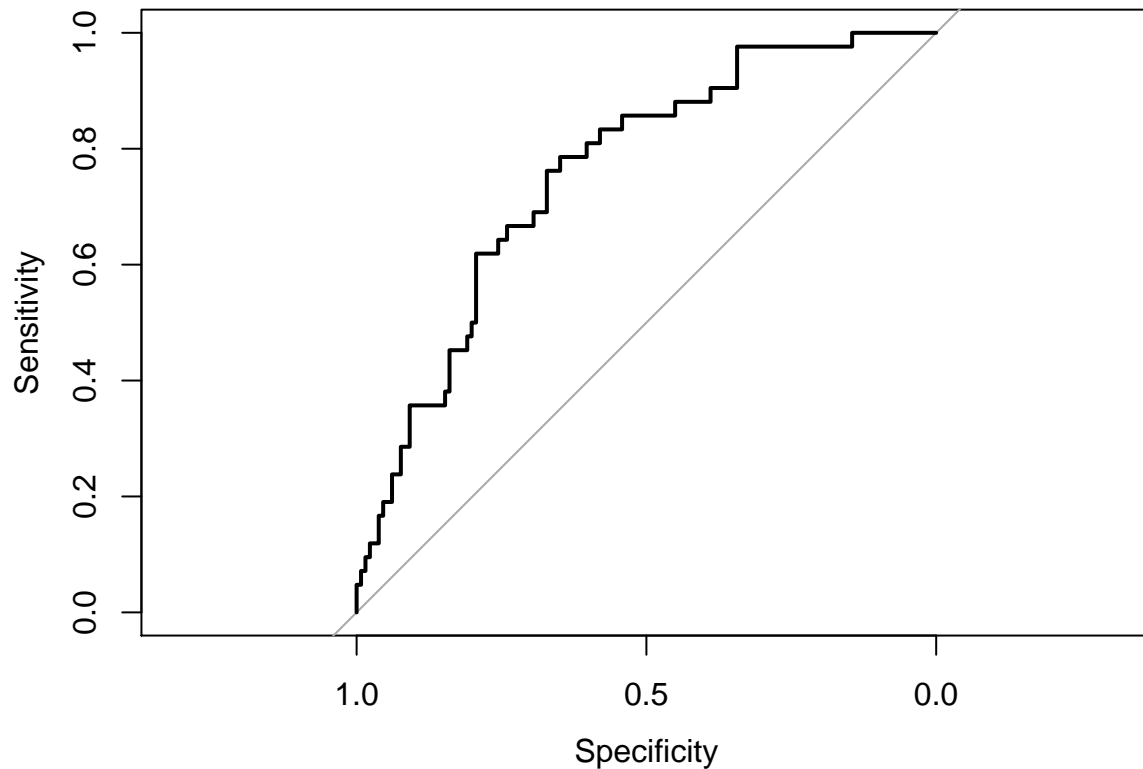
## TIF -7.092e-02 9.436e-02 -0.752 0.45231
## OLDCLAIM 6.430e-06 2.198e-05 0.293 0.76985
## CLM_FREQ -8.460e-01 9.978e-01 -0.848 0.39648
## MVR_PTS 1.588e-01 1.651e-01 0.962 0.33624
## CAR_AGE -4.210e-02 7.184e-02 -0.586 0.55784
## PARENT1_Yes 7.639e-01 4.649e-01 1.643 0.10036
## MSTATUS_Yes -6.501e-01 3.356e-01 -1.937 0.05275 .
## SEX_z_F -7.076e-01 4.499e-01 -1.573 0.11573
## EDUCATION_.High.School -1.476e-01 8.391e-01 -0.176 0.86040
## EDUCATION_Bachelors 2.645e-01 7.148e-01 0.370 0.71137
## EDUCATION_Masters 1.335e+00 6.447e-01 2.071 0.03832 *
## EDUCATION_z_High.School 4.581e-01 7.677e-01 0.597 0.55063
## JOB_ -7.490e-01 6.728e-01 -1.113 0.26563
## JOB_Clerical 6.044e-02 4.896e-01 0.123 0.90176
## JOB_Doctor 2.389e-01 1.023e+00 0.234 0.81528
## JOB_Home.Maker -1.339e-01 6.565e-01 -0.204 0.83835
## JOB_Lawyer -6.379e-01 6.269e-01 -1.017 0.30894
## JOB_Manager -6.874e-01 4.650e-01 -1.478 0.13933
## JOB_Student -6.249e-01 7.293e-01 -0.857 0.39152
## JOB_z_Blue.Collar -3.714e-02 4.485e-01 -0.083 0.93401
## CAR_USE_Commercial 3.939e-01 3.473e-01 1.134 0.25674
## CAR_TYPE_Panel.Truck 1.403e-01 6.132e-01 0.229 0.81900
## CAR_TYPE_Pickup 8.491e-01 4.067e-01 2.088 0.03683 *
## CAR_TYPE_Sports.Car 1.567e+00 5.208e-01 3.008 0.00263 **
## CAR_TYPE_Van 3.677e-01 4.723e-01 0.778 0.43634
## CAR_TYPE_z_SUV 1.419e+00 4.571e-01 3.105 0.00191 **
## RED_CAR_no -2.195e-01 3.306e-01 -0.664 0.50667
## REVOKED_Yes 4.290e-01 4.005e-01 1.071 0.28400
## URBANICITY_z_Highly.Rural..Rural -1.990e+00 3.969e-01 -5.014 5.33e-07 ***
## YOJ_NA -3.313e-01 4.145e-01 -0.799 0.42407
## INCOME_NA 2.515e-01 6.094e-01 0.413 0.67981
## CAR_AGE_NA -2.752e-01 4.887e-01 -0.563 0.57332
## HOME_VAL_NA -6.462e-02 2.877e-01 -0.225 0.82228
## ageSquared 1.978e-03 1.102e-03 1.796 0.07256 .
## yojSquared 1.092e-02 6.944e-03 1.573 0.11581
## income_log 2.392e-01 3.000e-01 0.797 0.42524
## homeval_log -1.317e+00 1.443e+00 -0.912 0.36152
## travtime_log -2.790e-02 6.286e-01 -0.044 0.96460
## bluebook_log 1.011e-01 4.846e-01 0.209 0.83470
## carage_log -1.341e-01 4.611e-01 -0.291 0.77114
## oldclaim_log -9.236e-02 1.657e-01 -0.557 0.57732
## clm_freq_log 2.823e+00 3.073e+00 0.919 0.35814
## mvr_pts_log -3.452e-01 4.958e-01 -0.696 0.48630
## tif_log 1.065e-01 5.244e-01 0.203 0.83905
## kidsdriv_log 1.147e+00 2.086e+00 0.550 0.58231
## homekids_log 9.798e-01 1.422e+00 0.689 0.49094
## inter 4.024e-02 3.512e-02 1.146 0.25189
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 748.85 on 642 degrees of freedom
## Residual deviance: 553.51 on 587 degrees of freedom

```

```

## AIC: 665.51
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 122  32
##           1   9  10
##
##           Accuracy : 0.763
##           95% CI : (0.6925, 0.8242)
##           No Information Rate : 0.7572
##           P-Value [Acc > NIR] : 0.4707206
##
##           Kappa : 0.2081
##
## Mcnemar's Test P-Value : 0.0005908
##
##           Sensitivity : 0.9313
##           Specificity : 0.2381
##           Pos Pred Value : 0.7922
##           Neg Pred Value : 0.5263
##           Prevalence : 0.7572
##           Detection Rate : 0.7052
##           Detection Prevalence : 0.8902
##           Balanced Accuracy : 0.5847
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.757724463831334"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.7577
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0412  -0.6464  -0.3503  -0.1032   2.9714
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.177e+01  1.590e+01   0.740  0.459264
## KIDSDRIV      -4.211e+00  2.201e+00  -1.913  0.055769 .
## AGE           -2.819e-01  1.092e-01  -2.582  0.009836 **
## HOMEKIDS      -5.269e-01  6.660e-01  -0.791  0.428856
## YOJ           -2.405e-02  1.339e-01  -0.180  0.857457
## INCOME        -2.443e-05  1.434e-05  -1.704  0.088451 .
## HOME_VAL       1.070e-05  9.514e-06   1.125  0.260642
## TRAVTIME       3.281e-02  2.228e-02   1.472  0.140984
## BLUEBOOK       2.977e-05  4.253e-05   0.700  0.483964
```

```

## TIF -3.140e-02 9.779e-02 -0.321 0.748145
## OLDCLAIM -1.417e-05 2.242e-05 -0.632 0.527181
## CLM_FREQ 1.326e-01 1.085e+00 0.122 0.902732
## MVR_PTS 1.570e-01 1.793e-01 0.876 0.381140
## CAR_AGE 2.021e-02 7.079e-02 0.286 0.775232
## PARENT1_Yes 2.483e-01 4.803e-01 0.517 0.605204
## MSTATUS_Yes -6.047e-01 3.422e-01 -1.767 0.077179 .
## SEX_z_F -8.345e-01 4.909e-01 -1.700 0.089165 .
## EDUCATION_.High.School -1.964e-01 8.203e-01 -0.239 0.810828
## EDUCATION_Bachelors -1.312e-01 6.741e-01 -0.195 0.845637
## EDUCATION_Masters 5.103e-01 5.622e-01 0.908 0.364033
## EDUCATION_z_High.School 1.475e-01 7.359e-01 0.201 0.841086
## JOB_ -1.422e+00 7.089e-01 -2.005 0.044931 *
## JOB_Clerical -1.200e-01 4.854e-01 -0.247 0.804738
## JOB_Doctor -1.424e+00 1.277e+00 -1.116 0.264599
## JOB_Home.Maker 7.292e-01 6.701e-01 1.088 0.276478
## JOB_Lawyer -8.979e-01 6.550e-01 -1.371 0.170472
## JOB_Manager -1.004e+00 5.278e-01 -1.903 0.057097 .
## JOB_Student -4.799e-01 7.165e-01 -0.670 0.503009
## JOB_z_Blue.Collar -2.647e-01 4.572e-01 -0.579 0.562631
## CAR_USE_Commercial 8.468e-01 3.589e-01 2.360 0.018288 *
## CAR_TYPE_Panel.Truck 2.520e-01 6.275e-01 0.402 0.688047
## CAR_TYPE_Pickup 9.444e-01 4.285e-01 2.204 0.027528 *
## CAR_TYPE_Sports.Car 2.174e+00 5.656e-01 3.844 0.000121 ***
## CAR_TYPE_Van 6.076e-01 4.843e-01 1.255 0.209640
## CAR_TYPE_z_SUV 1.935e+00 4.971e-01 3.893 9.89e-05 ***
## RED_CAR_no -3.629e-01 3.507e-01 -1.035 0.300767
## REVOKED_Yes 9.897e-01 4.016e-01 2.464 0.013726 *
## URBANICITY_z_Highly.Rural..Rural -2.433e+00 4.408e-01 -5.520 3.39e-08 ***
## YOJ_NA -3.133e-01 4.483e-01 -0.699 0.484646
## INCOME_NA 3.654e-01 5.673e-01 0.644 0.519503
## CAR_AGE_NA -5.738e-01 5.033e-01 -1.140 0.254308
## HOME_VAL_NA -1.019e-01 2.986e-01 -0.341 0.732886
## ageSquared 2.914e-03 1.179e-03 2.472 0.013434 *
## yojSquared 9.873e-04 6.894e-03 0.143 0.886133
## income_log 1.731e-01 3.220e-01 0.537 0.590979
## homeval_log -5.560e-01 1.479e+00 -0.376 0.706898
## travtime_log -3.611e-01 6.236e-01 -0.579 0.562590
## bluebook_log -1.506e-01 5.182e-01 -0.291 0.771304
## carage_log -4.326e-01 4.647e-01 -0.931 0.351912
## oldclaim_log 3.290e-02 1.773e-01 0.186 0.852787
## clm_freq_log -1.010e-01 3.312e+00 -0.031 0.975662
## mvr_pts_log -4.755e-01 5.227e-01 -0.910 0.362962
## tif_log 9.754e-02 5.440e-01 0.179 0.857714
## kidsdriv_log 2.599e+00 2.291e+00 1.134 0.256748
## homekids_log 1.098e+00 1.450e+00 0.757 0.448942
## inter 8.093e-02 4.201e-02 1.926 0.054046 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 707.95 on 642 degrees of freedom
## Residual deviance: 518.34 on 587 degrees of freedom

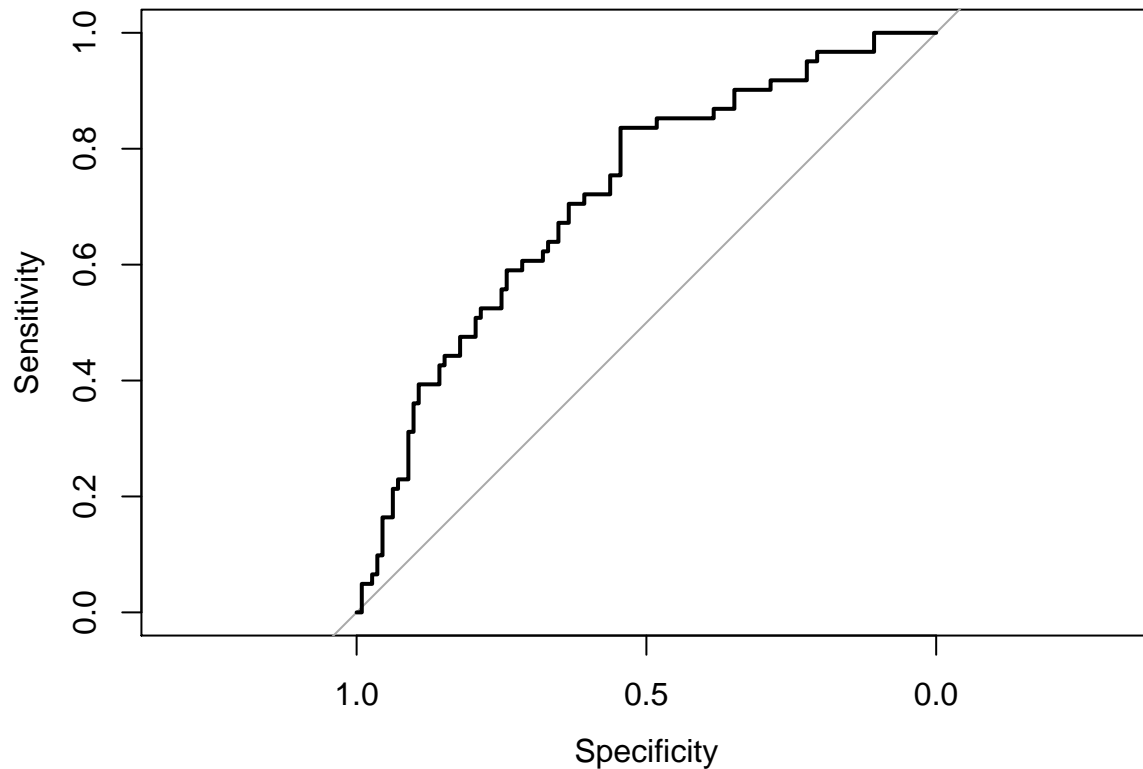
```



```

## (1 observation deleted due to missingness)
## AIC: 630.34
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 101  42
##           1   11  19
##
##           Accuracy : 0.6936
##           95% CI : (0.6192, 0.7614)
##           No Information Rate : 0.6474
##           P-Value [Acc > NIR] : 0.1156
##
##           Kappa : 0.2412
##
## Mcnemar's Test P-Value : 3.775e-05
##
##           Sensitivity : 0.9018
##           Specificity : 0.3115
##           Pos Pred Value : 0.7063
##           Neg Pred Value : 0.6333
##           Prevalence : 0.6474
##           Detection Rate : 0.5838
##           Detection Prevalence : 0.8266
##           Balanced Accuracy : 0.6066
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.71882318501171"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 112 controls (dfPred_raw$class 0) < 61 cases (dfPred_raw$class 1).
## Area under the curve: 0.7188
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9552  -0.6382  -0.3290   0.4491   3.0785
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.281e+01  1.477e+01   2.221 0.026321 *
## KIDSDRIV      -2.131e-01  1.860e+00  -0.115 0.908786
## AGE           -1.403e-01  1.104e-01  -1.271 0.203740
## HOMEKIDS      -1.596e+00  7.764e-01  -2.056 0.039808 *
## YOJ           -2.087e-01  1.364e-01  -1.530 0.126062
## INCOME        -1.911e-05  1.333e-05  -1.433 0.151744
## HOME_VAL       1.534e-05  8.845e-06   1.734 0.082846 .
## TRAVTIME       2.916e-02  2.278e-02   1.280 0.200464
## BLUEBOOK       5.796e-05  3.887e-05   1.491 0.135911
```

```

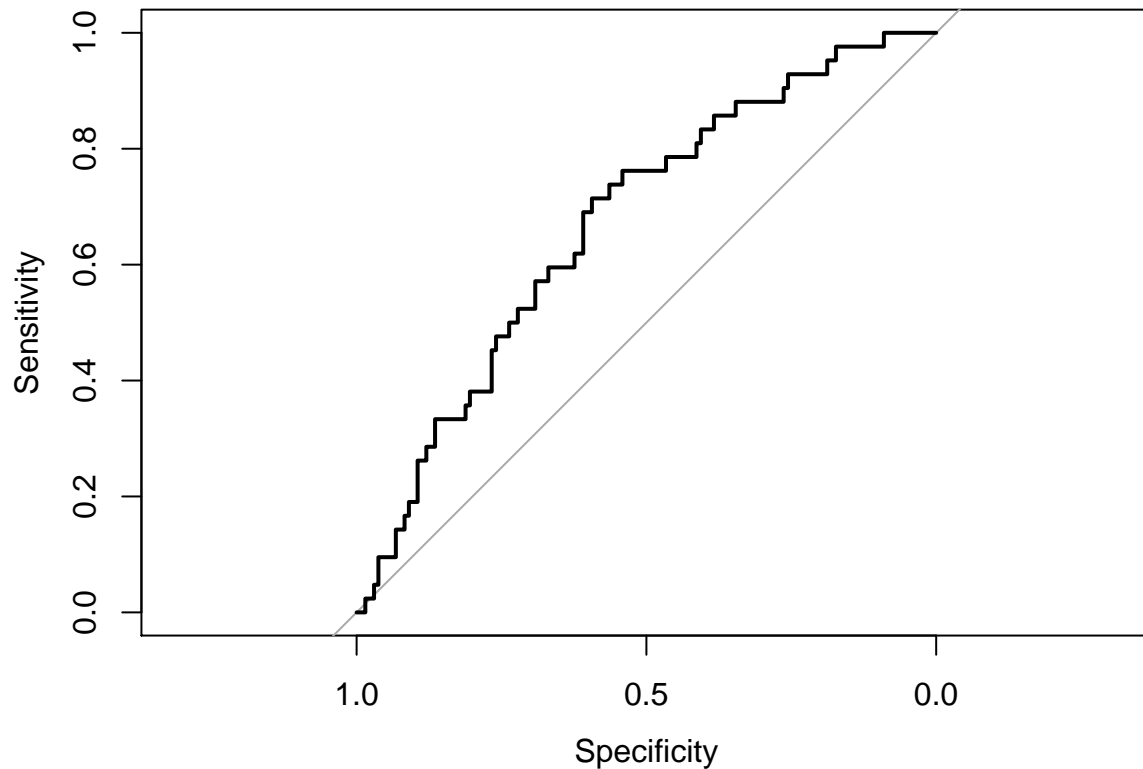
## TIF -1.387e-03 9.206e-02 -0.015 0.987978
## OLDCLAIM -3.455e-05 2.221e-05 -1.556 0.119711
## CLM_FREQ 5.995e-01 1.004e+00 0.597 0.550497
## MVRPTS 2.905e-02 1.739e-01 0.167 0.867350
## CAR_AGE -3.136e-02 7.004e-02 -0.448 0.654375
## PARENT1_Yes 1.071e+00 4.988e-01 2.147 0.031801 *
## MSTATUS_Yes -4.267e-01 3.395e-01 -1.257 0.208767
## SEX_z_F -1.039e+00 4.791e-01 -2.168 0.030121 *
## EDUCATION_High.School 9.600e-02 8.117e-01 0.118 0.905856
## EDUCATION_Bachelors 1.277e-01 6.840e-01 0.187 0.851943
## EDUCATION_Masters 7.091e-01 5.721e-01 1.240 0.215126
## EDUCATION_z_High.School 1.279e-01 7.472e-01 0.171 0.864065
## JOB_ -6.459e-01 6.831e-01 -0.945 0.344407
## JOB_Clerical -4.710e-01 4.840e-01 -0.973 0.330459
## JOB_Doctor 3.264e-02 8.948e-01 0.036 0.970902
## JOB_Home.Maker -1.239e+00 6.985e-01 -1.773 0.076147 .
## JOB_Lawyer -8.959e-01 6.664e-01 -1.344 0.178833
## JOB_Manager -8.520e-01 4.932e-01 -1.727 0.084107 .
## JOB_Student -1.188e+00 7.650e-01 -1.553 0.120442
## JOB_z_Blue.Collar -4.230e-01 4.509e-01 -0.938 0.348219
## CAR_USE_Commercial 5.498e-01 3.640e-01 1.510 0.130997
## CAR_TYPE_Panel.Truck 7.002e-01 6.335e-01 1.105 0.269027
## CAR_TYPE_Pickup 9.296e-01 4.241e-01 2.192 0.028384 *
## CAR_TYPE_Sports.Car 1.902e+00 5.530e-01 3.440 0.000583 ***
## CAR_TYPE_Van 6.574e-01 4.861e-01 1.352 0.176286
## CAR_TYPE_z_SUV 2.006e+00 4.857e-01 4.130 3.63e-05 ***
## RED_CAR_no 2.641e-01 3.411e-01 0.774 0.438715
## REVOKED_Yes 9.122e-01 4.197e-01 2.174 0.029737 *
## URBANICITY_z_Highly.Rural..Rural -2.631e+00 4.435e-01 -5.932 2.99e-09 ***
## YOJ_NA 3.862e-02 4.754e-01 0.081 0.935264
## INCOME_NA 4.445e-01 5.901e-01 0.753 0.451235
## CAR_AGE_NA 4.239e-01 5.888e-01 0.720 0.471500
## HOME_VAL_NA -1.820e-01 2.938e-01 -0.619 0.535661
## ageSquared 1.555e-03 1.186e-03 1.311 0.189835
## yojSquared 9.935e-03 6.952e-03 1.429 0.152936
## income_log -2.003e-01 3.312e-01 -0.605 0.545349
## homeval_log -2.228e+00 1.402e+00 -1.589 0.111959
## travtime_log -3.434e-02 6.474e-01 -0.053 0.957697
## bluebook_log -5.179e-01 4.629e-01 -1.119 0.263270
## carage_log -2.238e-01 4.638e-01 -0.482 0.629455
## oldclaim_log 1.817e-01 1.628e-01 1.116 0.264372
## clm_freq_log -1.781e+00 3.067e+00 -0.580 0.561579
## mvr_pts_log 7.880e-02 5.129e-01 0.154 0.877886
## tif_log -2.835e-01 5.220e-01 -0.543 0.587115
## kidsdriv_log -5.290e-01 2.410e+00 -0.219 0.826282
## homekids_log 3.250e+00 1.647e+00 1.974 0.048398 *
## inter 2.553e-02 3.466e-02 0.737 0.461419
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 747.60 on 640 degrees of freedom
## Residual deviance: 523.79 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 635.79
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 111  28
##           1  22  14
##
##           Accuracy : 0.7143
##           95% CI : (0.6412, 0.7799)
##           No Information Rate : 0.76
##           P-Value [Acc > NIR] : 0.9315
##
##           Kappa : 0.1765
##
## Mcnemar's Test P-Value : 0.4795
##
##           Sensitivity : 0.8346
##           Specificity : 0.3333
##           Pos Pred Value : 0.7986
##           Neg Pred Value : 0.3889
##           Prevalence : 0.7600
##           Detection Rate : 0.6343
##           Detection Prevalence : 0.7943
##           Balanced Accuracy : 0.5840
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.671142141066953"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.6711
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2830  -0.6576  -0.3427   0.4321   3.1471
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.005e+01  1.605e+01   0.626 0.531112
## KIDSDRIV      -2.060e+00  1.892e+00  -1.089 0.276179
## AGE           -8.499e-02  1.106e-01  -0.768 0.442208
## HOMEKIDS      -5.126e-01  6.850e-01  -0.748 0.454258
## YOJ           -2.715e-01  1.389e-01  -1.955 0.050594 .
## INCOME        -2.150e-05  1.371e-05  -1.568 0.116823
## HOME_VAL       9.904e-06  9.558e-06   1.036 0.300065
## TRAVTIME       4.467e-02  2.148e-02   2.079 0.037582 *
## BLUEBOOK       7.051e-05  3.808e-05   1.851 0.064108 .
```

```

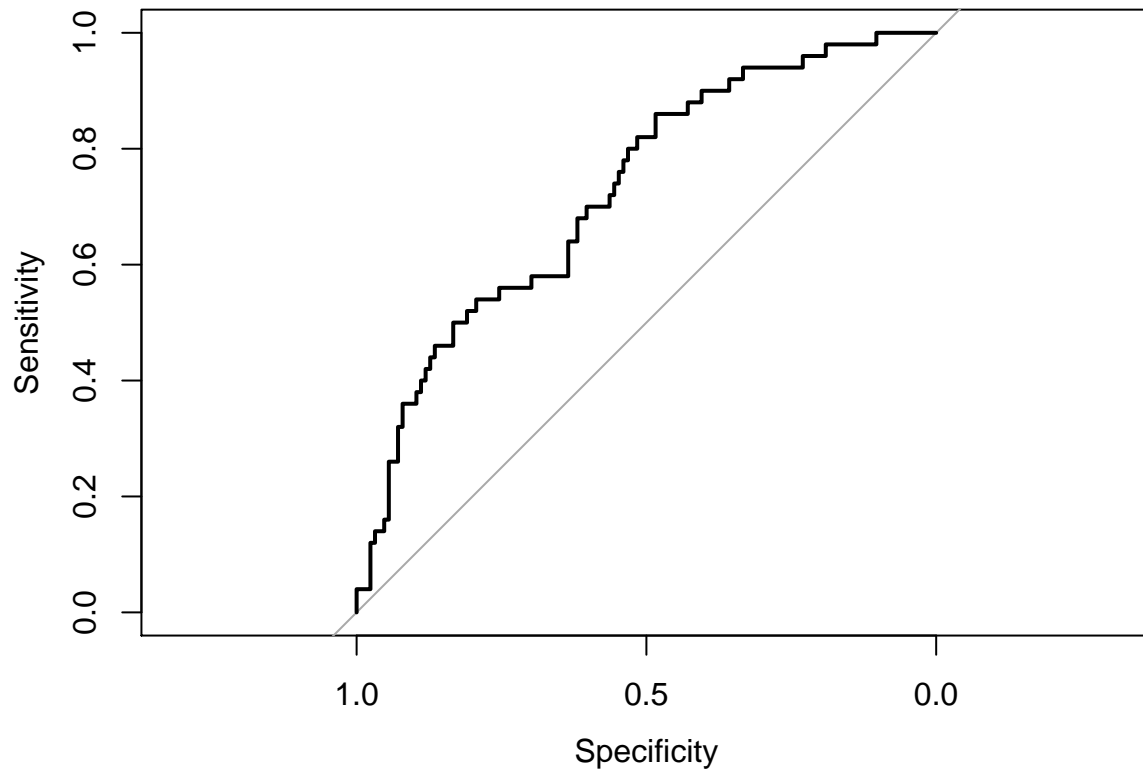
## TIF 1.838e-02 9.086e-02 0.202 0.839688
## OLDCLAIM 1.799e-05 2.130e-05 0.845 0.398242
## CLM_FREQ -1.232e-01 9.648e-01 -0.128 0.898415
## MVRPTS 1.065e-01 1.704e-01 0.625 0.531961
## CAR_AGE -5.809e-02 6.997e-02 -0.830 0.406412
## PARENT1_Yes 4.366e-01 4.796e-01 0.910 0.362636
## MSTATUS_Yes -6.808e-01 3.459e-01 -1.968 0.049055 *
## SEX_z_F -5.392e-01 4.792e-01 -1.125 0.260466
## EDUCATION_.High.School 1.145e+00 7.938e-01 1.443 0.149128
## EDUCATION_Bachelors 3.742e-01 6.705e-01 0.558 0.576818
## EDUCATION_Masters 1.215e+00 5.807e-01 2.093 0.036358 *
## EDUCATION_z_High.School 1.228e+00 7.267e-01 1.689 0.091144 .
## JOB_ -8.266e-01 6.931e-01 -1.193 0.233022
## JOB_Clerical -5.251e-01 4.991e-01 -1.052 0.292792
## JOB_Doctor 3.721e-01 9.407e-01 0.396 0.692465
## JOB_Home.Maker 2.546e-02 6.818e-01 0.037 0.970216
## JOB_Lawyer -8.846e-01 6.783e-01 -1.304 0.192167
## JOB_Manager -1.166e+00 5.107e-01 -2.282 0.022464 *
## JOB_Student -5.239e-01 7.654e-01 -0.684 0.493663
## JOB_z_Blue.Collar -5.219e-01 4.715e-01 -1.107 0.268358
## CAR_USE_Commercial 1.021e+00 3.643e-01 2.804 0.005047 **
## CAR_TYPE_Panel.Truck -5.491e-02 6.262e-01 -0.088 0.930123
## CAR_TYPE_Pickup 7.748e-01 4.205e-01 1.843 0.065396 .
## CAR_TYPE_Sports.Car 1.799e+00 5.503e-01 3.269 0.001078 **
## CAR_TYPE_Van 2.292e-01 4.841e-01 0.473 0.635862
## CAR_TYPE_z_SUV 1.752e+00 4.841e-01 3.618 0.000297 ***
## RED_CAR_no -3.403e-01 3.479e-01 -0.978 0.328040
## REVOKED_Yes 4.782e-01 4.307e-01 1.110 0.266841
## URBANICITY_z_Highly.Rural..Rural -2.719e+00 4.626e-01 -5.877 4.17e-09 ***
## YOJ_NA -8.530e-02 4.361e-01 -0.196 0.844921
## INCOME_NA 3.241e-01 5.744e-01 0.564 0.572591
## CAR_AGE_NA -2.835e-01 5.009e-01 -0.566 0.571462
## HOME_VAL_NA -2.163e-01 2.894e-01 -0.747 0.454801
## ageSquared 7.943e-04 1.206e-03 0.659 0.510174
## yojSquared 1.434e-02 7.127e-03 2.012 0.044270 *
## income_log -1.157e-02 3.005e-01 -0.039 0.969273
## homeval_log -3.206e-01 1.538e+00 -0.209 0.834824
## travtime_log -7.515e-01 6.042e-01 -1.244 0.213570
## bluebook_log -5.217e-01 4.506e-01 -1.158 0.246952
## carage_log 2.228e-01 4.635e-01 0.481 0.630758
## oldclaim_log -3.811e-02 1.659e-01 -0.230 0.818281
## clm_freq_log 7.943e-01 2.995e+00 0.265 0.790863
## mvr_pts_log -1.811e-01 5.034e-01 -0.360 0.718950
## tif_log -4.650e-01 5.184e-01 -0.897 0.369731
## kidsdriv_log 1.991e+00 2.258e+00 0.882 0.377813
## homekids_log 1.149e+00 1.508e+00 0.762 0.446083
## inter 3.744e-02 3.605e-02 1.039 0.298964
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 730.57 on 639 degrees of freedom
## Residual deviance: 526.17 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 638.17
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 110  29
##           1   16  21
##
##           Accuracy : 0.7443
##           95% CI : (0.6732, 0.807)
##           No Information Rate : 0.7159
##           P-Value [Acc > NIR] : 0.22760
##
##           Kappa : 0.3179
##
## Mcnemar's Test P-Value : 0.07364
##
##           Sensitivity : 0.8730
##           Specificity : 0.4200
##           Pos Pred Value : 0.7914
##           Neg Pred Value : 0.5676
##           Prevalence : 0.7159
##           Detection Rate : 0.6250
##           Detection Prevalence : 0.7898
##           Balanced Accuracy : 0.6465
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.726349206349206"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 126 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7263
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1133  -0.6504  -0.3462   0.2552   2.8999
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.305e+01  1.573e+01   1.466 0.142720
## KIDSDRIV      -4.322e-01  1.754e+00  -0.246 0.805407
## AGE           -1.689e-01  1.053e-01  -1.604 0.108606
## HOMEKIDS      -5.443e-01  7.007e-01  -0.777 0.437329
## YOJ           -1.682e-01  1.336e-01  -1.258 0.208222
## INCOME        -3.667e-05  1.470e-05  -2.495 0.012600 *
## HOME_VAL       1.723e-05  9.506e-06   1.813 0.069838 .
## TRAVTIME       3.911e-02  2.220e-02   1.762 0.078037 .
## BLUEBOOK       1.942e-05  4.268e-05   0.455 0.649164
```



```

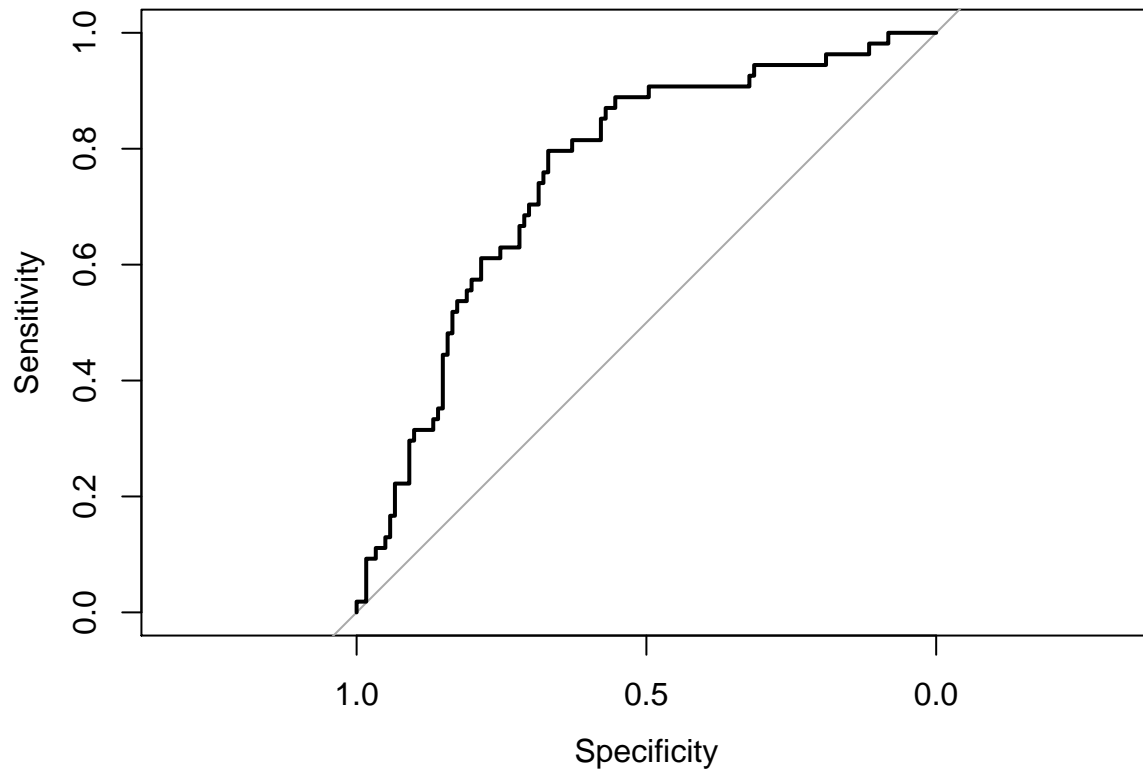
## TIF -3.498e-02 9.545e-02 -0.366 0.714014
## OLDCLAIM -2.114e-06 2.156e-05 -0.098 0.921892
## CLM_FREQ -3.945e-01 9.991e-01 -0.395 0.692946
## MVR_PTS 1.344e-01 1.726e-01 0.779 0.436264
## CAR_AGE -8.236e-02 7.048e-02 -1.168 0.242612
## PARENT1_Yes 5.415e-01 4.843e-01 1.118 0.263523
## MSTATUS_Yes -4.344e-01 3.469e-01 -1.252 0.210555
## SEX_z_F -1.043e+00 4.747e-01 -2.196 0.028074 *
## EDUCATION_.High.School -7.462e-01 7.827e-01 -0.953 0.340425
## EDUCATION_Bachelors -6.039e-01 6.614e-01 -0.913 0.361224
## EDUCATION_Masters 5.954e-01 5.681e-01 1.048 0.294617
## EDUCATION_z_High.School -3.379e-01 7.124e-01 -0.474 0.635293
## JOB_ -9.590e-01 6.662e-01 -1.440 0.150000
## JOB_Clerical -1.610e-01 4.892e-01 -0.329 0.742105
## JOB_Doctor -1.866e-01 9.658e-01 -0.193 0.846798
## JOB_Home.Maker 4.437e-02 6.540e-01 0.068 0.945912
## JOB_Lawyer -9.315e-01 6.444e-01 -1.446 0.148306
## JOB_Manager -6.700e-01 4.795e-01 -1.397 0.162290
## JOB_Student -8.165e-01 7.375e-01 -1.107 0.268265
## JOB_z_Blue.Collar -2.071e-01 4.674e-01 -0.443 0.657665
## CAR_USE_Commercial 3.915e-01 3.542e-01 1.105 0.269027
## CAR_TYPE_Panel.Truck 4.291e-01 6.315e-01 0.680 0.496754
## CAR_TYPE_Pickup 1.199e+00 4.134e-01 2.900 0.003734 **
## CAR_TYPE_Sports.Car 1.905e+00 5.661e-01 3.365 0.000765 ***
## CAR_TYPE_Van 1.473e-01 5.175e-01 0.285 0.775867
## CAR_TYPE_z_SUV 2.015e+00 4.907e-01 4.107 4.00e-05 ***
## RED_CAR_no -3.478e-03 3.442e-01 -0.010 0.991939
## REVOKED_Yes 8.369e-01 4.010e-01 2.087 0.036875 *
## URBANICITY_z_Highly.Rural..Rural -2.036e+00 4.111e-01 -4.952 7.35e-07 ***
## YOJ_NA -2.698e-02 4.286e-01 -0.063 0.949803
## INCOME_NA 5.265e-01 5.908e-01 0.891 0.372903
## CAR_AGE_NA -2.205e-01 4.882e-01 -0.452 0.651519
## HOME_VAL_NA -1.461e-01 2.995e-01 -0.488 0.625714
## ageSquared 1.696e-03 1.136e-03 1.493 0.135402
## yojSquared 7.535e-03 6.805e-03 1.107 0.268144
## income_log 7.224e-01 3.660e-01 1.974 0.048397 *
## homeval_log -2.379e+00 1.511e+00 -1.574 0.115436
## travtime_log -4.587e-01 6.273e-01 -0.731 0.464598
## bluebook_log 7.707e-02 5.319e-01 0.145 0.884782
## carage_log 1.679e-01 4.614e-01 0.364 0.715954
## oldclaim_log -2.102e-02 1.678e-01 -0.125 0.900266
## clm_freq_log 1.492e+00 3.082e+00 0.484 0.628366
## mvr_pts_log -2.945e-01 5.112e-01 -0.576 0.564511
## tif_log -1.584e-01 5.342e-01 -0.296 0.766884
## kidsdriv_log -1.004e+00 2.281e+00 -0.440 0.659690
## homekids_log 1.396e+00 1.508e+00 0.926 0.354352
## inter 3.412e-02 3.139e-02 1.087 0.277065
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 722.56 on 640 degrees of freedom
## Residual deviance: 528.70 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 640.7
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 112  42
##           1   9  12
##
##           Accuracy : 0.7086
##           95% CI : (0.6352, 0.7747)
##           No Information Rate : 0.6914
##           P-Value [Acc > NIR] : 0.3445
##
##           Kappa : 0.1779
##
## Mcnemar's Test P-Value : 7.433e-06
##
##           Sensitivity : 0.9256
##           Specificity : 0.2222
##           Pos Pred Value : 0.7273
##           Neg Pred Value : 0.5714
##           Prevalence : 0.6914
##           Detection Rate : 0.6400
##           Detection Prevalence : 0.8800
##           Balanced Accuracy : 0.5739
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.756351392715029"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 121 controls (dfPred_raw$class 0) < 54 cases (dfPred_raw$class 1).
## Area under the curve: 0.7564
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1209  -0.6568  -0.3610   0.5016   3.1135
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    9.130e+00  1.537e+01   0.594 0.552480
## KIDSDRIV     -1.699e+00  1.872e+00  -0.908 0.364012
## AGE          -1.604e-01  1.096e-01  -1.463 0.143369
## HOMEKIDS     -1.046e+00  6.860e-01  -1.525 0.127369
## YOJ          -1.564e-01  1.319e-01  -1.186 0.235753
## INCOME       -1.230e-05  1.330e-05  -0.925 0.354842
## HOME_VAL      3.640e-06  9.058e-06   0.402 0.687751
## TRAVTIME      3.741e-02  2.235e-02   1.674 0.094147
## BLUEBOOK      6.355e-05  4.026e-05   1.578 0.114480
```

```

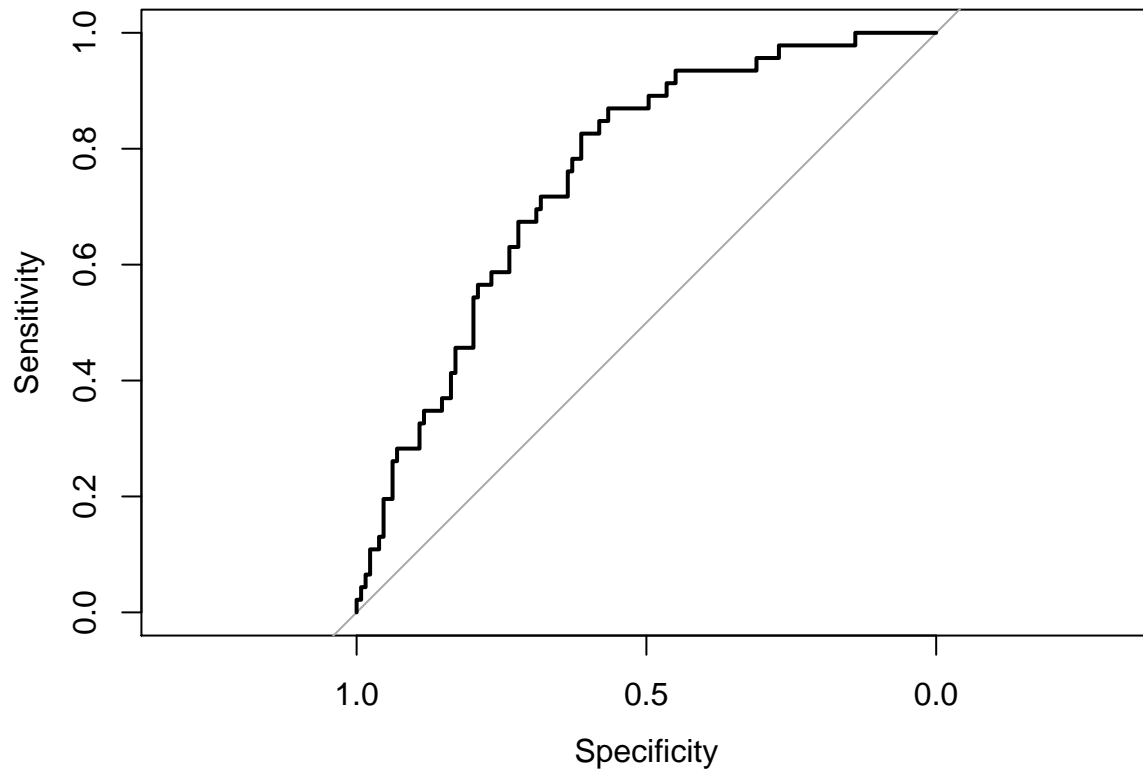
## TIF -9.096e-02 9.798e-02 -0.928 0.353217
## OLDCLAIM -4.899e-06 2.121e-05 -0.231 0.817318
## CLM_FREQ 5.520e-02 9.390e-01 0.059 0.953124
## MVRPTS 8.183e-02 1.739e-01 0.471 0.637934
## CAR_AGE -4.557e-02 6.954e-02 -0.655 0.512222
## PARENT1_Yes 1.643e-01 4.778e-01 0.344 0.730896
## MSTATUS_Yes -6.535e-01 3.381e-01 -1.933 0.053289 .
## SEX_z_F -1.352e+00 4.917e-01 -2.750 0.005957 **
## EDUCATION_.High.School -1.029e+00 7.793e-01 -1.321 0.186519
## EDUCATION_Bachelors -8.877e-01 6.693e-01 -1.326 0.184721
## EDUCATION_Masters 1.571e-01 5.568e-01 0.282 0.777846
## EDUCATION_z_High.School -7.319e-01 7.172e-01 -1.021 0.307471
## JOB_ -1.347e+00 6.755e-01 -1.995 0.046093 *
## JOB_Clerical 2.937e-01 4.872e-01 0.603 0.546588
## JOB_Doctor -4.241e-01 8.966e-01 -0.473 0.636223
## JOB_Home.Maker 2.881e-01 6.556e-01 0.439 0.660337
## JOB_Lawyer -8.244e-01 6.697e-01 -1.231 0.218313
## JOB_Manager -7.519e-01 4.961e-01 -1.516 0.129625
## JOB_Student -6.089e-01 7.179e-01 -0.848 0.396374
## JOB_z_Blue.Collar 2.546e-01 4.410e-01 0.577 0.563696
## CAR_USE_Commercial 5.730e-01 3.508e-01 1.633 0.102423
## CAR_TYPE_Panel.Truck 1.402e-02 6.055e-01 0.023 0.981532
## CAR_TYPE_Pickup 1.567e+00 4.123e-01 3.800 0.000145 ***
## CAR_TYPE_Sports.Car 2.216e+00 5.663e-01 3.913 9.12e-05 ***
## CAR_TYPE_Van 7.068e-01 4.806e-01 1.471 0.141346
## CAR_TYPE_z_SUV 2.417e+00 5.221e-01 4.629 3.68e-06 ***
## RED_CAR_no -1.570e-02 3.382e-01 -0.046 0.962982
## REVOKED_Yes 9.047e-01 4.035e-01 2.242 0.024955 *
## URBANICITY_z_Highly.Rural..Rural -2.334e+00 4.105e-01 -5.687 1.29e-08 ***
## YOJ_NA -1.413e-01 4.407e-01 -0.321 0.748459
## INCOME_NA 1.811e-01 5.542e-01 0.327 0.743783
## CAR_AGE_NA 6.814e-02 5.121e-01 0.133 0.894131
## HOME_VAL_NA -2.254e-01 2.870e-01 -0.785 0.432238
## ageSquared 1.719e-03 1.189e-03 1.446 0.148245
## yojSquared 7.196e-03 6.723e-03 1.070 0.284473
## income_log -6.077e-02 2.929e-01 -0.207 0.835647
## homeval_log -2.064e-01 1.422e+00 -0.145 0.884591
## travtime_log -2.862e-01 6.285e-01 -0.455 0.648849
## bluebook_log -3.896e-01 4.899e-01 -0.795 0.426435
## carage_log 5.520e-02 4.548e-01 0.121 0.903408
## oldclaim_log 5.568e-02 1.583e-01 0.352 0.725110
## clm_freq_log 3.870e-02 2.903e+00 0.013 0.989363
## mvr_pts_log -2.503e-01 5.116e-01 -0.489 0.624685
## tif_log 2.467e-01 5.380e-01 0.458 0.646621
## kidsdriv_log 9.807e-02 2.231e+00 0.044 0.964941
## homekids_log 2.400e+00 1.496e+00 1.605 0.108592
## inter 5.464e-02 3.750e-02 1.457 0.145110
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 739.51 on 640 degrees of freedom
## Residual deviance: 540.65 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 652.65
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  29
##           1  21  17
##
##           Accuracy : 0.7143
##           95% CI : (0.6412, 0.7799)
##           No Information Rate : 0.7371
##           P-Value [Acc > NIR] : 0.7818
##
##           Kappa : 0.219
##
## Mcnemar's Test P-Value : 0.3222
##
##           Sensitivity : 0.8372
##           Specificity : 0.3696
##           Pos Pred Value : 0.7883
##           Neg Pred Value : 0.4474
##           Prevalence : 0.7371
##           Detection Rate : 0.6171
##           Detection Prevalence : 0.7829
##           Balanced Accuracy : 0.6034
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.756319514661274"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 129 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7563
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4459  -0.6559  -0.3644   0.5604   3.0805
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    5.598e+00  1.490e+01   0.376 0.707121
## KIDSDRIV     -1.236e+00  1.649e+00  -0.749 0.453670
## AGE          -1.677e-01  1.076e-01  -1.559 0.119030
## HOMEKIDS     -2.658e-01  6.651e-01  -0.400 0.689474
## YOJ          -1.649e-01  1.325e-01  -1.244 0.213455
## INCOME       -1.938e-05  1.407e-05  -1.377 0.168436
## HOME_VAL      8.222e-06  9.028e-06   0.911 0.362412
## TRAVTIME      1.591e-02  2.119e-02   0.751 0.452614
## BLUEBOOK      1.937e-05  4.080e-05   0.475 0.634909
```

```

## TIF -1.075e-01 9.335e-02 -1.151 0.249716
## OLDCLAIM -1.217e-05 2.129e-05 -0.572 0.567533
## CLM_FREQ 2.684e-01 9.849e-01 0.273 0.785192
## MVR_PTS -1.365e-02 1.694e-01 -0.081 0.935776
## CAR_AGE -7.063e-02 7.370e-02 -0.958 0.337835
## PARENT1_Yes 3.891e-01 4.891e-01 0.796 0.426312
## MSTATUS_Yes -7.502e-01 3.335e-01 -2.249 0.024487 *
## SEX_z_F -8.611e-01 4.652e-01 -1.851 0.064146 .
## EDUCATION_.High.School 4.468e-01 8.033e-01 0.556 0.578122
## EDUCATION_Bachelors -2.886e-02 6.778e-01 -0.043 0.966041
## EDUCATION_Masters 1.127e+00 6.060e-01 1.860 0.062940 .
## EDUCATION_z_High.School 4.663e-01 7.331e-01 0.636 0.524753
## JOB_ -1.437e+00 6.784e-01 -2.119 0.034129 *
## JOB_Clerical -4.522e-01 4.862e-01 -0.930 0.352324
## JOB_Doctor -1.327e-02 9.168e-01 -0.014 0.988451
## JOB_Home.Maker -1.269e-01 6.497e-01 -0.195 0.845164
## JOB_Lawyer -1.091e+00 6.392e-01 -1.706 0.087931 .
## JOB_Manager -7.494e-01 4.625e-01 -1.620 0.105152
## JOB_Student -3.127e-01 7.345e-01 -0.426 0.670297
## JOB_z_Blue.Collar -5.410e-01 4.482e-01 -1.207 0.227393
## CAR_USE_Commercial 7.443e-01 3.547e-01 2.098 0.035884 *
## CAR_TYPE_Panel.Truck -3.622e-01 6.325e-01 -0.573 0.566931
## CAR_TYPE_Pickup 1.030e+00 4.177e-01 2.467 0.013621 *
## CAR_TYPE_Sports.Car 1.756e+00 5.536e-01 3.172 0.001514 **
## CAR_TYPE_Van 2.249e-01 4.743e-01 0.474 0.635400
## CAR_TYPE_z_SUV 1.875e+00 4.834e-01 3.880 0.000105 ***
## RED_CAR_no -1.320e-01 3.404e-01 -0.388 0.698145
## REVOKED_Yes 6.215e-01 4.067e-01 1.528 0.126475
## URBANICITY_z_Highly.Rural..Rural -2.494e+00 4.358e-01 -5.723 1.05e-08 ***
## YOJ_NA -2.188e-01 4.521e-01 -0.484 0.628409
## INCOME_NA -3.434e-01 5.522e-01 -0.622 0.534061
## CAR_AGE_NA -6.313e-01 4.713e-01 -1.340 0.180383
## HOME_VAL_NA -5.136e-02 2.924e-01 -0.176 0.860572
## ageSquared 1.690e-03 1.160e-03 1.457 0.145198
## yojSquared 8.982e-03 6.831e-03 1.315 0.188579
## income_log 1.108e-01 3.281e-01 0.338 0.735522
## homeval_log -4.883e-01 1.402e+00 -0.348 0.727643
## travtime_log 1.625e-01 5.987e-01 0.271 0.786041
## bluebook_log 1.536e-01 5.099e-01 0.301 0.763269
## carage_log 1.464e-01 4.735e-01 0.309 0.757179
## oldclaim_log 9.548e-02 1.666e-01 0.573 0.566671
## clm_freq_log -7.163e-01 3.054e+00 -0.235 0.814562
## mvr_pts_log 2.391e-01 4.999e-01 0.478 0.632399
## tif_log 4.373e-01 5.248e-01 0.833 0.404741
## kidsdriv_log 2.023e+00 2.134e+00 0.948 0.343007
## homekids_log 4.629e-01 1.461e+00 0.317 0.751309
## inter 2.177e-02 3.105e-02 0.701 0.483234
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 740.32 on 638 degrees of freedom
## Residual deviance: 542.69 on 583 degrees of freedom

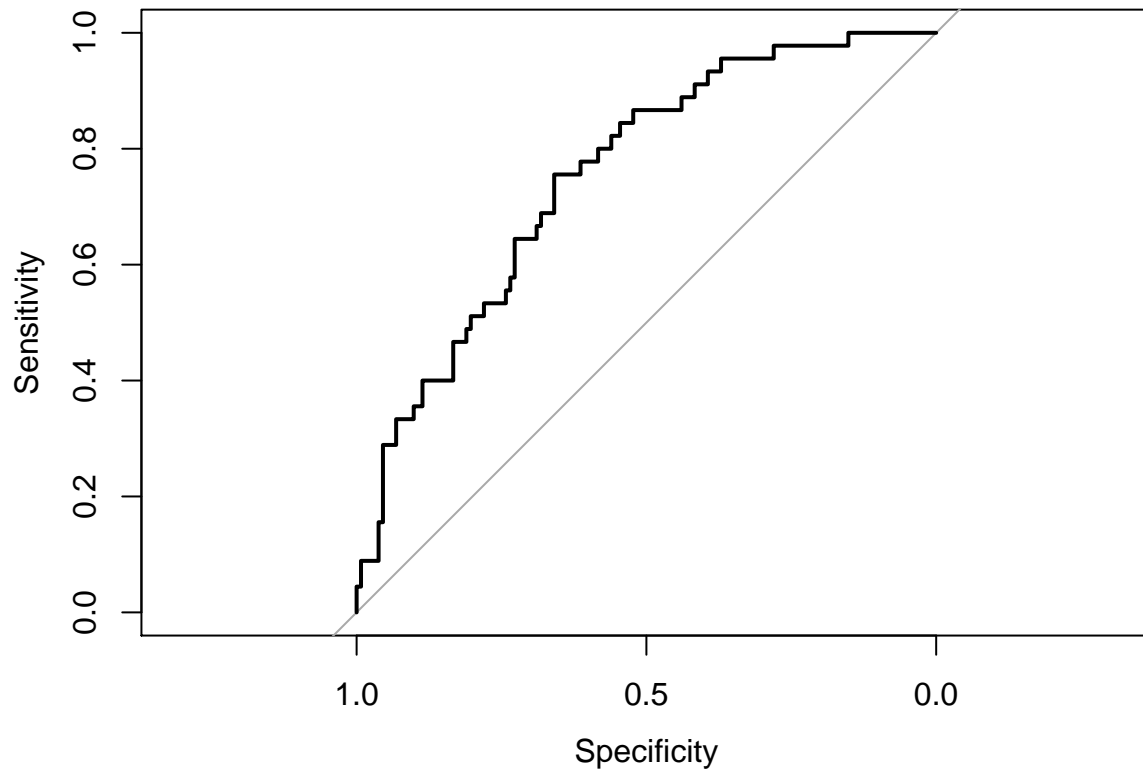
```

```

## (1 observation deleted due to missingness)
## AIC: 654.69
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  27
##           1   16  18
##
##           Accuracy : 0.7571
##           95% CI : (0.687, 0.8183)
##           No Information Rate : 0.7458
##           P-Value [Acc > NIR] : 0.4030
##
##           Kappa : 0.3032
##
## Mcnemar's Test P-Value : 0.1273
##
##           Sensitivity : 0.8788
##           Specificity : 0.4000
##           Pos Pred Value : 0.8112
##           Neg Pred Value : 0.5294
##           Prevalence : 0.7458
##           Detection Rate : 0.6554
##           Detection Prevalence : 0.8079
##           Balanced Accuracy : 0.6394
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.753703703703704"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 132 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7537
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1116  -0.6321  -0.3462   0.3659   3.2702
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.498e+01  1.501e+01   0.998  0.31822
## KIDSDRIV      -1.360e+00  1.908e+00  -0.713  0.47588
## AGE           -1.778e-01  1.153e-01  -1.542  0.12299
## HOMEKIDS      -8.867e-01  6.821e-01  -1.300  0.19360
## YOJ           -1.996e-01  1.339e-01  -1.491  0.13584
## INCOME        -7.706e-06  1.343e-05  -0.574  0.56601
## HOME_VAL       5.990e-06  9.015e-06   0.664  0.50640
## TRAVTIME       4.331e-02  2.311e-02   1.874  0.06095
## BLUEBOOK       5.078e-05  3.974e-05   1.278  0.20132
```

```

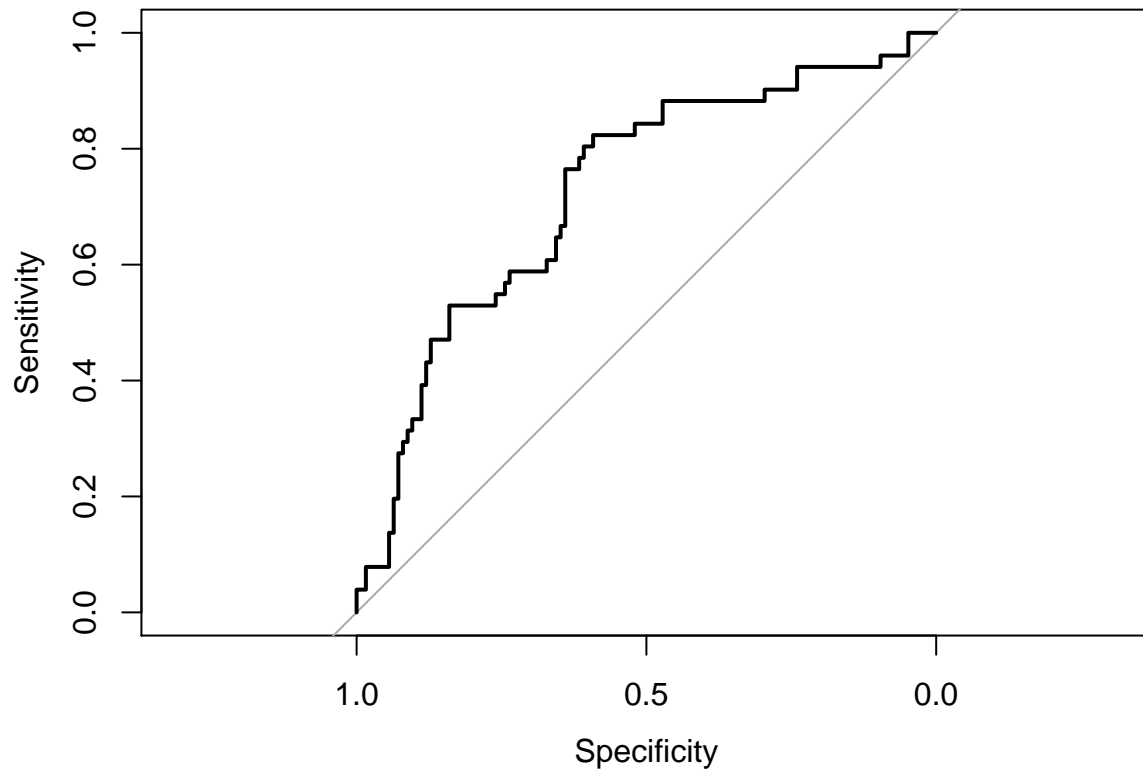
## TIF -8.963e-02 9.614e-02 -0.932 0.35119
## OLDCLAIM 1.444e-05 2.161e-05 0.669 0.50377
## CLM_FREQ 2.443e-01 1.066e+00 0.229 0.81871
## MVR_PTS 1.605e-02 1.802e-01 0.089 0.92902
## CAR_AGE -4.097e-02 7.493e-02 -0.547 0.58456
## PARENT1_Yes 2.893e-01 4.784e-01 0.605 0.54543
## MSTATUS_Yes -6.315e-01 3.381e-01 -1.868 0.06182 .
## SEX_z_F -9.026e-01 4.750e-01 -1.900 0.05742 .
## EDUCATION_.High.School 3.453e-01 7.810e-01 0.442 0.65834
## EDUCATION_Bachelors -4.319e-01 6.514e-01 -0.663 0.50729
## EDUCATION_Masters 5.652e-01 5.531e-01 1.022 0.30680
## EDUCATION_z_High.School 5.483e-01 7.045e-01 0.778 0.43643
## JOB_ -1.492e+00 7.345e-01 -2.032 0.04220 *
## JOB_Clerical -3.878e-01 5.072e-01 -0.765 0.44455
## JOB_Doctor -1.934e-01 9.474e-01 -0.204 0.83821
## JOB_Home.Maker -1.172e-01 6.957e-01 -0.169 0.86617
## JOB_Lawyer -7.588e-01 6.850e-01 -1.108 0.26801
## JOB_Manager -9.636e-01 5.068e-01 -1.901 0.05724 .
## JOB_Student -1.158e+00 7.671e-01 -1.509 0.13124
## JOB_z_Blue.Collar -3.805e-01 4.667e-01 -0.815 0.41489
## CAR_USE_Commercial 1.171e+00 3.657e-01 3.201 0.00137 **
## CAR_TYPE_Panel.Truck -7.979e-01 6.544e-01 -1.219 0.22276
## CAR_TYPE_Pickup 9.496e-01 4.057e-01 2.340 0.01926 *
## CAR_TYPE_Sports.Car 1.740e+00 5.552e-01 3.134 0.00173 **
## CAR_TYPE_Van 3.882e-01 4.735e-01 0.820 0.41235
## CAR_TYPE_z_SUV 2.001e+00 4.940e-01 4.052 5.08e-05 ***
## RED_CAR_no -2.849e-01 3.421e-01 -0.833 0.40490
## REVOKED_Yes 1.480e-01 4.273e-01 0.346 0.72900
## URBANICITY_z_Highly.Rural..Rural -2.746e+00 4.453e-01 -6.166 7.00e-10 ***
## YOJ_NA -3.476e-01 4.478e-01 -0.776 0.43760
## INCOME_NA 5.605e-02 5.404e-01 0.104 0.91739
## CAR_AGE_NA -2.984e-01 4.503e-01 -0.663 0.50752
## HOME_VAL_NA -2.585e-01 2.866e-01 -0.902 0.36722
## ageSquared 1.683e-03 1.247e-03 1.350 0.17703
## yojSquared 1.041e-02 7.019e-03 1.483 0.13803
## income_log -6.997e-02 2.653e-01 -0.264 0.79197
## homeval_log -6.659e-01 1.366e+00 -0.488 0.62584
## travtime_log -6.312e-01 6.267e-01 -1.007 0.31381
## bluebook_log -2.363e-01 4.812e-01 -0.491 0.62334
## carage_log 9.982e-02 4.828e-01 0.207 0.83620
## oldclaim_log 1.143e-01 1.735e-01 0.659 0.50996
## clm_freq_log -1.221e+00 3.266e+00 -0.374 0.70858
## mvr_pts_log -7.974e-02 5.211e-01 -0.153 0.87840
## tif_log 3.281e-01 5.387e-01 0.609 0.54245
## kidsdriv_log -1.278e+00 2.371e+00 -0.539 0.58982
## homekids_log 1.849e+00 1.492e+00 1.239 0.21518
## inter 7.289e-02 3.862e-02 1.887 0.05912 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 728.45 on 639 degrees of freedom
## Residual deviance: 526.52 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 638.52
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 111  32
##           1  14  19
##
##           Accuracy : 0.7386
##           95% CI : (0.6672, 0.8019)
##           No Information Rate : 0.7102
##           P-Value [Acc > NIR] : 0.22892
##
##           Kappa : 0.2909
##
## Mcnemar's Test P-Value : 0.01219
##
##           Sensitivity : 0.8880
##           Specificity : 0.3725
##           Pos Pred Value : 0.7762
##           Neg Pred Value : 0.5758
##           Prevalence : 0.7102
##           Detection Rate : 0.6307
##           Detection Prevalence : 0.8125
##           Balanced Accuracy : 0.6303
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.726745098039216"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 125 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.7267
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2504  -0.7139  -0.3919   0.6486   2.9616
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.057e+01  1.436e+01   0.736  0.46170
## KIDSDRIV      -1.376e+00  1.851e+00  -0.744  0.45705
## AGE           -7.084e-02  1.068e-01  -0.663  0.50710
## HOMEKIDS      -6.657e-01  6.300e-01  -1.057  0.29067
## YOJ           -1.614e-01  1.319e-01  -1.224  0.22107
## INCOME        -1.514e-05  1.261e-05  -1.201  0.22975
## HOME_VAL       7.580e-06  8.532e-06   0.888  0.37434
## TRAVTIME      1.918e-02  2.122e-02   0.904  0.36613
## BLUEBOOK      3.902e-05  3.815e-05   1.023  0.30641
```

```

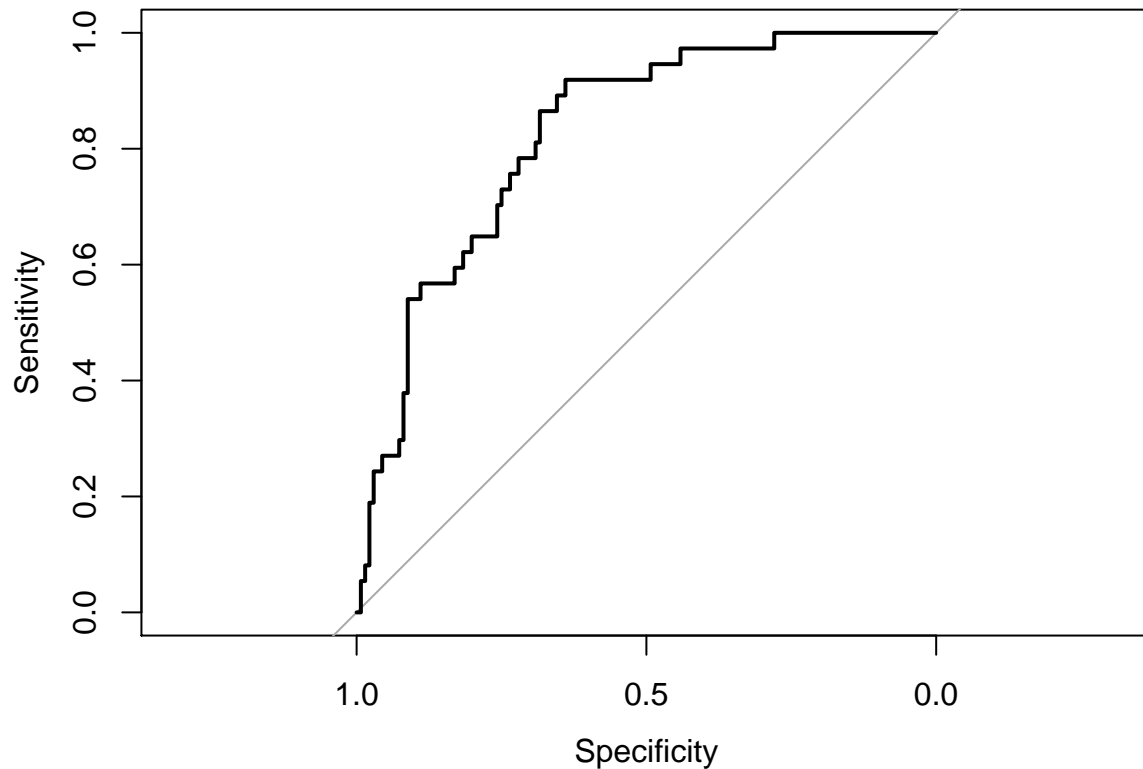
## TIF -6.300e-02 9.083e-02 -0.694 0.48794
## OLDCLAIM -1.627e-05 2.209e-05 -0.736 0.46161
## CLM_FREQ 2.756e-01 9.033e-01 0.305 0.76027
## MVR_PTS 4.762e-02 1.636e-01 0.291 0.77093
## CAR_AGE -6.431e-02 7.070e-02 -0.910 0.36302
## PARENT1_Yes 3.934e-01 4.579e-01 0.859 0.39035
## MSTATUS_Yes -6.512e-01 3.247e-01 -2.005 0.04492 *
## SEX_z_F -7.714e-01 4.435e-01 -1.739 0.08196 .
## EDUCATION_.High.School -1.141e-01 7.671e-01 -0.149 0.88178
## EDUCATION_Bachelors -1.103e-01 6.513e-01 -0.169 0.86553
## EDUCATION_Masters 8.989e-01 5.675e-01 1.584 0.11318
## EDUCATION_z_High.School 3.560e-01 7.020e-01 0.507 0.61204
## JOB_ -1.069e+00 6.710e-01 -1.593 0.11113
## JOB_Clerical -5.584e-02 4.769e-01 -0.117 0.90679
## JOB_Doctor 1.998e-02 9.134e-01 0.022 0.98255
## JOB_Home.Maker -3.896e-01 6.691e-01 -0.582 0.56043
## JOB_Lawyer -7.154e-01 6.180e-01 -1.158 0.24699
## JOB_Manager -7.768e-01 4.598e-01 -1.689 0.09117 .
## JOB_Student -6.448e-01 7.234e-01 -0.891 0.37275
## JOB_z_Blue.Collar -2.112e-01 4.358e-01 -0.485 0.62797
## CAR_USE_Commercial 5.494e-01 3.379e-01 1.626 0.10400
## CAR_TYPE_Panel.Truck -1.621e-01 6.100e-01 -0.266 0.79042
## CAR_TYPE_Pickup 9.407e-01 3.865e-01 2.434 0.01492 *
## CAR_TYPE_Sports.Car 1.545e+00 5.264e-01 2.935 0.00334 **
## CAR_TYPE_Van 5.238e-01 4.513e-01 1.161 0.24584
## CAR_TYPE_z_SUV 1.463e+00 4.573e-01 3.198 0.00138 **
## RED_CAR_no -1.533e-01 3.217e-01 -0.477 0.63365
## REVOKED_Yes 6.127e-01 4.146e-01 1.478 0.13943
## URBANICITY_z_Highly.Rural..Rural -2.159e+00 3.988e-01 -5.415 6.15e-08 ***
## YOJ_NA 9.044e-03 4.267e-01 0.021 0.98309
## INCOME_NA 1.130e-01 5.613e-01 0.201 0.84043
## CAR_AGE_NA -2.355e-01 4.411e-01 -0.534 0.59337
## HOME_VAL_NA -7.324e-02 2.800e-01 -0.262 0.79362
## ageSquared 7.292e-04 1.154e-03 0.632 0.52757
## yojSquared 7.215e-03 6.691e-03 1.078 0.28090
## income_log -7.844e-02 2.756e-01 -0.285 0.77591
## homeval_log -6.562e-01 1.326e+00 -0.495 0.62058
## travtime_log -8.524e-03 5.857e-01 -0.015 0.98839
## bluebook_log -2.382e-01 4.551e-01 -0.523 0.60072
## carage_log 1.493e-01 4.584e-01 0.326 0.74472
## oldclaim_log 1.011e-01 1.548e-01 0.653 0.51375
## clm_freq_log -8.049e-01 2.819e+00 -0.286 0.77522
## mvr_pts_log 5.577e-02 4.871e-01 0.115 0.90883
## tif_log 1.498e-01 5.088e-01 0.294 0.76849
## kidsdriv_log 4.110e-01 2.115e+00 0.194 0.84589
## homekids_log 1.752e+00 1.384e+00 1.266 0.20556
## inter 4.026e-02 3.707e-02 1.086 0.27739
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 758.65 on 642 degrees of freedom
## Residual deviance: 575.22 on 587 degrees of freedom

```

```

## AIC: 687.22
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 124  22
##           1  12  15
##
##           Accuracy : 0.8035
##           95% CI : (0.7363, 0.8599)
##           No Information Rate : 0.7861
##           P-Value [Acc > NIR] : 0.3266
##
##           Kappa : 0.3518
##
## Mcnemar's Test P-Value : 0.1227
##
##           Sensitivity : 0.9118
##           Specificity : 0.4054
##           Pos Pred Value : 0.8493
##           Neg Pred Value : 0.5556
##           Prevalence : 0.7861
##           Detection Rate : 0.7168
##           Detection Prevalence : 0.8439
##           Balanced Accuracy : 0.6586
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.825914149443561"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 136 controls (dfPred_raw$class 0) < 37 cases (dfPred_raw$class 1).
## Area under the curve: 0.8259
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0577  -0.6738  -0.3256   0.2264   2.7512
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.240e+01  1.487e+01   1.507 0.131915
## KIDSDRIV      -2.232e+00  1.814e+00  -1.230 0.218572
## AGE           -2.371e-01  1.123e-01  -2.110 0.034822 *
## HOMEKIDS       -7.177e-01  6.934e-01  -1.035 0.300646
## YOJ            -8.989e-02  1.396e-01  -0.644 0.519542
## INCOME         -1.885e-05  1.433e-05  -1.315 0.188389
## HOME_VAL        1.060e-05  9.287e-06   1.141 0.253867
## TRAVTIME        3.121e-02  2.179e-02   1.432 0.152159
## BLUEBOOK        5.746e-05  3.926e-05   1.464 0.143307
```

```

## TIF -7.834e-02 9.559e-02 -0.819 0.412509
## OLDCLAIM 6.966e-06 2.090e-05 0.333 0.738946
## CLM_FREQ -2.669e-01 9.409e-01 -0.284 0.776657
## MVRPTS 2.332e-02 1.788e-01 0.130 0.896259
## CAR_AGE -6.651e-02 7.120e-02 -0.934 0.350218
## PARENT1_Yes 1.196e-01 4.852e-01 0.246 0.805372
## MSTATUS_Yes -9.864e-01 3.494e-01 -2.823 0.004755 **
## SEX_z_F -6.320e-01 4.596e-01 -1.375 0.169164
## EDUCATION_.High.School -2.512e-01 7.498e-01 -0.335 0.737582
## EDUCATION_Bachelors -4.319e-01 6.135e-01 -0.704 0.481452
## EDUCATION_Masters 7.292e-01 5.375e-01 1.357 0.174931
## EDUCATION_z_High.School 2.228e-01 6.677e-01 0.334 0.738649
## JOB_ -1.199e+00 6.586e-01 -1.821 0.068617 .
## JOB_Clerical -4.776e-01 4.948e-01 -0.965 0.334406
## JOB_Doctor -6.412e-01 9.232e-01 -0.694 0.487371
## JOB_Home.Maker 1.471e-02 6.673e-01 0.022 0.982415
## JOB_Lawyer -1.428e+00 6.679e-01 -2.139 0.032474 *
## JOB_Manager -7.011e-01 4.673e-01 -1.500 0.133557
## JOB_Student -5.187e-01 7.321e-01 -0.708 0.478638
## JOB_z_Blue.Collar -8.510e-02 4.515e-01 -0.188 0.850515
## CAR_USE_Commercial 5.729e-01 3.507e-01 1.634 0.102291
## CAR_TYPE_Panel.Truck 7.374e-02 6.237e-01 0.118 0.905893
## CAR_TYPE_Pickup 1.184e+00 4.353e-01 2.719 0.006540 **
## CAR_TYPE_Sports.Car 2.024e+00 5.351e-01 3.782 0.000156 ***
## CAR_TYPE_Van 4.438e-01 4.699e-01 0.944 0.344946
## CAR_TYPE_z_SUV 2.100e+00 4.845e-01 4.334 1.47e-05 ***
## RED_CAR_no -2.914e-01 3.500e-01 -0.832 0.405154
## REVOKED_Yes 3.797e-01 4.046e-01 0.938 0.348060
## URBANICITY_z_Highly.Rural..Rural -2.864e+00 5.183e-01 -5.527 3.26e-08 ***
## YOJ_NA -3.761e-01 4.148e-01 -0.907 0.364552
## INCOME_NA 1.543e-01 5.623e-01 0.274 0.783745
## CAR_AGE_NA -5.386e-01 4.919e-01 -1.095 0.273522
## HOME_VAL_NA 1.734e-01 3.056e-01 0.567 0.570393
## ageSquared 2.470e-03 1.208e-03 2.044 0.040926 *
## yojSquared 5.263e-03 7.271e-03 0.724 0.469190
## income_log 3.853e-01 3.410e-01 1.130 0.258623
## homeval_log -1.603e+00 1.448e+00 -1.107 0.268315
## travtime_log -2.922e-01 6.038e-01 -0.484 0.628466
## bluebook_log -3.832e-01 4.977e-01 -0.770 0.441352
## carage_log 1.905e-01 4.695e-01 0.406 0.684942
## oldclaim_log -1.231e-02 1.631e-01 -0.075 0.939867
## clm_freq_log 1.036e+00 2.940e+00 0.352 0.724631
## mvr_pts_log -2.277e-02 5.216e-01 -0.044 0.965181
## tif_log 1.854e-01 5.352e-01 0.346 0.729090
## kidsdriv_log 1.033e+00 2.239e+00 0.461 0.644654
## homekids_log 1.520e+00 1.491e+00 1.020 0.307825
## inter 5.447e-02 3.459e-02 1.575 0.115312
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 723.13 on 641 degrees of freedom
## Residual deviance: 528.10 on 586 degrees of freedom

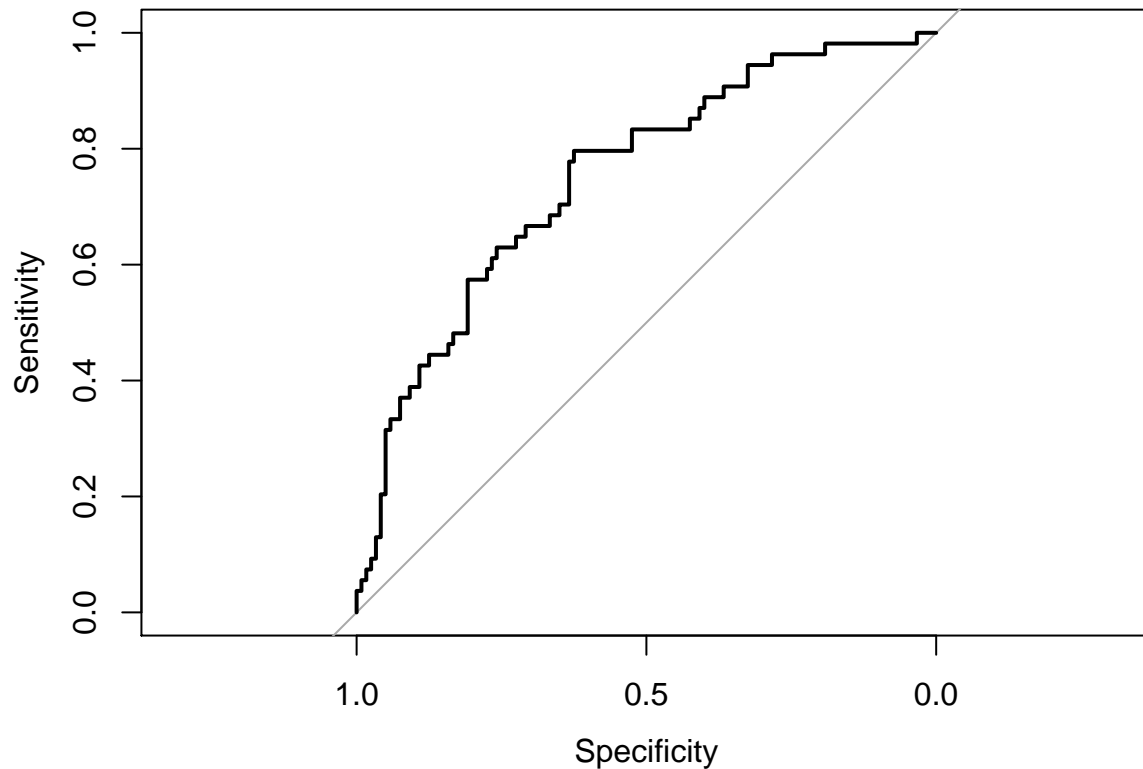
```



```

## (1 observation deleted due to missingness)
## AIC: 640.1
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 105  30
##           1  15  24
##
##           Accuracy : 0.7414
##           95% CI : (0.6697, 0.8047)
##           No Information Rate : 0.6897
##           P-Value [Acc > NIR] : 0.08018
##
##           Kappa : 0.3459
##
## Mcnemar's Test P-Value : 0.03689
##
##           Sensitivity : 0.8750
##           Specificity : 0.4444
##           Pos Pred Value : 0.7778
##           Neg Pred Value : 0.6154
##           Prevalence : 0.6897
##           Detection Rate : 0.6034
##           Detection Prevalence : 0.7759
##           Balanced Accuracy : 0.6597
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.750154320987654"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 120 controls (dfPred_raw$class 0) < 54 cases (dfPred_raw$class 1).
## Area under the curve: 0.7502
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2122  -0.7093  -0.3663   0.4802   3.0994
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.074e+01  1.503e+01   0.715  0.474911
## KIDSDRIV      -8.203e-01  1.870e+00  -0.439  0.660950
## AGE           -1.144e-01  1.036e-01  -1.104  0.269566
## HOMEKIDS      -6.683e-01  6.904e-01  -0.968  0.333035
## YOJ           -2.612e-01  1.262e-01  -2.070  0.038474 *
## INCOME        -6.742e-06  1.309e-05  -0.515  0.606424
## HOME_VAL       2.977e-06  8.749e-06   0.340  0.733617
## TRAVTIME       2.925e-02  2.138e-02   1.368  0.171186
## BLUEBOOK       2.178e-05  4.090e-05   0.533  0.594268
```

```

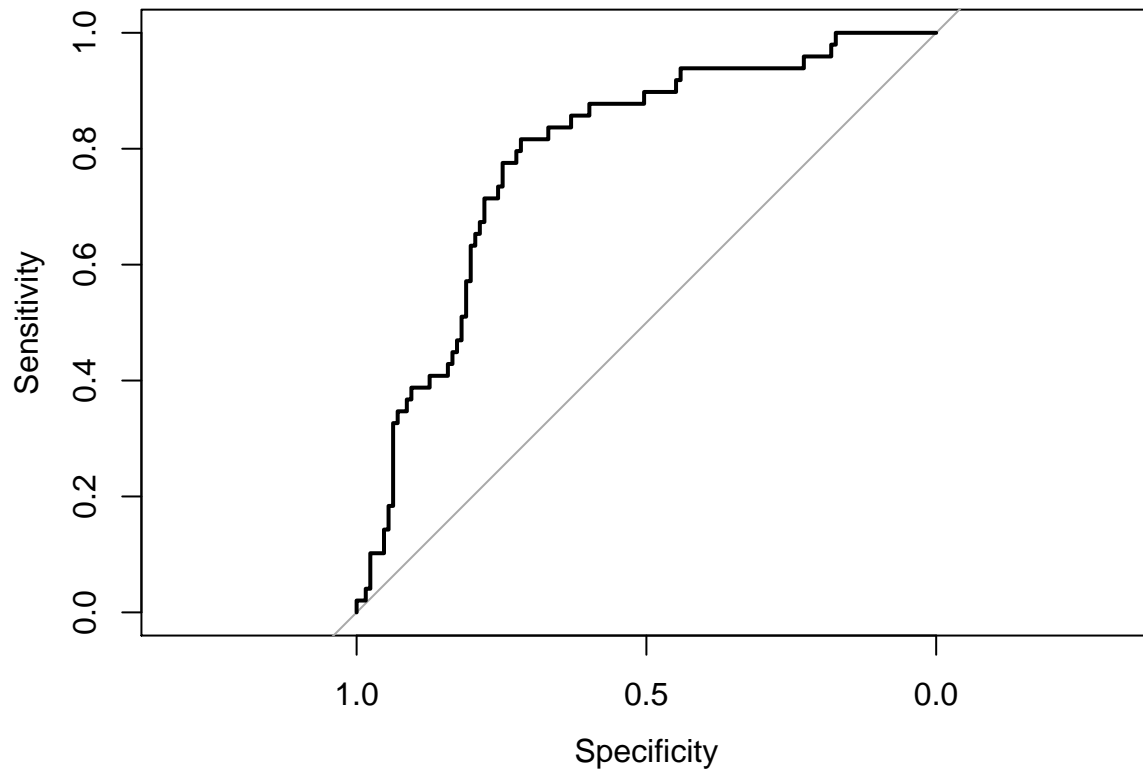
## TIF -3.419e-02 8.637e-02 -0.396 0.692190
## OLDCLAIM -2.987e-06 2.032e-05 -0.147 0.883173
## CLM_FREQ -5.452e-02 9.771e-01 -0.056 0.955499
## MVRPTS 1.242e-02 1.623e-01 0.077 0.938981
## CAR_AGE -9.703e-03 6.690e-02 -0.145 0.884672
## PARENT1_Yes 4.174e-01 4.831e-01 0.864 0.387614
## MSTATUS_Yes -6.551e-01 3.295e-01 -1.988 0.046788 *
## SEX_z_F -6.684e-01 4.780e-01 -1.398 0.162009
## EDUCATION_.High.School -3.183e-01 7.766e-01 -0.410 0.681861
## EDUCATION_Bachelors -2.985e-01 6.306e-01 -0.473 0.635995
## EDUCATION_Masters 5.769e-01 5.375e-01 1.073 0.283079
## EDUCATION_z_High.School -9.296e-02 6.928e-01 -0.134 0.893260
## JOB_ -9.211e-01 6.358e-01 -1.449 0.147387
## JOB_Clerical -2.530e-01 4.975e-01 -0.509 0.611055
## JOB_Doctor -7.061e-01 9.093e-01 -0.777 0.437411
## JOB_Home.Maker -1.729e-01 6.416e-01 -0.270 0.787537
## JOB_Lawyer -1.129e+00 6.195e-01 -1.822 0.068500 .
## JOB_Manager -7.036e-01 4.497e-01 -1.565 0.117690
## JOB_Student -7.121e-01 7.303e-01 -0.975 0.329500
## JOB_z_Blue.Collar -1.296e-01 4.562e-01 -0.284 0.776286
## CAR_USE_Commercial 5.505e-01 3.614e-01 1.523 0.127670
## CAR_TYPE_Panel.Truck 6.317e-02 6.065e-01 0.104 0.917042
## CAR_TYPE_Pickup 6.976e-01 4.104e-01 1.700 0.089189 .
## CAR_TYPE_Sports.Car 1.452e+00 5.259e-01 2.761 0.005764 **
## CAR_TYPE_Van 2.592e-01 4.762e-01 0.544 0.586260
## CAR_TYPE_z_SUV 1.635e+00 4.662e-01 3.506 0.000455 ***
## RED_CAR_no -1.067e-01 3.443e-01 -0.310 0.756574
## REVOKED_Yes 5.067e-01 3.978e-01 1.274 0.202832
## URBANICITY_z_Highly.Rural..Rural -2.719e+00 4.855e-01 -5.599 2.15e-08 ***
## YOJ_NA -4.038e-01 4.275e-01 -0.944 0.344944
## INCOME_NA -4.782e-02 4.993e-01 -0.096 0.923705
## CAR_AGE_NA -8.986e-02 5.297e-01 -0.170 0.865279
## HOME_VAL_NA -1.426e-01 2.892e-01 -0.493 0.622086
## ageSquared 1.240e-03 1.135e-03 1.093 0.274602
## yojSquared 1.333e-02 6.618e-03 2.014 0.044032 *
## income_log -5.304e-02 2.897e-01 -0.183 0.854752
## homeval_log -6.803e-01 1.401e+00 -0.486 0.627243
## travtime_log -2.176e-01 5.914e-01 -0.368 0.712964
## bluebook_log 1.194e-01 5.005e-01 0.239 0.811421
## carage_log -2.747e-01 4.470e-01 -0.615 0.538834
## oldclaim_log 2.986e-02 1.606e-01 0.186 0.852536
## clm_freq_log 3.621e-01 3.004e+00 0.121 0.904060
## mvr_pts_log -2.775e-02 4.879e-01 -0.057 0.954637
## tif_log 5.349e-02 5.037e-01 0.106 0.915437
## kidsdriv_log 2.329e-01 2.276e+00 0.102 0.918488
## homekids_log 1.426e+00 1.508e+00 0.946 0.344106
## inter 2.698e-02 3.516e-02 0.767 0.443007
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 732.67 on 639 degrees of freedom
## Residual deviance: 548.29 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 660.29
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 115  31
##           1  12  18
##
##           Accuracy : 0.7557
##           95% CI : (0.6853, 0.8172)
##           No Information Rate : 0.7216
##           P-Value [Acc > NIR] : 0.177944
##
##           Kappa : 0.3097
##
## Mcnemar's Test P-Value : 0.006052
##
##           Sensitivity : 0.9055
##           Specificity : 0.3673
##           Pos Pred Value : 0.7877
##           Neg Pred Value : 0.6000
##           Prevalence : 0.7216
##           Detection Rate : 0.6534
##           Detection Prevalence : 0.8295
##           Balanced Accuracy : 0.6364
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.78740157480315"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 127 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.7874
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3367  -0.6508  -0.3396   0.5118   2.7401
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.069e+01  1.470e+01   1.408 0.159135
## KIDSDRIV      -3.547e-01  1.777e+00  -0.200 0.841778
## AGE           -1.251e-01  1.095e-01  -1.142 0.253341
## HOMEKIDS      -1.399e+00  7.595e-01  -1.842 0.065456
## YOJ           -1.743e-01  1.324e-01  -1.316 0.188199
## INCOME        -4.143e-06  1.342e-05  -0.309 0.757611
## HOME_VAL       6.220e-06  8.703e-06   0.715 0.474823
## TRAVTIME      2.234e-02  2.211e-02   1.010 0.312353
## BLUEBOOK      4.768e-05  4.070e-05   1.171 0.241457
```

```

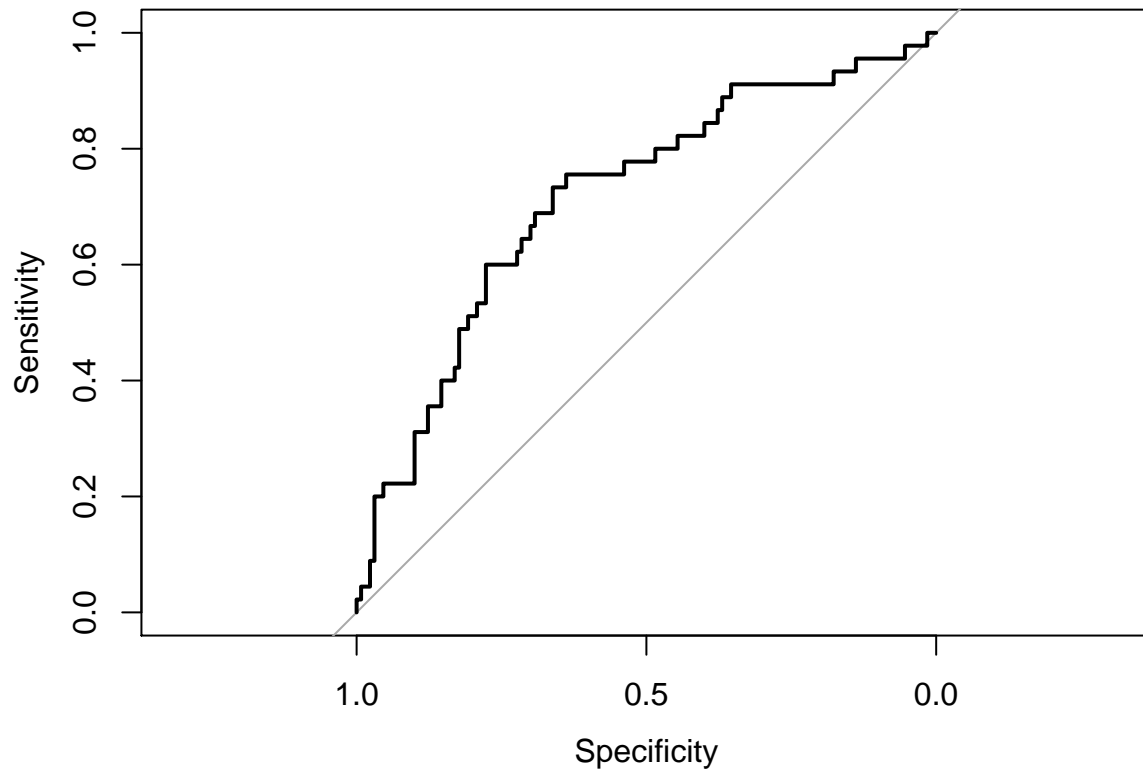
## TIF -1.037e-01 9.435e-02 -1.099 0.271680
## OLDCLAIM -2.053e-05 2.240e-05 -0.916 0.359503
## CLM_FREQ 3.752e-01 1.033e+00 0.363 0.716508
## MVR_PTS -1.477e-01 1.749e-01 -0.844 0.398531
## CAR_AGE -1.030e-02 7.528e-02 -0.137 0.891151
## PARENT1_Yes 9.018e-01 4.913e-01 1.836 0.066407 .
## MSTATUS_Yes -4.865e-01 3.338e-01 -1.457 0.144999
## SEX_z_F -5.325e-01 4.886e-01 -1.090 0.275807
## EDUCATION_.High.School -1.229e-01 7.621e-01 -0.161 0.871872
## EDUCATION_Bachelors -4.329e-01 6.166e-01 -0.702 0.482606
## EDUCATION_Masters 4.634e-01 5.301e-01 0.874 0.382015
## EDUCATION_z_High.School 9.542e-02 6.762e-01 0.141 0.887773
## JOB_ -9.478e-01 6.710e-01 -1.413 0.157797
## JOB_Clerical -4.446e-01 4.917e-01 -0.904 0.365858
## JOB_Doctor -7.002e-01 8.467e-01 -0.827 0.408266
## JOB_Home.Maker -4.872e-01 6.855e-01 -0.711 0.477208
## JOB_Lawyer -1.117e+00 6.434e-01 -1.736 0.082582 .
## JOB_Manager -9.325e-01 4.652e-01 -2.005 0.045002 *
## JOB_Student -8.034e-01 7.581e-01 -1.060 0.289245
## JOB_z_Blue.Collar -2.752e-01 4.729e-01 -0.582 0.560635
## CAR_USE_Commercial 5.745e-01 3.616e-01 1.589 0.112124
## CAR_TYPE_Panel.Truck -1.968e-01 6.750e-01 -0.292 0.770643
## CAR_TYPE_Pickup 1.040e+00 4.041e-01 2.574 0.010053 *
## CAR_TYPE_Sports.Car 1.017e+00 5.432e-01 1.873 0.061133 .
## CAR_TYPE_Van 2.105e-01 4.540e-01 0.464 0.642924
## CAR_TYPE_z_SUV 1.622e+00 4.690e-01 3.459 0.000543 ***
## RED_CAR_no -2.491e-01 3.465e-01 -0.719 0.472262
## REVOKED_Yes 6.150e-01 3.941e-01 1.561 0.118607
## URBANICITY_z_Highly.Rural..Rural -2.820e+00 4.664e-01 -6.046 1.48e-09 ***
## YOJ_NA -2.604e-01 4.351e-01 -0.599 0.549439
## INCOME_NA -9.512e-02 5.161e-01 -0.184 0.853770
## CAR_AGE_NA -2.026e-02 4.904e-01 -0.041 0.967055
## HOME_VAL_NA -1.753e-01 2.946e-01 -0.595 0.551889
## ageSquared 1.425e-03 1.185e-03 1.203 0.229031
## yojSquared 8.630e-03 7.028e-03 1.228 0.219446
## income_log -1.049e-01 2.876e-01 -0.365 0.715439
## homeval_log -1.308e+00 1.346e+00 -0.972 0.331059
## travtime_log 9.325e-02 6.208e-01 0.150 0.880604
## bluebook_log -3.285e-01 4.891e-01 -0.672 0.501862
## carage_log -2.989e-01 4.886e-01 -0.612 0.540637
## oldclaim_log 1.795e-01 1.682e-01 1.067 0.286118
## clm_freq_log -1.432e+00 3.151e+00 -0.454 0.649476
## mvr_pts_log 4.927e-01 5.173e-01 0.952 0.340894
## tif_log 2.664e-01 5.316e-01 0.501 0.616306
## kidsdriv_log 9.255e-01 2.285e+00 0.405 0.685456
## homekids_log 2.632e+00 1.594e+00 1.651 0.098814 .
## inter 1.390e-02 3.406e-02 0.408 0.683203
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 741.56 on 640 degrees of freedom
## Residual deviance: 533.79 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 645.79
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 111  29
##           1  19  16
##
##           Accuracy : 0.7257
##           95% CI : (0.6533, 0.7903)
##           No Information Rate : 0.7429
##           P-Value [Acc > NIR] : 0.7304
##
##           Kappa : 0.2258
##
## Mcnemar's Test P-Value : 0.1939
##
##           Sensitivity : 0.8538
##           Specificity : 0.3556
##           Pos Pred Value : 0.7929
##           Neg Pred Value : 0.4571
##           Prevalence : 0.7429
##           Detection Rate : 0.6343
##           Detection Prevalence : 0.8000
##           Balanced Accuracy : 0.6047
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.715213675213675"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 130 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7152
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1496  -0.6702  -0.3730   0.5228   2.8779
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.647e+00  1.609e+01  -0.102  0.918504
## KIDSDRIV       -1.020e+00  1.786e+00  -0.571  0.568007
## AGE            -1.323e-01  1.044e-01  -1.267  0.205170
## HOMEKIDS       -1.255e+00  6.814e-01  -1.843  0.065394
## YOJ            -1.820e-01  1.345e-01  -1.354  0.175783
## INCOME         -1.285e-05  1.306e-05  -0.985  0.324833
## HOME_VAL        2.780e-06  9.252e-06   0.300  0.763798
## TRAVTIME        2.344e-02  2.242e-02   1.046  0.295783
## BLUEBOOK        6.442e-05  4.028e-05   1.600  0.109703
```



```

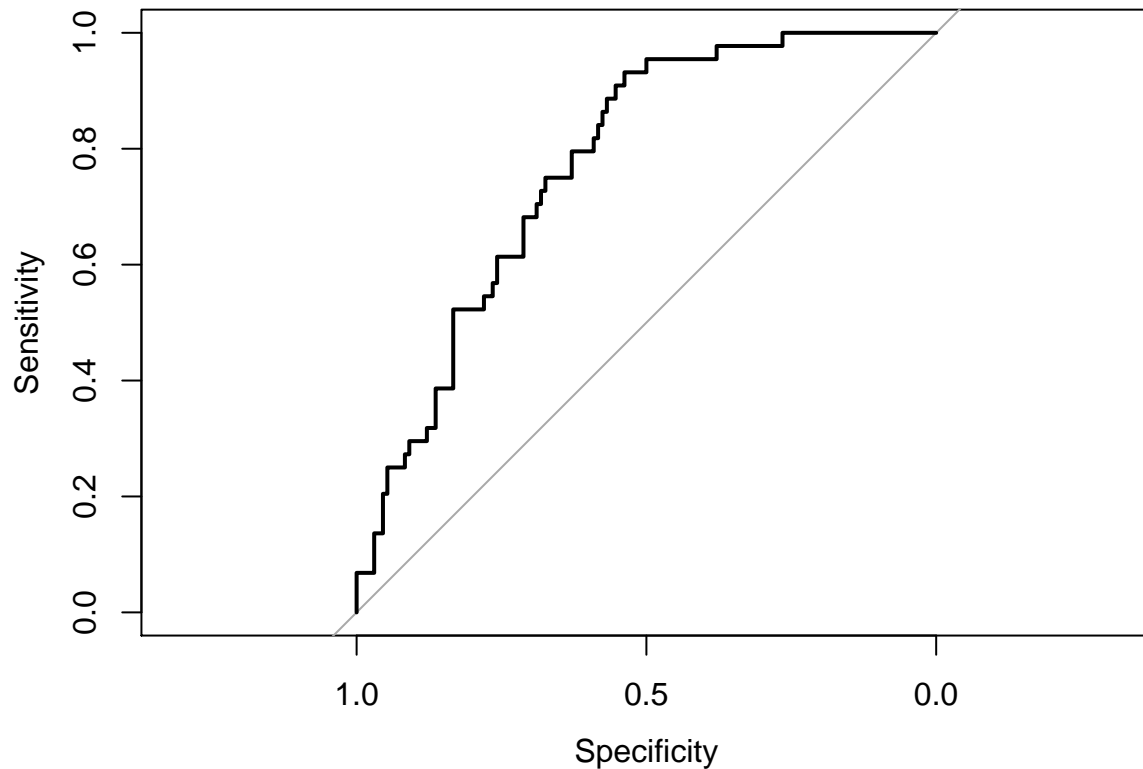
## TIF -8.544e-02 9.588e-02 -0.891 0.372862
## OLDCLAIM 1.224e-05 2.137e-05 0.573 0.566697
## CLM_FREQ -1.963e-01 1.011e+00 -0.194 0.846037
## MVRPTS 5.869e-02 1.622e-01 0.362 0.717509
## CAR_AGE -3.619e-02 6.770e-02 -0.535 0.592943
## PARENT1_Yes 7.619e-01 4.704e-01 1.620 0.105280
## MSTATUS_Yes -4.204e-01 3.349e-01 -1.255 0.209423
## SEX_z_F -1.076e+00 4.803e-01 -2.240 0.025091 *
## EDUCATION_.High.School 3.929e-01 7.986e-01 0.492 0.622694
## EDUCATION_Bachelors -2.169e-02 6.863e-01 -0.032 0.974792
## EDUCATION_Masters 8.624e-01 5.958e-01 1.447 0.147802
## EDUCATION_z_High.School 4.917e-01 7.350e-01 0.669 0.503502
## JOB_ -4.277e-01 6.700e-01 -0.638 0.523221
## JOB_Clerical 1.221e-01 4.767e-01 0.256 0.797831
## JOB_Doctor -1.417e-01 9.407e-01 -0.151 0.880240
## JOB_Home.Maker 1.829e-01 6.481e-01 0.282 0.777790
## JOB_Lawyer -6.127e-01 6.513e-01 -0.941 0.346835
## JOB_Manager -6.775e-01 4.878e-01 -1.389 0.164852
## JOB_Student -1.956e-01 7.260e-01 -0.269 0.787648
## JOB_z_Blue.Collar -1.091e-02 4.426e-01 -0.025 0.980338
## CAR_USE_Commercial 5.183e-01 3.519e-01 1.473 0.140841
## CAR_TYPE_Panel.Truck -4.784e-01 6.249e-01 -0.766 0.443940
## CAR_TYPE_Pickup 1.133e+00 4.079e-01 2.777 0.005481 **
## CAR_TYPE_Sports.Car 1.909e+00 5.450e-01 3.502 0.000462 ***
## CAR_TYPE_Van 4.595e-01 4.677e-01 0.982 0.325865
## CAR_TYPE_z_SUV 2.134e+00 4.877e-01 4.376 1.21e-05 ***
## RED_CAR_no -2.501e-01 3.403e-01 -0.735 0.462418
## REVOKED_Yes 3.365e-01 4.132e-01 0.815 0.415354
## URBANICITY_z_Highly.Rural..Rural -2.281e+00 3.971e-01 -5.744 9.24e-09 ***
## YOJ_NA -5.563e-02 4.415e-01 -0.126 0.899722
## INCOME_NA 1.273e-01 5.693e-01 0.224 0.823055
## CAR_AGE_NA 3.492e-01 5.645e-01 0.619 0.536134
## HOME_VAL_NA -3.005e-01 2.868e-01 -1.048 0.294706
## ageSquared 1.539e-03 1.134e-03 1.357 0.174799
## yojSquared 9.575e-03 6.805e-03 1.407 0.159382
## income_log -8.741e-02 2.855e-01 -0.306 0.759505
## homeval_log 3.686e-01 1.499e+00 0.246 0.805755
## travtime_log 6.960e-02 6.438e-01 0.108 0.913916
## bluebook_log -2.106e-01 4.917e-01 -0.428 0.668433
## carage_log -3.604e-02 4.487e-01 -0.080 0.935981
## oldclaim_log 4.446e-02 1.653e-01 0.269 0.787911
## clm_freq_log 4.872e-01 3.095e+00 0.157 0.874925
## mvr_pts_log -1.639e-01 4.841e-01 -0.339 0.734898
## tif_log 1.929e-01 5.298e-01 0.364 0.715794
## kidsdriv_log 2.296e-01 2.123e+00 0.108 0.913848
## homekids_log 2.687e+00 1.460e+00 1.840 0.065780 .
## inter 3.784e-02 3.563e-02 1.062 0.288254
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 742.96 on 639 degrees of freedom
## Residual deviance: 551.09 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 663.09
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  27
##           1  18  17
##
##           Accuracy : 0.7443
##           95% CI : (0.6732, 0.807)
##           No Information Rate : 0.75
##           P-Value [Acc > NIR] : 0.6082
##
##           Kappa : 0.2683
##
## Mcnemar's Test P-Value : 0.2330
##
##           Sensitivity : 0.8636
##           Specificity : 0.3864
##           Pos Pred Value : 0.8085
##           Neg Pred Value : 0.4857
##           Prevalence : 0.7500
##           Detection Rate : 0.6477
##           Detection Prevalence : 0.8011
##           Balanced Accuracy : 0.6250
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.773071625344353"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 132 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7731
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3358  -0.6777  -0.3302   0.5696   2.5420
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.748e+01  1.476e+01   2.539  0.01110 *
## KIDSDRIV       1.344e-01  1.758e+00   0.076  0.93906
## AGE           -2.169e-01  1.043e-01  -2.079  0.03760 *
## HOMEKIDS       -2.609e-01  6.876e-01  -0.379  0.70435
## YOJ            -1.517e-01  1.361e-01  -1.114  0.26516
## INCOME          -4.355e-05  1.421e-05  -3.064  0.00218 **
## HOME_VAL        2.627e-05  9.185e-06   2.860  0.00424 **
## TRAVTIME        2.558e-02  2.157e-02   1.185  0.23583
## BLUEBOOK        5.979e-05  4.147e-05   1.442  0.14940
```

```

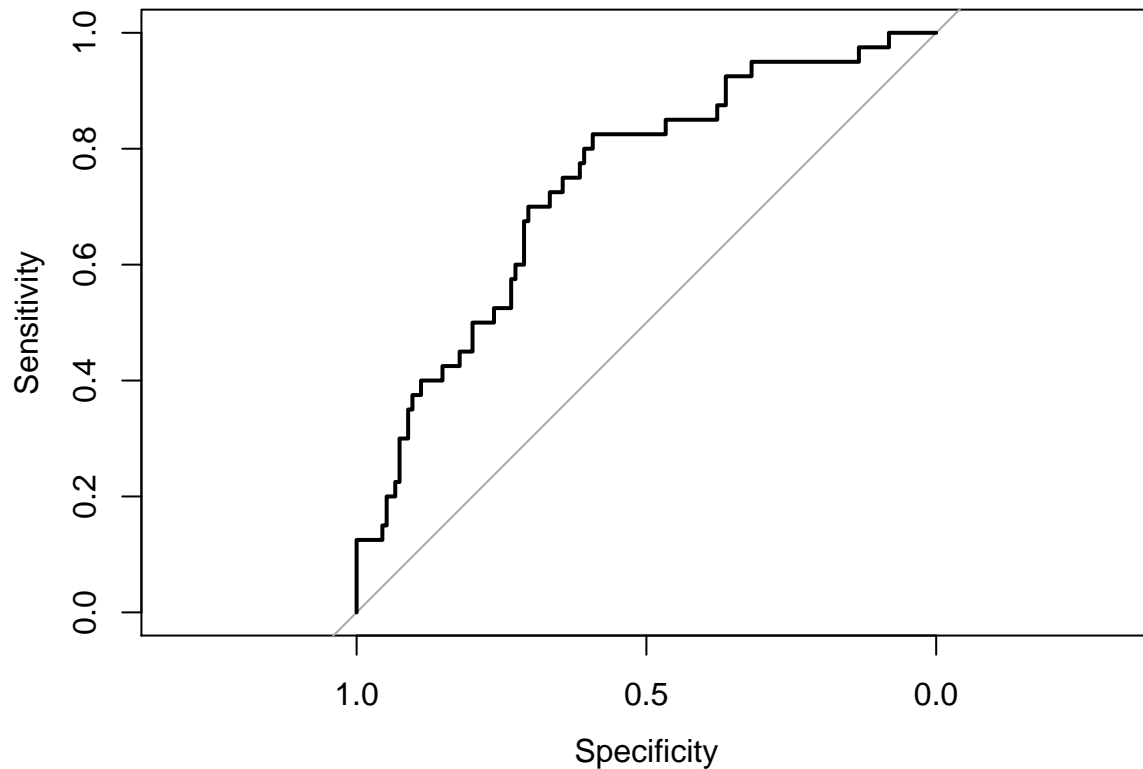
## TIF -1.065e-01 9.865e-02 -1.080 0.28027
## OLDCLAIM -8.180e-07 2.093e-05 -0.039 0.96883
## CLM_FREQ -5.476e-02 9.891e-01 -0.055 0.95585
## MVR_PTS -2.562e-02 1.772e-01 -0.145 0.88503
## CAR_AGE -3.738e-02 7.099e-02 -0.527 0.59850
## PARENT1_Yes 7.617e-01 4.746e-01 1.605 0.10854
## MSTATUS_Yes -8.023e-01 3.414e-01 -2.350 0.01877 *
## SEX_z_F -5.370e-01 4.574e-01 -1.174 0.24041
## EDUCATION_.High.School 3.177e-01 7.695e-01 0.413 0.67975
## EDUCATION_Bachelors 2.245e-01 6.510e-01 0.345 0.73019
## EDUCATION_Masters 1.201e+00 5.667e-01 2.119 0.03409 *
## EDUCATION_z_High.School 5.979e-01 7.058e-01 0.847 0.39691
## JOB_ -9.584e-01 6.786e-01 -1.412 0.15787
## JOB_Clerical -5.902e-01 4.928e-01 -1.198 0.23108
## JOB_Doctor 3.685e-02 9.269e-01 0.040 0.96829
## JOB_Home.Maker -4.280e-01 6.708e-01 -0.638 0.52338
## JOB_Lawyer -8.728e-01 6.562e-01 -1.330 0.18349
## JOB_Manager -8.683e-01 4.919e-01 -1.765 0.07755 .
## JOB_Student -9.845e-01 7.341e-01 -1.341 0.17987
## JOB_z_Blue.Collar -5.144e-01 4.685e-01 -1.098 0.27216
## CAR_USE_Commercial 6.367e-01 3.528e-01 1.805 0.07112 .
## CAR_TYPE_Panel.Truck 2.119e-01 6.359e-01 0.333 0.73897
## CAR_TYPE_Pickup 1.056e+00 4.292e-01 2.461 0.01385 *
## CAR_TYPE_Sports.Car 1.925e+00 5.285e-01 3.643 0.00027 ***
## CAR_TYPE_Van 7.233e-01 4.731e-01 1.529 0.12629
## CAR_TYPE_z_SUV 2.076e+00 4.738e-01 4.381 1.18e-05 ***
## RED_CAR_no -4.058e-01 3.411e-01 -1.190 0.23420
## REVOKED_Yes 3.301e-01 3.979e-01 0.830 0.40675
## URBANICITY_z_Highly.Rural..Rural -2.581e+00 4.432e-01 -5.823 5.77e-09 ***
## YOJ_NA -2.600e-01 4.242e-01 -0.613 0.53993
## INCOME_NA 1.430e-01 5.936e-01 0.241 0.80966
## CAR_AGE_NA -7.915e-01 4.937e-01 -1.603 0.10889
## HOME_VAL_NA 2.031e-02 2.992e-01 0.068 0.94587
## ageSquared 2.319e-03 1.124e-03 2.063 0.03909 *
## yojSquared 8.384e-03 7.115e-03 1.178 0.23871
## income_log 5.568e-01 3.231e-01 1.723 0.08486 .
## homeval_log -3.229e+00 1.428e+00 -2.260 0.02380 *
## travtime_log -1.345e-01 6.135e-01 -0.219 0.82649
## bluebook_log -3.349e-01 5.190e-01 -0.645 0.51882
## carage_log -1.514e-01 4.623e-01 -0.328 0.74326
## oldclaim_log 2.578e-02 1.659e-01 0.155 0.87651
## clm_freq_log 3.572e-01 3.060e+00 0.117 0.90706
## mvr_pts_log 2.446e-01 5.141e-01 0.476 0.63421
## tif_log 2.661e-01 5.402e-01 0.493 0.62226
## kidsdriv_log 4.594e-01 2.306e+00 0.199 0.84211
## homekids_log 5.550e-01 1.481e+00 0.375 0.70792
## inter 1.153e-02 3.069e-02 0.376 0.70713
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 751.55 on 640 degrees of freedom
## Residual deviance: 543.09 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 655.09
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 117  24
##           1  18  16
##
##           Accuracy : 0.76
##           95% CI : (0.6898, 0.8212)
##           No Information Rate : 0.7714
##           P-Value [Acc > NIR] : 0.6783
##
##           Kappa : 0.2815
##
## Mcnemar's Test P-Value : 0.4404
##
##           Sensitivity : 0.8667
##           Specificity : 0.4000
##           Pos Pred Value : 0.8298
##           Neg Pred Value : 0.4706
##           Prevalence : 0.7714
##           Detection Rate : 0.6686
##           Detection Prevalence : 0.8057
##           Balanced Accuracy : 0.6333
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.736851851851852"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 135 controls (dfPred_raw$class 0) < 40 cases (dfPred_raw$class 1).
## Area under the curve: 0.7369
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.28271  -0.64268  -0.34409  -0.06807   3.06582
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.119e+01  1.484e+01   0.754  0.45084
## KIDSDRIV      -6.032e-01  1.776e+00  -0.340  0.73419
## AGE           -1.312e-01  1.145e-01  -1.146  0.25187
## HOMEKIDS      -8.940e-01  6.891e-01  -1.297  0.19453
## YOJ           -1.861e-01  1.322e-01  -1.408  0.15914
## INCOME        -1.584e-05  1.358e-05  -1.167  0.24337
## HOME_VAL       1.255e-05  8.888e-06   1.412  0.15785
## TRAVTIME       1.493e-02  2.248e-02   0.664  0.50655
## BLUEBOOK      -1.587e-05  4.340e-05  -0.366  0.71459
```

```

## TIF -5.456e-02 9.529e-02 -0.573 0.56692
## OLDCLAIM -1.599e-06 2.267e-05 -0.071 0.94378
## CLM_FREQ -4.099e-01 1.088e+00 -0.377 0.70624
## MVR_PTS 1.688e-01 1.731e-01 0.975 0.32956
## CAR_AGE 7.113e-04 6.964e-02 0.010 0.99185
## PARENT1_Yes 6.491e-01 4.831e-01 1.344 0.17908
## MSTATUS_Yes -5.813e-01 3.438e-01 -1.691 0.09086 .
## SEX_z_F -8.604e-01 4.818e-01 -1.786 0.07411 .
## EDUCATION_.High.School 2.692e-01 7.988e-01 0.337 0.73615
## EDUCATION_Bachelors -3.128e-01 6.397e-01 -0.489 0.62485
## EDUCATION_Masters 7.502e-01 5.490e-01 1.366 0.17181
## EDUCATION_z_High.School 3.860e-01 6.983e-01 0.553 0.58045
## JOB_ -1.322e+00 6.916e-01 -1.912 0.05588 .
## JOB_Clerical -5.734e-01 5.032e-01 -1.140 0.25449
## JOB_Doctor -9.451e-01 9.461e-01 -0.999 0.31785
## JOB_Home.Maker -4.013e-02 6.608e-01 -0.061 0.95158
## JOB_Lawyer -1.343e+00 6.359e-01 -2.112 0.03465 *
## JOB_Manager -9.671e-01 4.990e-01 -1.938 0.05261 .
## JOB_Student -6.766e-01 7.309e-01 -0.926 0.35459
## JOB_z_Blue.Collar -4.529e-01 4.591e-01 -0.986 0.32390
## CAR_USE_Commercial 7.796e-01 3.583e-01 2.176 0.02957 *
## CAR_TYPE_Panel.Truck 1.795e-01 6.457e-01 0.278 0.78108
## CAR_TYPE_Pickup 1.114e+00 4.241e-01 2.626 0.00864 **
## CAR_TYPE_Sports.Car 2.222e+00 5.585e-01 3.979 6.93e-05 ***
## CAR_TYPE_Van 4.762e-01 4.818e-01 0.989 0.32290
## CAR_TYPE_z_SUV 2.202e+00 4.925e-01 4.471 7.80e-06 ***
## RED_CAR_no -1.349e-01 3.399e-01 -0.397 0.69143
## REVOKED_Yes 6.621e-01 4.019e-01 1.647 0.09948 .
## URBANICITY_z_Highly.Rural..Rural -2.582e+00 4.562e-01 -5.660 1.51e-08 ***
## YOJ_NA 4.632e-02 4.450e-01 0.104 0.91709
## INCOME_NA 3.234e-01 5.493e-01 0.589 0.55605
## CAR_AGE_NA -2.785e-01 4.952e-01 -0.562 0.57385
## HOME_VAL_NA -1.577e-01 2.917e-01 -0.541 0.58876
## ageSquared 1.331e-03 1.259e-03 1.058 0.29023
## yojSquared 1.132e-02 6.862e-03 1.650 0.09898 .
## income_log -7.860e-02 2.793e-01 -0.281 0.77836
## homeval_log -1.262e+00 1.370e+00 -0.921 0.35696
## travtime_log 1.819e-01 6.404e-01 0.284 0.77637
## bluebook_log 5.387e-01 5.570e-01 0.967 0.33347
## carage_log -1.578e-01 4.656e-01 -0.339 0.73461
## oldclaim_log -1.794e-02 1.714e-01 -0.105 0.91664
## clm_freq_log 1.405e+00 3.281e+00 0.428 0.66849
## mvr_pts_log -4.081e-01 5.116e-01 -0.798 0.42513
## tif_log -3.802e-02 5.320e-01 -0.071 0.94303
## kidsdriv_log 1.701e+00 2.290e+00 0.742 0.45783
## homekids_log 1.624e+00 1.475e+00 1.101 0.27103
## inter 8.911e-03 3.453e-02 0.258 0.79635
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 719.86 on 643 degrees of freedom
## Residual deviance: 528.00 on 588 degrees of freedom

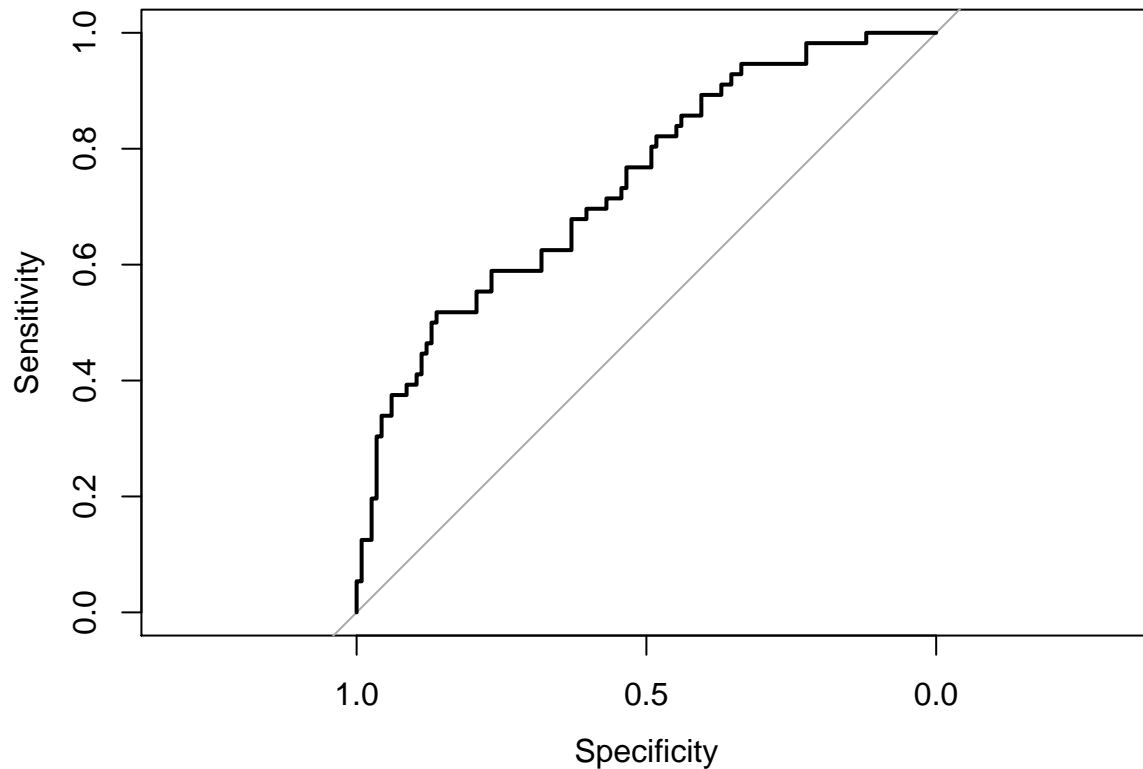
```

```

## (1 observation deleted due to missingness)
## AIC: 640
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  35
##           1   7  21
##
##           Accuracy : 0.7558
##           95% CI : (0.6846, 0.818)
##           No Information Rate : 0.6744
##           P-Value [Acc > NIR] : 0.01255
##
##           Kappa : 0.3614
##
## Mcnemar's Test P-Value : 3.097e-05
##
##           Sensitivity : 0.9397
##           Specificity : 0.3750
##           Pos Pred Value : 0.7569
##           Neg Pred Value : 0.7500
##           Prevalence : 0.6744
##           Detection Rate : 0.6337
##           Detection Prevalence : 0.8372
##           Balanced Accuracy : 0.6573
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.740455665024631"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 116 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7405
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0820  -0.6736  -0.3810   0.4023   3.1978
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    6.428e+00  1.583e+01   0.406  0.68468
## KIDSDRIV     -3.150e+00  1.931e+00  -1.631  0.10288
## AGE          -4.706e-02  1.082e-01  -0.435  0.66357
## HOMEKIDS     -2.747e-01  6.709e-01  -0.409  0.68224
## YOJ          -2.064e-01  1.322e-01  -1.561  0.11846
## INCOME       -1.879e-05  1.348e-05  -1.394  0.16337
## HOME_VAL      6.257e-06  9.582e-06   0.653  0.51374
## TRAVTIME      4.452e-02  2.112e-02   2.108  0.03502 *
## BLUEBOOK      1.979e-05  4.031e-05   0.491  0.62356
```

```

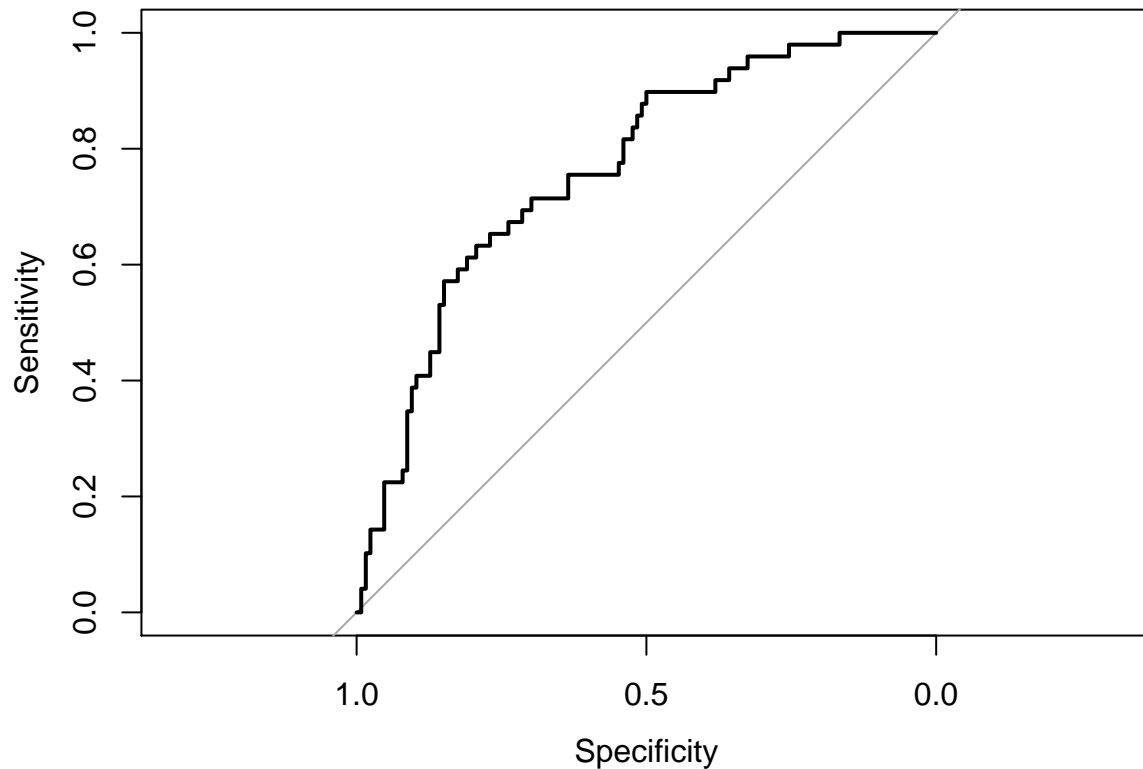
## TIF 3.285e-03 9.346e-02 0.035 0.97196
## OLDCLAIM 9.746e-06 2.159e-05 0.451 0.65170
## CLM_FREQ 3.735e-02 9.566e-01 0.039 0.96885
## MVR_PTS -1.770e-02 1.738e-01 -0.102 0.91888
## CAR_AGE -2.290e-02 6.807e-02 -0.336 0.73654
## PARENT1_Yes 7.323e-01 4.728e-01 1.549 0.12142
## MSTATUS_Yes -5.392e-01 3.370e-01 -1.600 0.10961
## SEX_z_F -3.088e-01 4.681e-01 -0.660 0.50942
## EDUCATION_.High.School 3.232e-01 8.022e-01 0.403 0.68700
## EDUCATION_Bachelors -1.084e-01 6.741e-01 -0.161 0.87228
## EDUCATION_Masters 8.118e-01 5.773e-01 1.406 0.15964
## EDUCATION_z_High.School 4.803e-01 7.213e-01 0.666 0.50549
## JOB_ -9.902e-01 6.750e-01 -1.467 0.14239
## JOB_Clerical -2.367e-01 4.855e-01 -0.488 0.62580
## JOB_Doctor 4.525e-03 9.575e-01 0.005 0.99623
## JOB_Home.Maker 1.920e-01 6.393e-01 0.300 0.76386
## JOB_Lawyer -4.360e-01 6.412e-01 -0.680 0.49652
## JOB_Manager -7.226e-01 4.773e-01 -1.514 0.13000
## JOB_Student -4.212e-01 7.203e-01 -0.585 0.55868
## JOB_z_Blue.Collar -2.425e-01 4.483e-01 -0.541 0.58861
## CAR_USE_Commercial 6.490e-01 3.507e-01 1.850 0.06425 .
## CAR_TYPE_Panel.Truck 4.739e-01 6.290e-01 0.753 0.45120
## CAR_TYPE_Pickup 1.048e+00 4.101e-01 2.554 0.01064 *
## CAR_TYPE_Sports.Car 1.675e+00 5.191e-01 3.227 0.00125 **
## CAR_TYPE_Van 4.314e-01 4.785e-01 0.902 0.36722
## CAR_TYPE_z_SUV 1.780e+00 4.496e-01 3.959 7.53e-05 ***
## RED_CAR_no -4.730e-01 3.581e-01 -1.321 0.18655
## REVOKED_Yes 6.373e-01 4.074e-01 1.564 0.11775
## URBANICITY_z_Highly.Rural..Rural -2.528e+00 4.410e-01 -5.733 9.89e-09 ***
## YOJ_NA -1.190e-03 4.507e-01 -0.003 0.99789
## INCOME_NA 1.111e-01 5.580e-01 0.199 0.84223
## CAR_AGE_NA 1.739e-01 4.986e-01 0.349 0.72728
## HOME_VAL_NA -1.771e-01 2.967e-01 -0.597 0.55054
## ageSquared 4.492e-04 1.186e-03 0.379 0.70498
## yojSquared 1.082e-02 6.780e-03 1.596 0.11058
## income_log 2.701e-01 2.826e-01 0.956 0.33907
## homeval_log -8.326e-01 1.490e+00 -0.559 0.57638
## travtime_log -5.386e-01 5.927e-01 -0.909 0.36346
## bluebook_log 1.018e-01 4.992e-01 0.204 0.83844
## carage_log -3.864e-02 4.548e-01 -0.085 0.93228
## oldclaim_log -8.133e-03 1.637e-01 -0.050 0.96038
## clm_freq_log 2.435e-01 2.970e+00 0.082 0.93465
## mvr_pts_log 2.880e-01 5.076e-01 0.567 0.57049
## tif_log -1.362e-01 5.264e-01 -0.259 0.79590
## kidsdriv_log 2.948e+00 2.172e+00 1.357 0.17468
## homekids_log 6.130e-01 1.442e+00 0.425 0.67082
## inter 4.838e-02 3.438e-02 1.407 0.15939
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.27 on 640 degrees of freedom
## Residual deviance: 538.54 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 650.54
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 110  29
##           1   16  20
##
##           Accuracy : 0.7429
##           95% CI : (0.6715, 0.8058)
##           No Information Rate : 0.72
##           P-Value [Acc > NIR] : 0.28064
##
##           Kappa : 0.306
##
## Mcnemar's Test P-Value : 0.07364
##
##           Sensitivity : 0.8730
##           Specificity : 0.4082
##           Pos Pred Value : 0.7914
##           Neg Pred Value : 0.5556
##           Prevalence : 0.7200
##           Detection Rate : 0.6286
##           Detection Prevalence : 0.7943
##           Balanced Accuracy : 0.6406
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.766115970197603"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 126 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.7661
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0653  -0.6660  -0.3249   0.5366   2.9645
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.074e+01  1.477e+01   2.081 0.037430 *
## KIDSDRIV      -1.010e+00  1.685e+00  -0.599 0.548875
## AGE           -1.507e-01  1.017e-01  -1.481 0.138611
## HOMEKIDS       1.604e-03  7.040e-01   0.002 0.998182
## YOJ            -1.995e-01  1.318e-01  -1.513 0.130275
## INCOME         -1.960e-05  1.422e-05  -1.378 0.168072
## HOME_VAL       1.215e-05  9.257e-06   1.312 0.189401
## TRAVTIME       3.609e-02  2.281e-02   1.583 0.113512
## BLUEBOOK       7.122e-05  3.706e-05   1.922 0.054664 .
```

```

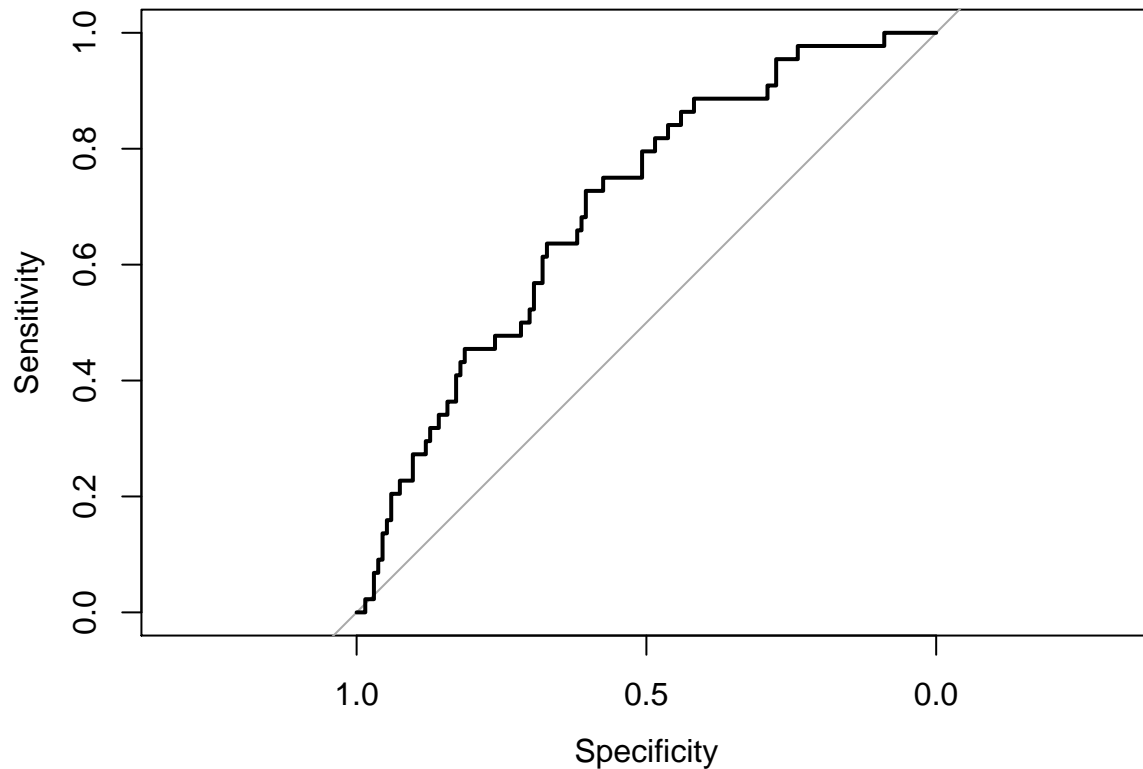
## TIF -5.307e-03 9.028e-02 -0.059 0.953121
## OLDCLAIM 3.717e-06 2.102e-05 0.177 0.859605
## CLM_FREQ 5.954e-01 9.766e-01 0.610 0.542059
## MVR_PTS 1.204e-01 1.683e-01 0.716 0.474281
## CAR_AGE -8.057e-02 7.103e-02 -1.134 0.256667
## PARENT1_Yes 7.258e-01 4.895e-01 1.483 0.138131
## MSTATUS_Yes -5.673e-01 3.381e-01 -1.678 0.093418 .
## SEX_z_F -9.499e-01 4.542e-01 -2.091 0.036506 *
## EDUCATION_.High.School -7.831e-01 7.917e-01 -0.989 0.322576
## EDUCATION_Bachelors -9.319e-01 6.487e-01 -1.437 0.150804
## EDUCATION_Masters 4.423e-01 5.486e-01 0.806 0.420120
## EDUCATION_z_High.School -4.646e-01 7.105e-01 -0.654 0.513178
## JOB_ -1.192e+00 6.787e-01 -1.756 0.079046 .
## JOB_Clerical -1.584e-01 4.994e-01 -0.317 0.751137
## JOB_Doctor -4.576e-01 9.050e-01 -0.506 0.613159
## JOB_Home.Maker -5.187e-01 6.699e-01 -0.774 0.438740
## JOB_Lawyer -1.396e+00 6.561e-01 -2.127 0.033394 *
## JOB_Manager -1.148e+00 4.955e-01 -2.316 0.020562 *
## JOB_Student -7.427e-01 7.363e-01 -1.009 0.313096
## JOB_z_Blue.Collar -6.225e-02 4.734e-01 -0.131 0.895392
## CAR_USE_Commercial 6.070e-01 3.550e-01 1.710 0.087306 .
## CAR_TYPE_Panel.Truck -1.370e-01 6.298e-01 -0.218 0.827761
## CAR_TYPE_Pickup 1.055e+00 4.063e-01 2.596 0.009418 **
## CAR_TYPE_Sports.Car 1.939e+00 5.337e-01 3.634 0.000279 ***
## CAR_TYPE_Van 2.231e-01 4.700e-01 0.475 0.634939
## CAR_TYPE_z_SUV 1.847e+00 4.686e-01 3.941 8.13e-05 ***
## RED_CAR_no -1.639e-01 3.404e-01 -0.481 0.630174
## REVOKED_Yes 5.565e-01 4.029e-01 1.381 0.167196
## URBANICITY_z_Highly.Rural..Rural -2.627e+00 4.638e-01 -5.664 1.48e-08 ***
## YOJ_NA -7.286e-01 4.290e-01 -1.698 0.089428 .
## INCOME_NA -4.223e-02 5.134e-01 -0.082 0.934453
## CAR_AGE_NA 7.412e-02 5.207e-01 0.142 0.886810
## HOME_VAL_NA -1.129e-01 2.884e-01 -0.391 0.695506
## ageSquared 1.642e-03 1.104e-03 1.487 0.136962
## yojSquared 1.096e-02 6.817e-03 1.608 0.107815
## income_log 5.804e-01 3.300e-01 1.759 0.078583 .
## homeval_log -2.545e+00 1.439e+00 -1.769 0.076928 .
## travtime_log -1.777e-01 6.421e-01 -0.277 0.782031
## bluebook_log -4.779e-01 4.422e-01 -1.081 0.279811
## carage_log 3.764e-01 4.682e-01 0.804 0.421448
## oldclaim_log 5.415e-02 1.653e-01 0.327 0.743316
## clm_freq_log -1.121e+00 3.010e+00 -0.372 0.709694
## mvr_pts_log -3.044e-01 5.007e-01 -0.608 0.543227
## tif_log -3.436e-01 5.193e-01 -0.662 0.508116
## kidsdriv_log 1.153e+00 2.175e+00 0.530 0.596045
## homekids_log -1.453e-01 1.522e+00 -0.095 0.923933
## inter 2.741e-02 3.057e-02 0.897 0.369829
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 741.72 on 637 degrees of freedom
## Residual deviance: 530.38 on 582 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 642.38
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  32
##           1  16  12
##
##           Accuracy : 0.7303
##           95% CI : (0.6588, 0.794)
##           No Information Rate : 0.7528
##           P-Value [Acc > NIR] : 0.78452
##
##           Kappa : 0.1747
##
## Mcnemar's Test P-Value : 0.03038
##
##           Sensitivity : 0.8806
##           Specificity : 0.2727
##           Pos Pred Value : 0.7867
##           Neg Pred Value : 0.4286
##           Prevalence : 0.7528
##           Detection Rate : 0.6629
##           Detection Prevalence : 0.8427
##           Balanced Accuracy : 0.5767
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.697930800542741"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.6979
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1257  -0.6708  -0.3693   0.4710   3.0489
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.596e+01  1.520e+01   1.708 0.087702 .
## KIDSDRIV      -1.005e+00  1.666e+00  -0.603 0.546225
## AGE           -2.144e-01  1.037e-01  -2.067 0.038744 *
## HOMEKIDS       -9.737e-01  6.814e-01  -1.429 0.153011
## YOJ            -2.055e-01  1.308e-01  -1.571 0.116096
## INCOME         -3.336e-05  1.435e-05  -2.324 0.020102 *
## HOME_VAL        1.493e-05  9.470e-06   1.577 0.114851
## TRAVTIME        3.294e-02  2.167e-02   1.520 0.128415
## BLUEBOOK        1.000e-04  3.854e-05   2.595 0.009447 **
```

```

## TIF -9.325e-02 9.395e-02 -0.993 0.320946
## OLDCLAIM 4.439e-06 2.218e-05 0.200 0.841347
## CLM_FREQ 7.002e-02 1.068e+00 0.066 0.947746
## MVRPTS 1.270e-01 1.732e-01 0.733 0.463338
## CAR_AGE -8.232e-02 7.121e-02 -1.156 0.247665
## PARENT1_Yes 3.228e-01 4.737e-01 0.682 0.495550
## MSTATUS_Yes -8.338e-01 3.310e-01 -2.519 0.011756 *
## SEX_z_F -1.051e+00 4.658e-01 -2.256 0.024080 *
## EDUCATION_.High.School 1.407e-02 8.011e-01 0.018 0.985992
## EDUCATION_Bachelors 2.065e-01 6.711e-01 0.308 0.758324
## EDUCATION_Masters 6.671e-01 5.728e-01 1.165 0.244177
## EDUCATION_z_High.School 6.341e-01 7.254e-01 0.874 0.382016
## JOB_ -2.230e-02 6.883e-01 -0.032 0.974157
## JOB_Clerical 3.125e-02 5.059e-01 0.062 0.950754
## JOB_Doctor 5.600e-01 9.410e-01 0.595 0.551769
## JOB_Home.Maker -2.642e-01 7.015e-01 -0.377 0.706398
## JOB_Lawyer -1.181e-01 6.656e-01 -0.177 0.859186
## JOB_Manager -4.856e-01 4.800e-01 -1.012 0.311760
## JOB_Student -5.779e-01 7.314e-01 -0.790 0.429472
## JOB_z_Blue.Collar 2.254e-02 4.919e-01 0.046 0.963456
## CAR_USE_Commercial 2.127e-01 3.652e-01 0.582 0.560290
## CAR_TYPE_Panel.Truck -2.582e-02 6.160e-01 -0.042 0.966563
## CAR_TYPE_Pickup 1.022e+00 4.107e-01 2.489 0.012805 *
## CAR_TYPE_Sports.Car 1.874e+00 5.485e-01 3.417 0.000634 ***
## CAR_TYPE_Van 4.315e-01 4.720e-01 0.914 0.360587
## CAR_TYPE_z_SUV 1.920e+00 4.876e-01 3.938 8.21e-05 ***
## RED_CAR_no -1.385e-01 3.308e-01 -0.419 0.675563
## REVOKED_Yes 2.907e-01 4.280e-01 0.679 0.496977
## URBANICITY_z_Highly.Rural..Rural -2.453e+00 4.157e-01 -5.901 3.61e-09 ***
## YOJ_NA -7.772e-01 4.168e-01 -1.865 0.062216 .
## INCOME_NA 2.536e-01 5.440e-01 0.466 0.641077
## CAR_AGE_NA -6.456e-01 4.808e-01 -1.343 0.179304
## HOME_VAL_NA -1.554e-01 2.895e-01 -0.537 0.591379
## ageSquared 2.106e-03 1.120e-03 1.881 0.060016 .
## yojSquared 1.026e-02 6.751e-03 1.520 0.128552
## income_log 5.554e-01 3.660e-01 1.518 0.129080
## homeval_log -1.821e+00 1.513e+00 -1.204 0.228774
## travtime_log -2.958e-01 6.087e-01 -0.486 0.626960
## bluebook_log -7.151e-01 4.531e-01 -1.578 0.114513
## carage_log 2.786e-01 4.659e-01 0.598 0.549776
## oldclaim_log 5.530e-03 1.746e-01 0.032 0.974733
## clm_freq_log -1.873e-01 3.263e+00 -0.057 0.954226
## mvr_pts_log -1.833e-01 5.069e-01 -0.362 0.717693
## tif_log 3.777e-01 5.353e-01 0.706 0.480407
## kidsdriv_log 1.644e+00 2.145e+00 0.766 0.443480
## homekids_log 2.121e+00 1.483e+00 1.431 0.152562
## inter 1.937e-02 3.033e-02 0.639 0.522984
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 734.16 on 638 degrees of freedom
## Residual deviance: 537.82 on 583 degrees of freedom

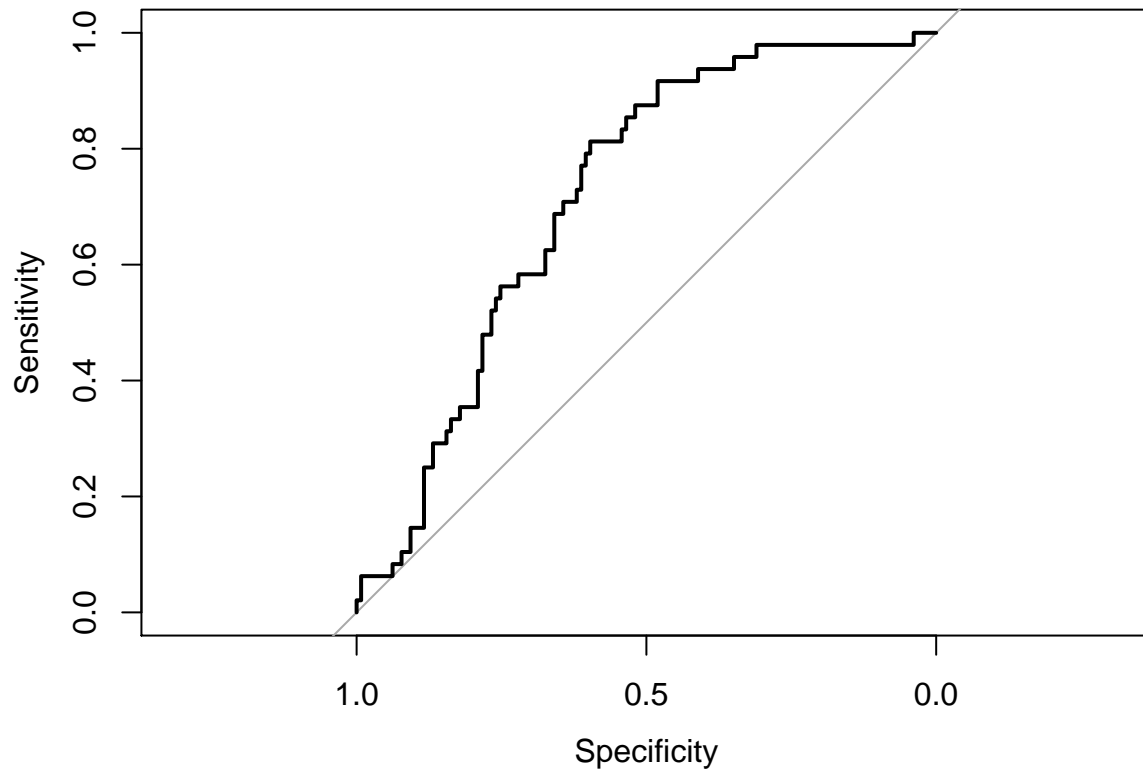
```



```

## (1 observation deleted due to missingness)
## AIC: 649.82
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 113  36
##           1  16  12
##
##           Accuracy : 0.7062
##           95% CI : (0.6332, 0.7722)
##           No Information Rate : 0.7288
##           P-Value [Acc > NIR] : 0.778219
##
##           Kappa : 0.1449
##
## Mcnemar's Test P-Value : 0.008418
##
##           Sensitivity : 0.8760
##           Specificity : 0.2500
##           Pos Pred Value : 0.7584
##           Neg Pred Value : 0.4286
##           Prevalence : 0.7288
##           Detection Rate : 0.6384
##           Detection Prevalence : 0.8418
##           Balanced Accuracy : 0.5630
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.718507751937984"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 129 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.7185
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0982  -0.6940  -0.3709   0.5842   2.9143
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.828e+01  1.423e+01   1.285 0.198898
## KIDSDRIV      -2.521e+00  1.964e+00  -1.284 0.199245
## AGE           -2.070e-01  1.064e-01  -1.945 0.051824 .
## HOMEKIDS      -9.337e-01  6.685e-01  -1.397 0.162492
## YOJ           -1.405e-01  1.323e-01  -1.062 0.288140
## INCOME        -2.109e-05  1.308e-05  -1.612 0.106923
## HOME_VAL       9.872e-06  8.613e-06   1.146 0.251750
## TRAVTIME       2.891e-02  2.153e-02   1.343 0.179350
## BLUEBOOK       6.978e-05  3.875e-05   1.801 0.071768 .
```

```

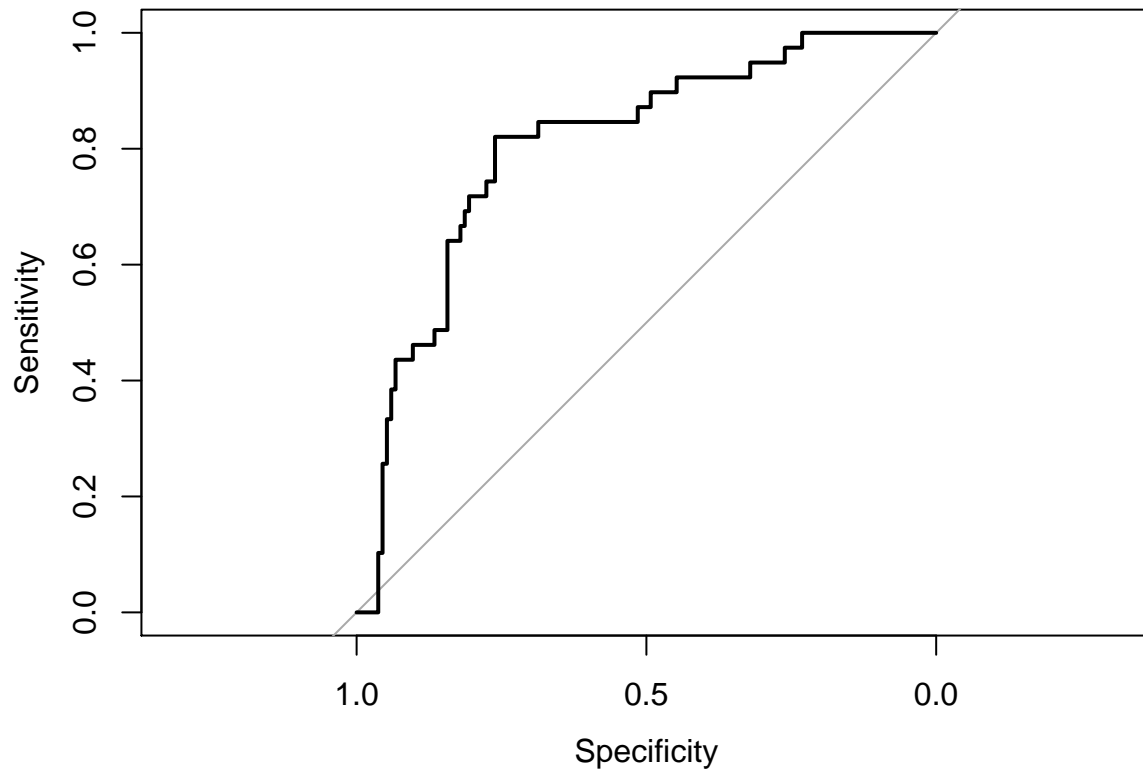
## TIF -8.347e-02 1.027e-01 -0.813 0.416505
## OLDCLAIM -4.113e-06 2.037e-05 -0.202 0.839985
## CLM_FREQ 1.154e-01 9.205e-01 0.125 0.900226
## MVRPTS 8.627e-02 1.637e-01 0.527 0.598078
## CAR_AGE -2.279e-03 7.025e-02 -0.032 0.974117
## PARENT1_Yes 2.788e-01 4.614e-01 0.604 0.545627
## MSTATUS_Yes -7.819e-01 3.348e-01 -2.335 0.019532 *
## SEX_z_F -4.075e-01 4.459e-01 -0.914 0.360700
## EDUCATION_.High.School 1.909e-01 7.385e-01 0.258 0.796045
## EDUCATION_Bachelors 1.037e-01 6.180e-01 0.168 0.866811
## EDUCATION_Masters 9.596e-01 5.402e-01 1.776 0.075656 .
## EDUCATION_z_High.School 3.101e-01 6.702e-01 0.463 0.643567
## JOB_ -8.448e-01 6.576e-01 -1.285 0.198915
## JOB_Clerical 7.094e-03 4.735e-01 0.015 0.988047
## JOB_Doctor 2.001e-01 9.246e-01 0.216 0.828621
## JOB_Home.Maker 3.672e-01 6.574e-01 0.559 0.576456
## JOB_Lawyer -5.774e-01 6.282e-01 -0.919 0.358055
## JOB_Manager -5.793e-01 4.709e-01 -1.230 0.218613
## JOB_Student -1.445e-01 7.071e-01 -0.204 0.838052
## JOB_z_Blue.Collar -4.445e-02 4.497e-01 -0.099 0.921265
## CAR_USE_Commercial 4.791e-01 3.383e-01 1.416 0.156811
## CAR_TYPE_Panel.Truck 1.995e-01 5.946e-01 0.336 0.737192
## CAR_TYPE_Pickup 1.013e+00 4.041e-01 2.507 0.012172 *
## CAR_TYPE_Sports.Car 1.665e+00 5.037e-01 3.305 0.000948 ***
## CAR_TYPE_Van 4.086e-01 4.544e-01 0.899 0.368515
## CAR_TYPE_z_SUV 1.442e+00 4.569e-01 3.157 0.001595 **
## RED_CAR_no -2.694e-01 3.359e-01 -0.802 0.422572
## REVOKED_Yes 5.791e-01 3.888e-01 1.489 0.136366
## URBANICITY_z_Highly.Rural..Rural -2.205e+00 4.121e-01 -5.350 8.79e-08 ***
## YOJ_NA -2.956e-01 4.076e-01 -0.725 0.468358
## INCOME_NA 3.965e-01 5.646e-01 0.702 0.482455
## CAR_AGE_NA -3.711e-01 5.029e-01 -0.738 0.460533
## HOME_VAL_NA -1.564e-02 2.884e-01 -0.054 0.956772
## ageSquared 2.209e-03 1.147e-03 1.926 0.054158 .
## yojSquared 6.821e-03 6.744e-03 1.011 0.311810
## income_log 2.890e-01 3.012e-01 0.960 0.337289
## homeval_log -1.041e+00 1.356e+00 -0.768 0.442605
## travtime_log -2.207e-01 5.964e-01 -0.370 0.711361
## bluebook_log -6.262e-01 4.799e-01 -1.305 0.191946
## carage_log -3.400e-01 4.546e-01 -0.748 0.454519
## oldclaim_log 6.970e-02 1.573e-01 0.443 0.657652
## clm_freq_log -2.873e-01 2.864e+00 -0.100 0.920079
## mvr_pts_log -9.858e-02 4.843e-01 -0.204 0.838703
## tif_log 1.820e-02 5.543e-01 0.033 0.973813
## kidsdriv_log 5.829e-01 2.171e+00 0.268 0.788343
## homekids_log 2.047e+00 1.433e+00 1.428 0.153233
## inter 6.638e-02 3.874e-02 1.713 0.086659 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 754.78 on 642 degrees of freedom
## Residual deviance: 565.19 on 587 degrees of freedom

```

```

## AIC: 677.19
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 121  22
##           1  13  17
##
##           Accuracy : 0.7977
##           95% CI : (0.73, 0.8549)
##       No Information Rate : 0.7746
##       P-Value [Acc > NIR] : 0.2654
##
##           Kappa : 0.3691
##
##  McNemar's Test P-Value : 0.1763
##
##           Sensitivity : 0.9030
##           Specificity : 0.4359
##       Pos Pred Value : 0.8462
##       Neg Pred Value : 0.5667
##           Prevalence : 0.7746
##       Detection Rate : 0.6994
##       Detection Prevalence : 0.8266
##       Balanced Accuracy : 0.6694
##
##       'Positive' Class : 0
##

```



```
## [1] "AUC: 0.80654420206659"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 39 cases (dfPred_raw$class 1).
## Area under the curve: 0.8065
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1602  -0.6727  -0.3630   0.6057   2.9397
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.630e+01  1.482e+01   1.100  0.271290
## KIDSDRIV      -5.122e+00  2.316e+00  -2.212  0.026981 *
## AGE           -5.893e-02  1.078e-01  -0.547  0.584466
## HOMEKIDS       8.425e-02  6.805e-01   0.124  0.901469
## YOJ           -2.514e-01  1.323e-01  -1.900  0.057389 .
## INCOME        -2.673e-05  1.348e-05  -1.982  0.047483 *
## HOME_VAL       1.310e-05  9.264e-06   1.414  0.157446
## TRAVTIME       4.395e-02  2.127e-02   2.066  0.038827 *
## BLUEBOOK       4.333e-05  3.891e-05   1.113  0.265522
```

```

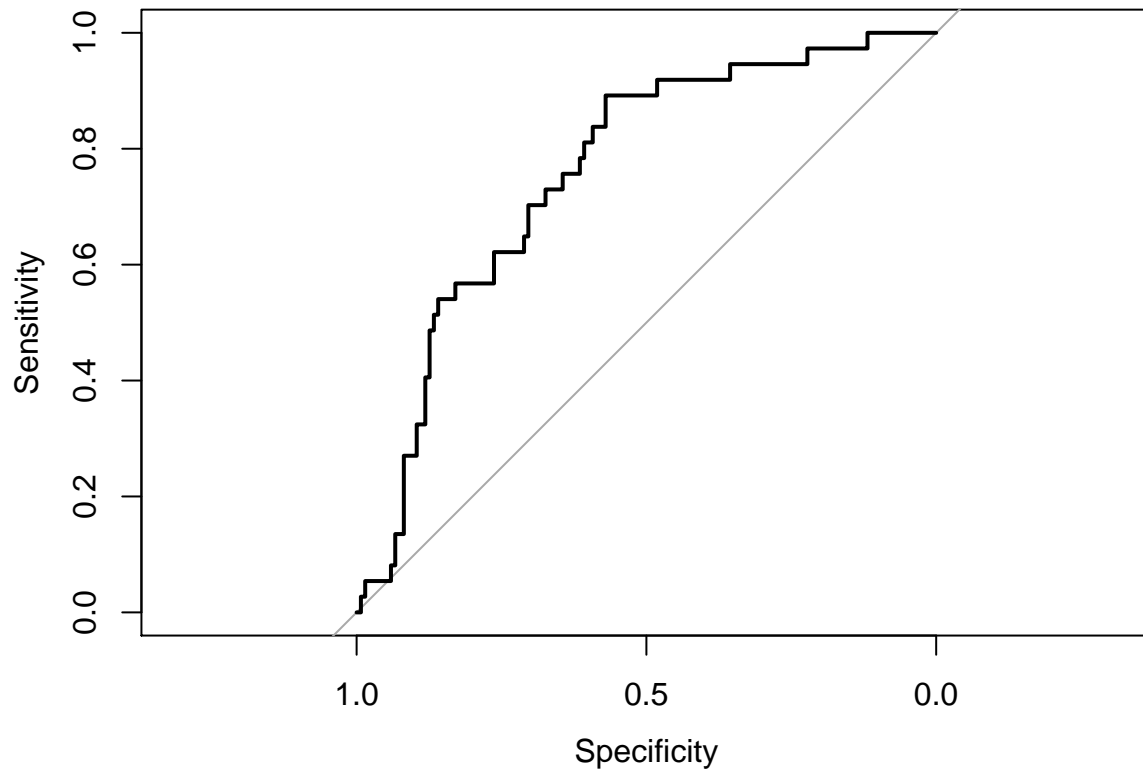
## TIF -8.581e-02 9.628e-02 -0.891 0.372844
## OLDCLAIM -1.227e-07 2.077e-05 -0.006 0.995286
## CLM_FREQ -2.514e-01 9.298e-01 -0.270 0.786889
## MVRPTS 5.188e-02 1.651e-01 0.314 0.753349
## CAR_AGE -8.039e-04 6.852e-02 -0.012 0.990640
## PARENT1_Yes 9.391e-01 4.633e-01 2.027 0.042691 *
## MSTATUS_Yes -6.800e-01 3.278e-01 -2.074 0.038048 *
## SEX_z_F -4.654e-01 4.392e-01 -1.060 0.289233
## EDUCATION_.High.School -7.529e-03 7.778e-01 -0.010 0.992277
## EDUCATION_Bachelors 8.920e-02 6.467e-01 0.138 0.890292
## EDUCATION_Masters 8.161e-01 5.707e-01 1.430 0.152693
## EDUCATION_z_High.School 3.132e-01 7.069e-01 0.443 0.657710
## JOB_ -5.651e-01 6.708e-01 -0.842 0.399519
## JOB_Clerical 2.092e-01 4.654e-01 0.449 0.653142
## JOB_Doctor 4.633e-02 8.716e-01 0.053 0.957605
## JOB_Home.Maker -2.509e-01 6.714e-01 -0.374 0.708635
## JOB_Lawyer -6.128e-01 6.388e-01 -0.959 0.337439
## JOB_Manager -5.566e-01 4.804e-01 -1.159 0.246565
## JOB_Student -3.758e-01 7.262e-01 -0.518 0.604755
## JOB_z_Blue.Collar -1.798e-01 4.475e-01 -0.402 0.687730
## CAR_USE_Commercial 6.004e-01 3.518e-01 1.707 0.087895 .
## CAR_TYPE_Panel.Truck 7.727e-02 6.250e-01 0.124 0.901615
## CAR_TYPE_Pickup 6.422e-01 4.046e-01 1.587 0.112414
## CAR_TYPE_Sports.Car 1.736e+00 5.006e-01 3.469 0.000523 ***
## CAR_TYPE_Van 1.452e-01 4.685e-01 0.310 0.756618
## CAR_TYPE_z_SUV 1.325e+00 4.414e-01 3.003 0.002673 **
## RED_CAR_no -3.086e-01 3.375e-01 -0.914 0.360498
## REVOKED_Yes 5.574e-01 4.081e-01 1.366 0.172048
## URBANICITY_z_Highly.Rural..Rural -2.333e+00 4.122e-01 -5.659 1.53e-08 ***
## YOJ_NA -4.860e-02 4.438e-01 -0.110 0.912798
## INCOME_NA 4.812e-01 5.594e-01 0.860 0.389658
## CAR_AGE_NA -3.145e-01 4.781e-01 -0.658 0.510624
## HOME_VAL_NA 1.487e-01 2.861e-01 0.520 0.603343
## ageSquared 6.637e-04 1.159e-03 0.573 0.566935
## yojSquared 1.260e-02 6.799e-03 1.854 0.063751 .
## income_log 2.320e-01 3.056e-01 0.759 0.447773
## homeval_log -1.442e+00 1.431e+00 -1.008 0.313439
## travtime_log -4.810e-01 5.956e-01 -0.808 0.419263
## bluebook_log -1.665e-01 4.635e-01 -0.359 0.719325
## carage_log -3.517e-01 4.497e-01 -0.782 0.434155
## oldclaim_log 3.326e-03 1.571e-01 0.021 0.983107
## clm_freq_log 9.383e-01 2.874e+00 0.327 0.744023
## mvr_pts_log 6.824e-02 4.858e-01 0.140 0.888293
## tif_log 1.580e-01 5.246e-01 0.301 0.763251
## kidsdriv_log 2.779e+00 2.317e+00 1.199 0.230396
## homekids_log -3.170e-02 1.478e+00 -0.021 0.982885
## inter 9.700e-02 4.773e-02 2.032 0.042121 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 759.30 on 643 degrees of freedom
## Residual deviance: 558.93 on 588 degrees of freedom

```

```

## AIC: 670.93
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 119  23
##           1  16  14
##
##           Accuracy : 0.7733
##           95% CI : (0.7033, 0.8335)
##           No Information Rate : 0.7849
##           P-Value [Acc > NIR] : 0.6835
##
##           Kappa : 0.279
##
## Mcnemar's Test P-Value : 0.3367
##
##           Sensitivity : 0.8815
##           Specificity : 0.3784
##           Pos Pred Value : 0.8380
##           Neg Pred Value : 0.4667
##           Prevalence : 0.7849
##           Detection Rate : 0.6919
##           Detection Prevalence : 0.8256
##           Balanced Accuracy : 0.6299
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.759159159159159"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 135 controls (dfPred_raw$class 0) < 37 cases (dfPred_raw$class 1).
## Area under the curve: 0.7592
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1146  -0.6371  -0.3664   0.4010   3.0451
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.833e+01  1.554e+01   1.180 0.238080
## KIDSDRIV      -1.294e+00  1.881e+00  -0.688 0.491327
## AGE           -6.219e-02  1.038e-01  -0.599 0.549213
## HOMEKIDS       -8.159e-01  6.888e-01  -1.184 0.236237
## YOJ            -1.772e-01  1.306e-01  -1.356 0.175081
## INCOME         -2.444e-05  1.395e-05  -1.752 0.079833 .
## HOME_VAL        1.229e-05  9.735e-06   1.263 0.206626
## TRAVTIME        4.625e-02  2.124e-02   2.178 0.029402 *
## BLUEBOOK        6.894e-05  3.921e-05   1.758 0.078703 .
```



```

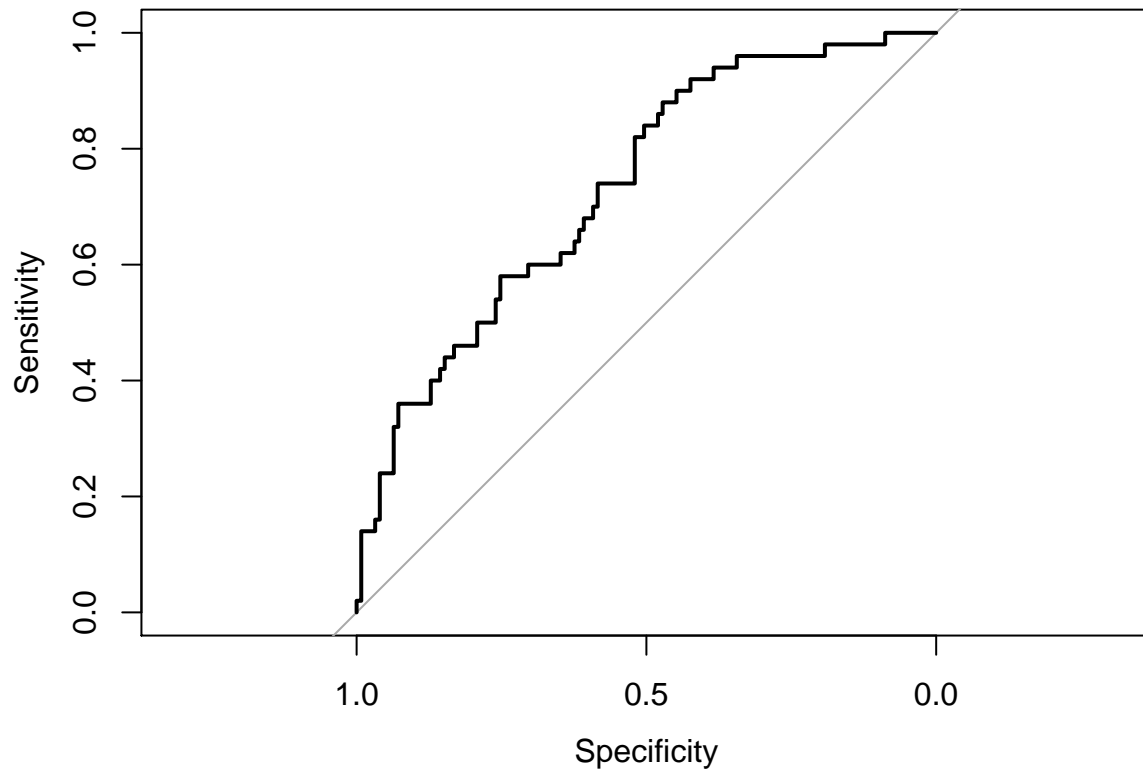
## TIF -5.735e-03 8.761e-02 -0.065 0.947805
## OLDCLAIM 6.071e-06 2.170e-05 0.280 0.779669
## CLM_FREQ -3.542e-01 1.018e+00 -0.348 0.727930
## MVRPTS 1.280e-01 1.773e-01 0.722 0.470342
## CAR_AGE -3.674e-02 6.974e-02 -0.527 0.598253
## PARENT1_Yes 5.688e-01 4.748e-01 1.198 0.230975
## MSTATUS_Yes -6.874e-01 3.353e-01 -2.050 0.040335 *
## SEX_z_F -8.143e-01 4.625e-01 -1.761 0.078276 .
## EDUCATION_.High.School -2.744e-01 8.485e-01 -0.323 0.746408
## EDUCATION_Bachelors 5.011e-02 7.063e-01 0.071 0.943435
## EDUCATION_Masters 1.001e+00 6.119e-01 1.637 0.101726
## EDUCATION_z_High.School 5.774e-01 7.638e-01 0.756 0.449681
## JOB_ -1.109e+00 7.069e-01 -1.569 0.116724
## JOB_Clerical -2.321e-01 4.972e-01 -0.467 0.640674
## JOB_Doctor 2.467e-01 9.012e-01 0.274 0.784305
## JOB_Home.Maker -4.068e-01 6.576e-01 -0.619 0.536208
## JOB_Lawyer -9.531e-01 6.583e-01 -1.448 0.147661
## JOB_Manager -1.065e+00 5.033e-01 -2.116 0.034338 *
## JOB_Student -8.944e-01 7.263e-01 -1.231 0.218153
## JOB_z_Blue.Collar -5.589e-01 4.827e-01 -1.158 0.246922
## CAR_USE_Commercial 8.372e-01 3.631e-01 2.306 0.021128 *
## CAR_TYPE_Panel.Truck 1.171e-01 6.315e-01 0.185 0.852936
## CAR_TYPE_Pickup 7.945e-01 4.047e-01 1.963 0.049600 *
## CAR_TYPE_Sports.Car 1.931e+00 5.523e-01 3.497 0.000471 ***
## CAR_TYPE_Van 4.967e-01 4.840e-01 1.026 0.304846
## CAR_TYPE_z_SUV 1.907e+00 4.871e-01 3.914 9.07e-05 ***
## RED_CAR_no -2.168e-01 3.368e-01 -0.644 0.519835
## REVOKED_Yes 3.141e-01 4.583e-01 0.685 0.493135
## URBANICITY_z_Highly.Rural..Rural -2.213e+00 4.039e-01 -5.479 4.28e-08 ***
## YOJ_NA 2.341e-01 4.699e-01 0.498 0.618331
## INCOME_NA 4.847e-01 5.449e-01 0.890 0.373689
## CAR_AGE_NA -9.738e-02 4.752e-01 -0.205 0.837615
## HOME_VAL_NA -1.757e-01 2.821e-01 -0.623 0.533377
## ageSquared 5.727e-04 1.126e-03 0.509 0.611058
## yojSquared 9.425e-03 6.766e-03 1.393 0.163632
## income_log 4.255e-02 2.951e-01 0.144 0.885351
## homeval_log -1.256e+00 1.507e+00 -0.833 0.404763
## travtime_log -7.110e-01 5.946e-01 -1.196 0.231771
## bluebook_log -4.587e-01 4.634e-01 -0.990 0.322300
## carage_log 2.013e-02 4.626e-01 0.044 0.965290
## oldclaim_log 2.312e-02 1.683e-01 0.137 0.890734
## clm_freq_log 9.291e-01 3.119e+00 0.298 0.765812
## mvr_pts_log -1.715e-01 5.123e-01 -0.335 0.737842
## tif_log -5.266e-02 5.017e-01 -0.105 0.916416
## kidsdriv_log 7.743e-02 2.277e+00 0.034 0.972866
## homekids_log 1.879e+00 1.506e+00 1.248 0.212082
## inter 4.461e-02 3.264e-02 1.367 0.171757
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 731.16 on 640 degrees of freedom
## Residual deviance: 532.92 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 644.92
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  32
##           1  11  18
##
##           Accuracy : 0.7543
##           95% CI : (0.6836, 0.8161)
##           No Information Rate : 0.7143
##           P-Value [Acc > NIR] : 0.137803
##
##           Kappa : 0.3112
##
## Mcnemar's Test P-Value : 0.002289
##
##           Sensitivity : 0.9120
##           Specificity : 0.3600
##           Pos Pred Value : 0.7808
##           Neg Pred Value : 0.6207
##           Prevalence : 0.7143
##           Detection Rate : 0.6514
##           Detection Prevalence : 0.8343
##           Balanced Accuracy : 0.6360
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.73248"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 125 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7325
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1587  -0.6641  -0.3406   0.4675   2.5950
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.271e+01  1.494e+01   0.851 0.394809
## KIDSDRIV      -4.175e+00  2.303e+00  -1.813 0.069878 .
## AGE          -1.631e-01  1.078e-01  -1.513 0.130317
## HOMEKIDS     -4.290e-01  6.559e-01  -0.654 0.513060
## YOJ          -2.316e-01  1.390e-01  -1.667 0.095548 .
## INCOME       -2.615e-05  1.338e-05  -1.955 0.050587 .
## HOME_VAL      1.152e-05  8.971e-06   1.284 0.199308
## TRAVTIME      2.301e-02  2.199e-02   1.046 0.295475
## BLUEBOOK      2.727e-05  4.088e-05   0.667 0.504813
```

```

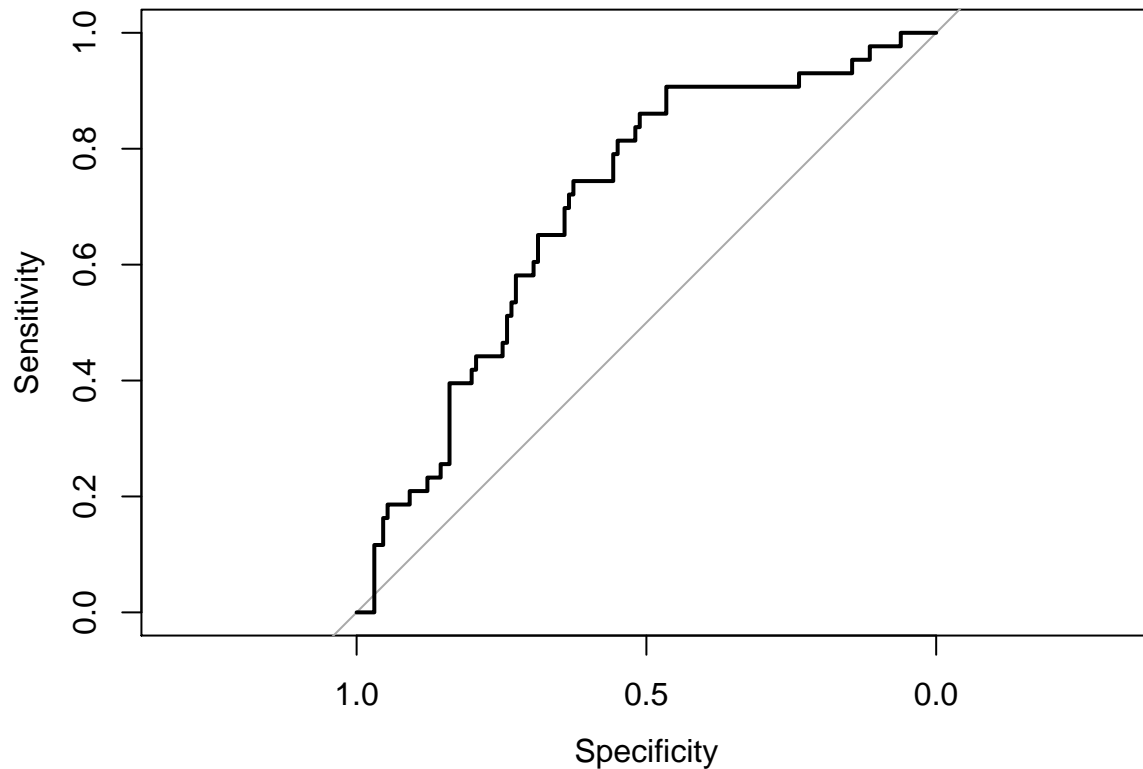
## TIF -1.074e-01 9.843e-02 -1.091 0.275209
## OLDCLAIM 1.252e-05 2.316e-05 0.540 0.588945
## CLM_FREQ -1.277e-01 1.063e+00 -0.120 0.904338
## MVRPTS 4.290e-02 1.726e-01 0.249 0.803718
## CAR_AGE -5.395e-03 7.044e-02 -0.077 0.938952
## PARENT1_Yes 4.316e-01 4.706e-01 0.917 0.359116
## MSTATUS_Yes -7.445e-01 3.411e-01 -2.183 0.029059 *
## SEX_z_F -5.461e-01 4.737e-01 -1.153 0.248996
## EDUCATION_.High.School 1.155e+00 8.462e-01 1.365 0.172227
## EDUCATION_Bachelors 6.002e-01 7.334e-01 0.818 0.413175
## EDUCATION_Masters 1.171e+00 6.509e-01 1.799 0.071963 .
## EDUCATION_z_High.School 1.078e+00 7.864e-01 1.371 0.170333
## JOB_ -7.341e-01 7.071e-01 -1.038 0.299196
## JOB_Clerical -4.931e-01 4.962e-01 -0.994 0.320327
## JOB_Doctor 9.189e-01 9.978e-01 0.921 0.357103
## JOB_Home.Maker -5.370e-01 7.001e-01 -0.767 0.443019
## JOB_Lawyer -5.872e-01 6.972e-01 -0.842 0.399674
## JOB_Manager -4.777e-01 4.958e-01 -0.964 0.335294
## JOB_Student -1.181e+00 7.549e-01 -1.565 0.117596
## JOB_z_Blue.Collar -1.569e-01 4.555e-01 -0.345 0.730467
## CAR_USE_Commercial 7.586e-01 3.589e-01 2.114 0.034557 *
## CAR_TYPE_Panel.Truck 1.577e-01 6.295e-01 0.251 0.802118
## CAR_TYPE_Pickup 1.031e+00 4.063e-01 2.537 0.011171 *
## CAR_TYPE_Sports.Car 1.390e+00 5.562e-01 2.499 0.012445 *
## CAR_TYPE_Van 5.873e-01 4.895e-01 1.200 0.230277
## CAR_TYPE_z_SUV 1.754e+00 4.745e-01 3.697 0.000218 ***
## RED_CAR_no -3.361e-01 3.486e-01 -0.964 0.335062
## REVOKED_Yes 6.342e-01 4.492e-01 1.412 0.157984
## URBANICITY_z_Highly.Rural..Rural -2.705e+00 4.408e-01 -6.137 8.42e-10 ***
## YOJ_NA -4.546e-01 4.503e-01 -1.010 0.312678
## INCOME_NA 6.213e-01 5.915e-01 1.050 0.293526
## CAR_AGE_NA -8.202e-01 5.075e-01 -1.616 0.106073
## HOME_VAL_NA 2.057e-01 2.926e-01 0.703 0.482081
## ageSquared 1.849e-03 1.160e-03 1.594 0.111023
## yojSquared 8.299e-03 7.063e-03 1.175 0.239972
## income_log -3.655e-03 2.905e-01 -0.013 0.989962
## homeval_log -7.854e-01 1.395e+00 -0.563 0.573396
## travtime_log 5.806e-02 6.303e-01 0.092 0.926607
## bluebook_log -1.751e-01 4.755e-01 -0.368 0.712741
## carage_log -2.387e-01 4.594e-01 -0.520 0.603272
## oldclaim_log -5.404e-02 1.784e-01 -0.303 0.761923
## clm_freq_log 7.245e-01 3.273e+00 0.221 0.824824
## mvr_pts_log 1.429e-02 5.049e-01 0.028 0.977423
## tif_log 3.332e-01 5.425e-01 0.614 0.539111
## kidsdriv_log 1.111e+00 2.327e+00 0.477 0.633160
## homekids_log 1.323e+00 1.423e+00 0.929 0.352643
## inter 9.891e-02 5.032e-02 1.966 0.049351 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 746.22 on 641 degrees of freedom
## Residual deviance: 535.20 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 647.2
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 110  31
##           1  21  12
##
##           Accuracy : 0.7011
##           95% CI : (0.6272, 0.7681)
##           No Information Rate : 0.7529
##           P-Value [Acc > NIR] : 0.9502
##
##           Kappa : 0.1288
##
## Mcnemar's Test P-Value : 0.2120
##
##           Sensitivity : 0.8397
##           Specificity : 0.2791
##           Pos Pred Value : 0.7801
##           Neg Pred Value : 0.3636
##           Prevalence : 0.7529
##           Detection Rate : 0.6322
##           Detection Prevalence : 0.8103
##           Balanced Accuracy : 0.5594
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.701935025741168"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7019
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0809  -0.6702  -0.3499   0.5078   2.9643
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.099e+01  1.470e+01   1.428  0.15322
## KIDSDRIV       1.480e+00  1.818e+00   0.814  0.41562
## AGE          -1.814e-01  1.035e-01  -1.753  0.07962 .
## HOMEKIDS      -1.292e+00  6.889e-01  -1.876  0.06070 .
## YOJ           -1.282e-01  1.272e-01  -1.008  0.31354
## INCOME        -2.508e-05  1.291e-05  -1.942  0.05210 .
## HOME_VAL       1.521e-05  8.627e-06   1.763  0.07789 .
## TRAVTIME       3.357e-02  2.266e-02   1.481  0.13848
## BLUEBOOK       7.965e-05  4.025e-05   1.979  0.04780 *
```

```

## TIF                3.517e-02  8.985e-02   0.391  0.69551
## OLDCLAIM           1.277e-05  2.100e-05   0.608  0.54324
## CLM_FREQ          -2.826e-01  1.003e+00  -0.282  0.77813
## MVR_PTS            1.142e-01  1.703e-01   0.670  0.50268
## CAR_AGE            -2.285e-02  6.962e-02  -0.328  0.74276
## PARENT1_Yes        7.421e-01  4.759e-01   1.559  0.11895
## MSTATUS_Yes        -6.873e-01  3.402e-01  -2.021  0.04331 *
## SEX_z_F            -1.098e+00  4.738e-01  -2.316  0.02054 *
## EDUCATION_.High.School  5.643e-01  8.147e-01   0.693  0.48859
## EDUCATION_Bachelors -4.268e-02  6.888e-01  -0.062  0.95059
## EDUCATION_Masters   8.252e-01  5.934e-01   1.391  0.16433
## EDUCATION_z_High.School  5.902e-01  7.416e-01   0.796  0.42607
## JOB_               -1.222e+00  6.917e-01  -1.766  0.07739 .
## JOB_Clerical        -2.603e-01  4.851e-01  -0.536  0.59162
## JOB_Doctor          -7.600e-01  1.003e+00  -0.758  0.44842
## JOB_Home.Maker      -2.000e-01  6.749e-01  -0.296  0.76698
## JOB_Lawyer          -1.044e+00  6.493e-01  -1.607  0.10804
## JOB_Manager         -9.915e-01  4.838e-01  -2.049  0.04042 *
## JOB_Student         -4.102e-01  7.212e-01  -0.569  0.56949
## JOB_z_Blue.Collar   -3.656e-01  4.554e-01  -0.803  0.42206
## CAR_USE_Commercial   2.411e-01  3.430e-01   0.703  0.48207
## CAR_TYPE_Panel.Truck  2.611e-01  6.252e-01   0.418  0.67627
## CAR_TYPE_Pickup      1.316e+00  4.136e-01   3.181  0.00147 **
## CAR_TYPE_Sports.Car  2.398e+00  5.621e-01   4.266  1.99e-05 ***
## CAR_TYPE_Van         7.752e-01  4.760e-01   1.629  0.10340
## CAR_TYPE_z_SUV       2.373e+00  5.051e-01   4.699  2.61e-06 ***
## RED_CAR_no          -1.813e-01  3.306e-01  -0.548  0.58349
## REVOKED_Yes         1.331e-01  4.157e-01   0.320  0.74883
## URBANICITY_z_Highly.Rural..Rural -2.590e+00  4.278e-01  -6.055  1.41e-09 ***
## YOJ_NA              5.523e-03  4.514e-01   0.012  0.99024
## INCOME_NA           2.870e-01  5.109e-01   0.562  0.57428
## CAR_AGE_NA          -1.398e-01  5.198e-01  -0.269  0.78800
## HOME_VAL_NA         -3.099e-02  2.917e-01  -0.106  0.91541
## ageSquared          1.901e-03  1.113e-03   1.707  0.08779 .
## yojSquared          5.783e-03  6.562e-03   0.881  0.37813
## income_log          1.525e-01  3.043e-01   0.501  0.61636
## homeval_log         -1.345e+00  1.375e+00  -0.978  0.32800
## travtime_log        -2.654e-01  6.346e-01  -0.418  0.67577
## bluebook_log        -5.727e-01  4.995e-01  -1.147  0.25156
## carage_log          -2.934e-02  4.605e-01  -0.064  0.94919
## oldclaim_log        -1.267e-02  1.638e-01  -0.077  0.93838
## clm_freq_log         1.035e+00  3.086e+00   0.336  0.73724
## mvr_pts_log         -2.066e-01  4.984e-01  -0.414  0.67855
## tif_log             -5.192e-01  5.166e-01  -1.005  0.31491
## kidsdriv_log        -2.841e+00  2.305e+00  -1.233  0.21763
## homekids_log         2.663e+00  1.475e+00   1.805  0.07113 .
## inter               1.842e-02  3.276e-02   0.562  0.57390
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 746.85 on 642 degrees of freedom
## Residual deviance: 543.05 on 587 degrees of freedom

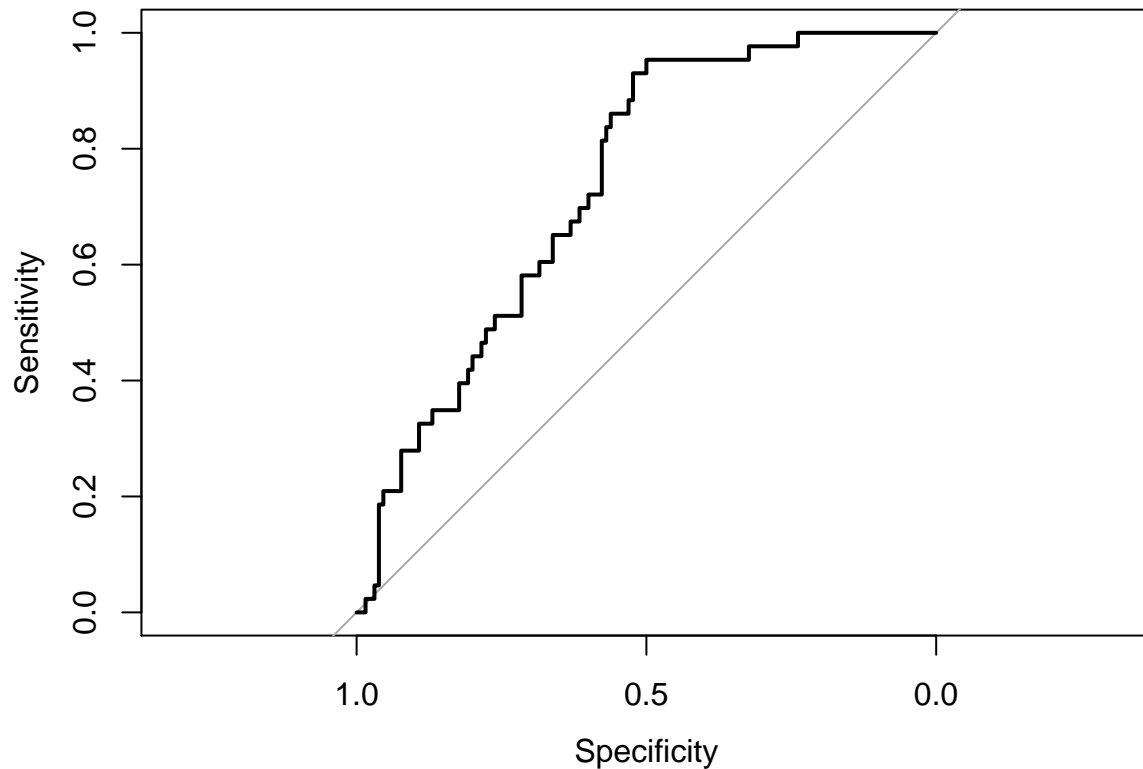
```

```

## (1 observation deleted due to missingness)
## AIC: 655.05
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  28
##           1  21  15
##
##           Accuracy : 0.7168
##           95% CI : (0.6434, 0.7825)
##           No Information Rate : 0.7514
##           P-Value [Acc > NIR] : 0.8728
##
##           Kappa : 0.1981
##
## Mcnemar's Test P-Value : 0.3914
##
##           Sensitivity : 0.8385
##           Specificity : 0.3488
##           Pos Pred Value : 0.7956
##           Neg Pred Value : 0.4167
##           Prevalence : 0.7514
##           Detection Rate : 0.6301
##           Detection Prevalence : 0.7919
##           Balanced Accuracy : 0.5936
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.738461538461539"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 130 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7385
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5097  -0.7057  -0.3424   0.5215   2.6346
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.256e+01  1.506e+01   1.499  0.133961
## KIDSDRIV      -1.952e+00  1.821e+00  -1.072  0.283647
## AGE           -1.351e-01  1.035e-01  -1.305  0.191747
## HOMEKIDS      -2.256e-01  6.708e-01  -0.336  0.736671
## YOJ           -1.856e-01  1.312e-01  -1.414  0.157279
## INCOME        -2.212e-05  1.412e-05  -1.566  0.117348
## HOME_VAL       1.225e-05  9.468e-06   1.294  0.195815
## TRAVTIME       2.793e-02  2.208e-02   1.265  0.205970
## BLUEBOOK       5.305e-05  3.786e-05   1.401  0.161157
```

```

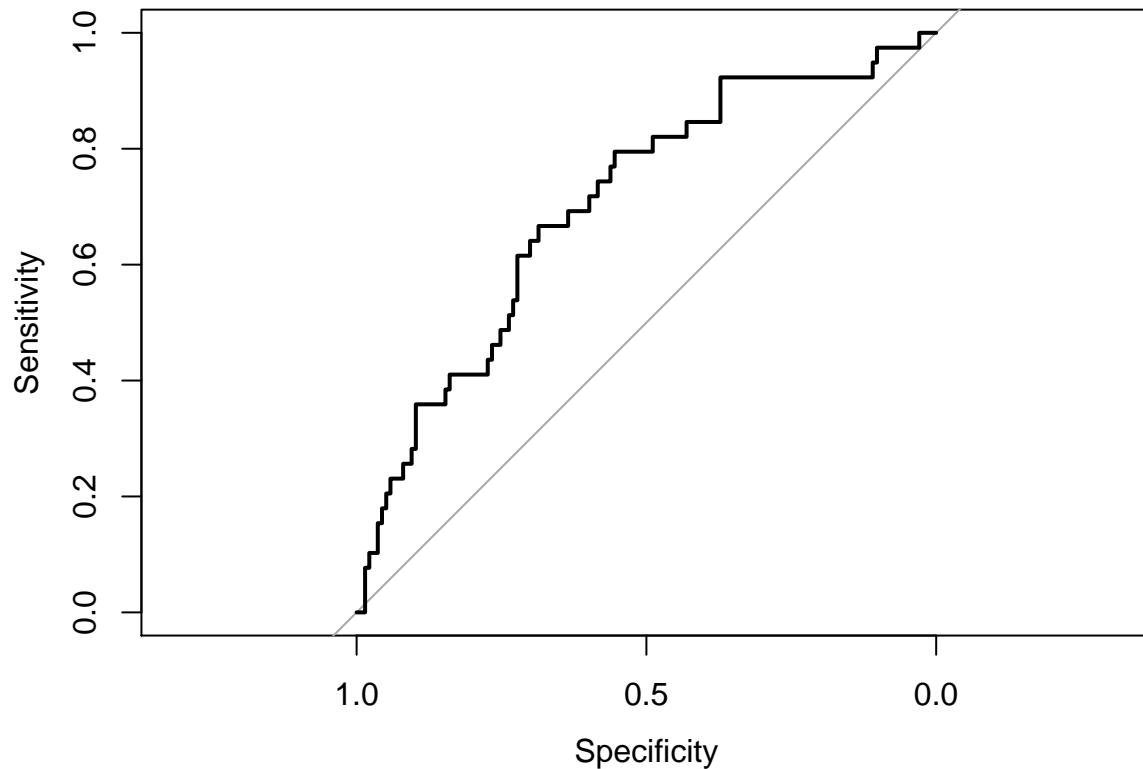
## TIF -3.664e-02 9.184e-02 -0.399 0.689947
## OLDCLAIM -1.389e-05 2.145e-05 -0.648 0.517300
## CLM_FREQ 7.904e-01 9.970e-01 0.793 0.427915
## MVR_PTS 3.105e-02 1.818e-01 0.171 0.864366
## CAR_AGE -4.207e-02 6.950e-02 -0.605 0.545007
## PARENT1_Yes 9.229e-01 4.856e-01 1.901 0.057351 .
## MSTATUS_Yes -9.212e-01 3.351e-01 -2.749 0.005973 **
## SEX_z_F -4.907e-01 4.667e-01 -1.051 0.293062
## EDUCATION_.High.School -1.918e-01 8.156e-01 -0.235 0.814123
## EDUCATION_Bachelors -4.637e-01 6.661e-01 -0.696 0.486327
## EDUCATION_Masters 4.389e-01 5.610e-01 0.782 0.433981
## EDUCATION_z_High.School 3.633e-02 7.335e-01 0.050 0.960497
## JOB_ -1.225e+00 6.783e-01 -1.806 0.070916 .
## JOB_Clerical -4.154e-01 4.701e-01 -0.884 0.376847
## JOB_Doctor -5.263e-01 8.772e-01 -0.600 0.548539
## JOB_Home.Maker -5.477e-01 6.718e-01 -0.815 0.414904
## JOB_Lawyer -8.760e-01 6.534e-01 -1.341 0.180033
## JOB_Manager -6.809e-01 4.852e-01 -1.403 0.160514
## JOB_Student -4.472e-01 7.249e-01 -0.617 0.537272
## JOB_z_Blue.Collar -4.985e-01 4.428e-01 -1.126 0.260280
## CAR_USE_Commercial 8.796e-01 3.541e-01 2.484 0.012988 *
## CAR_TYPE_Panel.Truck 3.769e-01 6.342e-01 0.594 0.552356
## CAR_TYPE_Pickup 1.061e+00 4.124e-01 2.572 0.010121 *
## CAR_TYPE_Sports.Car 1.812e+00 5.223e-01 3.470 0.000521 ***
## CAR_TYPE_Van 5.382e-01 4.689e-01 1.148 0.251070
## CAR_TYPE_z_SUV 1.686e+00 4.694e-01 3.591 0.000329 ***
## RED_CAR_no -1.644e-01 3.437e-01 -0.478 0.632396
## REVOKED_Yes 5.823e-01 4.215e-01 1.381 0.167132
## URBANICITY_z_Highly.Rural..Rural -2.473e+00 4.688e-01 -5.275 1.33e-07 ***
## YOJ_NA -8.066e-01 4.456e-01 -1.810 0.070285 .
## INCOME_NA -1.168e-01 5.040e-01 -0.232 0.816795
## CAR_AGE_NA -7.046e-01 5.233e-01 -1.346 0.178169
## HOME_VAL_NA 3.399e-01 2.942e-01 1.155 0.247958
## ageSquared 1.505e-03 1.122e-03 1.341 0.179885
## yojSquared 1.008e-02 6.890e-03 1.463 0.143367
## income_log 2.642e-01 3.055e-01 0.865 0.387099
## homeval_log -1.624e+00 1.465e+00 -1.109 0.267607
## travtime_log 8.441e-02 6.313e-01 0.134 0.893641
## bluebook_log -5.557e-01 4.453e-01 -1.248 0.212115
## carage_log 1.151e-01 4.619e-01 0.249 0.803122
## oldclaim_log 1.134e-01 1.641e-01 0.691 0.489701
## clm_freq_log -1.730e+00 3.048e+00 -0.568 0.570312
## mvr_pts_log 1.995e-01 5.212e-01 0.383 0.701912
## tif_log 1.025e-01 5.212e-01 0.197 0.844118
## kidsdriv_log 2.257e+00 2.199e+00 1.026 0.304684
## homekids_log 1.993e-01 1.450e+00 0.137 0.890665
## inter 3.379e-02 3.251e-02 1.039 0.298602
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 752.86 on 639 degrees of freedom
## Residual deviance: 539.48 on 584 degrees of freedom

```

```

## AIC: 651.48
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  25
##           1  21  14
##
##           Accuracy : 0.7386
##           95% CI : (0.6672, 0.8019)
##           No Information Rate : 0.7784
##           P-Value [Acc > NIR] : 0.9113
##
##           Kappa : 0.2135
##
## Mcnemar's Test P-Value : 0.6583
##
##           Sensitivity : 0.8467
##           Specificity : 0.3590
##           Pos Pred Value : 0.8227
##           Neg Pred Value : 0.4000
##           Prevalence : 0.7784
##           Detection Rate : 0.6591
##           Detection Prevalence : 0.8011
##           Balanced Accuracy : 0.6028
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.703537338573835"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 137 controls (dfPred_raw$class 0) < 39 cases (dfPred_raw$class 1).
## Area under the curve: 0.7035
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1970  -0.6921  -0.3599   0.5893   2.7224
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.899e+01  1.440e+01   1.319  0.18717
## KIDSDRIV      -1.551e+00  1.700e+00  -0.912  0.36161
## AGE           -1.288e-01  1.049e-01  -1.228  0.21962
## HOMEKIDS      -5.179e-01  6.493e-01  -0.798  0.42513
## YOJ           -1.693e-01  1.343e-01  -1.260  0.20761
## INCOME        -5.710e-06  1.284e-05  -0.445  0.65640
## HOME_VAL       4.716e-06  8.623e-06   0.547  0.58443
## TRAVTIME       2.576e-02  2.283e-02   1.128  0.25928
## BLUEBOOK       5.227e-05  3.749e-05   1.394  0.16318
```

```

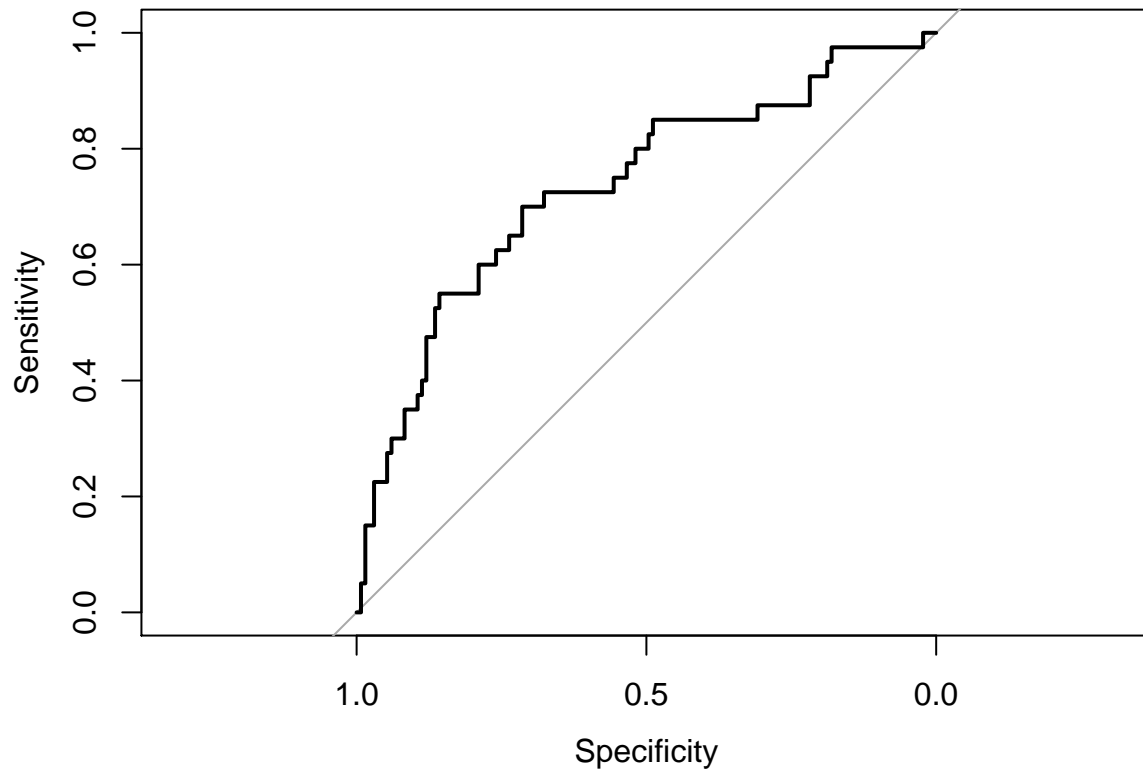
## TIF -2.492e-02 8.564e-02 -0.291 0.77105
## OLDCLAIM 9.397e-06 2.094e-05 0.449 0.65355
## CLM_FREQ -3.122e-01 9.468e-01 -0.330 0.74160
## MVR_PTS -1.439e-02 1.686e-01 -0.085 0.93195
## CAR_AGE -1.012e-02 6.907e-02 -0.146 0.88357
## PARENT1_Yes 4.715e-01 4.648e-01 1.014 0.31038
## MSTATUS_Yes -5.331e-01 3.310e-01 -1.611 0.10724
## SEX_z_F -5.724e-01 4.451e-01 -1.286 0.19841
## EDUCATION_.High.School 7.573e-02 7.814e-01 0.097 0.92280
## EDUCATION_Bachelors -5.763e-02 6.579e-01 -0.088 0.93020
## EDUCATION_Masters 8.014e-01 5.683e-01 1.410 0.15843
## EDUCATION_z_High.School 3.519e-01 7.114e-01 0.495 0.62081
## JOB_ -7.313e-01 6.490e-01 -1.127 0.25977
## JOB_Clerical 9.957e-02 4.845e-01 0.205 0.83718
## JOB_Doctor -7.662e-02 9.499e-01 -0.081 0.93571
## JOB_Home.Maker -3.842e-01 6.995e-01 -0.549 0.58289
## JOB_Lawyer -7.782e-01 6.554e-01 -1.187 0.23508
## JOB_Manager -8.052e-01 4.749e-01 -1.695 0.09000
## JOB_Student -5.454e-01 7.295e-01 -0.748 0.45468
## JOB_z_Blue.Collar -1.418e-02 4.484e-01 -0.032 0.97477
## CAR_USE_Commercial 4.726e-01 3.368e-01 1.403 0.16060
## CAR_TYPE_Panel.Truck 1.330e-01 6.203e-01 0.214 0.83022
## CAR_TYPE_Pickup 9.980e-01 3.910e-01 2.553 0.01069 *
## CAR_TYPE_Sports.Car 1.082e+00 5.215e-01 2.075 0.03795 *
## CAR_TYPE_Van 3.296e-01 4.527e-01 0.728 0.46660
## CAR_TYPE_z_SUV 1.251e+00 4.482e-01 2.790 0.00527 **
## RED_CAR_no -2.891e-01 3.245e-01 -0.891 0.37293
## REVOKED_Yes 1.942e-01 4.075e-01 0.477 0.63364
## URBANICITY_z_Highly.Rural..Rural -2.877e+00 4.900e-01 -5.872 4.31e-09 ***
## YOJ_NA -2.968e-01 4.317e-01 -0.688 0.49167
## INCOME_NA 5.539e-02 5.345e-01 0.104 0.91746
## CAR_AGE_NA -2.410e-01 4.847e-01 -0.497 0.61897
## HOME_VAL_NA 1.861e-02 2.899e-01 0.064 0.94882
## ageSquared 1.396e-03 1.151e-03 1.213 0.22513
## yojSquared 6.392e-03 6.886e-03 0.928 0.35323
## income_log -8.282e-03 2.831e-01 -0.029 0.97666
## homeval_log -1.080e+00 1.367e+00 -0.790 0.42949
## travtime_log 2.511e-02 6.317e-01 0.040 0.96829
## bluebook_log -4.693e-01 4.401e-01 -1.066 0.28628
## carage_log -2.887e-01 4.579e-01 -0.630 0.52839
## oldclaim_log -3.322e-02 1.598e-01 -0.208 0.83536
## clm_freq_log 1.223e+00 2.944e+00 0.415 0.67786
## mvr_pts_log 1.897e-01 4.967e-01 0.382 0.70245
## tif_log -1.011e-01 4.967e-01 -0.204 0.83868
## kidsdriv_log 1.300e+00 2.061e+00 0.631 0.52812
## homekids_log 1.331e+00 1.415e+00 0.940 0.34704
## inter 3.291e-02 3.133e-02 1.050 0.29362
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 752.82 on 642 degrees of freedom
## Residual deviance: 553.56 on 587 degrees of freedom

```

```

## AIC: 665.56
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  24
##           1  15  16
##
##           Accuracy : 0.7746
##           95% CI : (0.705, 0.8345)
##           No Information Rate : 0.7688
##           P-Value [Acc > NIR] : 0.4705
##
##           Kappa : 0.3117
##
## Mcnemar's Test P-Value : 0.2002
##
##           Sensitivity : 0.8872
##           Specificity : 0.4000
##           Pos Pred Value : 0.8310
##           Neg Pred Value : 0.5161
##           Prevalence : 0.7688
##           Detection Rate : 0.6821
##           Detection Prevalence : 0.8208
##           Balanced Accuracy : 0.6436
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.735526315789474"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 40 cases (dfPred_raw$class 1).
## Area under the curve: 0.7355
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1703  -0.6634  -0.3439   0.3589   3.1152
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.534e+01  1.440e+01   1.065  0.28692
## KIDSDRIV      -9.100e-01  1.700e+00  -0.535  0.59234
## AGE           -2.830e-02  1.150e-01  -0.246  0.80561
## HOMEKIDS      -1.431e+00  7.005e-01  -2.043  0.04107 *
## YOJ           -2.350e-01  1.318e-01  -1.783  0.07461 .
## INCOME        -1.455e-05  1.398e-05  -1.041  0.29809
## HOME_VAL       9.716e-06  9.065e-06   1.072  0.28377
## TRAVTIME      2.212e-02  2.196e-02   1.008  0.31366
## BLUEBOOK      4.952e-05  3.922e-05   1.262  0.20678
```

```

## TIF -1.033e-01 9.211e-02 -1.122 0.26192
## OLDCLAIM 1.053e-05 2.121e-05 0.496 0.61955
## CLM_FREQ -1.783e-01 1.001e+00 -0.178 0.85864
## MVRPTS 1.568e-01 1.752e-01 0.895 0.37092
## CAR_AGE -6.627e-02 6.980e-02 -0.949 0.34240
## PARENT1_Yes 4.291e-01 4.856e-01 0.884 0.37685
## MSTATUS_Yes -8.250e-01 3.373e-01 -2.446 0.01446 *
## SEX_z_F -1.272e+00 4.753e-01 -2.676 0.00745 **
## EDUCATION_.High.School 8.923e-02 8.077e-01 0.110 0.91203
## EDUCATION_Bachelors -4.952e-01 6.568e-01 -0.754 0.45091
## EDUCATION_Masters 3.548e-01 5.550e-01 0.639 0.52257
## EDUCATION_z_High.School 1.851e-01 7.103e-01 0.261 0.79440
## JOB_ -9.917e-01 6.742e-01 -1.471 0.14131
## JOB_Clerical -8.029e-01 5.059e-01 -1.587 0.11250
## JOB_Doctor -3.803e-01 8.784e-01 -0.433 0.66500
## JOB_Home.Maker -5.570e-01 6.538e-01 -0.852 0.39426
## JOB_Lawyer -1.252e+00 6.742e-01 -1.857 0.06330 .
## JOB_Manager -6.853e-01 4.812e-01 -1.424 0.15441
## JOB_Student -9.089e-01 7.344e-01 -1.238 0.21588
## JOB_z_Blue.Collar -2.851e-01 4.535e-01 -0.629 0.52962
## CAR_USE_Commercial 5.710e-01 3.603e-01 1.585 0.11298
## CAR_TYPE_Panel.Truck -4.295e-01 6.294e-01 -0.682 0.49506
## CAR_TYPE_Pickup 1.128e+00 4.118e-01 2.740 0.00614 **
## CAR_TYPE_Sports.Car 2.666e+00 5.656e-01 4.713 2.44e-06 ***
## CAR_TYPE_Van 4.353e-01 4.687e-01 0.929 0.35306
## CAR_TYPE_z_SUV 2.517e+00 5.140e-01 4.898 9.70e-07 ***
## RED_CAR_no -5.970e-02 3.393e-01 -0.176 0.86035
## REVOKED_Yes -5.764e-03 4.251e-01 -0.014 0.98918
## URBANICITY_z_Highly.Rural..Rural -2.445e+00 4.391e-01 -5.568 2.58e-08 ***
## YOJ_NA -3.947e-01 4.276e-01 -0.923 0.35602
## INCOME_NA 8.902e-02 5.300e-01 0.168 0.86662
## CAR_AGE_NA 2.142e-02 4.835e-01 0.044 0.96467
## HOME_VAL_NA -4.849e-02 2.878e-01 -0.169 0.86617
## ageSquared 1.358e-04 1.258e-03 0.108 0.91405
## yojSquared 1.438e-02 6.880e-03 2.090 0.03665 *
## income_log 8.300e-03 3.015e-01 0.028 0.97804
## homeval_log -1.431e+00 1.366e+00 -1.047 0.29505
## travtime_log -1.396e-02 6.170e-01 -0.023 0.98195
## bluebook_log 9.024e-03 4.891e-01 0.018 0.98528
## carage_log 2.980e-01 4.646e-01 0.641 0.52130
## oldclaim_log 4.486e-02 1.639e-01 0.274 0.78434
## clm_freq_log 4.274e-01 3.059e+00 0.140 0.88890
## mvr_pts_log -4.164e-01 5.100e-01 -0.817 0.41419
## tif_log 3.851e-01 5.264e-01 0.732 0.46442
## kidsdriv_log 3.155e-01 2.220e+00 0.142 0.88701
## homekids_log 2.647e+00 1.486e+00 1.781 0.07490 .
## inter 3.750e-02 3.176e-02 1.181 0.23762
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 724.73 on 640 degrees of freedom
## Residual deviance: 533.51 on 585 degrees of freedom

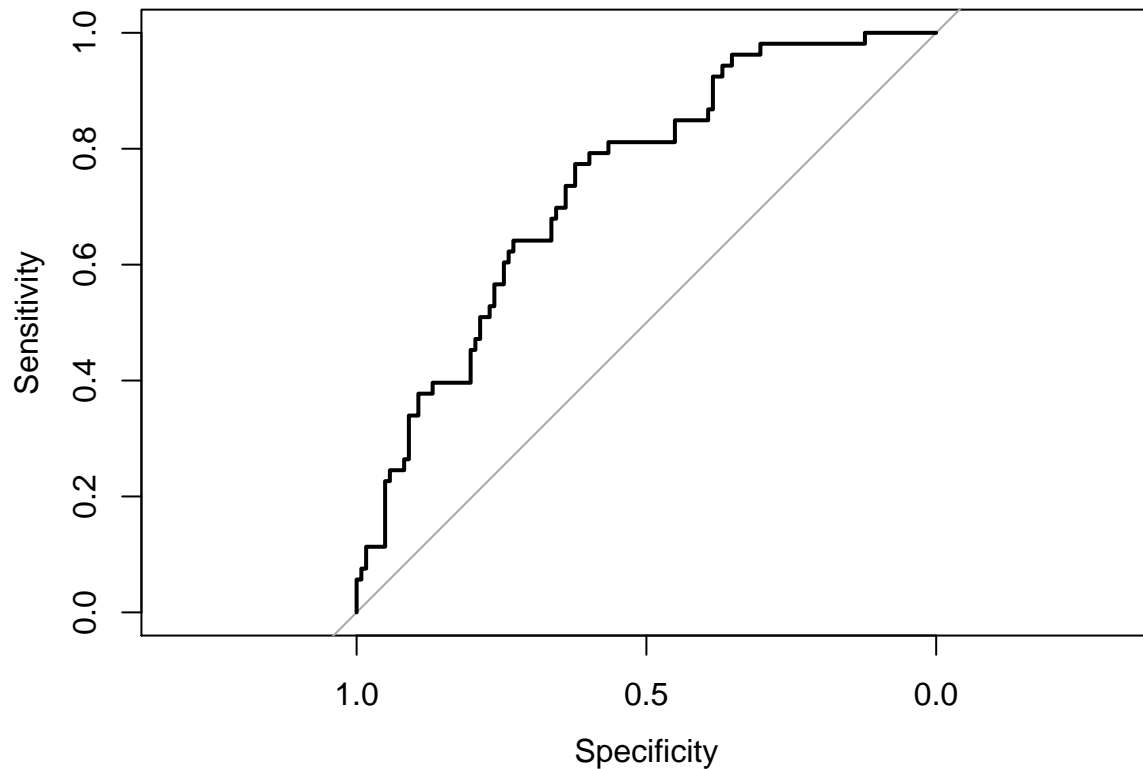
```



```

## (1 observation deleted due to missingness)
## AIC: 645.51
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  33
##           1  14  20
##
##           Accuracy : 0.7314
##           95% CI : (0.6593, 0.7955)
##           No Information Rate : 0.6971
##           P-Value [Acc > NIR] : 0.18330
##
##           Kappa : 0.2922
##
## Mcnemar's Test P-Value : 0.00865
##
##           Sensitivity : 0.8852
##           Specificity : 0.3774
##           Pos Pred Value : 0.7660
##           Neg Pred Value : 0.5882
##           Prevalence : 0.6971
##           Detection Rate : 0.6171
##           Detection Prevalence : 0.8057
##           Balanced Accuracy : 0.6313
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.73816888339004"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 122 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.7382
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1029  -0.6812  -0.3615   0.5521   3.0382
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.666e+01  1.454e+01   1.834 0.066708 .
## KIDSDRIV      -1.982e-01  1.838e+00  -0.108 0.914101
## AGE           -1.353e-02  1.085e-01  -0.125 0.900687
## HOMEKIDS       -6.857e-01  6.829e-01  -1.004 0.315326
## YOJ            -1.269e-01  1.294e-01  -0.981 0.326715
## INCOME         -2.549e-05  1.279e-05  -1.992 0.046357 *
## HOME_VAL        1.835e-05  8.586e-06   2.137 0.032595 *
## TRAVTIME        2.296e-02  2.045e-02   1.122 0.261737
## BLUEBOOK        3.668e-05  4.076e-05   0.900 0.368204
```

```

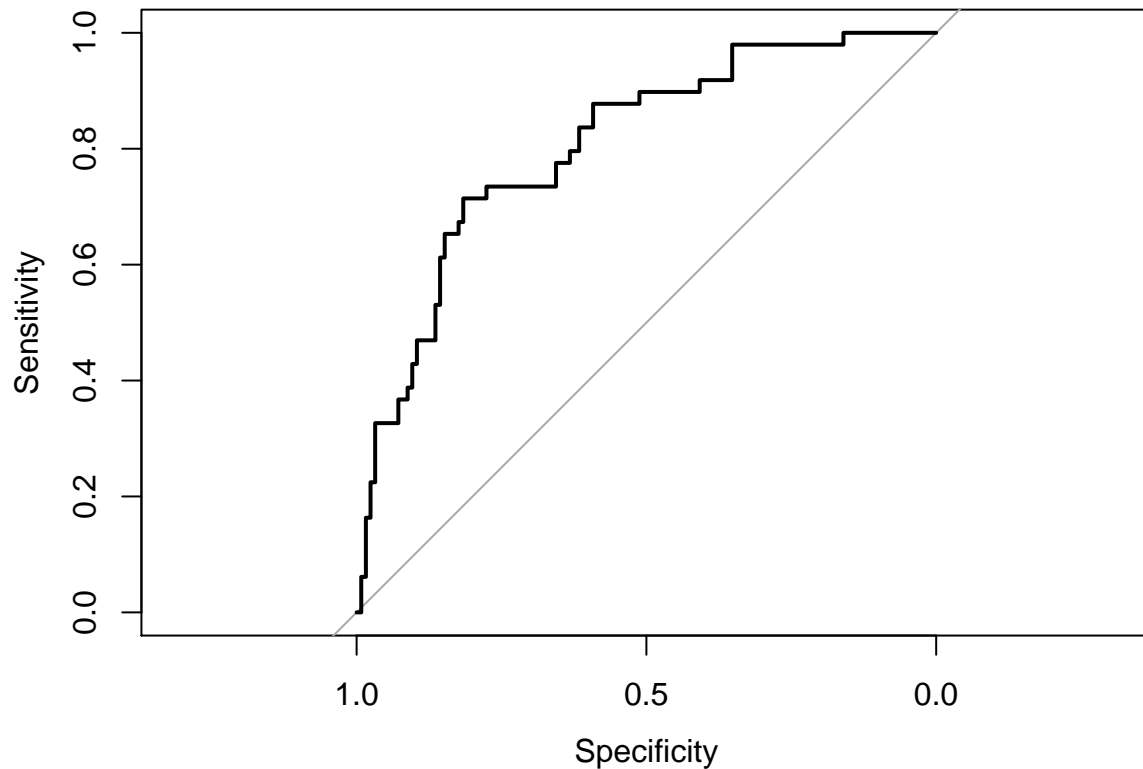
## TIF -5.502e-02 9.335e-02 -0.589 0.555567
## OLDCLAIM -9.705e-06 2.123e-05 -0.457 0.647491
## CLM_FREQ 3.942e-01 9.386e-01 0.420 0.674488
## MVRPTS 5.540e-03 1.692e-01 0.033 0.973886
## CAR_AGE 5.436e-02 6.801e-02 0.799 0.424112
## PARENT1_Yes 7.906e-01 4.621e-01 1.711 0.087078 .
## MSTATUS_Yes -5.699e-01 3.327e-01 -1.713 0.086678 .
## SEX_z_F -6.301e-01 4.680e-01 -1.346 0.178175
## EDUCATION_.High.School 3.267e-01 7.279e-01 0.449 0.653599
## EDUCATION_Bachelors 8.377e-02 6.082e-01 0.138 0.890463
## EDUCATION_Masters 8.365e-01 5.172e-01 1.617 0.105790
## EDUCATION_z_High.School 5.504e-01 6.612e-01 0.832 0.405190
## JOB_ -8.165e-01 6.746e-01 -1.210 0.226136
## JOB_Clerical -3.113e-01 4.805e-01 -0.648 0.517153
## JOB_Doctor -8.376e-01 1.006e+00 -0.833 0.404924
## JOB_Home.Maker -2.851e-01 6.489e-01 -0.439 0.660448
## JOB_Lawyer -9.007e-01 6.158e-01 -1.463 0.143558
## JOB_Manager -8.598e-01 4.623e-01 -1.860 0.062889 .
## JOB_Student -7.810e-01 6.955e-01 -1.123 0.261469
## JOB_z_Blue.Collar -2.856e-01 4.429e-01 -0.645 0.519139
## CAR_USE_Commercial 4.886e-01 3.353e-01 1.457 0.145007
## CAR_TYPE_Panel.Truck 3.786e-01 6.233e-01 0.607 0.543534
## CAR_TYPE_Pickup 1.174e+00 4.078e-01 2.879 0.003994 **
## CAR_TYPE_Sports.Car 1.642e+00 5.339e-01 3.075 0.002107 **
## CAR_TYPE_Van 4.189e-01 4.612e-01 0.908 0.363705
## CAR_TYPE_z_SUV 1.739e+00 4.716e-01 3.688 0.000226 ***
## RED_CAR_no -8.861e-02 3.369e-01 -0.263 0.792516
## REVOKED_Yes 7.604e-01 4.045e-01 1.880 0.060118 .
## URBANICITY_z_Highly.Rural..Rural -2.424e+00 4.357e-01 -5.564 2.64e-08 ***
## YOJ_NA 1.545e-02 4.413e-01 0.035 0.972069
## INCOME_NA 1.909e-01 5.559e-01 0.343 0.731283
## CAR_AGE_NA -1.856e-02 4.674e-01 -0.040 0.968320
## HOME_VAL_NA -1.097e-01 2.891e-01 -0.380 0.704208
## ageSquared 1.971e-04 1.173e-03 0.168 0.866613
## yojSquared 7.202e-03 6.698e-03 1.075 0.282258
## income_log 1.754e-01 2.637e-01 0.665 0.506009
## homeval_log -2.485e+00 1.325e+00 -1.875 0.060725 .
## travtime_log -2.025e-01 5.675e-01 -0.357 0.721262
## bluebook_log -1.875e-01 5.135e-01 -0.365 0.715053
## carage_log -5.085e-01 4.492e-01 -1.132 0.257674
## oldclaim_log 7.105e-02 1.565e-01 0.454 0.649785
## clm_freq_log -7.121e-01 2.887e+00 -0.247 0.805191
## mvr_pts_log 1.651e-01 5.025e-01 0.329 0.742477
## tif_log -8.140e-02 5.195e-01 -0.157 0.875502
## kidsdriv_log 1.566e+00 2.169e+00 0.722 0.470434
## homekids_log 1.394e+00 1.458e+00 0.956 0.338886
## inter -3.319e-03 3.473e-02 -0.096 0.923863
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.87 on 641 degrees of freedom
## Residual deviance: 555.45 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 667.45
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 113  29
##           1  12  20
##
##           Accuracy : 0.7644
##           95% CI : (0.6942, 0.8253)
##           No Information Rate : 0.7184
##           P-Value [Acc > NIR] : 0.10161
##
##           Kappa : 0.349
##
## Mcnemar's Test P-Value : 0.01246
##
##           Sensitivity : 0.9040
##           Specificity : 0.4082
##           Pos Pred Value : 0.7958
##           Neg Pred Value : 0.6250
##           Prevalence : 0.7184
##           Detection Rate : 0.6494
##           Detection Prevalence : 0.8161
##           Balanced Accuracy : 0.6561
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.805551020408163"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 125 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.8056
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1785  -0.6895  -0.3787   0.6681   2.8801
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.820e+01  1.391e+01   1.309 0.190687
## KIDSDRIV       7.711e-01  1.799e+00   0.429 0.668121
## AGE          -6.054e-02  1.082e-01  -0.559 0.575907
## HOMEKIDS     -7.613e-01  6.785e-01  -1.122 0.261861
## YOJ         -1.251e-01  1.282e-01  -0.976 0.329263
## INCOME      -2.828e-05  1.323e-05  -2.138 0.032529 *
## HOME_VAL     1.525e-05  8.659e-06   1.761 0.078284 .
## TRAVTIME     1.294e-02  2.141e-02   0.605 0.545456
## BLUEBOOK     4.300e-05  3.996e-05   1.076 0.281949
```

```

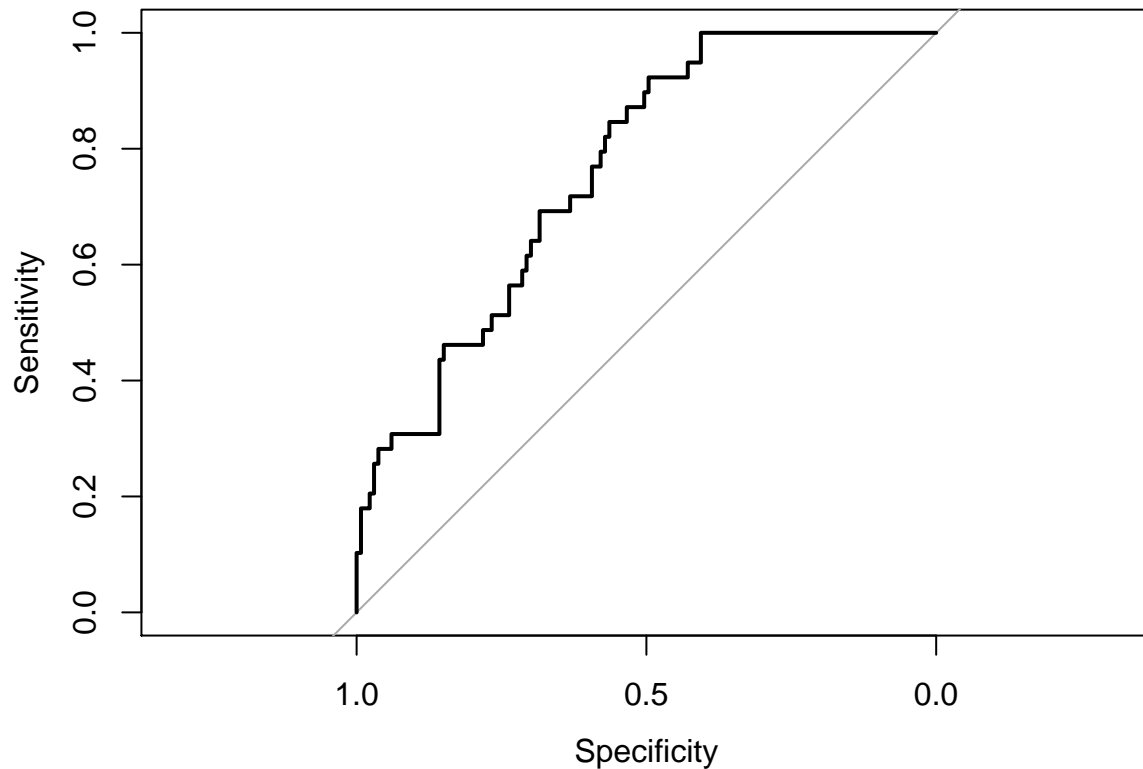
## TIF -3.246e-02 8.602e-02 -0.377 0.705874
## OLDCLAIM 1.726e-05 2.243e-05 0.770 0.441537
## CLM_FREQ -1.011e+00 1.010e+00 -1.002 0.316521
## MVRPTS 1.216e-01 1.629e-01 0.746 0.455400
## CAR_AGE -1.731e-02 7.176e-02 -0.241 0.809396
## PARENT1_Yes 7.210e-01 4.534e-01 1.590 0.111798
## MSTATUS_Yes -5.126e-01 3.248e-01 -1.578 0.114477
## SEX_z_F -8.892e-01 4.540e-01 -1.959 0.050134 .
## EDUCATION_High.School -3.267e-01 8.047e-01 -0.406 0.684717
## EDUCATION_Bachelors -1.967e-01 6.841e-01 -0.287 0.773756
## EDUCATION_Masters 8.443e-01 5.867e-01 1.439 0.150118
## EDUCATION_z_High.School 3.437e-01 7.305e-01 0.470 0.638020
## JOB_ -1.116e+00 6.856e-01 -1.628 0.103430
## JOB_Clerical -1.854e-01 4.737e-01 -0.391 0.695511
## JOB_Doctor -3.853e-01 9.872e-01 -0.390 0.696284
## JOB_Home.Maker -4.427e-01 6.319e-01 -0.701 0.483595
## JOB_Lawyer -9.091e-01 6.344e-01 -1.433 0.151817
## JOB_Manager -9.057e-01 4.764e-01 -1.901 0.057277 .
## JOB_Student -8.526e-01 7.002e-01 -1.218 0.223356
## JOB_z_Blue.Collar -1.458e-01 4.426e-01 -0.329 0.741904
## CAR_USE_Commercial 1.685e-01 3.352e-01 0.503 0.615282
## CAR_TYPE_Panel.Truck 3.612e-01 5.940e-01 0.608 0.543186
## CAR_TYPE_Pickup 1.447e+00 3.974e-01 3.641 0.000271 ***
## CAR_TYPE_Sports.Car 1.760e+00 5.361e-01 3.282 0.001029 **
## CAR_TYPE_Van 3.625e-01 4.686e-01 0.774 0.439093
## CAR_TYPE_z_SUV 1.983e+00 4.685e-01 4.233 2.31e-05 ***
## RED_CAR_no -3.411e-02 3.227e-01 -0.106 0.915824
## REVOKED_Yes 2.995e-01 4.100e-01 0.731 0.465083
## URBANICITY_z_Highly.Rural..Rural -2.260e+00 4.018e-01 -5.623 1.88e-08 ***
## YOJ_NA -1.727e-01 4.430e-01 -0.390 0.696654
## INCOME_NA 1.285e-01 5.497e-01 0.234 0.815208
## CAR_AGE_NA -1.892e-01 4.616e-01 -0.410 0.681879
## HOME_VAL_NA -1.748e-02 2.799e-01 -0.062 0.950208
## ageSquared 6.366e-04 1.192e-03 0.534 0.593368
## yojSquared 6.216e-03 6.556e-03 0.948 0.343073
## income_log 2.207e-01 2.851e-01 0.774 0.438828
## homeval_log -1.741e+00 1.299e+00 -1.339 0.180424
## travtime_log 1.493e-01 6.101e-01 0.245 0.806666
## bluebook_log -1.073e-01 4.979e-01 -0.216 0.829310
## carage_log -1.702e-01 4.550e-01 -0.374 0.708384
## oldclaim_log -1.869e-01 1.681e-01 -1.112 0.266200
## clm_freq_log 3.611e+00 3.098e+00 1.166 0.243742
## mvr_pts_log -1.867e-01 4.852e-01 -0.385 0.700301
## tif_log -1.336e-01 4.896e-01 -0.273 0.784899
## kidsdriv_log -1.091e+00 2.307e+00 -0.473 0.636296
## homekids_log 1.752e+00 1.450e+00 1.208 0.226874
## inter 1.064e-02 3.135e-02 0.339 0.734427
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 755.42 on 643 degrees of freedom
## Residual deviance: 567.87 on 588 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 679.87
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 125  28
##           1   8  11
##
##           Accuracy : 0.7907
##           95% CI : (0.7222, 0.8489)
##           No Information Rate : 0.7733
##           P-Value [Acc > NIR] : 0.329321
##
##           Kappa : 0.271
##
## Mcnemar's Test P-Value : 0.001542
##
##           Sensitivity : 0.9398
##           Specificity : 0.2821
##           Pos Pred Value : 0.8170
##           Neg Pred Value : 0.5789
##           Prevalence : 0.7733
##           Detection Rate : 0.7267
##           Detection Prevalence : 0.8895
##           Balanced Accuracy : 0.6110
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.762868710237131"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 39 cases (dfPred_raw$class 1).
## Area under the curve: 0.7629
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2398  -0.6014  -0.3314   0.4406   2.7079
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.512e+00  1.543e+01   0.098  0.921940
## KIDSDRIV      -7.290e-01  1.719e+00  -0.424  0.671446
## AGE           -1.815e-01  1.091e-01  -1.664  0.096048 .
## HOMEKIDS      -1.061e+00  6.835e-01  -1.552  0.120625
## YOJ           -2.865e-01  1.365e-01  -2.099  0.035816 *
## INCOME        -1.222e-05  1.419e-05  -0.861  0.389400
## HOME_VAL       3.716e-06  9.652e-06   0.385  0.700211
## TRAVTIME       3.301e-02  2.239e-02   1.475  0.140326
## BLUEBOOK       3.368e-05  4.111e-05   0.819  0.412615
```



```

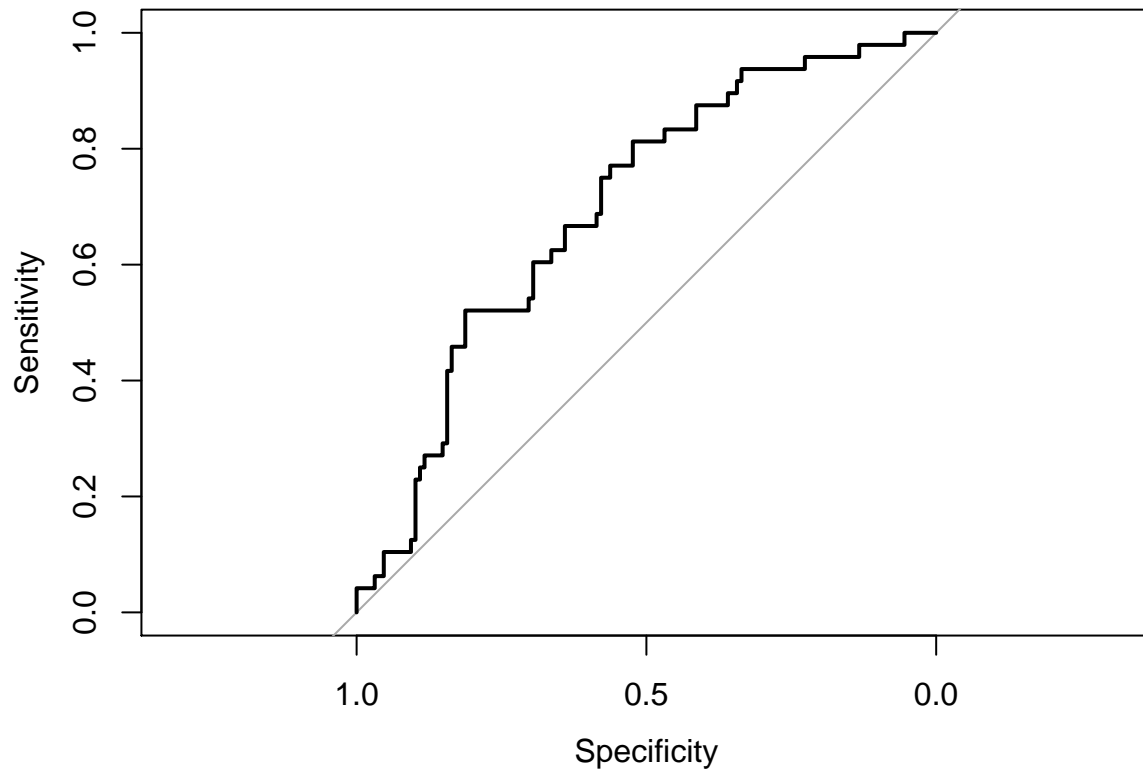
## TIF -1.171e-01 9.499e-02 -1.233 0.217717
## OLDCLAIM 3.769e-05 2.223e-05 1.696 0.089967
## CLM_FREQ -9.502e-01 1.081e+00 -0.879 0.379581
## MVR_PTS 4.435e-03 1.802e-01 0.025 0.980362
## CAR_AGE -2.635e-02 7.276e-02 -0.362 0.717260
## PARENT1_Yes 3.275e-01 4.909e-01 0.667 0.504711
## MSTATUS_Yes -8.719e-01 3.486e-01 -2.501 0.012371 *
## SEX_z_F -7.329e-01 4.785e-01 -1.532 0.125621
## EDUCATION_.High.School 1.354e+00 8.617e-01 1.571 0.116174
## EDUCATION_Bachelors 2.652e-01 7.412e-01 0.358 0.720477
## EDUCATION_Masters 9.884e-01 6.809e-01 1.452 0.146609
## EDUCATION_z_High.School 9.740e-01 7.916e-01 1.230 0.218530
## JOB_ -1.095e+00 7.034e-01 -1.557 0.119584
## JOB_Clerical -3.034e-01 5.065e-01 -0.599 0.549175
## JOB_Doctor 2.981e-01 9.988e-01 0.298 0.765380
## JOB_Home.Maker -2.763e-01 6.836e-01 -0.404 0.686070
## JOB_Lawyer -8.549e-01 6.929e-01 -1.234 0.217307
## JOB_Manager -4.036e-01 4.758e-01 -0.848 0.396204
## JOB_Student -8.305e-01 7.643e-01 -1.087 0.277213
## JOB_z_Blue.Collar -5.935e-01 4.604e-01 -1.289 0.197377
## CAR_USE_Commercial 1.009e+00 3.623e-01 2.785 0.005359 **
## CAR_TYPE_Panel.Truck -5.423e-01 6.623e-01 -0.819 0.412828
## CAR_TYPE_Pickup 9.729e-01 4.044e-01 2.406 0.016125 *
## CAR_TYPE_Sports.Car 1.697e+00 5.482e-01 3.095 0.001970 **
## CAR_TYPE_Van 2.660e-01 4.904e-01 0.542 0.587479
## CAR_TYPE_z_SUV 1.779e+00 4.875e-01 3.649 0.000263 ***
## RED_CAR_no -1.993e-01 3.450e-01 -0.578 0.563469
## REVOKED_Yes -7.618e-02 4.469e-01 -0.170 0.864648
## URBANICITY_z_Highly.Rural..Rural -2.773e+00 4.505e-01 -6.156 7.48e-10 ***
## YOJ_NA 8.458e-02 4.717e-01 0.179 0.857714
## INCOME_NA 2.476e-02 5.277e-01 0.047 0.962583
## CAR_AGE_NA 1.703e-02 5.174e-01 0.033 0.973747
## HOME_VAL_NA -5.060e-02 2.941e-01 -0.172 0.863399
## ageSquared 1.818e-03 1.191e-03 1.526 0.126973
## yojSquared 1.532e-02 7.207e-03 2.125 0.033575 *
## income_log -1.199e-01 2.836e-01 -0.423 0.672521
## homeval_log 5.977e-02 1.446e+00 0.041 0.967040
## travtime_log -2.501e-01 6.255e-01 -0.400 0.689220
## bluebook_log 1.883e-01 5.010e-01 0.376 0.707091
## carage_log -8.254e-02 4.742e-01 -0.174 0.861827
## oldclaim_log -1.196e-01 1.783e-01 -0.671 0.502341
## clm_freq_log 3.024e+00 3.316e+00 0.912 0.361928
## mvr_pts_log 1.300e-01 5.204e-01 0.250 0.802744
## tif_log 4.576e-01 5.359e-01 0.854 0.393111
## kidsdriv_log 2.320e+00 2.189e+00 1.060 0.289216
## homekids_log 2.115e+00 1.471e+00 1.437 0.150592
## inter 8.580e-03 3.230e-02 0.266 0.790521
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 734.77 on 639 degrees of freedom
## Residual deviance: 519.85 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 631.85
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  30
##           1  20  18
##
##           Accuracy : 0.7159
##           95% CI : (0.6432, 0.7812)
##           No Information Rate : 0.7273
##           P-Value [Acc > NIR] : 0.6677
##
##           Kappa : 0.234
##
## Mcnemar's Test P-Value : 0.2031
##
##           Sensitivity : 0.8438
##           Specificity : 0.3750
##           Pos Pred Value : 0.7826
##           Neg Pred Value : 0.4737
##           Prevalence : 0.7273
##           Detection Rate : 0.6136
##           Detection Prevalence : 0.7841
##           Balanced Accuracy : 0.6094
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.697591145833333"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 128 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.6976
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1967  -0.6714  -0.3373   0.5407   2.7610
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.415e+01  1.438e+01   1.679  0.09316 .
## KIDSDRIV       1.379e-01  1.749e+00   0.079  0.93715
## AGE          -6.609e-02  1.043e-01  -0.634  0.52619
## HOMEKIDS      -8.731e-01  7.059e-01  -1.237  0.21613
## YOJ          -2.251e-01  1.286e-01  -1.751  0.07992 .
## INCOME       -2.736e-05  1.351e-05  -2.025  0.04285 *
## HOME_VAL      1.436e-05  8.831e-06   1.626  0.10399
## TRAVTIME      1.869e-02  2.120e-02   0.881  0.37810
## BLUEBOOK      5.473e-05  4.032e-05   1.357  0.17471
```

```

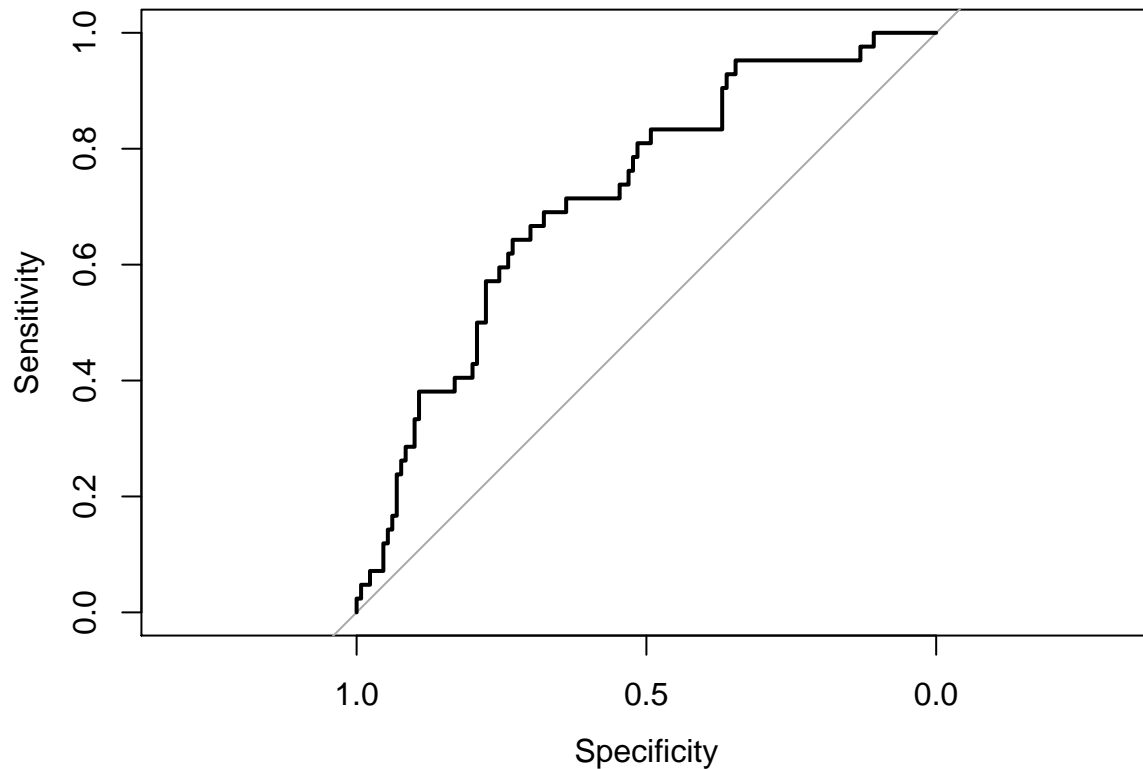
## TIF -6.361e-02 9.604e-02 -0.662 0.50778
## OLDCLAIM -1.369e-05 2.255e-05 -0.607 0.54385
## CLM_FREQ -3.566e-01 1.055e+00 -0.338 0.73539
## MVR_PTS 6.289e-02 1.705e-01 0.369 0.71219
## CAR_AGE -8.610e-03 7.234e-02 -0.119 0.90525
## PARENT1_Yes 1.034e+00 4.709e-01 2.195 0.02816 *
## MSTATUS_Yes -6.747e-01 3.316e-01 -2.035 0.04190 *
## SEX_z_F -1.028e+00 4.715e-01 -2.180 0.02929 *
## EDUCATION_.High.School -4.167e-01 7.801e-01 -0.534 0.59324
## EDUCATION_Bachelors -4.834e-01 6.541e-01 -0.739 0.45985
## EDUCATION_Masters 6.529e-01 5.793e-01 1.127 0.25974
## EDUCATION_z_High.School -7.661e-02 7.038e-01 -0.109 0.91333
## JOB_ -1.268e+00 6.949e-01 -1.824 0.06814 .
## JOB_Clerical -3.554e-01 4.807e-01 -0.739 0.45963
## JOB_Doctor -1.073e+00 9.517e-01 -1.128 0.25933
## JOB_Home.Maker -3.505e-01 6.378e-01 -0.550 0.58266
## JOB_Lawyer -1.341e+00 6.324e-01 -2.121 0.03392 *
## JOB_Manager -7.377e-01 4.655e-01 -1.585 0.11299
## JOB_Student -9.336e-01 7.038e-01 -1.327 0.18467
## JOB_z_Blue.Collar -2.856e-01 4.478e-01 -0.638 0.52366
## CAR_USE_Commercial 5.308e-01 3.454e-01 1.537 0.12435
## CAR_TYPE_Panel.Truck -2.918e-01 6.212e-01 -0.470 0.63854
## CAR_TYPE_Pickup 9.464e-01 4.011e-01 2.359 0.01830 *
## CAR_TYPE_Sports.Car 1.600e+00 5.419e-01 2.953 0.00314 **
## CAR_TYPE_Van -2.299e-02 4.696e-01 -0.049 0.96096
## CAR_TYPE_z_SUV 1.904e+00 4.766e-01 3.995 6.48e-05 ***
## RED_CAR_no -4.535e-02 3.389e-01 -0.134 0.89353
## REVOKED_Yes 8.649e-01 3.985e-01 2.170 0.03000 *
## URBANICITY_z_Highly.Rural..Rural -2.075e+00 4.003e-01 -5.184 2.18e-07 ***
## YOJ_NA -8.223e-03 4.442e-01 -0.019 0.98523
## INCOME_NA 1.420e-02 5.381e-01 0.026 0.97894
## CAR_AGE_NA -2.913e-01 5.026e-01 -0.579 0.56226
## HOME_VAL_NA -1.129e-02 2.873e-01 -0.039 0.96867
## ageSquared 7.967e-04 1.122e-03 0.710 0.47754
## yojSquared 1.282e-02 6.629e-03 1.934 0.05307 .
## income_log 3.771e-01 3.003e-01 1.256 0.20917
## homeval_log -2.167e+00 1.354e+00 -1.601 0.10933
## travtime_log 1.413e-01 6.095e-01 0.232 0.81665
## bluebook_log -3.092e-01 5.060e-01 -0.611 0.54120
## carage_log -2.296e-01 4.649e-01 -0.494 0.62132
## oldclaim_log 9.812e-04 1.733e-01 0.006 0.99548
## clm_freq_log 1.369e+00 3.219e+00 0.425 0.67066
## mvr_pts_log 1.728e-02 5.026e-01 0.034 0.97258
## tif_log -7.922e-02 5.313e-01 -0.149 0.88146
## kidsdriv_log -4.717e-01 2.171e+00 -0.217 0.82798
## homekids_log 1.542e+00 1.502e+00 1.027 0.30465
## inter 1.725e-02 3.359e-02 0.513 0.60764
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 749.48 on 643 degrees of freedom
## Residual deviance: 544.12 on 588 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 656.12
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  26
##           1  14  16
##
##           Accuracy : 0.7674
##           95% CI : (0.6971, 0.8284)
##           No Information Rate : 0.7558
##           P-Value [Acc > NIR] : 0.40054
##
##           Kappa : 0.3025
##
## Mcnemar's Test P-Value : 0.08199
##
##           Sensitivity : 0.8923
##           Specificity : 0.3810
##           Pos Pred Value : 0.8169
##           Neg Pred Value : 0.5333
##           Prevalence : 0.7558
##           Detection Rate : 0.6744
##           Detection Prevalence : 0.8256
##           Balanced Accuracy : 0.6366
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.719413919413919"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 130 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.7194
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1379  -0.6738  -0.3611   0.5637   3.1272
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.257e+01  1.392e+01   1.622  0.10486
## KIDSDRIV      -7.115e-01  1.670e+00  -0.426  0.67009
## AGE          -1.491e-01  1.084e-01  -1.375  0.16927
## HOMEKIDS      -6.469e-01  6.951e-01  -0.931  0.35203
## YOJ           -1.452e-01  1.262e-01  -1.151  0.24969
## INCOME        -6.969e-06  1.261e-05  -0.552  0.58061
## HOME_VAL       7.664e-06  8.353e-06   0.918  0.35886
## TRAVTIME       3.140e-02  2.180e-02   1.440  0.14973
## BLUEBOOK       5.630e-05  3.914e-05   1.439  0.15028
```

```

## TIF -9.547e-02 8.896e-02 -1.073 0.28318
## OLDCLAIM 1.264e-05 2.123e-05 0.595 0.55156
## CLM_FREQ 1.130e-01 9.950e-01 0.114 0.90958
## MVRPTS 2.969e-02 1.759e-01 0.169 0.86594
## CAR_AGE -8.858e-03 7.162e-02 -0.124 0.90157
## PARENT1_Yes 5.048e-01 4.724e-01 1.068 0.28531
## MSTATUS_Yes -8.270e-01 3.255e-01 -2.540 0.01107 *
## SEX_z_F -1.253e+00 4.868e-01 -2.575 0.01003 *
## EDUCATION_.High.School -2.123e-01 7.652e-01 -0.277 0.78143
## EDUCATION_Bachelors -4.865e-01 6.400e-01 -0.760 0.44714
## EDUCATION_Masters 5.700e-01 5.310e-01 1.073 0.28309
## EDUCATION_z_High.School -9.053e-02 6.947e-01 -0.130 0.89631
## JOB_ -1.190e+00 6.779e-01 -1.756 0.07914 .
## JOB_Clerical -7.437e-02 4.904e-01 -0.152 0.87945
## JOB_Doctor -6.816e-01 8.617e-01 -0.791 0.42895
## JOB_Home.Maker -4.290e-01 6.591e-01 -0.651 0.51514
## JOB_Lawyer -7.472e-01 6.377e-01 -1.172 0.24131
## JOB_Manager -6.523e-01 4.709e-01 -1.385 0.16602
## JOB_Student -7.388e-01 7.198e-01 -1.026 0.30467
## JOB_z_Blue.Collar 1.785e-01 4.435e-01 0.403 0.68724
## CAR_USE_Commercial 2.158e-01 3.578e-01 0.603 0.54650
## CAR_TYPE_Panel.Truck 1.691e-01 6.239e-01 0.271 0.78637
## CAR_TYPE_Pickup 1.300e+00 4.116e-01 3.159 0.00158 **
## CAR_TYPE_Sports.Car 2.324e+00 5.486e-01 4.237 2.27e-05 ***
## CAR_TYPE_Van 7.695e-01 4.707e-01 1.635 0.10208
## CAR_TYPE_z_SUV 2.138e+00 5.010e-01 4.269 1.97e-05 ***
## RED_CAR_no 3.775e-02 3.362e-01 0.112 0.91060
## REVOKED_Yes 2.606e-01 4.104e-01 0.635 0.52546
## URBANICITY_z_Highly.Rural..Rural -2.639e+00 4.331e-01 -6.092 1.11e-09 ***
## YOJ_NA -3.167e-01 4.339e-01 -0.730 0.46543
## INCOME_NA 2.194e-01 5.168e-01 0.425 0.67113
## CAR_AGE_NA 4.018e-01 5.061e-01 0.794 0.42723
## HOME_VAL_NA -3.001e-01 2.798e-01 -1.073 0.28350
## ageSquared 1.704e-03 1.166e-03 1.461 0.14391
## yojSquared 8.463e-03 6.549e-03 1.292 0.19631
## income_log -6.144e-02 2.766e-01 -0.222 0.82419
## homeval_log -1.596e+00 1.308e+00 -1.220 0.22248
## travtime_log -6.635e-02 6.109e-01 -0.109 0.91351
## bluebook_log -2.885e-01 4.817e-01 -0.599 0.54926
## carage_log -1.252e-01 4.661e-01 -0.269 0.78819
## oldclaim_log 6.345e-02 1.628e-01 0.390 0.69676
## clm_freq_log -4.419e-01 3.052e+00 -0.145 0.88487
## mvr_pts_log -7.202e-05 5.069e-01 0.000 0.99989
## tif_log 2.921e-01 5.083e-01 0.575 0.56547
## kidsdriv_log 8.856e-01 2.173e+00 0.408 0.68355
## homekids_log 1.497e+00 1.498e+00 0.999 0.31774
## inter 1.907e-02 3.112e-02 0.613 0.54015
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 743.59 on 640 degrees of freedom
## Residual deviance: 547.90 on 585 degrees of freedom

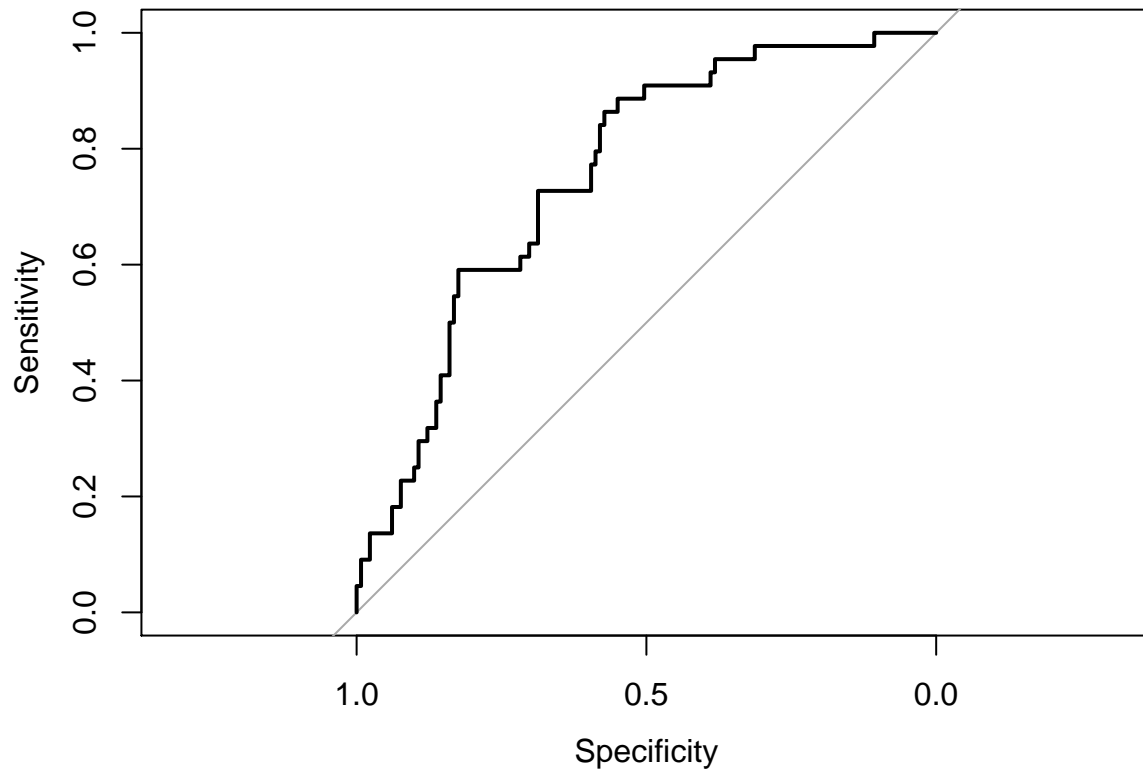
```

```

## (1 observation deleted due to missingness)
## AIC: 659.9
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 113  29
##           1  18  15
##
##           Accuracy : 0.7314
##           95% CI : (0.6593, 0.7955)
##           No Information Rate : 0.7486
##           P-Value [Acc > NIR] : 0.7320
##
##           Kappa : 0.2219
##
## Mcnemar's Test P-Value : 0.1447
##
##           Sensitivity : 0.8626
##           Specificity : 0.3409
##           Pos Pred Value : 0.7958
##           Neg Pred Value : 0.4545
##           Prevalence : 0.7486
##           Detection Rate : 0.6457
##           Detection Prevalence : 0.8114
##           Balanced Accuracy : 0.6018
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.755898681471201"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7559
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4296  -0.6545  -0.3576   0.5305   2.8736
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.481e+01  1.498e+01   0.989 0.322718
## KIDSDRIV      -9.098e-01  1.805e+00  -0.504 0.614267
## AGE          -1.512e-01  1.081e-01  -1.399 0.161921
## HOMEKIDS      -7.864e-01  6.830e-01  -1.151 0.249591
## YOJ           -2.227e-01  1.345e-01  -1.656 0.097771 .
## INCOME        -2.457e-05  1.432e-05  -1.716 0.086219 .
## HOME_VAL       1.410e-05  9.182e-06   1.536 0.124624
## TRAVTIME       3.505e-02  2.201e-02   1.592 0.111310
## BLUEBOOK       3.995e-05  3.918e-05   1.020 0.307903
```

```

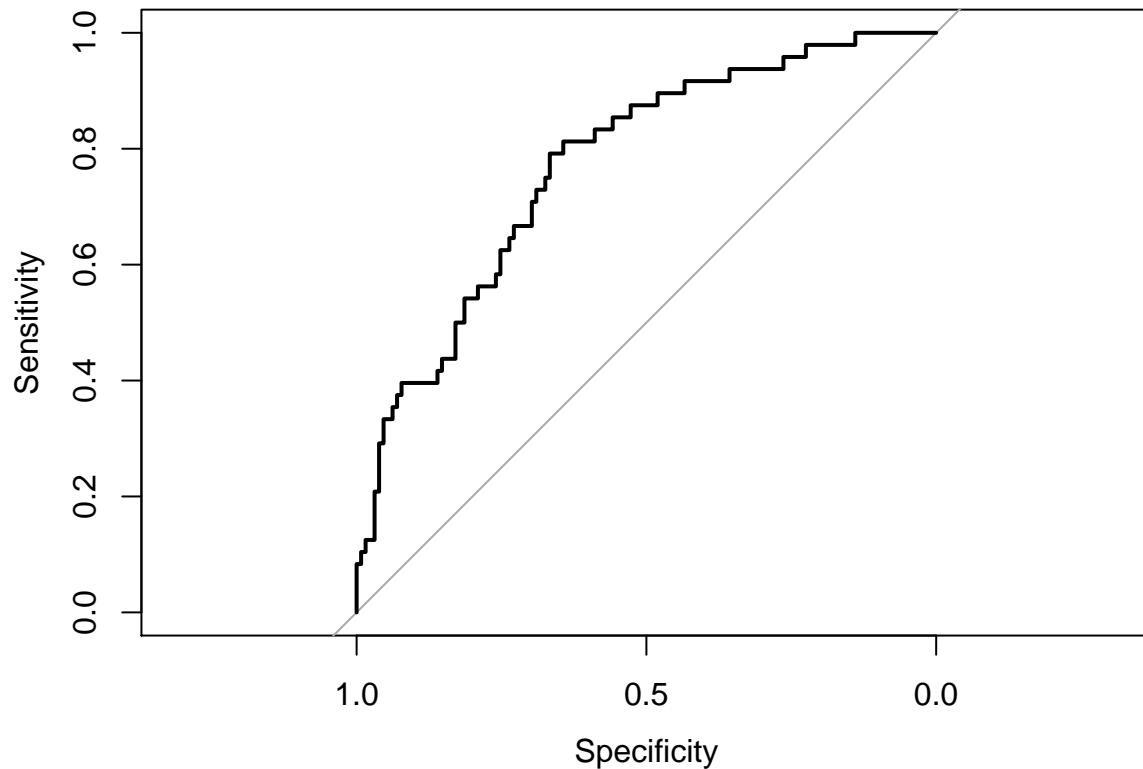
## TIF -8.150e-02 9.213e-02 -0.885 0.376374
## OLDCLAIM -1.138e-05 2.078e-05 -0.548 0.583796
## CLM_FREQ 9.588e-02 9.402e-01 0.102 0.918767
## MVR_PTS 1.170e-01 1.669e-01 0.701 0.483241
## CAR_AGE -1.055e-01 7.233e-02 -1.459 0.144655
## PARENT1_Yes 4.616e-01 4.806e-01 0.961 0.336764
## MSTATUS_Yes -6.740e-01 3.358e-01 -2.007 0.044756 *
## SEX_z_F -6.081e-01 4.559e-01 -1.334 0.182203
## EDUCATION_.High.School 5.643e-01 7.752e-01 0.728 0.466673
## EDUCATION_Bachelors 7.716e-02 6.411e-01 0.120 0.904196
## EDUCATION_Masters 9.695e-01 5.479e-01 1.769 0.076811 .
## EDUCATION_z_High.School 6.421e-01 6.979e-01 0.920 0.357564
## JOB_ -7.963e-01 6.611e-01 -1.205 0.228348
## JOB_Clerical -5.897e-01 4.830e-01 -1.221 0.222071
## JOB_Doctor -2.255e-01 9.012e-01 -0.250 0.802387
## JOB_Home.Maker -6.421e-01 6.862e-01 -0.936 0.349413
## JOB_Lawyer -7.449e-01 6.434e-01 -1.158 0.246932
## JOB_Manager -9.461e-01 4.648e-01 -2.036 0.041791 *
## JOB_Student -6.880e-01 7.734e-01 -0.890 0.373686
## JOB_z_Blue.Collar -6.664e-01 4.494e-01 -1.483 0.138104
## CAR_USE_Commercial 6.435e-01 3.587e-01 1.794 0.072770 .
## CAR_TYPE_Panel.Truck -2.269e-01 6.428e-01 -0.353 0.724106
## CAR_TYPE_Pickup 9.183e-01 4.211e-01 2.181 0.029197 *
## CAR_TYPE_Sports.Car 2.029e+00 5.359e-01 3.787 0.000152 ***
## CAR_TYPE_Van 4.682e-01 4.659e-01 1.005 0.314939
## CAR_TYPE_z_SUV 1.833e+00 4.747e-01 3.861 0.000113 ***
## RED_CAR_no -4.253e-01 3.372e-01 -1.261 0.207164
## REVOKED_Yes 3.070e-01 3.950e-01 0.777 0.437009
## URBANICITY_z_Highly.Rural..Rural -2.187e+00 4.221e-01 -5.180 2.22e-07 ***
## YOJ_NA -5.856e-02 4.357e-01 -0.134 0.893098
## INCOME_NA -1.859e-01 5.781e-01 -0.322 0.747732
## CAR_AGE_NA -2.013e-01 4.715e-01 -0.427 0.669366
## HOME_VAL_NA -2.355e-01 2.888e-01 -0.815 0.414898
## ageSquared 1.470e-03 1.156e-03 1.271 0.203558
## yojSquared 1.352e-02 6.923e-03 1.953 0.050837 .
## income_log 2.075e-02 3.495e-01 0.059 0.952649
## homeval_log -1.038e+00 1.426e+00 -0.728 0.466805
## travtime_log -4.080e-01 6.093e-01 -0.670 0.503088
## bluebook_log -5.820e-02 4.857e-01 -0.120 0.904607
## carage_log 3.478e-01 4.731e-01 0.735 0.462271
## oldclaim_log 1.194e-01 1.581e-01 0.755 0.450293
## clm_freq_log -3.227e-01 2.917e+00 -0.111 0.911912
## mvr_pts_log -3.372e-01 4.950e-01 -0.681 0.495809
## tif_log 1.106e-01 5.192e-01 0.213 0.831268
## kidsdriv_log 1.097e+00 2.326e+00 0.472 0.637160
## homekids_log 1.575e+00 1.471e+00 1.071 0.284350
## inter 2.466e-02 3.352e-02 0.736 0.461810
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 734.16 on 638 degrees of freedom
## Residual deviance: 545.65 on 583 degrees of freedom

```

```

## AIC: 657.65
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  29
##           1  15  19
##
##           Accuracy : 0.7514
##           95% CI : (0.681, 0.8132)
##           No Information Rate : 0.7288
##           P-Value [Acc > NIR] : 0.27990
##
##           Kappa : 0.3077
##
## Mcnemar's Test P-Value : 0.05002
##
##           Sensitivity : 0.8837
##           Specificity : 0.3958
##           Pos Pred Value : 0.7972
##           Neg Pred Value : 0.5588
##           Prevalence : 0.7288
##           Detection Rate : 0.6441
##           Detection Prevalence : 0.8079
##           Balanced Accuracy : 0.6398
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.771963824289406"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 129 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.772
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1036  -0.6877  -0.3619   0.5961   3.0636
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.455e+01  1.427e+01   1.020  0.30787
## KIDSDRIV      -1.536e+00  1.830e+00  -0.839  0.40132
## AGE           -1.456e-01  1.096e-01  -1.329  0.18388
## HOMEKIDS      -9.025e-01  6.735e-01  -1.340  0.18027
## YOJ           -6.873e-02  1.300e-01  -0.529  0.59713
## INCOME        -9.987e-06  1.277e-05  -0.782  0.43398
## HOME_VAL       6.963e-06  8.454e-06   0.824  0.41018
## TRAVTIME      2.753e-02  2.257e-02   1.220  0.22246
## BLUEBOOK      5.849e-05  3.789e-05   1.544  0.12268
```

```

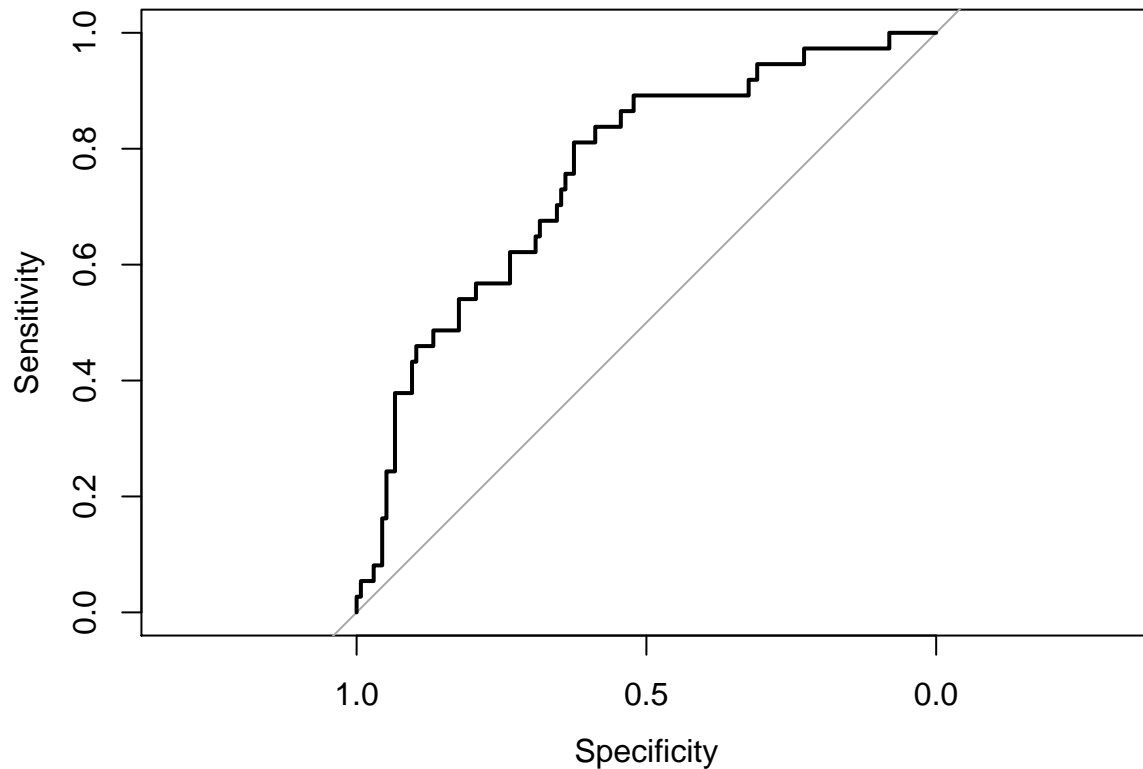
## TIF -9.476e-02 9.435e-02 -1.004 0.31524
## OLDCLAIM -8.179e-06 2.080e-05 -0.393 0.69411
## CLM_FREQ 7.018e-01 9.570e-01 0.733 0.46334
## MVR_PTS 6.423e-02 1.709e-01 0.376 0.70699
## CAR_AGE -2.910e-02 7.505e-02 -0.388 0.69816
## PARENT1_Yes 5.688e-01 4.721e-01 1.205 0.22829
## MSTATUS_Yes -6.693e-01 3.295e-01 -2.031 0.04225 *
## SEX_z_F -7.052e-01 4.461e-01 -1.581 0.11393
## EDUCATION_.High.School 1.393e-01 7.730e-01 0.180 0.85697
## EDUCATION_Bachelors -1.638e-01 6.370e-01 -0.257 0.79712
## EDUCATION_Masters 9.328e-01 5.457e-01 1.709 0.08739 .
## EDUCATION_z_High.School 5.301e-01 6.952e-01 0.762 0.44579
## JOB_ -9.253e-01 6.862e-01 -1.348 0.17750
## JOB_Clerical 2.216e-01 4.739e-01 0.468 0.64005
## JOB_Doctor -3.792e-01 9.105e-01 -0.417 0.67703
## JOB_Home.Maker 1.853e-01 6.850e-01 0.270 0.78678
## JOB_Lawyer -6.294e-01 6.490e-01 -0.970 0.33215
## JOB_Manager -8.057e-01 4.930e-01 -1.634 0.10223
## JOB_Student 1.761e-01 7.260e-01 0.243 0.80835
## JOB_z_Blue.Collar -3.599e-02 4.409e-01 -0.082 0.93495
## CAR_USE_Commercial 3.798e-01 3.427e-01 1.108 0.26772
## CAR_TYPE_Panel.Truck -9.784e-02 6.325e-01 -0.155 0.87706
## CAR_TYPE_Pickup 1.105e+00 4.025e-01 2.746 0.00604 **
## CAR_TYPE_Sports.Car 1.648e+00 5.270e-01 3.126 0.00177 **
## CAR_TYPE_Van 4.692e-01 4.560e-01 1.029 0.30352
## CAR_TYPE_z_SUV 1.507e+00 4.646e-01 3.244 0.00118 **
## RED_CAR_no -2.853e-01 3.294e-01 -0.866 0.38645
## REVOKED_Yes 4.478e-01 3.997e-01 1.120 0.26265
## URBANICITY_z_Highly.Rural..Rural -2.423e+00 4.241e-01 -5.715 1.1e-08 ***
## YOJ_NA -2.311e-01 4.303e-01 -0.537 0.59130
## INCOME_NA 6.758e-02 5.137e-01 0.132 0.89533
## CAR_AGE_NA -4.547e-01 4.652e-01 -0.977 0.32836
## HOME_VAL_NA -1.501e-01 2.800e-01 -0.536 0.59188
## ageSquared 1.589e-03 1.174e-03 1.353 0.17608
## yojSquared 2.830e-03 6.711e-03 0.422 0.67323
## income_log 6.450e-03 2.876e-01 0.022 0.98211
## homeval_log -7.330e-01 1.335e+00 -0.549 0.58303
## travtime_log -2.175e-02 6.197e-01 -0.035 0.97200
## bluebook_log -5.554e-01 4.657e-01 -1.193 0.23298
## carage_log -3.884e-02 4.780e-01 -0.081 0.93524
## oldclaim_log 1.544e-01 1.586e-01 0.974 0.33028
## clm_freq_log -1.984e+00 2.965e+00 -0.669 0.50342
## mvr_pts_log -4.086e-02 4.970e-01 -0.082 0.93447
## tif_log 2.222e-01 5.246e-01 0.423 0.67197
## kidsdriv_log -2.553e-01 2.173e+00 -0.118 0.90645
## homekids_log 1.957e+00 1.457e+00 1.343 0.17916
## inter 5.560e-02 3.606e-02 1.542 0.12311
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 758.65 on 642 degrees of freedom
## Residual deviance: 554.27 on 587 degrees of freedom

```

```

## AIC: 666.27
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 122  20
##           1  14  17
##
##           Accuracy : 0.8035
##           95% CI : (0.7363, 0.8599)
##           No Information Rate : 0.7861
##           P-Value [Acc > NIR] : 0.3266
##
##           Kappa : 0.3789
##
## Mcnemar's Test P-Value : 0.3912
##
##           Sensitivity : 0.8971
##           Specificity : 0.4595
##           Pos Pred Value : 0.8592
##           Neg Pred Value : 0.5484
##           Prevalence : 0.7861
##           Detection Rate : 0.7052
##           Detection Prevalence : 0.8208
##           Balanced Accuracy : 0.6783
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.756558028616852"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 136 controls (dfPred_raw$class 0) < 37 cases (dfPred_raw$class 1).
## Area under the curve: 0.7566
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1090  -0.6852  -0.3618   0.5403   2.9536
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.949e+01  1.544e+01   1.262  0.2069
## KIDSDRIV      -6.758e-01  1.750e+00  -0.386  0.6993
## AGE           -1.902e-01  1.032e-01  -1.842  0.0655 .
## HOMEKIDS      -7.236e-01  6.401e-01  -1.131  0.2582
## YOJ           -1.011e-01  1.373e-01  -0.736  0.4617
## INCOME        -2.715e-05  1.349e-05  -2.013  0.0441 *
## HOME_VAL       1.563e-05  9.106e-06   1.716  0.0861 .
## TRAVTIME       2.315e-02  2.196e-02   1.054  0.2919
## BLUEBOOK       6.829e-05  4.103e-05   1.664  0.0960 .
```

```

## TIF 5.908e-03 9.516e-02 0.062 0.9505
## OLDCLAIM 2.559e-06 2.028e-05 0.126 0.8996
## CLM_FREQ -6.038e-02 9.367e-01 -0.064 0.9486
## MVR_PTS 1.077e-01 1.687e-01 0.639 0.5230
## CAR_AGE -4.328e-02 6.765e-02 -0.640 0.5224
## PARENT1_Yes 1.724e-01 4.657e-01 0.370 0.7113
## MSTATUS_Yes -6.115e-01 3.415e-01 -1.791 0.0733 .
## SEX_z_F -7.768e-01 4.552e-01 -1.707 0.0879 .
## EDUCATION_.High.School 5.670e-01 7.636e-01 0.743 0.4578
## EDUCATION_Bachelors 1.555e-01 6.501e-01 0.239 0.8109
## EDUCATION_Masters 1.153e+00 5.478e-01 2.105 0.0353 *
## EDUCATION_z_High.School 4.482e-01 6.962e-01 0.644 0.5197
## JOB_ -9.504e-01 6.544e-01 -1.452 0.1464
## JOB_Clerical -3.161e-01 4.971e-01 -0.636 0.5248
## JOB_Doctor 1.473e-01 9.363e-01 0.157 0.8750
## JOB_Home.Maker 1.391e-01 6.641e-01 0.209 0.8341
## JOB_Lawyer -6.899e-01 6.511e-01 -1.060 0.2893
## JOB_Manager -1.006e+00 4.918e-01 -2.046 0.0408 *
## JOB_Student -2.976e-01 7.203e-01 -0.413 0.6795
## JOB_z_Blue.Collar -2.965e-01 4.605e-01 -0.644 0.5197
## CAR_USE_Commercial 4.572e-01 3.409e-01 1.341 0.1799
## CAR_TYPE_Panel.Truck 1.695e-01 6.137e-01 0.276 0.7824
## CAR_TYPE_Pickup 1.352e+00 4.141e-01 3.264 0.0011 **
## CAR_TYPE_Sports.Car 2.299e+00 5.455e-01 4.214 2.51e-05 ***
## CAR_TYPE_Van 7.900e-01 4.667e-01 1.693 0.0905 .
## CAR_TYPE_z_SUV 2.070e+00 4.894e-01 4.231 2.33e-05 ***
## RED_CAR_no -4.211e-01 3.360e-01 -1.253 0.2101
## REVOKED_Yes 5.478e-01 3.869e-01 1.416 0.1569
## URBANICITY_z_Highly.Rural..Rural -2.459e+00 4.416e-01 -5.569 2.56e-08 ***
## YOJ_NA 2.299e-02 4.232e-01 0.054 0.9567
## INCOME_NA 2.059e-01 6.086e-01 0.338 0.7352
## CAR_AGE_NA -1.834e-01 4.869e-01 -0.377 0.7065
## HOME_VAL_NA -2.905e-01 2.896e-01 -1.003 0.3157
## ageSquared 1.996e-03 1.117e-03 1.787 0.0740 .
## yojSquared 5.736e-03 6.901e-03 0.831 0.4059
## income_log 2.110e-01 3.137e-01 0.673 0.5012
## homeval_log -1.473e+00 1.476e+00 -0.998 0.3181
## travtime_log -1.363e-01 6.112e-01 -0.223 0.8235
## bluebook_log -3.071e-01 5.116e-01 -0.600 0.5484
## carage_log -3.979e-02 4.478e-01 -0.089 0.9292
## oldclaim_log 1.654e-02 1.586e-01 0.104 0.9169
## clm_freq_log 3.303e-01 2.913e+00 0.113 0.9097
## mvr_pts_log -1.552e-01 4.942e-01 -0.314 0.7535
## tif_log -4.146e-01 5.293e-01 -0.783 0.4335
## kidsdriv_log 6.216e-01 2.192e+00 0.284 0.7767
## homekids_log 1.795e+00 1.391e+00 1.290 0.1970
## inter 2.202e-02 3.050e-02 0.722 0.4703
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 735.37 on 640 degrees of freedom
## Residual deviance: 552.99 on 585 degrees of freedom

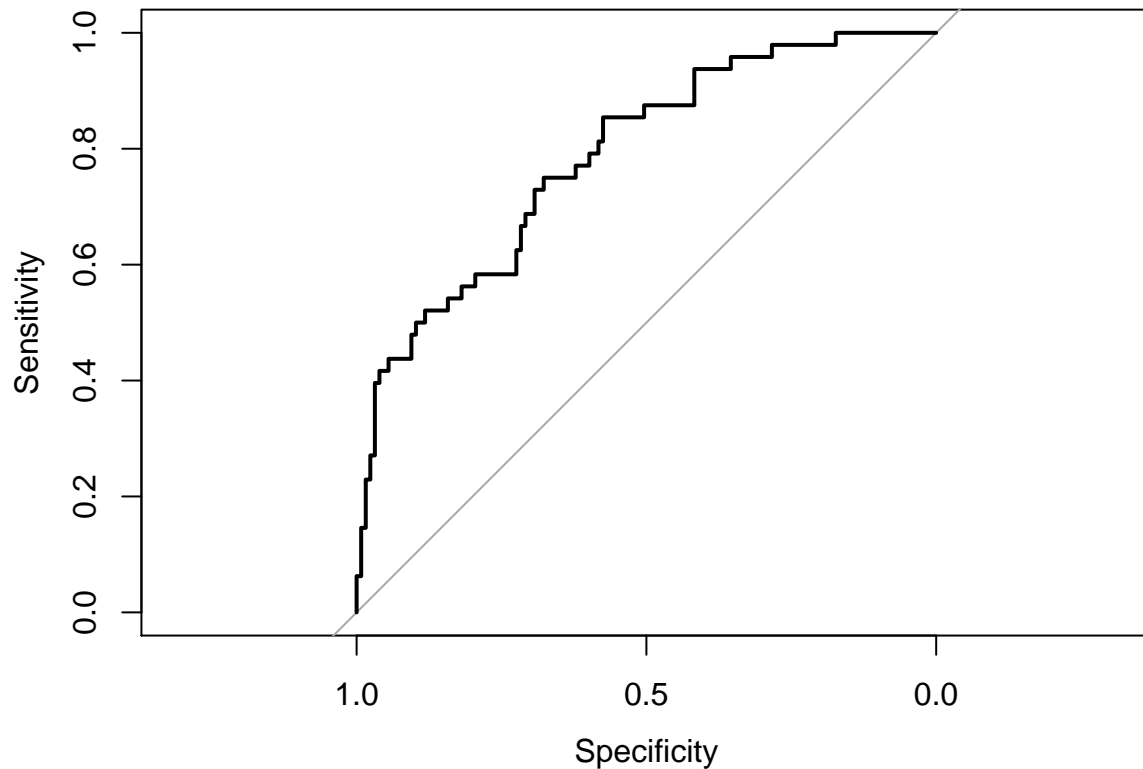
```



```

## AIC: 664.99
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 115  27
##           1  12  21
##
##           Accuracy : 0.7771
##           95% CI : (0.7082, 0.8365)
##           No Information Rate : 0.7257
##           P-Value [Acc > NIR] : 0.07273
##
##           Kappa : 0.3799
##
## Mcnemar's Test P-Value : 0.02497
##
##           Sensitivity : 0.9055
##           Specificity : 0.4375
##           Pos Pred Value : 0.8099
##           Neg Pred Value : 0.6364
##           Prevalence : 0.7257
##           Detection Rate : 0.6571
##           Detection Prevalence : 0.8114
##           Balanced Accuracy : 0.6715
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.78740157480315"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 127 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.7874
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8859  -0.6386  -0.3367   0.4755   3.0462
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    4.307e+01  1.478e+01   2.914 0.003570 **
## KIDSDRIV      -1.000e+00  1.670e+00  -0.599 0.549223
## AGE           -8.703e-02  1.070e-01  -0.813 0.416085
## HOMEKIDS      -5.243e-01  7.067e-01  -0.742 0.458101
## YOJ           -1.774e-01  1.308e-01  -1.357 0.174875
## INCOME        -4.092e-05  1.379e-05  -2.967 0.003011 **
## HOME_VAL       2.765e-05  9.195e-06   3.007 0.002641 **
## TRAVTIME       4.067e-02  2.112e-02   1.926 0.054109 .
## BLUEBOOK       7.120e-05  3.963e-05   1.797 0.072372 .
```

```

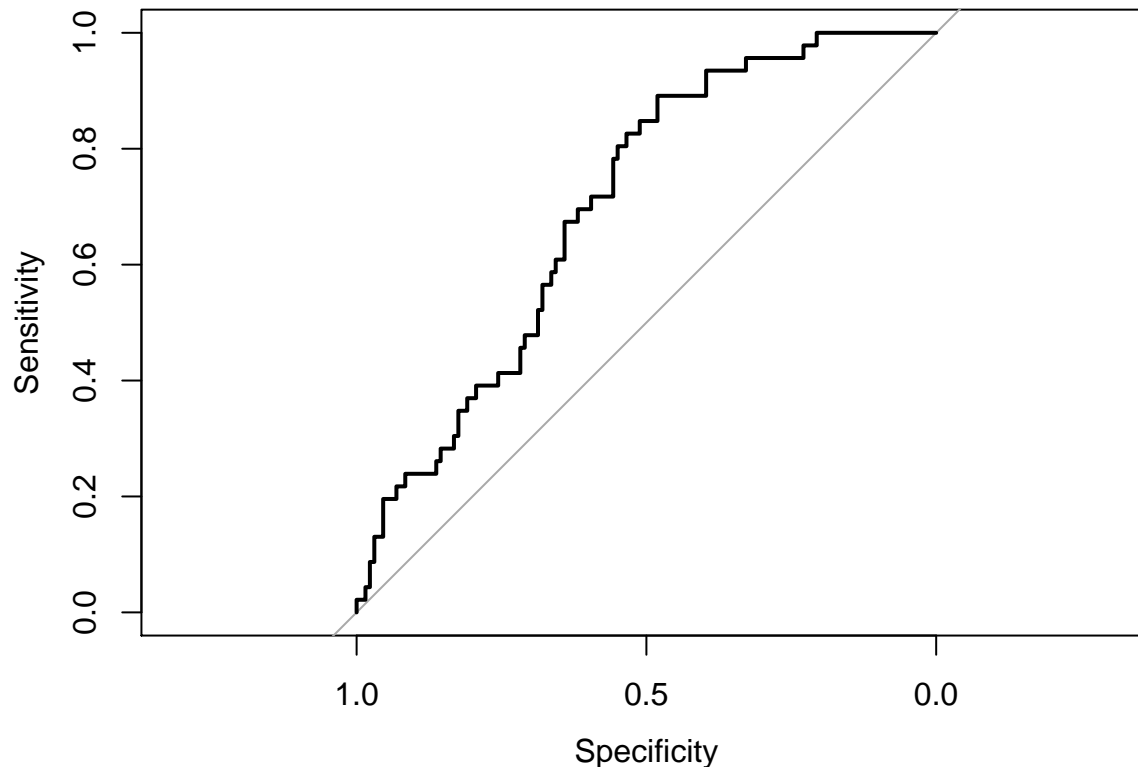
## TIF -9.281e-02 9.160e-02 -1.013 0.310973
## OLDCLAIM 1.285e-05 2.216e-05 0.580 0.561938
## CLM_FREQ 2.846e-01 9.690e-01 0.294 0.769017
## MVRPTS 2.457e-01 1.837e-01 1.338 0.181050
## CAR_AGE -3.762e-02 6.984e-02 -0.539 0.590094
## PARENT1_Yes 1.061e+00 4.825e-01 2.199 0.027869 *
## MSTATUS_Yes -7.773e-01 3.386e-01 -2.296 0.021688 *
## SEX_z_F -1.197e+00 4.729e-01 -2.531 0.011378 *
## EDUCATION_High.School 2.051e-01 8.031e-01 0.255 0.798417
## EDUCATION_Bachelors -1.161e-01 6.661e-01 -0.174 0.861692
## EDUCATION_Masters 8.769e-01 5.923e-01 1.481 0.138726
## EDUCATION_z_High.School 4.073e-01 7.237e-01 0.563 0.573621
## JOB_ -8.316e-01 6.872e-01 -1.210 0.226176
## JOB_Clerical -2.755e-01 4.869e-01 -0.566 0.571466
## JOB_Doctor -1.912e-01 9.268e-01 -0.206 0.836575
## JOB_Home.Maker -7.953e-01 6.577e-01 -1.209 0.226579
## JOB_Lawyer -8.526e-01 6.667e-01 -1.279 0.200980
## JOB_Manager -5.644e-01 4.839e-01 -1.166 0.243497
## JOB_Student -8.774e-01 7.244e-01 -1.211 0.225857
## JOB_z_Blue.Collar 1.349e-02 4.515e-01 0.030 0.976169
## CAR_USE_Commercial 2.725e-01 3.506e-01 0.777 0.437110
## CAR_TYPE_Panel.Truck 3.867e-01 6.394e-01 0.605 0.545348
## CAR_TYPE_Pickup 1.292e+00 4.070e-01 3.175 0.001501 **
## CAR_TYPE_Sports.Car 2.033e+00 5.581e-01 3.642 0.000271 ***
## CAR_TYPE_Van 4.803e-01 4.736e-01 1.014 0.310438
## CAR_TYPE_z_SUV 2.106e+00 4.963e-01 4.243 2.20e-05 ***
## RED_CAR_no 1.612e-01 3.337e-01 0.483 0.629001
## REVOKED_Yes -9.600e-02 4.841e-01 -0.198 0.842790
## URBANICITY_z_Highly.Rural..Rural -2.169e+00 4.013e-01 -5.404 6.52e-08 ***
## YOJ_NA -3.445e-01 4.429e-01 -0.778 0.436639
## INCOME_NA 1.549e-01 5.416e-01 0.286 0.774941
## CAR_AGE_NA 3.673e-01 5.148e-01 0.713 0.475538
## HOME_VAL_NA 3.461e-01 2.920e-01 1.185 0.236023
## ageSquared 8.772e-04 1.159e-03 0.757 0.448981
## yojSquared 9.279e-03 6.790e-03 1.367 0.171768
## income_log 4.630e-01 2.919e-01 1.586 0.112684
## homeval_log -3.772e+00 1.403e+00 -2.688 0.007179 **
## travtime_log -5.713e-01 5.924e-01 -0.964 0.334854
## bluebook_log -5.588e-01 4.755e-01 -1.175 0.239923
## carage_log 9.252e-02 4.581e-01 0.202 0.839942
## oldclaim_log 6.432e-02 1.607e-01 0.400 0.689062
## clm_freq_log -4.970e-01 2.973e+00 -0.167 0.867243
## mvr_pts_log -4.463e-01 5.332e-01 -0.837 0.402561
## tif_log 2.982e-01 5.219e-01 0.571 0.567689
## kidsdriv_log 1.705e+00 2.150e+00 0.793 0.427610
## homekids_log 1.006e+00 1.522e+00 0.661 0.508727
## inter 1.783e-02 3.046e-02 0.585 0.558236
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 738.28 on 638 degrees of freedom
## Residual deviance: 529.73 on 583 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 641.73
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  32
##           1  23  14
##
##           Accuracy : 0.6893
##           95% CI : (0.6155, 0.7566)
##           No Information Rate : 0.7401
##           P-Value [Acc > NIR] : 0.9461
##
##           Kappa : 0.1375
##
## Mcnemar's Test P-Value : 0.2807
##
##           Sensitivity : 0.8244
##           Specificity : 0.3043
##           Pos Pred Value : 0.7714
##           Neg Pred Value : 0.3784
##           Prevalence : 0.7401
##           Detection Rate : 0.6102
##           Detection Prevalence : 0.7910
##           Balanced Accuracy : 0.5644
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.701460338533024"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7015
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1922  -0.6645  -0.3302   0.5860   2.7997
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.767e+01  1.487e+01   1.861 0.062729 .
## KIDSDRIV      -1.392e+00  1.798e+00  -0.774 0.438849
## AGE           -5.510e-02  1.027e-01  -0.536 0.591767
## HOMEKIDS      -3.244e-01  6.990e-01  -0.464 0.642544
## YOJ           -1.676e-01  1.263e-01  -1.328 0.184274
## INCOME        -2.897e-05  1.358e-05  -2.132 0.032967 *
## HOME_VAL       1.741e-05  9.155e-06   1.901 0.057250 .
## TRAVTIME       1.840e-02  2.150e-02   0.856 0.391985
## BLUEBOOK       6.137e-05  3.978e-05   1.543 0.122913
```

```

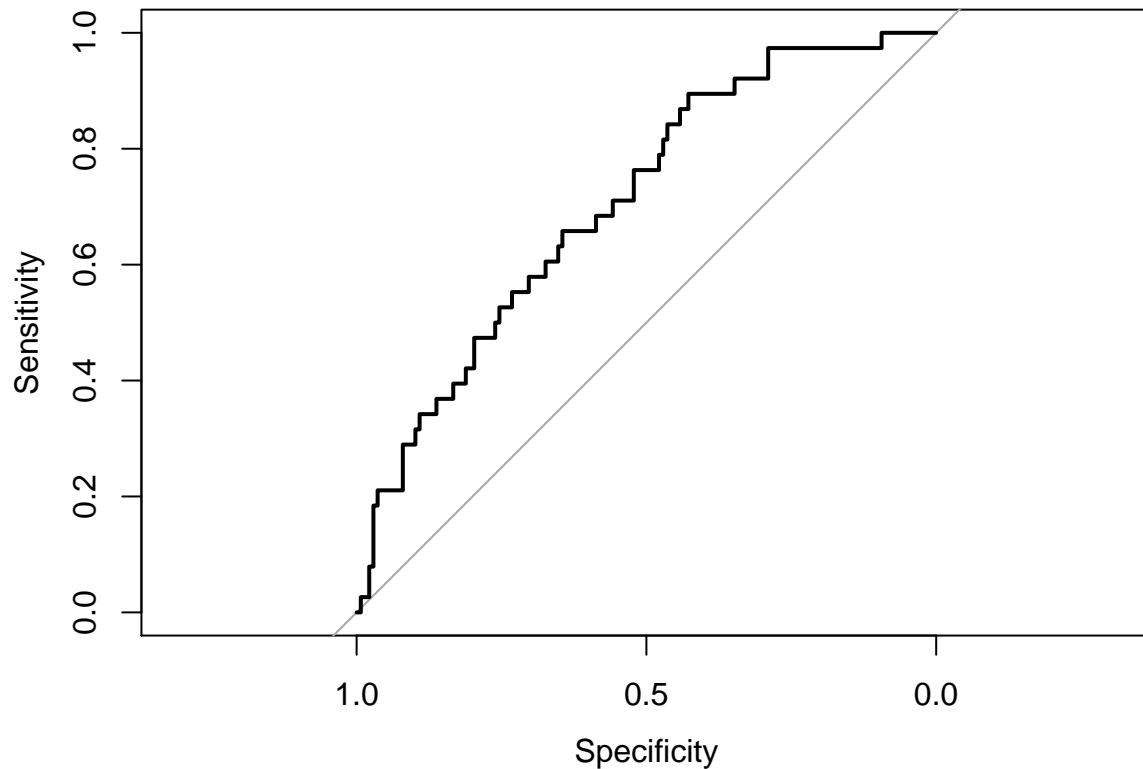
## TIF -1.028e-02 8.994e-02 -0.114 0.908997
## OLDCLAIM -3.625e-06 2.234e-05 -0.162 0.871124
## CLM_FREQ 1.777e-01 1.100e+00 0.162 0.871678
## MVRPTS 8.355e-02 1.729e-01 0.483 0.629041
## CAR_AGE 1.076e-02 7.452e-02 0.144 0.885216
## PARENT1_Yes 1.094e+00 4.744e-01 2.306 0.021096 *
## MSTATUS_Yes -6.298e-01 3.318e-01 -1.898 0.057678 .
## SEX_z_F -4.426e-01 4.665e-01 -0.949 0.342728
## EDUCATION_.High.School 2.006e-01 8.434e-01 0.238 0.812013
## EDUCATION_Bachelors 1.333e-01 7.025e-01 0.190 0.849493
## EDUCATION_Masters 1.145e+00 6.180e-01 1.853 0.063843 .
## EDUCATION_z_High.School 5.534e-01 7.620e-01 0.726 0.467665
## JOB_ -1.547e+00 6.956e-01 -2.224 0.026124 *
## JOB_Clerical -7.491e-01 4.787e-01 -1.565 0.117667
## JOB_Doctor -2.468e-01 9.565e-01 -0.258 0.796353
## JOB_Home.Maker -3.737e-01 6.340e-01 -0.589 0.555533
## JOB_Lawyer -1.003e+00 6.308e-01 -1.589 0.111979
## JOB_Manager -1.085e+00 4.802e-01 -2.260 0.023796 *
## JOB_Student -1.070e+00 7.052e-01 -1.517 0.129299
## JOB_z_Blue.Collar -7.318e-01 4.651e-01 -1.573 0.115641
## CAR_USE_Commercial 7.207e-01 3.498e-01 2.060 0.039352 *
## CAR_TYPE_Panel.Truck 3.652e-01 6.231e-01 0.586 0.557812
## CAR_TYPE_Pickup 9.119e-01 4.024e-01 2.266 0.023458 *
## CAR_TYPE_Sports.Car 1.540e+00 5.461e-01 2.821 0.004791 **
## CAR_TYPE_Van 3.484e-01 4.703e-01 0.741 0.458797
## CAR_TYPE_z_SUV 1.662e+00 4.635e-01 3.587 0.000335 ***
## RED_CAR_no -3.359e-01 3.382e-01 -0.993 0.320586
## REVOKED_Yes 6.464e-01 4.069e-01 1.589 0.112126
## URBANICITY_z_Highly.Rural..Rural -2.188e+00 4.015e-01 -5.450 5.04e-08 ***
## YOJ_NA -1.085e-01 4.603e-01 -0.236 0.813645
## INCOME_NA 2.275e-01 5.313e-01 0.428 0.668482
## CAR_AGE_NA -2.379e-01 4.928e-01 -0.483 0.629322
## HOME_VAL_NA -1.755e-01 2.828e-01 -0.621 0.534817
## ageSquared 6.913e-04 1.108e-03 0.624 0.532817
## yojSquared 9.632e-03 6.512e-03 1.479 0.139113
## income_log 2.252e-01 2.814e-01 0.800 0.423542
## homeval_log -2.334e+00 1.403e+00 -1.664 0.096134 .
## travtime_log 2.210e-01 6.378e-01 0.346 0.729030
## bluebook_log -4.718e-01 4.954e-01 -0.952 0.340876
## carage_log -4.047e-01 4.748e-01 -0.852 0.394052
## oldclaim_log 5.577e-02 1.772e-01 0.315 0.752975
## clm_freq_log -1.442e-01 3.325e+00 -0.043 0.965409
## mvr_pts_log -5.580e-02 5.026e-01 -0.111 0.911592
## tif_log -3.462e-01 5.100e-01 -0.679 0.497292
## kidsdriv_log 2.239e+00 2.192e+00 1.022 0.306919
## homekids_log 5.503e-01 1.496e+00 0.368 0.713064
## inter 1.905e-02 3.242e-02 0.588 0.556694
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 754.79 on 639 degrees of freedom
## Residual deviance: 542.49 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 654.49
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 121  25
##           1  17  13
##
##           Accuracy : 0.7614
##           95% CI : (0.6914, 0.8223)
##           No Information Rate : 0.7841
##           P-Value [Acc > NIR] : 0.7967
##
##           Kappa : 0.237
##
## Mcnemar's Test P-Value : 0.2801
##
##           Sensitivity : 0.8768
##           Specificity : 0.3421
##           Pos Pred Value : 0.8288
##           Neg Pred Value : 0.4333
##           Prevalence : 0.7841
##           Detection Rate : 0.6875
##           Detection Prevalence : 0.8295
##           Balanced Accuracy : 0.6095
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.706903127383677"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 38 cases (dfPred_raw$class 1).
## Area under the curve: 0.7069
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1806  -0.6352  -0.3309   0.3591   3.0508
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.437e+01  1.525e+01   1.598  0.110112
## KIDSDRIV       7.672e-01  1.814e+00   0.423  0.672270
## AGE          -1.256e-01  1.078e-01  -1.165  0.244035
## HOMEKIDS      -1.993e+00  7.952e-01  -2.506  0.012203 *
## YOJ           -2.440e-01  1.324e-01  -1.843  0.065363 .
## INCOME        -1.452e-05  1.353e-05  -1.074  0.283033
## HOME_VAL       1.224e-05  9.020e-06   1.357  0.174889
## TRAVTIME       3.981e-02  2.126e-02   1.873  0.061064 .
## BLUEBOOK       7.366e-05  3.988e-05   1.847  0.064747 .
```



```

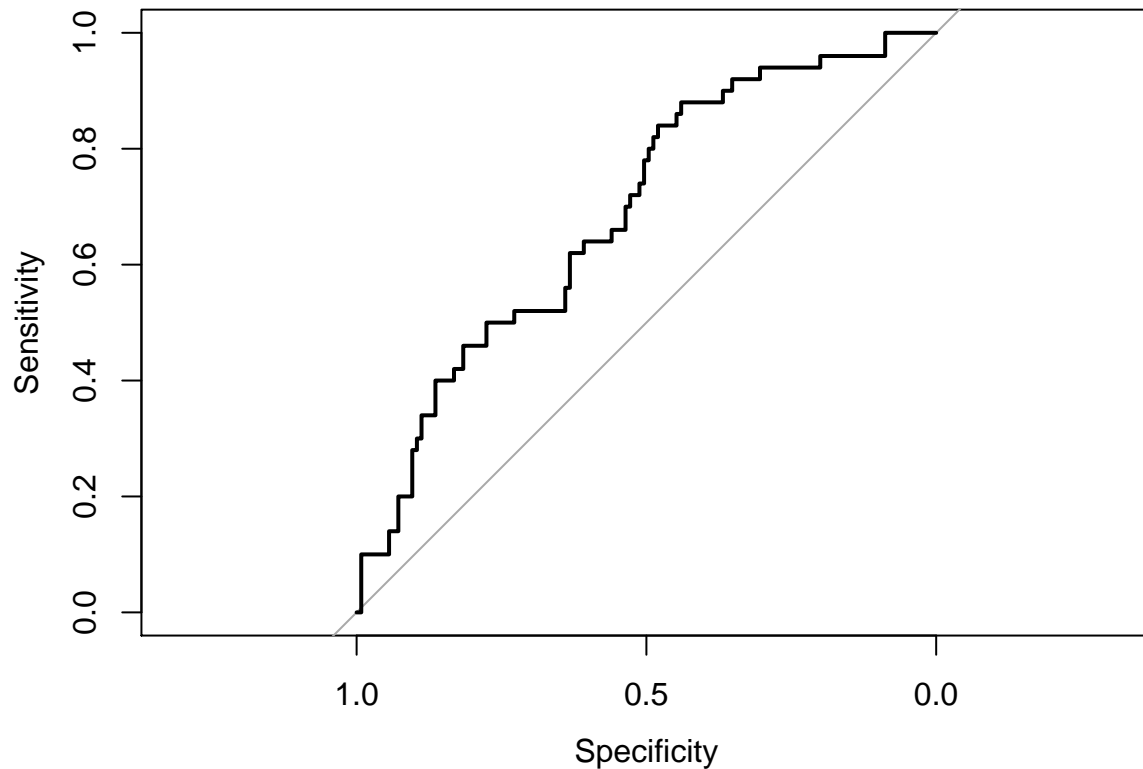
## TIF -7.219e-02 9.173e-02 -0.787 0.431289
## OLDCLAIM 1.522e-06 2.249e-05 0.068 0.946043
## CLM_FREQ 8.449e-02 1.011e+00 0.084 0.933391
## MVRPTS 8.451e-02 1.753e-01 0.482 0.629680
## CAR_AGE -9.223e-02 7.649e-02 -1.206 0.227915
## PARENT1_Yes 7.541e-01 4.982e-01 1.514 0.130123
## MSTATUS_Yes -7.352e-01 3.360e-01 -2.188 0.028663 *
## SEX_z_F -8.801e-01 5.045e-01 -1.744 0.081114 .
## EDUCATION_.High.School 5.604e-01 7.996e-01 0.701 0.483381
## EDUCATION_Bachelors -4.562e-02 6.587e-01 -0.069 0.944778
## EDUCATION_Masters 7.924e-01 5.453e-01 1.453 0.146244
## EDUCATION_z_High.School 8.320e-01 7.187e-01 1.158 0.246963
## JOB_ -6.328e-01 7.190e-01 -0.880 0.378784
## JOB_Clerical -4.900e-01 4.992e-01 -0.982 0.326326
## JOB_Doctor -2.262e-01 8.738e-01 -0.259 0.795758
## JOB_Home.Maker -6.902e-01 7.044e-01 -0.980 0.327196
## JOB_Lawyer -3.352e-01 6.592e-01 -0.508 0.611125
## JOB_Manager -9.296e-01 4.934e-01 -1.884 0.059585 .
## JOB_Student -6.146e-01 7.744e-01 -0.794 0.427427
## JOB_z_Blue.Collar -4.895e-01 4.707e-01 -1.040 0.298358
## CAR_USE_Commercial 5.772e-01 3.638e-01 1.586 0.112642
## CAR_TYPE_Panel.Truck -5.053e-01 6.637e-01 -0.761 0.446482
## CAR_TYPE_Pickup 1.001e+00 4.203e-01 2.381 0.017250 *
## CAR_TYPE_Sports.Car 2.076e+00 5.588e-01 3.715 0.000203 ***
## CAR_TYPE_Van 4.314e-01 4.683e-01 0.921 0.356906
## CAR_TYPE_z_SUV 1.971e+00 5.121e-01 3.849 0.000119 ***
## RED_CAR_no -3.694e-01 3.448e-01 -1.071 0.284077
## REVOKED_Yes 4.287e-03 4.415e-01 0.010 0.992254
## URBANICITY_z_Highly.Rural..Rural -2.532e+00 4.528e-01 -5.592 2.24e-08 ***
## YOJ_NA -3.346e-01 4.208e-01 -0.795 0.426565
## INCOME_NA 5.119e-02 5.265e-01 0.097 0.922552
## CAR_AGE_NA 2.549e-01 5.349e-01 0.476 0.633746
## HOME_VAL_NA -3.667e-01 2.921e-01 -1.255 0.209313
## ageSquared 1.325e-03 1.153e-03 1.149 0.250471
## yojSquared 1.404e-02 6.918e-03 2.030 0.042340 *
## income_log -2.338e-01 3.033e-01 -0.771 0.440696
## homeval_log -1.323e+00 1.435e+00 -0.922 0.356702
## travtime_log -4.601e-01 5.918e-01 -0.777 0.436882
## bluebook_log -6.145e-01 4.958e-01 -1.239 0.215204
## carage_log 4.424e-01 4.880e-01 0.906 0.364689
## oldclaim_log 1.340e-01 1.645e-01 0.815 0.415220
## clm_freq_log -4.994e-01 3.091e+00 -0.162 0.871657
## mvr_pts_log -2.249e-01 5.134e-01 -0.438 0.661374
## tif_log 7.599e-02 5.242e-01 0.145 0.884737
## kidsdriv_log -7.691e-01 2.321e+00 -0.331 0.740375
## homekids_log 3.715e+00 1.662e+00 2.235 0.025386 *
## inter 7.261e-03 3.472e-02 0.209 0.834341
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 731.16 on 640 degrees of freedom
## Residual deviance: 521.79 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 633.79
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  30
##           1  17  20
##
##           Accuracy : 0.7314
##           95% CI : (0.6593, 0.7955)
##           No Information Rate : 0.7143
##           P-Value [Acc > NIR] : 0.34155
##
##           Kappa : 0.2863
##
## Mcnemar's Test P-Value : 0.08005
##
##           Sensitivity : 0.8640
##           Specificity : 0.4000
##           Pos Pred Value : 0.7826
##           Neg Pred Value : 0.5405
##           Prevalence : 0.7143
##           Detection Rate : 0.6171
##           Detection Prevalence : 0.7886
##           Balanced Accuracy : 0.6320
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.68944"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 125 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.6894
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0928  -0.6931  -0.3679   0.6506   2.9180
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.010e+01  1.383e+01   2.176  0.02953 *
## KIDSDRIV     -3.339e-01  1.834e+00  -0.182  0.85554
## AGE          -2.033e-01  1.043e-01  -1.950  0.05118 .
## HOMEKIDS     -1.085e+00  7.073e-01  -1.535  0.12489
## YOJ          -1.956e-01  1.263e-01  -1.549  0.12127
## INCOME       -2.138e-05  1.293e-05  -1.653  0.09832 .
## HOME_VAL      1.563e-05  8.585e-06   1.820  0.06875 .
## TRAVTIME      3.393e-02  2.128e-02   1.595  0.11074
## BLUEBOOK      4.547e-05  3.861e-05   1.178  0.23889
```

```

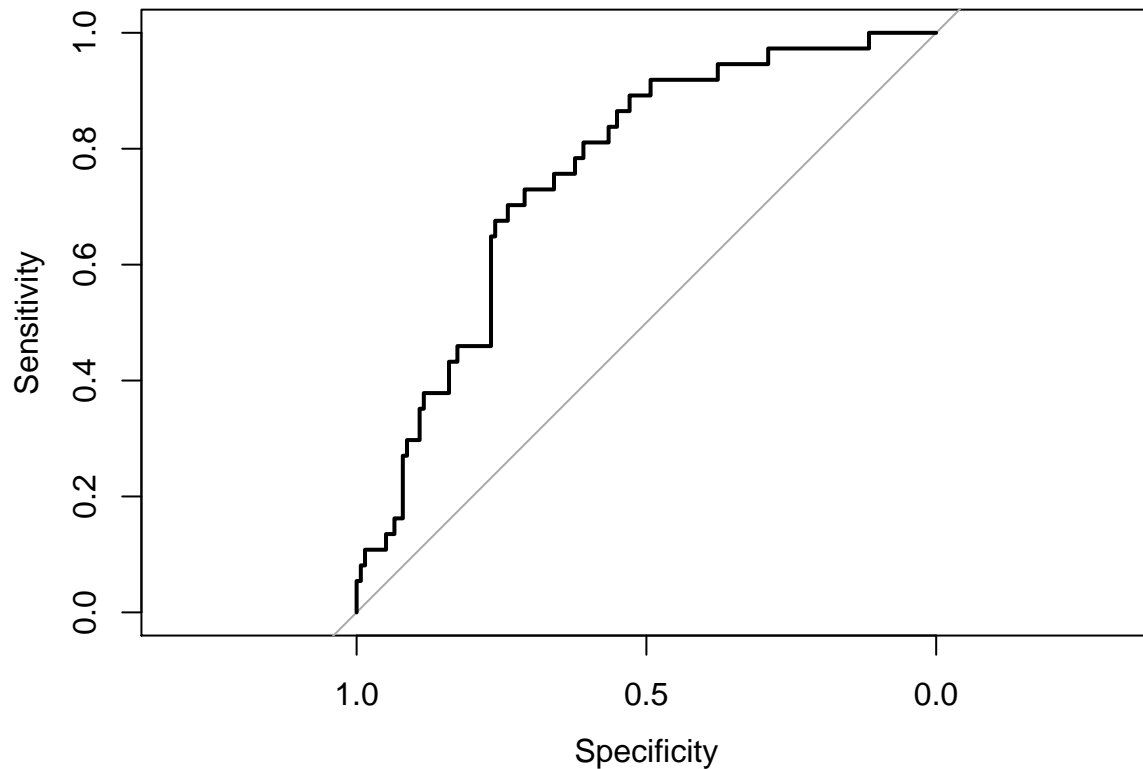
## TIF -2.000e-02 8.833e-02 -0.226 0.82089
## OLDCLAIM 5.353e-07 2.147e-05 0.025 0.98011
## CLM_FREQ 2.524e-01 9.791e-01 0.258 0.79661
## MVR_PTS 6.596e-02 1.709e-01 0.386 0.69953
## CAR_AGE 5.501e-02 7.131e-02 0.771 0.44046
## PARENT1_Yes 4.228e-01 4.557e-01 0.928 0.35354
## MSTATUS_Yes -7.531e-01 3.278e-01 -2.298 0.02158 *
## SEX_z_F -4.924e-01 4.583e-01 -1.074 0.28262
## EDUCATION_.High.School 2.981e-01 7.889e-01 0.378 0.70551
## EDUCATION_Bachelors 4.475e-02 6.679e-01 0.067 0.94658
## EDUCATION_Masters 6.794e-01 5.553e-01 1.224 0.22113
## EDUCATION_z_High.School 5.674e-01 7.195e-01 0.789 0.43039
## JOB_ -1.342e+00 7.197e-01 -1.865 0.06214 .
## JOB_Clerical -5.134e-01 4.744e-01 -1.082 0.27922
## JOB_Doctor -1.446e-01 9.546e-01 -0.151 0.87963
## JOB_Home.Maker -5.363e-01 6.458e-01 -0.830 0.40628
## JOB_Lawyer -9.065e-01 6.519e-01 -1.391 0.16435
## JOB_Manager -1.125e+00 4.941e-01 -2.278 0.02274 *
## JOB_Student -1.677e+00 7.180e-01 -2.336 0.01948 *
## JOB_z_Blue.Collar -3.403e-01 4.514e-01 -0.754 0.45091
## CAR_USE_Commercial 5.047e-01 3.508e-01 1.439 0.15029
## CAR_TYPE_Panel.Truck 4.387e-01 5.971e-01 0.735 0.46246
## CAR_TYPE_Pickup 1.281e+00 3.934e-01 3.257 0.00113 **
## CAR_TYPE_Sports.Car 1.668e+00 5.197e-01 3.209 0.00133 **
## CAR_TYPE_Van 5.857e-01 4.579e-01 1.279 0.20087
## CAR_TYPE_z_SUV 1.862e+00 4.635e-01 4.017 5.91e-05 ***
## RED_CAR_no -3.734e-01 3.275e-01 -1.140 0.25412
## REVOKED_Yes 4.568e-01 3.955e-01 1.155 0.24815
## URBANICITY_z_Highly.Rural..Rural -2.456e+00 4.120e-01 -5.960 2.52e-09 ***
## YOJ_NA -4.156e-01 4.140e-01 -1.004 0.31550
## INCOME_NA 5.845e-01 5.309e-01 1.101 0.27089
## CAR_AGE_NA 9.133e-03 4.811e-01 0.019 0.98485
## HOME_VAL_NA -2.384e-01 2.778e-01 -0.858 0.39074
## ageSquared 2.119e-03 1.140e-03 1.859 0.06305 .
## yojSquared 1.001e-02 6.494e-03 1.541 0.12321
## income_log 1.054e-01 2.694e-01 0.391 0.69555
## homeval_log -2.207e+00 1.301e+00 -1.697 0.08973 .
## travtime_log -3.307e-01 6.004e-01 -0.551 0.58170
## bluebook_log -1.680e-01 4.657e-01 -0.361 0.71825
## carage_log -6.229e-01 4.572e-01 -1.363 0.17299
## oldclaim_log 2.942e-02 1.611e-01 0.183 0.85509
## clm_freq_log -5.080e-01 2.990e+00 -0.170 0.86508
## mvr_pts_log -1.221e-01 5.003e-01 -0.244 0.80713
## tif_log -2.441e-01 5.033e-01 -0.485 0.62770
## kidsdriv_log 1.061e+00 2.316e+00 0.458 0.64680
## homekids_log 2.219e+00 1.506e+00 1.473 0.14064
## inter 1.174e-02 3.244e-02 0.362 0.71753
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 757.35 on 640 degrees of freedom
## Residual deviance: 560.17 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 672.17
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 123  24
##           1  15  13
##
##           Accuracy : 0.7771
##           95% CI : (0.7082, 0.8365)
##           No Information Rate : 0.7886
##           P-Value [Acc > NIR] : 0.6832
##
##           Kappa : 0.2664
##
## Mcnemar's Test P-Value : 0.2002
##
##           Sensitivity : 0.8913
##           Specificity : 0.3514
##           Pos Pred Value : 0.8367
##           Neg Pred Value : 0.4643
##           Prevalence : 0.7886
##           Detection Rate : 0.7029
##           Detection Prevalence : 0.8400
##           Balanced Accuracy : 0.6213
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.757540148844497"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 37 cases (dfPred_raw$class 1).
## Area under the curve: 0.7575
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0095  -0.6666  -0.3349   0.4658   3.1071
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.522e+01  1.530e+01   0.995 0.319790
## KIDSDRIV      -2.846e+00  2.009e+00  -1.417 0.156614
## AGE           -1.291e-01  1.076e-01  -1.200 0.230041
## HOMEKIDS      -3.613e-01  6.834e-01  -0.529 0.596974
## YOJ           -1.579e-01  1.340e-01  -1.178 0.238712
## INCOME        -2.361e-05  1.354e-05  -1.744 0.081108 .
## HOME_VAL       1.088e-05  9.277e-06   1.172 0.241000
## TRAVTIME       4.198e-02  2.221e-02   1.890 0.058750 .
## BLUEBOOK       2.820e-05  3.952e-05   0.714 0.475500
```

```

## TIF -3.040e-02 9.613e-02 -0.316 0.751819
## OLDCLAIM 4.589e-06 2.418e-05 0.190 0.849486
## CLM_FREQ 2.977e-02 1.024e+00 0.029 0.976810
## MVR_PTS 4.310e-02 1.750e-01 0.246 0.805500
## CAR_AGE -7.569e-02 7.261e-02 -1.042 0.297195
## PARENT1_Yes 9.757e-01 4.746e-01 2.056 0.039796 *
## MSTATUS_Yes -5.531e-01 3.374e-01 -1.639 0.101135
## SEX_z_F -6.694e-01 4.616e-01 -1.450 0.147079
## EDUCATION_.High.School 1.145e-02 8.626e-01 0.013 0.989413
## EDUCATION_Bachelors -8.266e-02 7.378e-01 -0.112 0.910788
## EDUCATION_Masters 1.044e+00 6.587e-01 1.585 0.113037
## EDUCATION_z_High.School 5.606e-01 7.926e-01 0.707 0.479387
## JOB_ -1.185e+00 7.123e-01 -1.664 0.096051 .
## JOB_Clerical -2.300e-01 4.881e-01 -0.471 0.637479
## JOB_Doctor 1.232e-01 1.045e+00 0.118 0.906176
## JOB_Home.Maker -4.524e-01 6.902e-01 -0.655 0.512175
## JOB_Lawyer -5.516e-01 6.707e-01 -0.822 0.410869
## JOB_Manager -7.979e-01 4.934e-01 -1.617 0.105863
## JOB_Student -1.002e+00 7.459e-01 -1.344 0.179099
## JOB_z_Blue.Collar -2.262e-01 4.412e-01 -0.513 0.608108
## CAR_USE_Commercial 7.319e-01 3.436e-01 2.130 0.033164 *
## CAR_TYPE_Panel.Truck 4.257e-01 6.319e-01 0.674 0.500585
## CAR_TYPE_Pickup 1.192e+00 4.086e-01 2.919 0.003515 **
## CAR_TYPE_Sports.Car 1.655e+00 5.529e-01 2.993 0.002763 **
## CAR_TYPE_Van 6.847e-01 4.806e-01 1.425 0.154273
## CAR_TYPE_z_SUV 1.805e+00 4.687e-01 3.850 0.000118 ***
## RED_CAR_no -1.075e-01 3.426e-01 -0.314 0.753629
## REVOKED_Yes 5.619e-01 4.491e-01 1.251 0.210833
## URBANICITY_z_Highly.Rural..Rural -2.291e+00 4.070e-01 -5.630 1.8e-08 ***
## YOJ_NA 3.502e-02 4.745e-01 0.074 0.941178
## INCOME_NA 1.419e-01 5.655e-01 0.251 0.801836
## CAR_AGE_NA 1.173e-01 4.859e-01 0.241 0.809249
## HOME_VAL_NA 3.755e-03 2.899e-01 0.013 0.989662
## ageSquared 1.395e-03 1.160e-03 1.202 0.229258
## yojSquared 7.353e-03 6.801e-03 1.081 0.279686
## income_log 7.568e-02 2.833e-01 0.267 0.789342
## homeval_log -1.110e+00 1.416e+00 -0.784 0.433021
## travtime_log -3.903e-01 6.354e-01 -0.614 0.539090
## bluebook_log -2.464e-01 4.805e-01 -0.513 0.608175
## carage_log 2.026e-01 4.666e-01 0.434 0.664189
## oldclaim_log 4.664e-02 1.736e-01 0.269 0.788155
## clm_freq_log 1.010e-01 3.173e+00 0.032 0.974604
## mvr_pts_log 2.955e-02 5.123e-01 0.058 0.953994
## tif_log 3.844e-02 5.339e-01 0.072 0.942608
## kidsdriv_log 9.491e-01 2.289e+00 0.415 0.678385
## homekids_log 9.177e-01 1.475e+00 0.622 0.533793
## inter 7.013e-02 3.835e-02 1.829 0.067473 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 746.97 on 639 degrees of freedom
## Residual deviance: 528.89 on 584 degrees of freedom

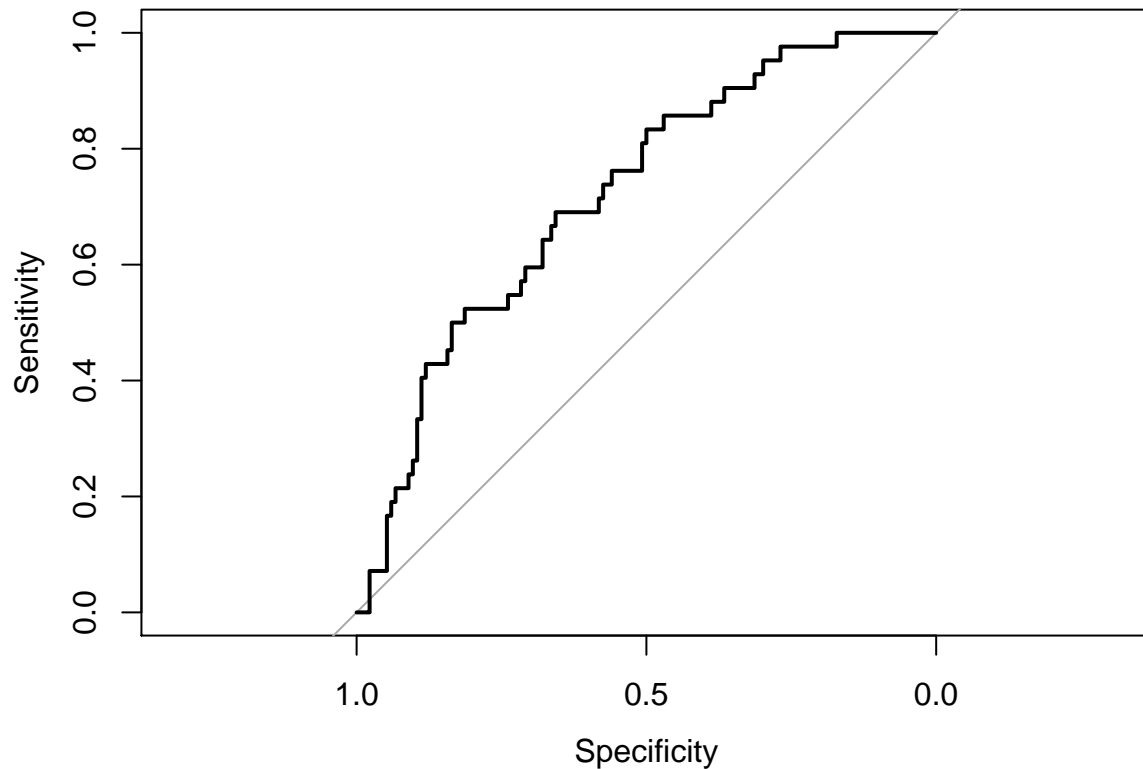
```

```

## (1 observation deleted due to missingness)
## AIC: 640.89
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 119  27
##           1  15  15
##
##           Accuracy : 0.7614
##           95% CI : (0.6914, 0.8223)
##           No Information Rate : 0.7614
##           P-Value [Acc > NIR] : 0.54130
##
##           Kappa : 0.2719
##
## Mcnemar's Test P-Value : 0.08963
##
##           Sensitivity : 0.8881
##           Specificity : 0.3571
##           Pos Pred Value : 0.8151
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7614
##           Detection Rate : 0.6761
##           Detection Prevalence : 0.8295
##           Balanced Accuracy : 0.6226
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.721926083866382"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.7219
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9984  -0.6632  -0.3336  -0.0707   3.1671
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.410e+01  1.500e+01   1.606 0.108268
## KIDSDRIV      -3.901e-01  1.757e+00  -0.222 0.824300
## AGE           -1.645e-01  1.072e-01  -1.535 0.124854
## HOMEKIDS      -1.083e+00  6.982e-01  -1.551 0.120931
## YOJ           -1.227e-01  1.309e-01  -0.937 0.348719
## INCOME        -2.399e-05  1.394e-05  -1.721 0.085275 .
## HOME_VAL       1.549e-05  9.165e-06   1.690 0.091060 .
## TRAVTIME       5.382e-02  2.110e-02   2.550 0.010769 *
## BLUEBOOK       6.385e-05  3.886e-05   1.643 0.100328
```

```

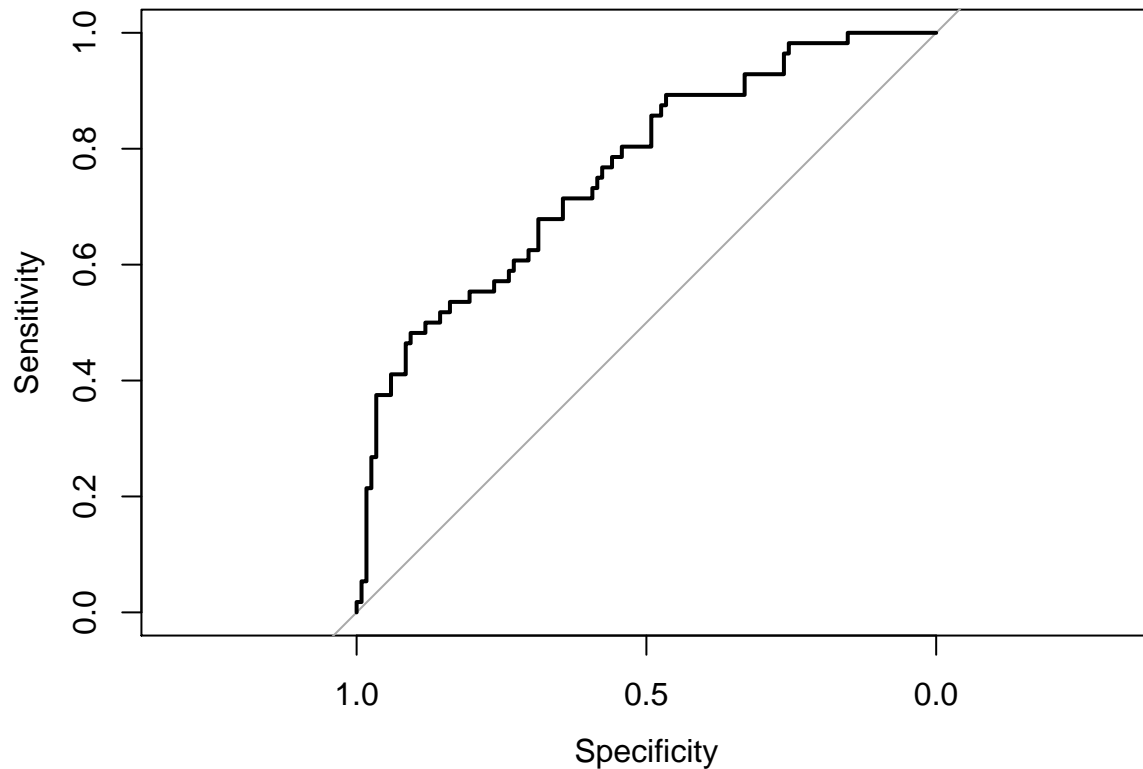
## TIF -2.303e-02 9.463e-02 -0.243 0.807729
## OLDCLAIM 5.413e-06 2.053e-05 0.264 0.792035
## CLM_FREQ 6.002e-02 9.384e-01 0.064 0.948997
## MVR_PTS -3.848e-03 1.692e-01 -0.023 0.981860
## CAR_AGE 2.603e-02 7.114e-02 0.366 0.714480
## PARENT1_Yes 4.400e-01 4.752e-01 0.926 0.354527
## MSTATUS_Yes -5.122e-01 3.421e-01 -1.498 0.134256
## SEX_z_F -7.236e-01 4.776e-01 -1.515 0.129780
## EDUCATION_.High.School 6.442e-01 7.976e-01 0.808 0.419316
## EDUCATION_Bachelors 3.521e-01 6.675e-01 0.527 0.597871
## EDUCATION_Masters 1.089e+00 5.677e-01 1.919 0.055013
## EDUCATION_z_High.School 9.139e-01 7.222e-01 1.265 0.205732
## JOB_ -8.919e-01 7.179e-01 -1.242 0.214142
## JOB_Clerical -4.151e-01 4.974e-01 -0.835 0.403910
## JOB_Doctor 2.010e-01 9.373e-01 0.214 0.830224
## JOB_Home.Maker -1.440e-01 6.695e-01 -0.215 0.829653
## JOB_Lawyer -6.951e-01 6.585e-01 -1.056 0.291176
## JOB_Manager -1.273e+00 5.113e-01 -2.491 0.012753 *
## JOB_Student -8.789e-01 7.307e-01 -1.203 0.228990
## JOB_z_Blue.Collar -4.231e-01 4.609e-01 -0.918 0.358634
## CAR_USE_Commercial 7.463e-01 3.504e-01 2.130 0.033161 *
## CAR_TYPE_Panel.Truck -2.129e-01 6.333e-01 -0.336 0.736722
## CAR_TYPE_Pickup 1.056e+00 4.090e-01 2.582 0.009816 **
## CAR_TYPE_Sports.Car 1.925e+00 5.419e-01 3.553 0.000381 ***
## CAR_TYPE_Van 2.560e-01 4.862e-01 0.526 0.598577
## CAR_TYPE_z_SUV 1.785e+00 4.881e-01 3.657 0.000255 ***
## RED_CAR_no -1.520e-01 3.468e-01 -0.438 0.661273
## REVOKED_Yes 6.075e-01 3.974e-01 1.529 0.126280
## URBANICITY_z_Highly.Rural..Rural -2.621e+00 4.846e-01 -5.408 6.38e-08 ***
## YOJ_NA -4.559e-01 4.281e-01 -1.065 0.286938
## INCOME_NA 3.312e-01 5.534e-01 0.598 0.549528
## CAR_AGE_NA -5.754e-01 4.644e-01 -1.239 0.215323
## HOME_VAL_NA -2.146e-01 2.908e-01 -0.738 0.460464
## ageSquared 1.743e-03 1.161e-03 1.502 0.133093
## yojSquared 5.419e-03 6.826e-03 0.794 0.427252
## income_log 1.107e-01 3.001e-01 0.369 0.712242
## homeval_log -1.473e+00 1.412e+00 -1.044 0.296705
## travtime_log -1.155e+00 5.742e-01 -2.011 0.044302 *
## bluebook_log -4.533e-01 4.639e-01 -0.977 0.328478
## carage_log -4.150e-01 4.668e-01 -0.889 0.374023
## oldclaim_log 1.432e-02 1.608e-01 0.089 0.929072
## clm_freq_log 3.417e-01 2.928e+00 0.117 0.907090
## mvr_pts_log 1.232e-01 4.993e-01 0.247 0.805095
## tif_log -6.963e-02 5.326e-01 -0.131 0.895974
## kidsdriv_log -2.611e-01 2.229e+00 -0.117 0.906778
## homekids_log 2.221e+00 1.494e+00 1.487 0.136951
## inter 2.814e-02 3.155e-02 0.892 0.372409
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 718.72 on 641 degrees of freedom
## Residual deviance: 530.71 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 642.71
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  31
##           1  10  25
##
##           Accuracy : 0.7644
##           95% CI : (0.6942, 0.8253)
##           No Information Rate : 0.6782
##           P-Value [Acc > NIR] : 0.008070
##
##           Kappa : 0.4012
##
## Mcnemar's Test P-Value : 0.001787
##
##           Sensitivity : 0.9153
##           Specificity : 0.4464
##           Pos Pred Value : 0.7770
##           Neg Pred Value : 0.7143
##           Prevalence : 0.6782
##           Detection Rate : 0.6207
##           Detection Prevalence : 0.7989
##           Balanced Accuracy : 0.6808
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.76089588377724"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 118 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7609
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1445  -0.6568  -0.3637   0.6093   3.1770
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    6.955e+00  1.510e+01   0.461  0.645143
## KIDSDRIV     -2.315e+00  1.855e+00  -1.248  0.212143
## AGE          -5.860e-02  1.132e-01  -0.518  0.604550
## HOMEKIDS     -1.270e-01  6.262e-01  -0.203  0.839271
## YOJ          -1.468e-01  1.302e-01  -1.128  0.259485
## INCOME       -7.438e-06  1.349e-05  -0.551  0.581314
## HOME_VAL     -2.491e-07  9.496e-06  -0.026  0.979075
## TRAVTIME      2.272e-02  2.214e-02   1.026  0.304812
## BLUEBOOK      5.883e-05  3.962e-05   1.485  0.137521
```

```

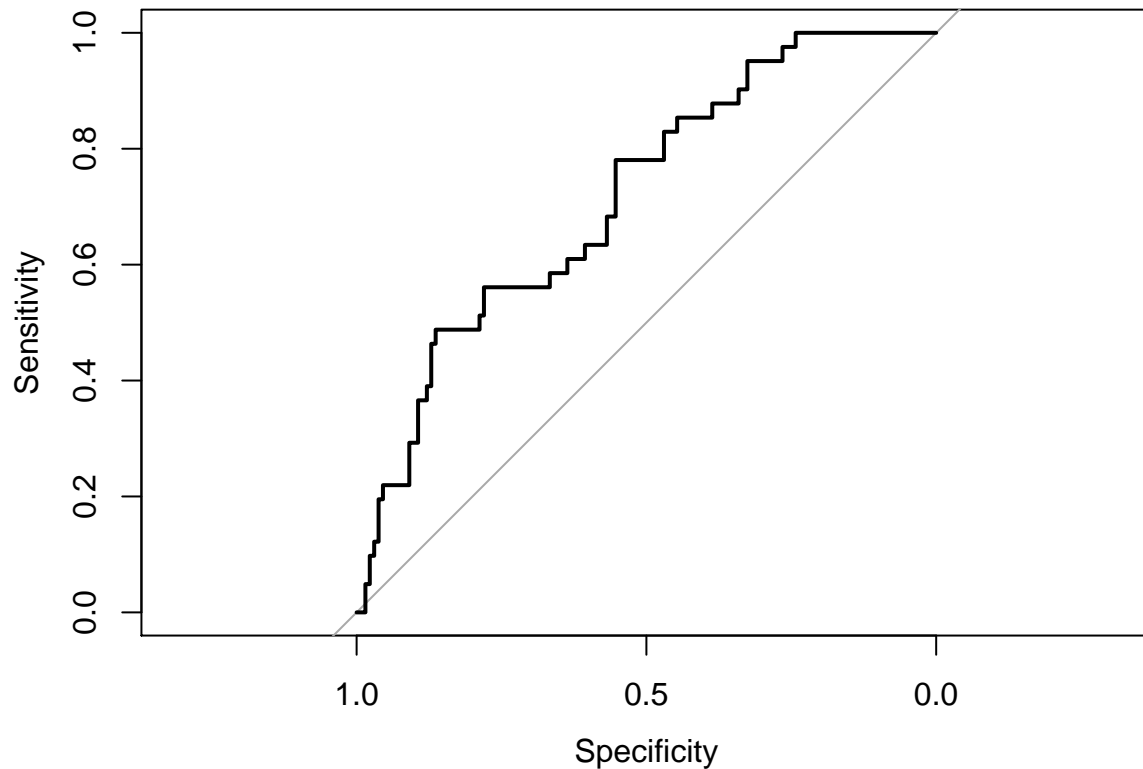
## TIF -4.734e-02 8.818e-02 -0.537 0.591364
## OLDCLAIM 1.686e-05 2.039e-05 0.827 0.408329
## CLM_FREQ 3.467e-01 9.649e-01 0.359 0.719369
## MVR_PTS -6.811e-02 1.668e-01 -0.408 0.682929
## CAR_AGE -2.446e-02 6.987e-02 -0.350 0.726276
## PARENT1_Yes 5.653e-01 4.663e-01 1.212 0.225344
## MSTATUS_Yes -5.907e-01 3.311e-01 -1.784 0.074404
## SEX_z_F -6.497e-01 4.629e-01 -1.403 0.160472
## EDUCATION_.High.School -5.568e-01 7.927e-01 -0.702 0.482394
## EDUCATION_Bachelors -6.956e-01 6.677e-01 -1.042 0.297458
## EDUCATION_Masters 6.098e-01 5.849e-01 1.043 0.297165
## EDUCATION_z_High.School -1.358e-01 7.188e-01 -0.189 0.850108
## JOB_ -1.349e+00 6.757e-01 -1.996 0.045964 *
## JOB_Clerical 1.588e-02 4.829e-01 0.033 0.973759
## JOB_Doctor -4.713e-01 9.785e-01 -0.482 0.630034
## JOB_Home.Maker 4.848e-01 6.421e-01 0.755 0.450264
## JOB_Lawyer -8.131e-01 6.312e-01 -1.288 0.197667
## JOB_Manager -7.217e-01 4.945e-01 -1.459 0.144441
## JOB_Student -1.032e-01 7.116e-01 -0.145 0.884682
## JOB_z_Blue.Collar 3.984e-04 4.459e-01 0.001 0.999287
## CAR_USE_Commercial 7.130e-01 3.501e-01 2.036 0.041728 *
## CAR_TYPE_Panel.Truck 8.587e-02 6.280e-01 0.137 0.891243
## CAR_TYPE_Pickup 1.085e+00 4.099e-01 2.646 0.008143 **
## CAR_TYPE_Sports.Car 1.864e+00 5.375e-01 3.467 0.000526 ***
## CAR_TYPE_Van 2.274e-01 4.905e-01 0.464 0.642946
## CAR_TYPE_z_SUV 1.908e+00 4.702e-01 4.059 4.93e-05 ***
## RED_CAR_no -4.342e-01 3.512e-01 -1.236 0.216378
## REVOKED_Yes 3.586e-01 4.066e-01 0.882 0.377801
## URBANICITY_z_Highly.Rural..Rural -2.356e+00 4.050e-01 -5.816 6.03e-09 ***
## YOJ_NA -3.797e-01 4.501e-01 -0.844 0.398939
## INCOME_NA 2.713e-01 5.529e-01 0.491 0.623690
## CAR_AGE_NA -4.390e-01 4.532e-01 -0.969 0.332732
## HOME_VAL_NA 3.055e-02 2.898e-01 0.105 0.916039
## ageSquared 5.451e-04 1.244e-03 0.438 0.661169
## yojSquared 7.812e-03 6.745e-03 1.158 0.246778
## income_log 2.327e-01 2.819e-01 0.826 0.409051
## homeval_log -4.591e-01 1.447e+00 -0.317 0.750983
## travtime_log 1.707e-01 6.270e-01 0.272 0.785455
## bluebook_log -3.798e-01 4.856e-01 -0.782 0.434233
## carage_log -2.491e-01 4.559e-01 -0.546 0.584853
## oldclaim_log 3.602e-02 1.641e-01 0.220 0.826235
## clm_freq_log -6.380e-01 2.995e+00 -0.213 0.831318
## mvr_pts_log 3.249e-01 4.911e-01 0.661 0.508328
## tif_log 7.427e-02 5.053e-01 0.147 0.883144
## kidsdriv_log 1.704e+00 2.098e+00 0.813 0.416472
## homekids_log 3.800e-01 1.379e+00 0.276 0.782874
## inter 4.564e-02 3.378e-02 1.351 0.176698
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 750.85 on 642 degrees of freedom
## Residual deviance: 543.82 on 587 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 655.82
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  28
##           1  14  13
##
##           Accuracy : 0.7572
##           95% CI : (0.6863, 0.8191)
##           No Information Rate : 0.763
##           P-Value [Acc > NIR] : 0.61126
##
##           Kappa : 0.2392
##
## Mcnemar's Test P-Value : 0.04486
##
##           Sensitivity : 0.8939
##           Specificity : 0.3171
##           Pos Pred Value : 0.8082
##           Neg Pred Value : 0.4815
##           Prevalence : 0.7630
##           Detection Rate : 0.6821
##           Detection Prevalence : 0.8439
##           Balanced Accuracy : 0.6055
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.716555801921656"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 132 controls (dfPred_raw$class 0) < 41 cases (dfPred_raw$class 1).
## Area under the curve: 0.7166
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1060  -0.6704  -0.3769   0.4258   3.1415
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.713e+01  1.489e+01   1.150  0.250031
## KIDSDRIV      -1.667e+00  1.733e+00  -0.962  0.336084
## AGE           1.168e-02  1.158e-01   0.101  0.919676
## HOMEKIDS      -1.345e+00  7.058e-01  -1.906  0.056706 .
## YOJ           -1.594e-01  1.254e-01  -1.271  0.203685
## INCOME        -2.828e-05  1.365e-05  -2.072  0.038276 *
## HOME_VAL       1.444e-05  8.880e-06   1.626  0.103908
## TRAVTIME       3.596e-02  2.197e-02   1.636  0.101736
## BLUEBOOK       3.092e-05  4.057e-05   0.762  0.445901
```

```

## TIF -2.049e-02 8.908e-02 -0.230 0.818074
## OLDCLAIM -3.711e-06 2.218e-05 -0.167 0.867120
## CLM_FREQ -2.029e-01 1.086e+00 -0.187 0.851853
## MVRPTS 1.287e-01 1.764e-01 0.729 0.465751
## CAR_AGE -3.997e-02 7.330e-02 -0.545 0.585586
## PARENT1_Yes 4.583e-01 4.757e-01 0.963 0.335350
## MSTATUS_Yes -5.880e-01 3.246e-01 -1.811 0.070102 .
## SEX_z_F -1.279e+00 5.027e-01 -2.545 0.010935 *
## EDUCATION_.High.School -7.084e-01 7.982e-01 -0.888 0.374789
## EDUCATION_Bachelors -7.041e-01 6.589e-01 -1.068 0.285301
## EDUCATION_Masters -7.417e-03 5.484e-01 -0.014 0.989209
## EDUCATION_z_High.School -3.624e-01 7.154e-01 -0.507 0.612472
## JOB_ -1.127e+00 7.048e-01 -1.600 0.109686
## JOB_Clerical -4.241e-01 4.862e-01 -0.872 0.383054
## JOB_Doctor -4.991e-01 8.948e-01 -0.558 0.576961
## JOB_Home.Maker -1.995e-01 6.377e-01 -0.313 0.754463
## JOB_Lawyer -6.222e-01 6.453e-01 -0.964 0.334905
## JOB_Manager -8.525e-01 4.872e-01 -1.750 0.080178 .
## JOB_Student -3.185e-01 7.211e-01 -0.442 0.658735
## JOB_z_Blue.Collar 1.899e-01 4.657e-01 0.408 0.683454
## CAR_USE_Commercial 2.801e-01 3.660e-01 0.765 0.444192
## CAR_TYPE_Panel.Truck 3.992e-01 6.214e-01 0.642 0.520613
## CAR_TYPE_Pickup 1.385e+00 4.107e-01 3.373 0.000744 ***
## CAR_TYPE_Sports.Car 2.367e+00 5.641e-01 4.196 2.71e-05 ***
## CAR_TYPE_Van 3.969e-01 4.811e-01 0.825 0.409403
## CAR_TYPE_z_SUV 2.433e+00 5.082e-01 4.788 1.68e-06 ***
## RED_CAR_no 1.036e-02 3.393e-01 0.031 0.975643
## REVOKED_Yes 4.347e-01 4.112e-01 1.057 0.290347
## URBANICITY_z_Highly.Rural..Rural -2.421e+00 4.199e-01 -5.766 8.10e-09 ***
## YOJ_NA -5.342e-01 4.414e-01 -1.210 0.226176
## INCOME_NA 1.158e-01 5.250e-01 0.221 0.825364
## CAR_AGE_NA -1.672e-02 5.002e-01 -0.033 0.973329
## HOME_VAL_NA -1.646e-01 2.849e-01 -0.578 0.563337
## ageSquared -3.299e-04 1.278e-03 -0.258 0.796300
## yojSquared 8.195e-03 6.579e-03 1.246 0.212931
## income_log 3.512e-01 3.001e-01 1.170 0.241859
## homeval_log -1.751e+00 1.381e+00 -1.269 0.204574
## travtime_log -4.467e-01 6.202e-01 -0.720 0.471388
## bluebook_log -4.082e-02 4.846e-01 -0.084 0.932881
## carage_log 4.033e-02 4.751e-01 0.085 0.932361
## oldclaim_log -1.185e-02 1.715e-01 -0.069 0.944907
## clm_freq_log 6.512e-01 3.287e+00 0.198 0.842969
## mvr_pts_log -3.452e-01 5.133e-01 -0.672 0.501310
## tif_log 4.581e-02 5.141e-01 0.089 0.929003
## kidsdriv_log 7.946e-01 2.222e+00 0.358 0.720640
## homekids_log 2.692e+00 1.520e+00 1.772 0.076453 .
## inter 4.511e-02 3.274e-02 1.378 0.168201
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 732.35 on 642 degrees of freedom
## Residual deviance: 545.30 on 587 degrees of freedom

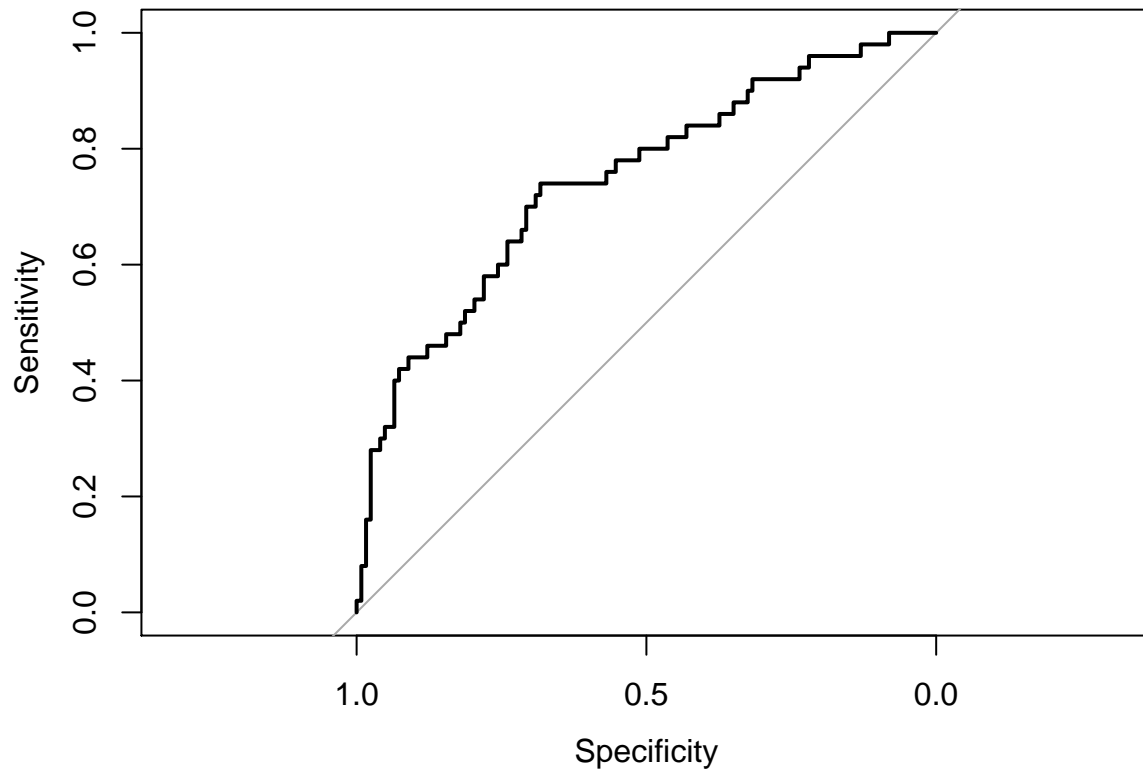
```



```

## (1 observation deleted due to missingness)
## AIC: 657.3
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 115  34
##           1   8  16
##
##           Accuracy : 0.7572
##           95% CI : (0.6863, 0.8191)
##           No Information Rate : 0.711
##           P-Value [Acc > NIR] : 0.1028125
##
##           Kappa : 0.3015
##
## Mcnemar's Test P-Value : 0.0001145
##
##           Sensitivity : 0.9350
##           Specificity : 0.3200
##           Pos Pred Value : 0.7718
##           Neg Pred Value : 0.6667
##           Prevalence : 0.7110
##           Detection Rate : 0.6647
##           Detection Prevalence : 0.8613
##           Balanced Accuracy : 0.6275
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.745365853658537"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 123 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7454
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2389  -0.6327  -0.3581   0.3736   3.2407
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.290e+01  1.557e+01  0.829  0.40734
## KIDSDRIV      -9.978e-01  1.774e+00 -0.563  0.57373
## AGE          -2.463e-01  1.070e-01 -2.302  0.02133 *
## HOMEKIDS     -1.272e+00  7.177e-01 -1.773  0.07626 .
## YOJ          -1.899e-01  1.342e-01 -1.415  0.15715
## INCOME       -1.955e-05  1.373e-05 -1.424  0.15447
## HOME_VAL      7.944e-06  9.238e-06  0.860  0.38983
## TRAVTIME      4.327e-02  2.179e-02  1.985  0.04710 *
## BLUEBOOK      6.871e-05  4.225e-05  1.626  0.10386
```

```

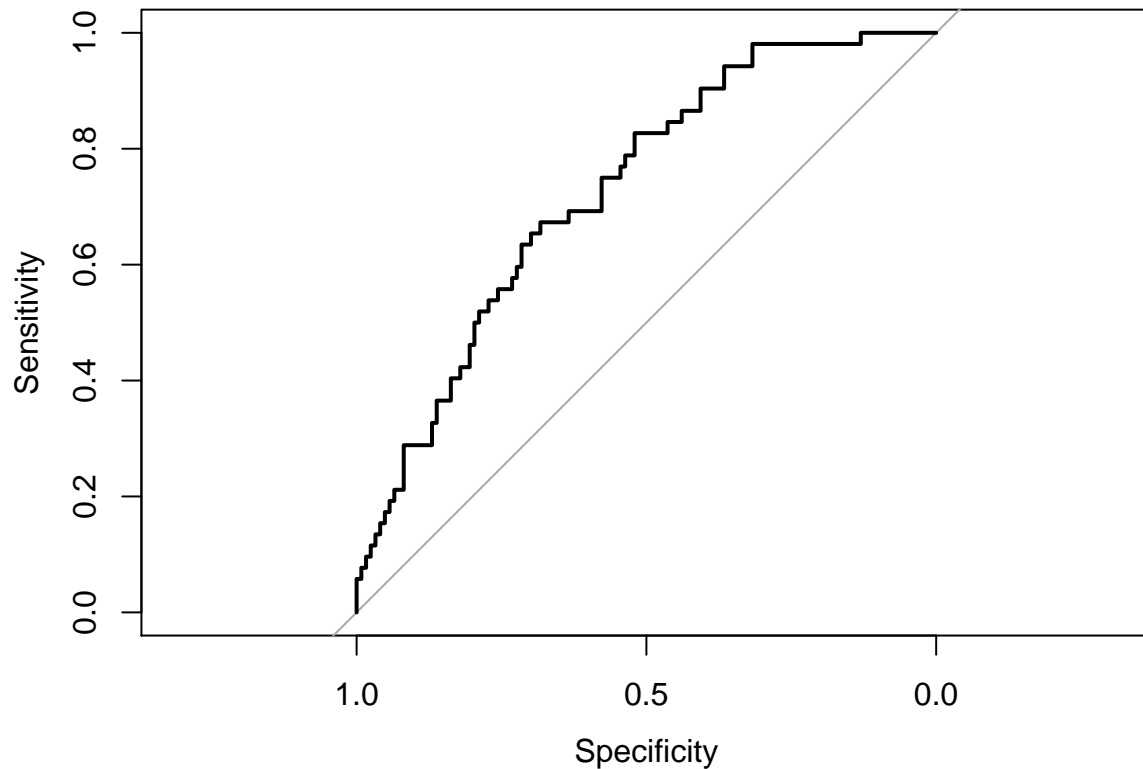
## TIF -9.193e-02 9.651e-02 -0.952 0.34086
## OLDCLAIM 2.235e-05 2.100e-05 1.065 0.28708
## CLM_FREQ -2.146e-01 1.009e+00 -0.213 0.83166
## MVR_PTS -5.284e-02 1.745e-01 -0.303 0.76198
## CAR_AGE -6.238e-02 7.313e-02 -0.853 0.39366
## PARENT1_Yes 4.983e-01 4.922e-01 1.013 0.31127
## MSTATUS_Yes -7.017e-01 3.377e-01 -2.078 0.03774 *
## SEX_z_F -1.083e+00 4.896e-01 -2.211 0.02702 *
## EDUCATION_.High.School 1.425e-01 8.068e-01 0.177 0.85979
## EDUCATION_Bachelors -8.145e-02 6.970e-01 -0.117 0.90697
## EDUCATION_Masters 7.772e-01 6.187e-01 1.256 0.20902
## EDUCATION_z_High.School 2.552e-01 7.395e-01 0.345 0.73006
## JOB_ -3.451e-01 7.011e-01 -0.492 0.62261
## JOB_Clerical 3.196e-01 5.060e-01 0.632 0.52769
## JOB_Doctor 3.684e-01 9.675e-01 0.381 0.70339
## JOB_Home.Maker 3.411e-01 6.943e-01 0.491 0.62324
## JOB_Lawyer 4.783e-02 6.937e-01 0.069 0.94504
## JOB_Manager -4.482e-01 4.936e-01 -0.908 0.36393
## JOB_Student 1.785e-01 7.309e-01 0.244 0.80706
## JOB_z_Blue.Collar 3.291e-01 4.786e-01 0.688 0.49168
## CAR_USE_Commercial 3.053e-01 3.545e-01 0.861 0.38916
## CAR_TYPE_Panel.Truck 2.154e-01 6.366e-01 0.338 0.73507
## CAR_TYPE_Pickup 1.371e+00 4.274e-01 3.208 0.00134 **
## CAR_TYPE_Sports.Car 2.336e+00 5.575e-01 4.189 2.80e-05 ***
## CAR_TYPE_Van 5.914e-01 5.092e-01 1.162 0.24542
## CAR_TYPE_z_SUV 2.287e+00 5.192e-01 4.405 1.06e-05 ***
## RED_CAR_no -1.577e-01 3.406e-01 -0.463 0.64345
## REVOKED_Yes -3.363e-02 4.391e-01 -0.077 0.93894
## URBANICITY_z_Highly.Rural..Rural -2.395e+00 4.096e-01 -5.847 5.02e-09 ***
## YOJ_NA -2.744e-01 4.493e-01 -0.611 0.54149
## INCOME_NA 2.623e-01 5.867e-01 0.447 0.65481
## CAR_AGE_NA 5.898e-02 5.104e-01 0.116 0.90800
## HOME_VAL_NA -1.906e-01 2.969e-01 -0.642 0.52084
## ageSquared 2.470e-03 1.155e-03 2.139 0.03245 *
## yojSquared 8.973e-03 6.950e-03 1.291 0.19666
## income_log 1.205e-01 3.113e-01 0.387 0.69877
## homeval_log -4.457e-01 1.469e+00 -0.303 0.76163
## travtime_log -5.772e-01 6.136e-01 -0.941 0.34690
## bluebook_log -4.903e-01 5.152e-01 -0.952 0.34121
## carage_log 4.567e-02 4.696e-01 0.097 0.92252
## oldclaim_log 4.233e-02 1.659e-01 0.255 0.79861
## clm_freq_log 2.960e-01 3.115e+00 0.095 0.92429
## mvr_pts_log 3.045e-01 5.073e-01 0.600 0.54834
## tif_log 3.101e-01 5.380e-01 0.576 0.56431
## kidsdriv_log -2.307e-01 2.250e+00 -0.103 0.91834
## homekids_log 2.487e+00 1.540e+00 1.615 0.10634
## inter 4.368e-02 3.249e-02 1.344 0.17883
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 726.89 on 640 degrees of freedom
## Residual deviance: 527.19 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 639.19
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  37
##           1  15  15
##
##           Accuracy : 0.7029
##           95% CI : (0.6292, 0.7695)
##           No Information Rate : 0.7029
##           P-Value [Acc > NIR] : 0.537366
##
##           Kappa : 0.1897
##
## Mcnemar's Test P-Value : 0.003589
##
##           Sensitivity : 0.8780
##           Specificity : 0.2885
##           Pos Pred Value : 0.7448
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7029
##           Detection Rate : 0.6171
##           Detection Prevalence : 0.8286
##           Balanced Accuracy : 0.5833
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.727485928705441"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 123 controls (dfPred_raw$class 0) < 52 cases (dfPred_raw$class 1).
## Area under the curve: 0.7275
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0129  -0.7129  -0.3758   0.5662   3.0619
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.169e+01  1.418e+01   1.530 0.125936
## KIDSDRIV      -6.894e-02  2.111e+00  -0.033 0.973947
## AGE          -8.853e-02  1.135e-01  -0.780 0.435344
## HOMEKIDS     -1.210e+00  6.907e-01  -1.752 0.079838 .
## YOJ          -1.091e-01  1.277e-01  -0.855 0.392574
## INCOME       -2.463e-05  1.278e-05  -1.928 0.053915 .
## HOME_VAL      1.483e-05  8.504e-06   1.744 0.081099 .
## TRAVTIME      3.396e-02  2.129e-02   1.595 0.110649
## BLUEBOOK      5.396e-05  3.950e-05   1.366 0.171916
```

```

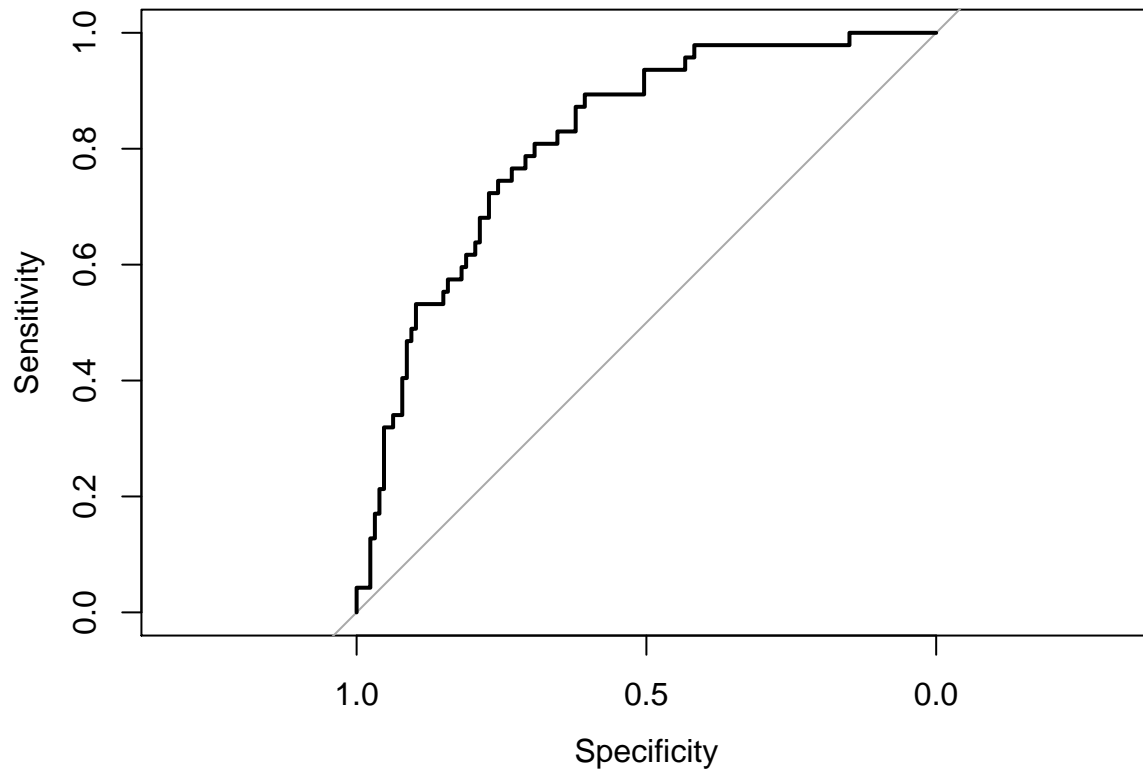
## TIF -1.560e-02 9.383e-02 -0.166 0.867943
## OLDCLAIM 7.130e-06 2.018e-05 0.353 0.723855
## CLM_FREQ 2.284e-01 9.102e-01 0.251 0.801863
## MVRPTS 1.706e-02 1.584e-01 0.108 0.914216
## CAR_AGE 2.649e-02 6.794e-02 0.390 0.696629
## PARENT1_Yes 4.859e-01 4.707e-01 1.032 0.301965
## MSTATUS_Yes -3.540e-01 3.308e-01 -1.070 0.284635
## SEX_z_F -5.485e-01 4.470e-01 -1.227 0.219830
## EDUCATION_.High.School 1.899e-01 7.611e-01 0.249 0.802983
## EDUCATION_Bachelors 6.181e-02 6.412e-01 0.096 0.923214
## EDUCATION_Masters 8.793e-01 5.544e-01 1.586 0.112753
## EDUCATION_z_High.School 4.056e-01 6.952e-01 0.583 0.559616
## JOB_ -9.629e-01 6.692e-01 -1.439 0.150170
## JOB_Clerical -2.315e-01 4.867e-01 -0.476 0.634297
## JOB_Doctor 4.893e-02 9.640e-01 0.051 0.959522
## JOB_Home.Maker 3.480e-02 6.400e-01 0.054 0.956633
## JOB_Lawyer -8.238e-01 6.461e-01 -1.275 0.202353
## JOB_Manager -1.018e+00 4.974e-01 -2.047 0.040653 *
## JOB_Student -6.953e-01 7.307e-01 -0.952 0.341289
## JOB_z_Blue.Collar -4.233e-02 4.562e-01 -0.093 0.926059
## CAR_USE_Commercial 5.100e-01 3.377e-01 1.510 0.130957
## CAR_TYPE_Panel.Truck 2.311e-01 6.008e-01 0.385 0.700532
## CAR_TYPE_Pickup 1.234e+00 4.007e-01 3.079 0.002080 **
## CAR_TYPE_Sports.Car 1.865e+00 5.117e-01 3.645 0.000267 ***
## CAR_TYPE_Van 6.016e-01 4.665e-01 1.290 0.197189
## CAR_TYPE_z_SUV 1.850e+00 4.671e-01 3.960 7.49e-05 ***
## RED_CAR_no -2.162e-01 3.342e-01 -0.647 0.517594
## REVOKED_Yes 3.016e-01 4.137e-01 0.729 0.465909
## URBANICITY_z_Highly.Rural..Rural -2.499e+00 4.525e-01 -5.522 3.35e-08 ***
## YOJ_NA -1.481e-01 4.317e-01 -0.343 0.731554
## INCOME_NA 2.650e-01 5.498e-01 0.482 0.629879
## CAR_AGE_NA -2.734e-01 4.574e-01 -0.598 0.549949
## HOME_VAL_NA -9.474e-02 2.924e-01 -0.324 0.745921
## ageSquared 8.700e-04 1.248e-03 0.697 0.485874
## yojSquared 5.178e-03 6.583e-03 0.787 0.431569
## income_log 1.920e-01 2.598e-01 0.739 0.459934
## homeval_log -1.661e+00 1.291e+00 -1.287 0.198243
## travtime_log -6.880e-01 5.923e-01 -1.162 0.245388
## bluebook_log -2.954e-01 4.791e-01 -0.617 0.537509
## carage_log -4.527e-01 4.482e-01 -1.010 0.312496
## oldclaim_log 1.681e-02 1.547e-01 0.109 0.913486
## clm_freq_log -1.549e-01 2.826e+00 -0.055 0.956286
## mvr_pts_log -8.729e-04 4.776e-01 -0.002 0.998542
## tif_log -2.264e-01 5.272e-01 -0.429 0.667673
## kidsdriv_log -1.840e+00 2.422e+00 -0.760 0.447416
## homekids_log 2.433e+00 1.461e+00 1.666 0.095739 .
## inter 4.355e-02 3.856e-02 1.129 0.258730
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 738.05 on 641 degrees of freedom
## Residual deviance: 562.95 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 674.95
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 115  25
##           1  12  22
##
##           Accuracy : 0.7874
##           95% CI : (0.719, 0.8456)
##           No Information Rate : 0.7299
##           P-Value [Acc > NIR] : 0.04983
##
##           Kappa : 0.4092
##
## Mcnemar's Test P-Value : 0.04852
##
##           Sensitivity : 0.9055
##           Specificity : 0.4681
##           Pos Pred Value : 0.8214
##           Neg Pred Value : 0.6471
##           Prevalence : 0.7299
##           Detection Rate : 0.6609
##           Detection Prevalence : 0.8046
##           Balanced Accuracy : 0.6868
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.815044396046239"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 127 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.815
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.58854  -0.63931  -0.35179  -0.09042   3.10394
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.050e+01  1.582e+01   1.296  0.19512
## KIDSDRIV      -3.632e+00  1.937e+00  -1.875  0.06074 .
## AGE          -1.234e-01  1.078e-01  -1.145  0.25222
## HOMEKIDS      -2.228e-01  6.706e-01  -0.332  0.73969
## YOJ           -2.376e-01  1.385e-01  -1.715  0.08626 .
## INCOME        -2.095e-05  1.382e-05  -1.516  0.12954
## HOME_VAL       1.259e-05  9.582e-06   1.314  0.18893
## TRAVTIME       3.052e-02  2.114e-02   1.444  0.14880
## BLUEBOOK       2.974e-05  4.207e-05   0.707  0.47970
```



```

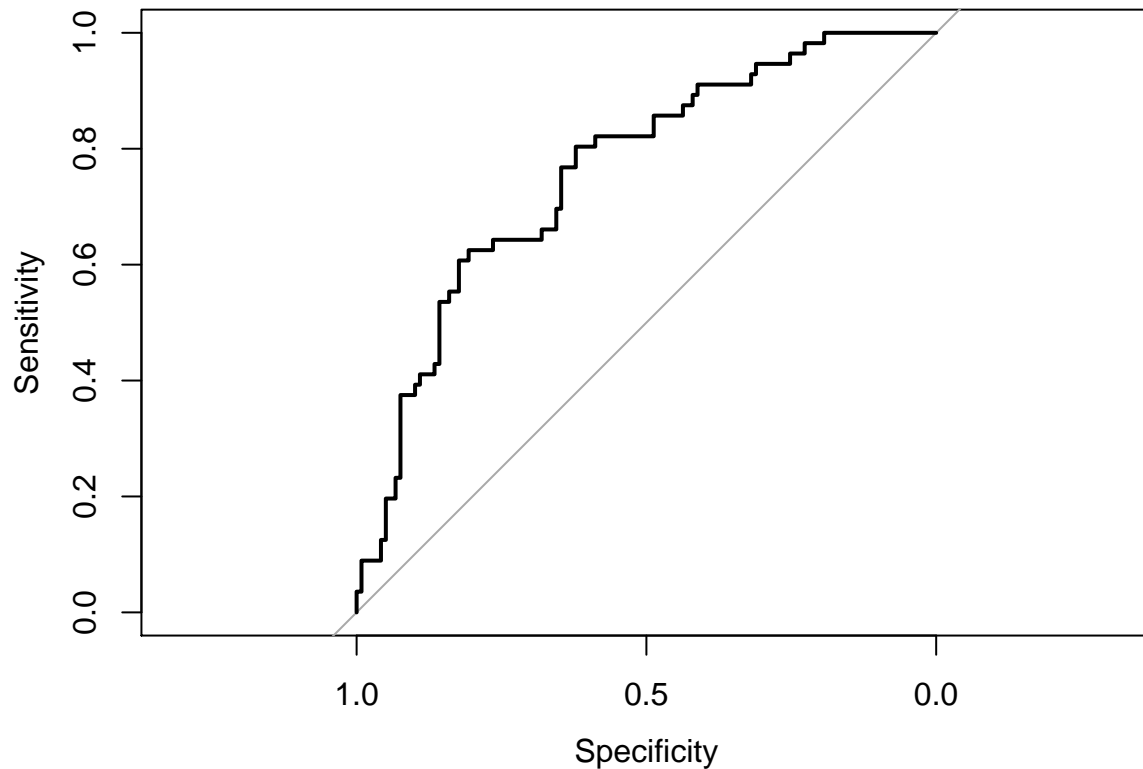
## TIF -7.349e-02 9.535e-02 -0.771 0.44083
## OLDCLAIM 2.626e-07 2.153e-05 0.012 0.99027
## CLM_FREQ -4.248e-01 9.456e-01 -0.449 0.65323
## MVRPTS 1.219e-01 1.717e-01 0.710 0.47774
## CAR_AGE 1.591e-02 7.041e-02 0.226 0.82125
## PARENT1_Yes 5.560e-01 4.805e-01 1.157 0.24723
## MSTATUS_Yes -7.187e-01 3.472e-01 -2.070 0.03846 *
## SEX_z_F -2.795e-01 4.685e-01 -0.597 0.55074
## EDUCATION_.High.School 5.829e-02 7.707e-01 0.076 0.93971
## EDUCATION_Bachelors 3.551e-02 6.324e-01 0.056 0.95522
## EDUCATION_Masters 1.001e+00 5.639e-01 1.776 0.07581 .
## EDUCATION_z_High.School 5.085e-01 6.911e-01 0.736 0.46181
## JOB_ -1.189e+00 6.816e-01 -1.744 0.08110 .
## JOB_Clerical -4.241e-01 5.003e-01 -0.848 0.39663
## JOB_Doctor -5.440e-01 1.013e+00 -0.537 0.59118
## JOB_Home.Maker -4.339e-02 6.514e-01 -0.067 0.94688
## JOB_Lawyer -1.081e+00 6.402e-01 -1.689 0.09127 .
## JOB_Manager -8.063e-01 4.858e-01 -1.660 0.09699 .
## JOB_Student -8.034e-01 7.226e-01 -1.112 0.26622
## JOB_z_Blue.Collar -3.897e-01 4.545e-01 -0.857 0.39126
## CAR_USE_Commercial 9.619e-01 3.529e-01 2.726 0.00642 **
## CAR_TYPE_Panel.Truck -1.238e-01 6.320e-01 -0.196 0.84469
## CAR_TYPE_Pickup 7.464e-01 4.261e-01 1.752 0.07980 .
## CAR_TYPE_Sports.Car 1.418e+00 5.365e-01 2.642 0.00823 **
## CAR_TYPE_Van 2.599e-01 4.816e-01 0.540 0.58944
## CAR_TYPE_z_SUV 1.467e+00 4.714e-01 3.111 0.00186 **
## RED_CAR_no -3.224e-01 3.553e-01 -0.907 0.36427
## REVOKED_Yes 5.951e-01 4.230e-01 1.407 0.15952
## URBANICITY_z_Highly.Rural..Rural -2.405e+00 4.570e-01 -5.262 1.43e-07 ***
## YOJ_NA 2.090e-01 4.493e-01 0.465 0.64177
## INCOME_NA 2.970e-01 6.074e-01 0.489 0.62487
## CAR_AGE_NA -3.694e-01 4.836e-01 -0.764 0.44494
## HOME_VAL_NA 4.281e-02 3.048e-01 0.140 0.88829
## ageSquared 1.223e-03 1.176e-03 1.040 0.29855
## yojSquared 1.306e-02 7.149e-03 1.827 0.06772 .
## income_log 1.382e-01 2.743e-01 0.504 0.61431
## homeval_log -1.649e+00 1.494e+00 -1.104 0.26958
## travtime_log -4.368e-01 5.803e-01 -0.753 0.45158
## bluebook_log -7.511e-02 5.376e-01 -0.140 0.88888
## carage_log -3.534e-01 4.664e-01 -0.758 0.44868
## oldclaim_log 1.416e-02 1.648e-01 0.086 0.93152
## clm_freq_log 1.260e+00 2.958e+00 0.426 0.67025
## mvr_pts_log -1.356e-01 5.106e-01 -0.266 0.79051
## tif_log 1.684e-01 5.322e-01 0.316 0.75175
## kidsdriv_log 4.089e+00 2.262e+00 1.808 0.07063 .
## homekids_log 3.563e-01 1.460e+00 0.244 0.80716
## inter 4.812e-02 3.435e-02 1.401 0.16131
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 718.15 on 640 degrees of freedom
## Residual deviance: 526.22 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 638.22
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  35
##           1  11  21
##
##           Accuracy : 0.7371
##           95% CI : (0.6654, 0.8007)
##           No Information Rate : 0.68
##           P-Value [Acc > NIR] : 0.060012
##
##           Kappa : 0.3187
##
## Mcnemar's Test P-Value : 0.000696
##
##           Sensitivity : 0.9076
##           Specificity : 0.3750
##           Pos Pred Value : 0.7552
##           Neg Pred Value : 0.6562
##           Prevalence : 0.6800
##           Detection Rate : 0.6171
##           Detection Prevalence : 0.8171
##           Balanced Accuracy : 0.6413
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.760504201680672"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 119 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7605
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1931  -0.6531  -0.3501  -0.0533   3.1998
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    9.702e+00  1.531e+01   0.634  0.5264
## KIDSDRIV     -2.006e+00  1.906e+00  -1.052  0.2927
## AGE           8.494e-03  1.133e-01   0.075  0.9403
## HOMEKIDS     -7.810e-01  6.995e-01  -1.116  0.2642
## YOJ          -1.253e-01  1.366e-01  -0.917  0.3590
## INCOME       -9.803e-06  1.382e-05  -0.709  0.4781
## HOME_VAL      5.798e-06  9.269e-06   0.625  0.5317
## TRAVTIME      2.945e-02  2.131e-02   1.382  0.1670
## BLUEBOOK      4.908e-05  3.731e-05   1.315  0.1884
```

```

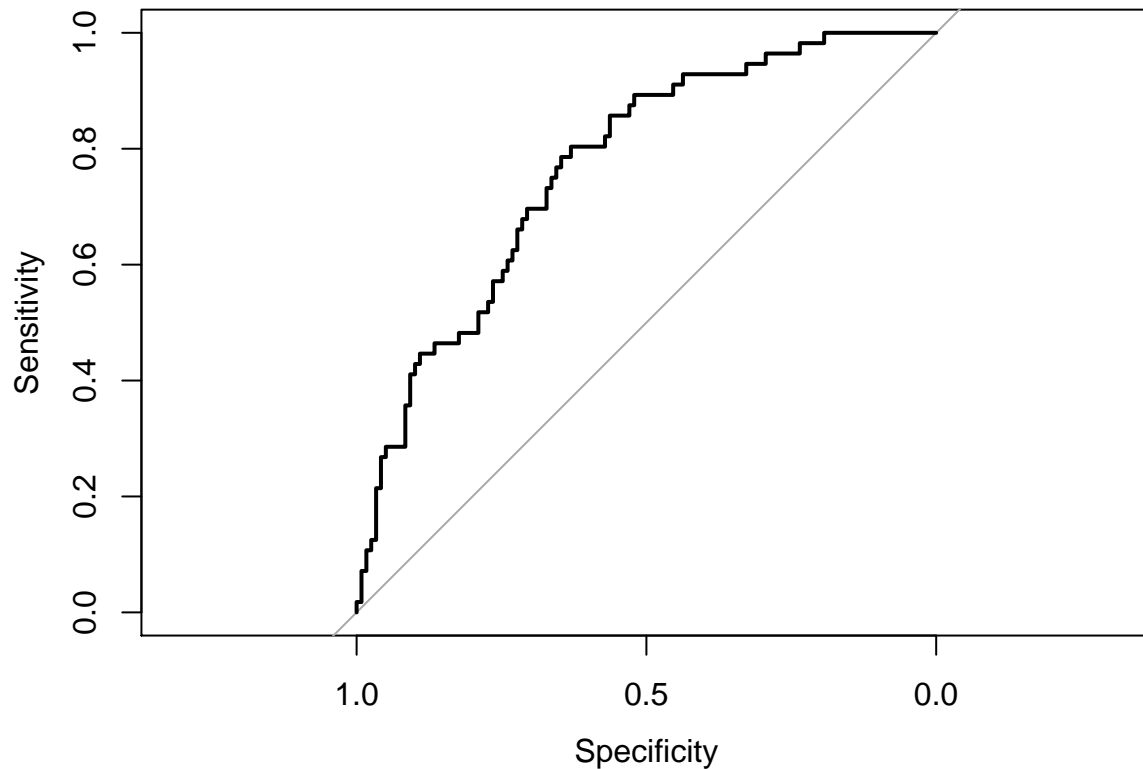
## TIF -1.722e-02 9.058e-02 -0.190 0.8492
## OLDCLAIM -6.099e-06 2.184e-05 -0.279 0.7800
## CLM_FREQ 1.110e-01 8.949e-01 0.124 0.9013
## MVR_PTS 1.142e-01 1.682e-01 0.679 0.4969
## CAR_AGE -1.029e-01 6.887e-02 -1.494 0.1351
## PARENT1_Yes 5.525e-01 4.826e-01 1.145 0.2523
## MSTATUS_Yes -6.295e-01 3.428e-01 -1.836 0.0663 .
## SEX_z_F -8.252e-01 4.627e-01 -1.784 0.0745 .
## EDUCATION_.High.School -5.995e-01 7.940e-01 -0.755 0.4502
## EDUCATION_Bachelors -5.602e-01 6.411e-01 -0.874 0.3822
## EDUCATION_Masters 6.242e-01 5.301e-01 1.178 0.2390
## EDUCATION_z_High.School 9.765e-02 6.929e-01 0.141 0.8879
## JOB_ -8.745e-01 6.494e-01 -1.347 0.1781
## JOB_Clerical -2.645e-01 4.931e-01 -0.536 0.5918
## JOB_Doctor -6.567e-01 8.683e-01 -0.756 0.4495
## JOB_Home.Maker -2.298e-01 6.577e-01 -0.349 0.7268
## JOB_Lawyer -1.392e+00 6.637e-01 -2.097 0.0360 *
## JOB_Manager -9.181e-01 4.854e-01 -1.891 0.0586 .
## JOB_Student -3.337e-01 7.469e-01 -0.447 0.6550
## JOB_z_Blue.Collar -1.329e-01 4.486e-01 -0.296 0.7671
## CAR_USE_Commercial 4.874e-01 3.578e-01 1.362 0.1731
## CAR_TYPE_Panel.Truck 2.411e-02 6.126e-01 0.039 0.9686
## CAR_TYPE_Pickup 1.012e+00 4.254e-01 2.378 0.0174 *
## CAR_TYPE_Sports.Car 2.168e+00 5.408e-01 4.008 6.13e-05 ***
## CAR_TYPE_Van 4.797e-01 4.756e-01 1.009 0.3131
## CAR_TYPE_z_SUV 1.949e+00 4.796e-01 4.065 4.80e-05 ***
## RED_CAR_no 7.491e-03 3.399e-01 0.022 0.9824
## REVOKED_Yes 5.686e-01 4.179e-01 1.361 0.1736
## URBANICITY_z_Highly.Rural..Rural -2.549e+00 5.097e-01 -5.001 5.69e-07 ***
## YOJ_NA -1.553e-01 4.256e-01 -0.365 0.7151
## INCOME_NA -7.972e-02 5.508e-01 -0.145 0.8849
## CAR_AGE_NA 7.198e-02 5.130e-01 0.140 0.8884
## HOME_VAL_NA -4.843e-02 2.908e-01 -0.167 0.8677
## ageSquared -1.248e-04 1.240e-03 -0.101 0.9198
## yojSquared 7.973e-03 6.982e-03 1.142 0.2535
## income_log 1.357e-02 3.241e-01 0.042 0.9666
## homeval_log -7.813e-01 1.474e+00 -0.530 0.5962
## travtime_log -2.666e-01 6.042e-01 -0.441 0.6591
## bluebook_log -3.212e-01 4.623e-01 -0.695 0.4872
## carage_log 5.693e-01 4.665e-01 1.220 0.2223
## oldclaim_log 3.616e-02 1.549e-01 0.233 0.8154
## clm_freq_log 2.180e-01 2.783e+00 0.078 0.9376
## mvr_pts_log -2.608e-01 5.010e-01 -0.521 0.6027
## tif_log -7.309e-02 5.228e-01 -0.140 0.8888
## kidsdriv_log 1.271e+00 2.313e+00 0.549 0.5827
## homekids_log 1.582e+00 1.507e+00 1.050 0.2939
## inter 4.552e-02 3.576e-02 1.273 0.2031
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 718.15 on 640 degrees of freedom
## Residual deviance: 531.84 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 643.84
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  37
##           1  10  19
##
##           Accuracy : 0.7314
##           95% CI : (0.6593, 0.7955)
##           No Information Rate : 0.68
##           P-Value [Acc > NIR] : 0.0826746
##
##           Kappa : 0.2926
##
## Mcnemar's Test P-Value : 0.0001491
##
##           Sensitivity : 0.9160
##           Specificity : 0.3393
##           Pos Pred Value : 0.7466
##           Neg Pred Value : 0.6552
##           Prevalence : 0.6800
##           Detection Rate : 0.6229
##           Detection Prevalence : 0.8343
##           Balanced Accuracy : 0.6276
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.768457382953181"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 119 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7685
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1309  -0.6655  -0.3615   0.5381   2.4175
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.433e+01  1.458e+01   1.669  0.09514 .
## KIDSDRIV      -3.232e-01  1.825e+00  -0.177  0.85947
## AGE           -1.420e-01  1.038e-01  -1.368  0.17123
## HOMEKIDS      -5.028e-01  6.815e-01  -0.738  0.46065
## YOJ           -2.164e-01  1.338e-01  -1.617  0.10589
## INCOME        -1.675e-05  1.281e-05  -1.307  0.19110
## HOME_VAL       1.046e-05  8.595e-06   1.217  0.22344
## TRAVTIME       2.499e-02  2.202e-02   1.135  0.25634
## BLUEBOOK       5.354e-05  3.994e-05   1.340  0.18012
```

```

## TIF -4.456e-02 9.090e-02 -0.490 0.62397
## OLDCLAIM 5.494e-06 2.172e-05 0.253 0.80033
## CLM_FREQ -3.372e-01 1.077e+00 -0.313 0.75419
## MVR_PTS -1.393e-01 1.751e-01 -0.795 0.42656
## CAR_AGE 6.482e-02 7.029e-02 0.922 0.35646
## PARENT1_Yes 8.239e-01 4.747e-01 1.736 0.08260 .
## MSTATUS_Yes -6.244e-01 3.296e-01 -1.894 0.05818 .
## SEX_z_F -9.413e-01 5.028e-01 -1.872 0.06120 .
## EDUCATION_.High.School -5.170e-01 8.156e-01 -0.634 0.52619
## EDUCATION_Bachelors -6.368e-01 6.987e-01 -0.911 0.36210
## EDUCATION_Masters 1.648e-01 5.830e-01 0.283 0.77738
## EDUCATION_z_High.School -4.483e-01 7.525e-01 -0.596 0.55138
## JOB_ -1.275e+00 7.099e-01 -1.796 0.07249 .
## JOB_Clerical -1.287e-01 4.725e-01 -0.272 0.78526
## JOB_Doctor -1.261e+00 9.857e-01 -1.279 0.20081
## JOB_Home.Maker -3.839e-01 6.632e-01 -0.579 0.56271
## JOB_Lawyer -1.007e+00 6.341e-01 -1.588 0.11223
## JOB_Manager -9.647e-01 4.875e-01 -1.979 0.04785 *
## JOB_Student -1.002e+00 7.219e-01 -1.387 0.16535
## JOB_z_Blue.Collar -1.009e-01 4.587e-01 -0.220 0.82587
## CAR_USE_Commercial 5.875e-01 3.573e-01 1.644 0.10014
## CAR_TYPE_Panel.Truck 1.230e-02 6.336e-01 0.019 0.98451
## CAR_TYPE_Pickup 8.045e-01 3.993e-01 2.015 0.04391 *
## CAR_TYPE_Sports.Car 1.747e+00 5.511e-01 3.170 0.00153 **
## CAR_TYPE_Van 2.558e-01 4.722e-01 0.542 0.58795
## CAR_TYPE_z_SUV 1.931e+00 4.883e-01 3.955 7.64e-05 ***
## RED_CAR_no -8.305e-02 3.309e-01 -0.251 0.80186
## REVOKED_Yes 6.287e-01 4.042e-01 1.555 0.11991
## URBANICITY_z_Highly.Rural..Rural -2.373e+00 4.086e-01 -5.809 6.30e-09 ***
## YOJ_NA -5.420e-01 4.363e-01 -1.242 0.21409
## INCOME_NA 5.098e-01 5.559e-01 0.917 0.35918
## CAR_AGE_NA 2.406e-01 5.415e-01 0.444 0.65684
## HOME_VAL_NA -1.729e-01 2.887e-01 -0.599 0.54931
## ageSquared 1.536e-03 1.133e-03 1.355 0.17550
## yojSquared 9.894e-03 6.993e-03 1.415 0.15713
## income_log 1.319e-01 2.702e-01 0.488 0.62543
## homeval_log -1.732e+00 1.346e+00 -1.287 0.19798
## travtime_log 1.002e-01 6.438e-01 0.156 0.87638
## bluebook_log -3.700e-01 4.728e-01 -0.783 0.43385
## carage_log -7.101e-01 4.614e-01 -1.539 0.12383
## oldclaim_log -3.151e-02 1.750e-01 -0.180 0.85710
## clm_freq_log 1.294e+00 3.285e+00 0.394 0.69373
## mvr_pts_log 5.334e-01 5.089e-01 1.048 0.29457
## tif_log -5.354e-02 5.154e-01 -0.104 0.91726
## kidsdriv_log -8.929e-02 2.216e+00 -0.040 0.96785
## homekids_log 1.038e+00 1.473e+00 0.704 0.48136
## inter 2.228e-02 3.410e-02 0.653 0.51356
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 748.85 on 642 degrees of freedom
## Residual deviance: 545.88 on 587 degrees of freedom

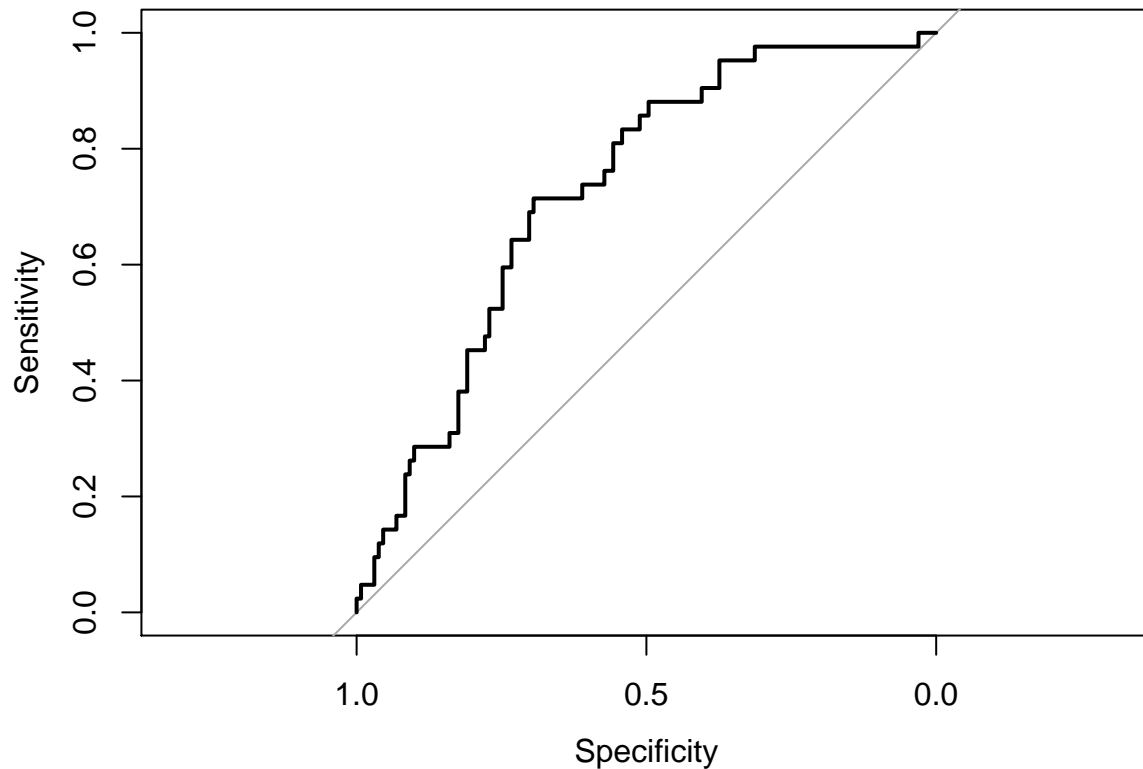
```

```

## (1 observation deleted due to missingness)
## AIC: 657.88
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  30
##           1  13  12
##
##           Accuracy : 0.7514
##           95% CI : (0.6802, 0.8139)
##           No Information Rate : 0.7572
##           P-Value [Acc > NIR] : 0.61026
##
##           Kappa : 0.2162
##
## Mcnemar's Test P-Value : 0.01469
##
##           Sensitivity : 0.9008
##           Specificity : 0.2857
##           Pos Pred Value : 0.7973
##           Neg Pred Value : 0.4800
##           Prevalence : 0.7572
##           Detection Rate : 0.6821
##           Detection Prevalence : 0.8555
##           Balanced Accuracy : 0.5932
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.727371864776445"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.7274
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2166  -0.6511  -0.3698   0.5100   3.1724
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.790e+01  1.471e+01   1.217  0.22378
## KIDSDRIV      -3.652e+00  1.940e+00  -1.883  0.05976 .
## AGE          -1.538e-01  1.130e-01  -1.361  0.17344
## HOMEKIDS      -3.093e-01  6.651e-01  -0.465  0.64190
## YOJ           -1.982e-01  1.306e-01  -1.517  0.12925
## INCOME        -1.542e-05  1.374e-05  -1.122  0.26191
## HOME_VAL       6.020e-06  9.276e-06   0.649  0.51632
## TRAVTIME      4.530e-02  2.169e-02   2.088  0.03679 *
## BLUEBOOK      4.987e-05  3.934e-05   1.267  0.20498
```

```

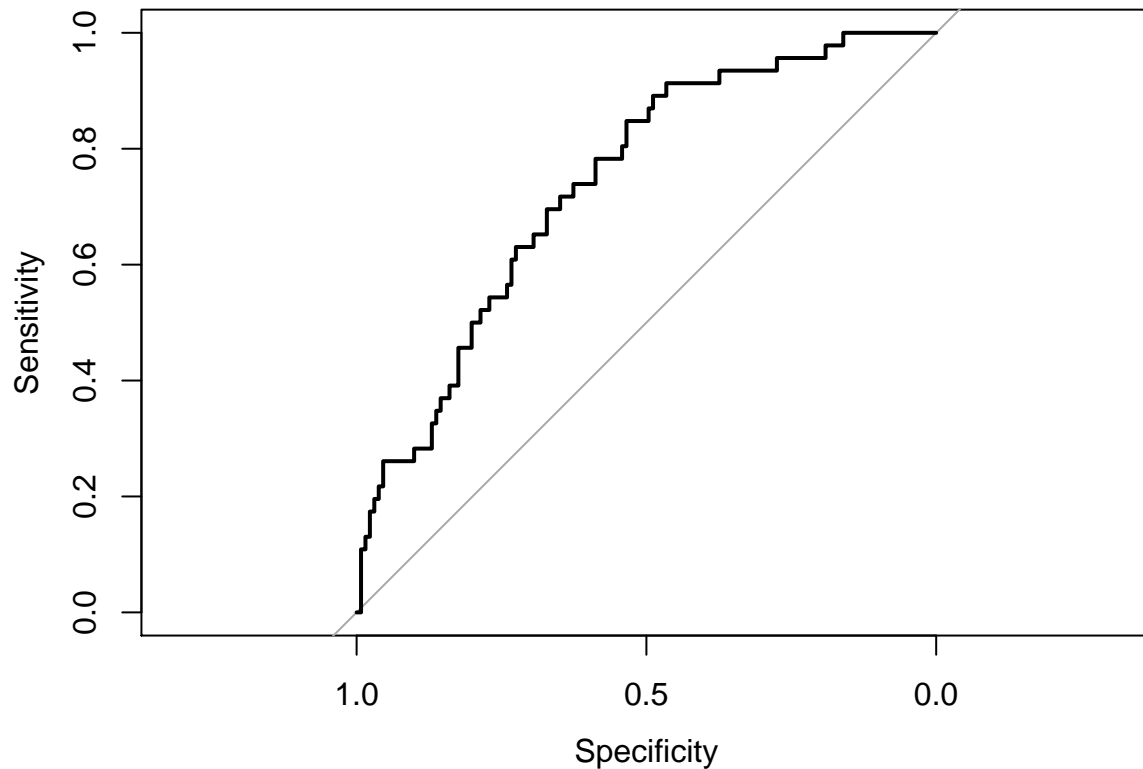
## TIF -6.096e-02 9.122e-02 -0.668 0.50398
## OLDCLAIM 5.374e-06 2.141e-05 0.251 0.80183
## CLM_FREQ 5.753e-01 9.444e-01 0.609 0.54244
## MVR_PTS -1.452e-02 1.690e-01 -0.086 0.93152
## CAR_AGE -2.703e-02 6.974e-02 -0.388 0.69831
## PARENT1_Yes 5.313e-01 4.748e-01 1.119 0.26316
## MSTATUS_Yes -4.481e-01 3.340e-01 -1.342 0.17972
## SEX_z_F -7.924e-01 4.660e-01 -1.701 0.08903
## EDUCATION_.High.School -4.640e-01 7.998e-01 -0.580 0.56181
## EDUCATION_Bachelors -3.958e-01 6.806e-01 -0.582 0.56085
## EDUCATION_Masters 9.106e-01 6.069e-01 1.500 0.13352
## EDUCATION_z_High.School 1.062e-01 7.277e-01 0.146 0.88396
## JOB_ -9.442e-01 6.924e-01 -1.364 0.17267
## JOB_Clerical 8.315e-02 4.892e-01 0.170 0.86504
## JOB_Doctor 3.587e-01 9.095e-01 0.394 0.69332
## JOB_Home.Maker -2.439e-01 6.693e-01 -0.364 0.71558
## JOB_Lawyer -7.036e-01 6.777e-01 -1.038 0.29915
## JOB_Manager -7.488e-01 5.100e-01 -1.468 0.14201
## JOB_Student -9.210e-01 7.408e-01 -1.243 0.21378
## JOB_z_Blue.Collar 4.288e-01 4.529e-01 0.947 0.34371
## CAR_USE_Commercial 5.582e-01 3.492e-01 1.598 0.10997
## CAR_TYPE_Panel.Truck -1.884e-01 6.383e-01 -0.295 0.76788
## CAR_TYPE_Pickup 1.196e+00 4.009e-01 2.983 0.00285 **
## CAR_TYPE_Sports.Car 1.787e+00 5.427e-01 3.293 0.00099 ***
## CAR_TYPE_Van 5.989e-01 4.717e-01 1.270 0.20422
## CAR_TYPE_z_SUV 2.013e+00 4.778e-01 4.213 2.52e-05 ***
## RED_CAR_no -4.259e-01 3.585e-01 -1.188 0.23490
## REVOKED_Yes 6.211e-01 4.149e-01 1.497 0.13439
## URBANICITY_z_Highly.Rural..Rural -2.423e+00 4.163e-01 -5.819 5.92e-09 ***
## YOJ_NA -2.071e-01 4.040e-01 -0.513 0.60827
## INCOME_NA 4.482e-01 5.408e-01 0.829 0.40724
## CAR_AGE_NA -1.650e-01 4.701e-01 -0.351 0.72559
## HOME_VAL_NA 5.060e-02 2.861e-01 0.177 0.85962
## ageSquared 1.639e-03 1.235e-03 1.327 0.18457
## yojSquared 7.643e-03 6.699e-03 1.141 0.25389
## income_log 1.631e-01 2.989e-01 0.546 0.58533
## homeval_log -1.121e+00 1.389e+00 -0.807 0.41960
## travtime_log -6.105e-01 6.012e-01 -1.015 0.30990
## bluebook_log -3.167e-01 4.601e-01 -0.688 0.49125
## carage_log -2.520e-01 4.572e-01 -0.551 0.58156
## oldclaim_log 1.336e-01 1.649e-01 0.810 0.41790
## clm_freq_log -1.802e+00 2.964e+00 -0.608 0.54307
## mvr_pts_log 2.126e-01 5.019e-01 0.423 0.67193
## tif_log 2.018e-01 5.199e-01 0.388 0.69794
## kidsdriv_log 1.028e+00 2.196e+00 0.468 0.63979
## homekids_log 9.529e-01 1.449e+00 0.657 0.51089
## inter 8.141e-02 3.581e-02 2.274 0.02298 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 738.28 on 638 degrees of freedom
## Residual deviance: 535.53 on 583 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 647.53
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  27
##           1  23  19
##
##           Accuracy : 0.7175
##           95% CI : (0.6451, 0.7825)
##           No Information Rate : 0.7401
##           P-Value [Acc > NIR] : 0.7813
##
##           Kappa : 0.2444
##
## Mcnemar's Test P-Value : 0.6714
##
##           Sensitivity : 0.8244
##           Specificity : 0.4130
##           Pos Pred Value : 0.8000
##           Neg Pred Value : 0.4524
##           Prevalence : 0.7401
##           Detection Rate : 0.6102
##           Detection Prevalence : 0.7627
##           Balanced Accuracy : 0.6187
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.740292067706605"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7403
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0636  -0.6541  -0.3432   0.3983   2.9418
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.243e+01  1.593e+01   0.780 0.435257
## KIDSDRIV      -2.582e-01  1.936e+00  -0.133 0.893926
## AGE           -1.043e-01  1.054e-01  -0.990 0.322193
## HOMEKIDS      -3.447e-01  6.703e-01  -0.514 0.607039
## YOJ           -2.077e-01  1.334e-01  -1.557 0.119372
## INCOME        -1.155e-05  1.350e-05  -0.856 0.392236
## HOME_VAL       6.441e-06  9.388e-06   0.686 0.492711
## TRAVTIME      2.981e-02  2.179e-02   1.368 0.171259
## BLUEBOOK      4.547e-05  4.016e-05   1.132 0.257552
```

```

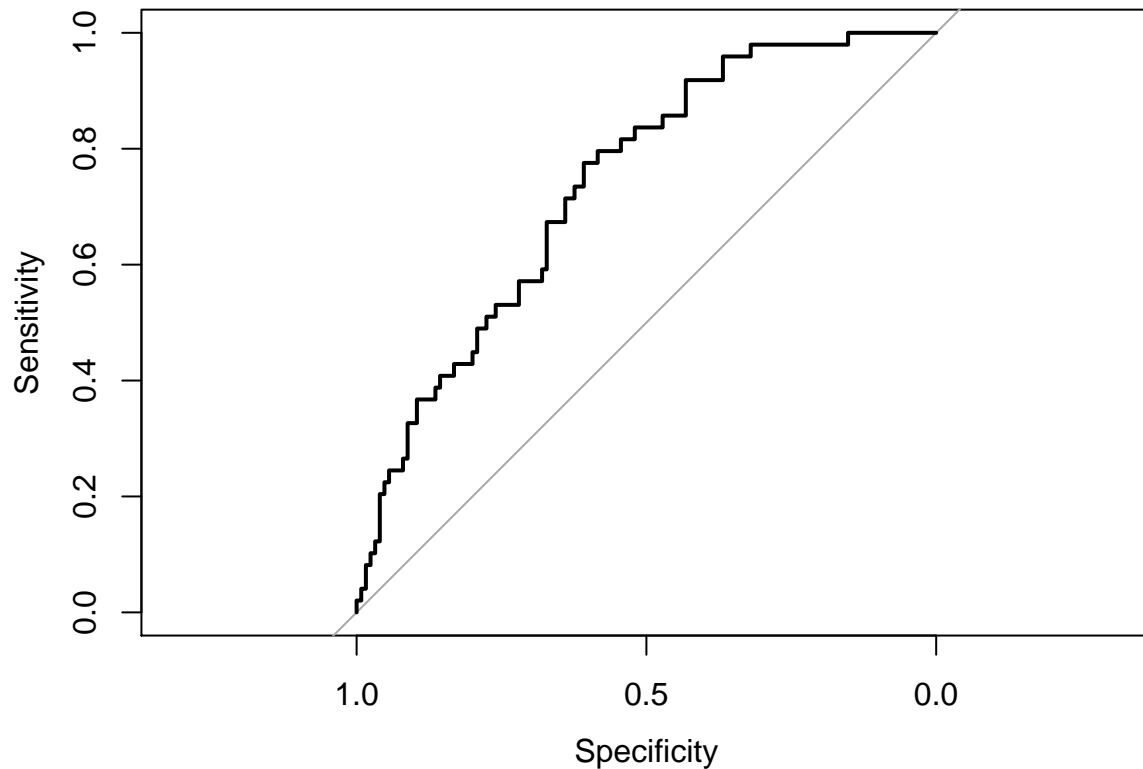
## TIF -4.475e-03 9.270e-02 -0.048 0.961495
## OLDCLAIM -5.793e-06 2.211e-05 -0.262 0.793303
## CLM_FREQ -1.861e-02 1.018e+00 -0.018 0.985419
## MVRPTS -8.380e-03 1.738e-01 -0.048 0.961549
## CAR_AGE -8.876e-03 7.172e-02 -0.124 0.901495
## PARENT1_Yes 9.383e-01 4.808e-01 1.952 0.050990 .
## MSTATUS_Yes -4.498e-01 3.404e-01 -1.321 0.186435
## SEX_z_F -1.141e+00 4.675e-01 -2.441 0.014638 *
## EDUCATION_.High.School 6.435e-02 8.423e-01 0.076 0.939097
## EDUCATION_Bachelors 6.420e-03 7.224e-01 0.009 0.992909
## EDUCATION_Masters 1.200e+00 6.430e-01 1.866 0.062096 .
## EDUCATION_z_High.School 2.653e-01 7.695e-01 0.345 0.730324
## JOB_ -1.151e+00 7.031e-01 -1.638 0.101486
## JOB_Clerical -2.942e-02 5.012e-01 -0.059 0.953184
## JOB_Doctor -3.105e-01 9.707e-01 -0.320 0.749076
## JOB_Home.Maker 3.117e-02 6.807e-01 0.046 0.963477
## JOB_Lawyer -1.480e+00 6.757e-01 -2.190 0.028544 *
## JOB_Manager -1.091e+00 4.887e-01 -2.232 0.025620 *
## JOB_Student -9.563e-01 7.371e-01 -1.297 0.194515
## JOB_z_Blue.Collar -4.578e-01 4.569e-01 -1.002 0.316366
## CAR_USE_Commercial 9.445e-01 3.603e-01 2.621 0.008760 **
## CAR_TYPE_Panel.Truck -4.557e-01 6.349e-01 -0.718 0.472841
## CAR_TYPE_Pickup 4.531e-01 4.042e-01 1.121 0.262384
## CAR_TYPE_Sports.Car 1.615e+00 5.440e-01 2.969 0.002984 **
## CAR_TYPE_Van 5.467e-02 4.835e-01 0.113 0.909983
## CAR_TYPE_z_SUV 1.599e+00 4.689e-01 3.411 0.000648 ***
## RED_CAR_no -6.247e-03 3.365e-01 -0.019 0.985190
## REVOKED_Yes 9.558e-01 4.059e-01 2.355 0.018524 *
## URBANICITY_z_Highly.Rural..Rural -2.456e+00 4.347e-01 -5.649 1.61e-08 ***
## YOJ_NA -1.314e-01 4.460e-01 -0.295 0.768314
## INCOME_NA 2.153e-01 5.647e-01 0.381 0.702972
## CAR_AGE_NA 1.313e-02 5.084e-01 0.026 0.979399
## HOME_VAL_NA -3.602e-01 2.882e-01 -1.250 0.211303
## ageSquared 1.273e-03 1.135e-03 1.121 0.262176
## yojSquared 1.099e-02 6.856e-03 1.602 0.109083
## income_log 4.356e-03 2.994e-01 0.015 0.988390
## homeval_log -8.849e-01 1.522e+00 -0.582 0.560901
## travtime_log -1.032e-01 6.166e-01 -0.167 0.867091
## bluebook_log -1.330e-01 4.862e-01 -0.274 0.784352
## carage_log -3.794e-01 4.686e-01 -0.810 0.418205
## oldclaim_log 4.349e-02 1.683e-01 0.258 0.796086
## clm_freq_log 4.120e-01 3.128e+00 0.132 0.895226
## mvr_pts_log 1.167e-01 5.106e-01 0.229 0.819252
## tif_log -2.693e-01 5.259e-01 -0.512 0.608619
## kidsdriv_log 1.227e+00 2.312e+00 0.530 0.595799
## homekids_log 8.766e-01 1.462e+00 0.600 0.548684
## inter 4.104e-03 3.723e-02 0.110 0.912209
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.87 on 641 degrees of freedom
## Residual deviance: 527.62 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 639.62
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  31
##           1  16  18
##
##           Accuracy : 0.7299
##           95% CI : (0.6575, 0.7943)
##           No Information Rate : 0.7184
##           P-Value [Acc > NIR] : 0.40472
##
##           Kappa : 0.2639
##
## Mcnemar's Test P-Value : 0.04114
##
##           Sensitivity : 0.8720
##           Specificity : 0.3673
##           Pos Pred Value : 0.7786
##           Neg Pred Value : 0.5294
##           Prevalence : 0.7184
##           Detection Rate : 0.6264
##           Detection Prevalence : 0.8046
##           Balanced Accuracy : 0.6197
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.736979591836735"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 125 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.737
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3246  -0.6608  -0.3630   0.5121   2.5221
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.379e+01  1.483e+01   1.604  0.108757
## KIDSDRIV      -1.291e+00  1.979e+00  -0.652  0.514171
## AGE           -3.935e-02  1.109e-01  -0.355  0.722726
## HOMEKIDS      -1.106e+00  6.603e-01  -1.674  0.094090 .
## YOJ           -2.003e-01  1.341e-01  -1.494  0.135164
## INCOME        -3.056e-05  1.335e-05  -2.288  0.022128 *
## HOME_VAL       1.840e-05  8.967e-06   2.052  0.040180 *
## TRAVTIME       3.488e-02  2.164e-02   1.612  0.106960
## BLUEBOOK       5.556e-05  3.894e-05   1.427  0.153610
```

```

## TIF 1.256e-03 9.159e-02 0.014 0.989061
## OLDCLAIM 6.085e-06 2.035e-05 0.299 0.764884
## CLM_FREQ 1.997e-01 9.439e-01 0.212 0.832412
## MVR_PTS 9.013e-02 1.763e-01 0.511 0.609246
## CAR_AGE 3.506e-02 7.098e-02 0.494 0.621310
## PARENT1_Yes 4.790e-01 4.704e-01 1.018 0.308540
## MSTATUS_Yes -7.483e-01 3.436e-01 -2.178 0.029394 *
## SEX_z_F -4.510e-01 4.715e-01 -0.957 0.338743
## EDUCATION_.High.School 9.263e-01 8.074e-01 1.147 0.251287
## EDUCATION_Bachelors 5.193e-01 6.935e-01 0.749 0.453980
## EDUCATION_Masters 1.146e+00 6.168e-01 1.858 0.063192 .
## EDUCATION_z_High.School 8.368e-01 7.462e-01 1.121 0.262160
## JOB_ -6.858e-01 7.090e-01 -0.967 0.333400
## JOB_Clerical -3.971e-01 4.911e-01 -0.809 0.418757
## JOB_Doctor 4.042e-01 9.702e-01 0.417 0.676957
## JOB_Home.Maker -6.010e-01 6.969e-01 -0.862 0.388486
## JOB_Lawyer -6.188e-01 6.960e-01 -0.889 0.373893
## JOB_Manager -8.761e-01 4.824e-01 -1.816 0.069356 .
## JOB_Student -8.721e-01 7.497e-01 -1.163 0.244736
## JOB_z_Blue.Collar -2.812e-01 4.553e-01 -0.618 0.536843
## CAR_USE_Commercial 7.229e-01 3.480e-01 2.077 0.037761 *
## CAR_TYPE_Panel.Truck 3.110e-01 6.339e-01 0.491 0.623727
## CAR_TYPE_Pickup 8.397e-01 4.065e-01 2.066 0.038858 *
## CAR_TYPE_Sports.Car 2.059e+00 5.196e-01 3.962 7.43e-05 ***
## CAR_TYPE_Van 6.876e-01 4.825e-01 1.425 0.154140
## CAR_TYPE_z_SUV 1.732e+00 4.797e-01 3.611 0.000305 ***
## RED_CAR_no -3.490e-01 3.418e-01 -1.021 0.307182
## REVOKED_Yes 4.376e-01 4.202e-01 1.041 0.297731
## URBANICITY_z_Highly.Rural..Rural -2.376e+00 4.364e-01 -5.446 5.16e-08 ***
## YOJ_NA -4.302e-01 4.342e-01 -0.991 0.321769
## INCOME_NA 4.785e-01 5.638e-01 0.849 0.396107
## CAR_AGE_NA -3.612e-01 5.096e-01 -0.709 0.478436
## HOME_VAL_NA -1.452e-01 2.892e-01 -0.502 0.615685
## ageSquared 3.178e-04 1.207e-03 0.263 0.792368
## yojSquared 8.802e-03 7.004e-03 1.257 0.208901
## income_log 1.012e-01 2.953e-01 0.343 0.731742
## homeval_log -1.826e+00 1.394e+00 -1.310 0.190241
## travtime_log -6.284e-01 5.951e-01 -1.056 0.291022
## bluebook_log -3.333e-01 4.514e-01 -0.738 0.460252
## carage_log -5.361e-01 4.640e-01 -1.155 0.247915
## oldclaim_log 8.533e-03 1.596e-01 0.053 0.957358
## clm_freq_log -1.385e-01 2.934e+00 -0.047 0.962348
## mvr_pts_log -7.301e-02 5.129e-01 -0.142 0.886810
## tif_log -2.245e-01 5.176e-01 -0.434 0.664460
## kidsdriv_log -4.636e-01 2.363e+00 -0.196 0.844443
## homekids_log 2.582e+00 1.421e+00 1.818 0.069110 .
## inter 4.895e-02 3.595e-02 1.362 0.173343
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 743.59 on 640 degrees of freedom
## Residual deviance: 546.02 on 585 degrees of freedom

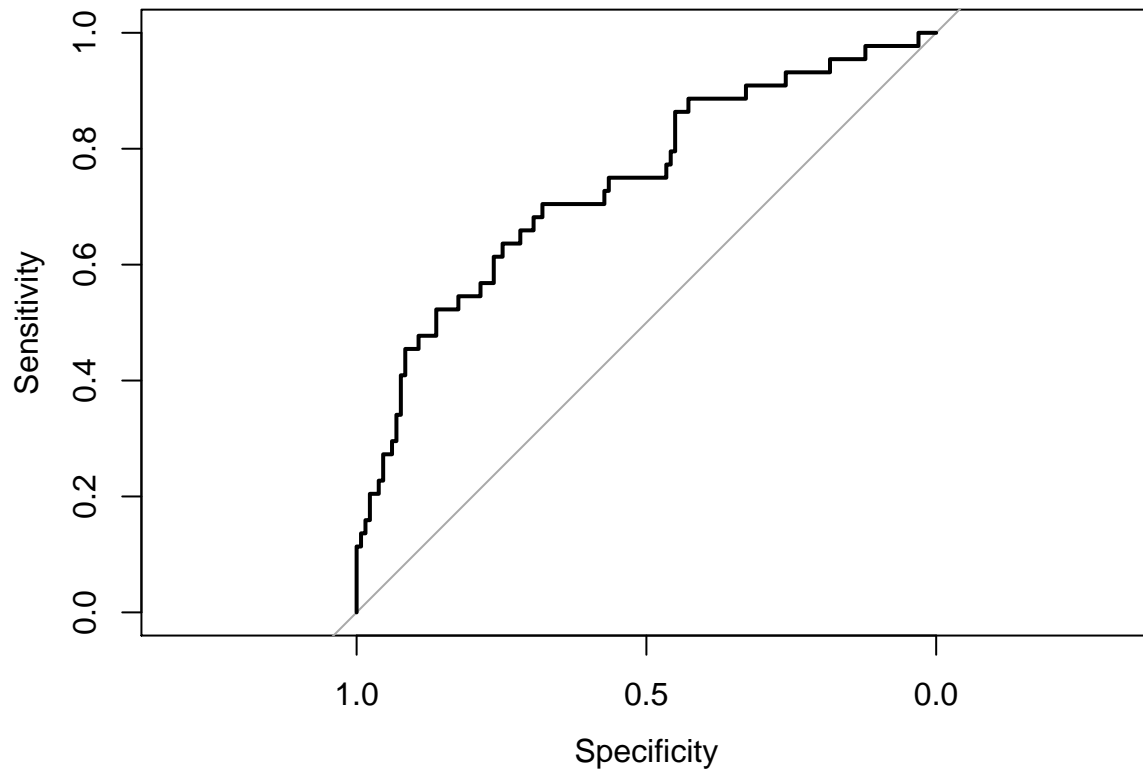
```



```

## AIC: 658.02
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  23
##           1  15  21
##
##           Accuracy : 0.7829
##           95% CI : (0.7144, 0.8415)
##           No Information Rate : 0.7486
##           P-Value [Acc > NIR] : 0.1692
##
##           Kappa : 0.3861
##
## Mcnemar's Test P-Value : 0.2561
##
##           Sensitivity : 0.8855
##           Specificity : 0.4773
##           Pos Pred Value : 0.8345
##           Neg Pred Value : 0.5833
##           Prevalence : 0.7486
##           Detection Rate : 0.6629
##           Detection Prevalence : 0.7943
##           Balanced Accuracy : 0.6814
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.740111034004164"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7401
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3571  -0.6349  -0.3452   0.5897   2.9253
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.164e+01  1.536e+01   1.409 0.158892
## KIDSDRIV      -1.382e+00  1.939e+00  -0.713 0.476070
## AGE           -6.322e-02  1.076e-01  -0.587 0.556938
## HOMEKIDS      -1.611e-02  7.146e-01  -0.023 0.982016
## YOJ           -2.714e-01  1.332e-01  -2.037 0.041618 *
## INCOME        -3.318e-05  1.396e-05  -2.377 0.017455 *
## HOME_VAL       1.822e-05  9.734e-06   1.872 0.061187 .
## TRAVTIME       2.626e-02  2.060e-02   1.275 0.202375
## BLUEBOOK       6.172e-05  3.994e-05   1.545 0.122249
```

```

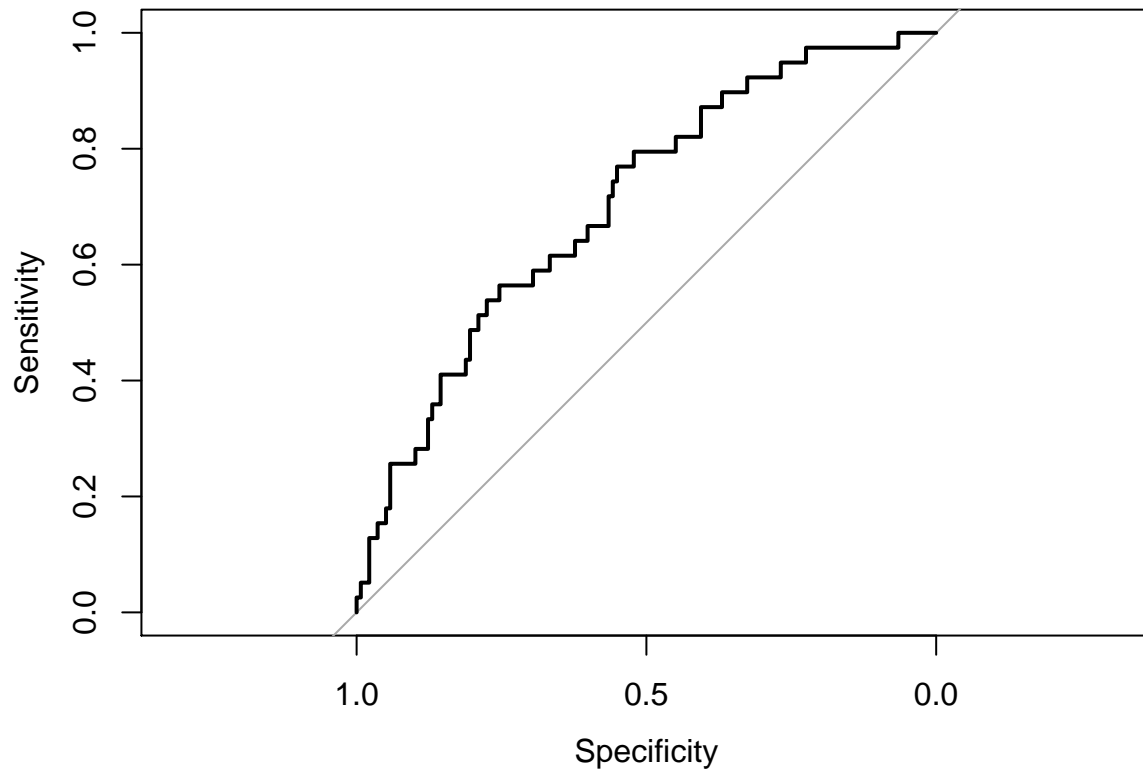
## TIF -3.416e-02 8.873e-02 -0.385 0.700281
## OLDCLAIM -3.225e-06 2.427e-05 -0.133 0.894300
## CLM_FREQ 7.077e-01 9.569e-01 0.740 0.459588
## MVRPTS 5.770e-02 1.740e-01 0.332 0.740233
## CAR_AGE -3.944e-02 7.834e-02 -0.503 0.614628
## PARENT1_Yes 8.626e-01 4.828e-01 1.787 0.074007 .
## MSTATUS_Yes -7.071e-01 3.293e-01 -2.147 0.031756 *
## SEX_z_F -5.004e-01 4.574e-01 -1.094 0.273956
## EDUCATION_.High.School 1.184e-01 8.601e-01 0.138 0.890508
## EDUCATION_Bachelors 1.490e-01 7.417e-01 0.201 0.840817
## EDUCATION_Masters 1.218e+00 6.312e-01 1.930 0.053552 .
## EDUCATION_z_High.School 7.514e-01 7.963e-01 0.944 0.345357
## JOB_ -1.428e+00 7.367e-01 -1.938 0.052594 .
## JOB_Clerical -4.823e-01 4.766e-01 -1.012 0.311567
## JOB_Doctor 1.896e-01 9.519e-01 0.199 0.842152
## JOB_Home.Maker -8.881e-01 6.603e-01 -1.345 0.178643
## JOB_Lawyer -1.136e+00 6.613e-01 -1.718 0.085879 .
## JOB_Manager -1.271e+00 5.034e-01 -2.524 0.011610 *
## JOB_Student -1.056e+00 7.570e-01 -1.395 0.163154
## JOB_z_Blue.Collar -6.772e-01 4.532e-01 -1.494 0.135113
## CAR_USE_Commercial 8.556e-01 3.564e-01 2.401 0.016368 *
## CAR_TYPE_Panel.Truck 6.814e-02 6.290e-01 0.108 0.913734
## CAR_TYPE_Pickup 9.107e-01 4.130e-01 2.205 0.027463 *
## CAR_TYPE_Sports.Car 1.731e+00 5.242e-01 3.302 0.000959 ***
## CAR_TYPE_Van 4.867e-01 4.727e-01 1.030 0.303131
## CAR_TYPE_z_SUV 1.651e+00 4.563e-01 3.617 0.000298 ***
## RED_CAR_no -1.994e-01 3.424e-01 -0.582 0.560387
## REVOKED_Yes 5.263e-01 4.536e-01 1.160 0.246023
## URBANICITY_z_Highly.Rural..Rural -2.268e+00 4.011e-01 -5.655 1.56e-08 ***
## YOJ_NA 9.144e-02 4.511e-01 0.203 0.839366
## INCOME_NA 5.213e-01 5.827e-01 0.895 0.370984
## CAR_AGE_NA 1.761e-01 4.800e-01 0.367 0.713669
## HOME_VAL_NA -2.966e-01 2.868e-01 -1.034 0.301096
## ageSquared 6.100e-04 1.158e-03 0.527 0.598433
## yojSquared 1.508e-02 6.768e-03 2.228 0.025858 *
## income_log 9.077e-02 2.948e-01 0.308 0.758160
## homeval_log -1.817e+00 1.493e+00 -1.217 0.223588
## travtime_log -4.850e-02 5.917e-01 -0.082 0.934672
## bluebook_log -3.293e-01 4.833e-01 -0.681 0.495655
## carage_log 2.053e-02 4.900e-01 0.042 0.966583
## oldclaim_log 1.399e-01 1.679e-01 0.833 0.404773
## clm_freq_log -1.889e+00 2.978e+00 -0.634 0.525810
## mvr_pts_log 1.097e-02 5.073e-01 0.022 0.982741
## tif_log -1.070e-01 5.054e-01 -0.212 0.832345
## kidsdriv_log 2.543e+00 2.339e+00 1.087 0.276957
## homekids_log -6.359e-02 1.541e+00 -0.041 0.967084
## inter 1.726e-02 3.278e-02 0.526 0.598596
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 752.21 on 638 degrees of freedom
## Residual deviance: 537.33 on 583 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 649.33
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 119  25
##           1  19  14
##
##           Accuracy : 0.7514
##           95% CI : (0.681, 0.8132)
##           No Information Rate : 0.7797
##           P-Value [Acc > NIR] : 0.8408
##
##           Kappa : 0.2342
##
## Mcnemar's Test P-Value : 0.4510
##
##           Sensitivity : 0.8623
##           Specificity : 0.3590
##           Pos Pred Value : 0.8264
##           Neg Pred Value : 0.4242
##           Prevalence : 0.7797
##           Detection Rate : 0.6723
##           Detection Prevalence : 0.8136
##           Balanced Accuracy : 0.6106
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.705128205128205"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 39 cases (dfPred_raw$class 1).
## Area under the curve: 0.7051
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1149  -0.6971  -0.3782   0.6085   2.9701
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.190e+01  1.422e+01   0.837  0.402718
## KIDSDRIV      -3.298e+00  2.029e+00  -1.625  0.104162
## AGE           -5.476e-02  1.046e-01  -0.523  0.600774
## HOMEKIDS       5.352e-02  6.663e-01   0.080  0.935974
## YOJ           -2.075e-01  1.233e-01  -1.684  0.092233 .
## INCOME        -5.874e-06  1.331e-05  -0.441  0.658901
## HOME_VAL      -2.361e-07  8.967e-06  -0.026  0.978995
## TRAVTIME       4.259e-02  2.249e-02   1.894  0.058286 .
## BLUEBOOK       5.378e-05  3.700e-05   1.454  0.146051
```

```

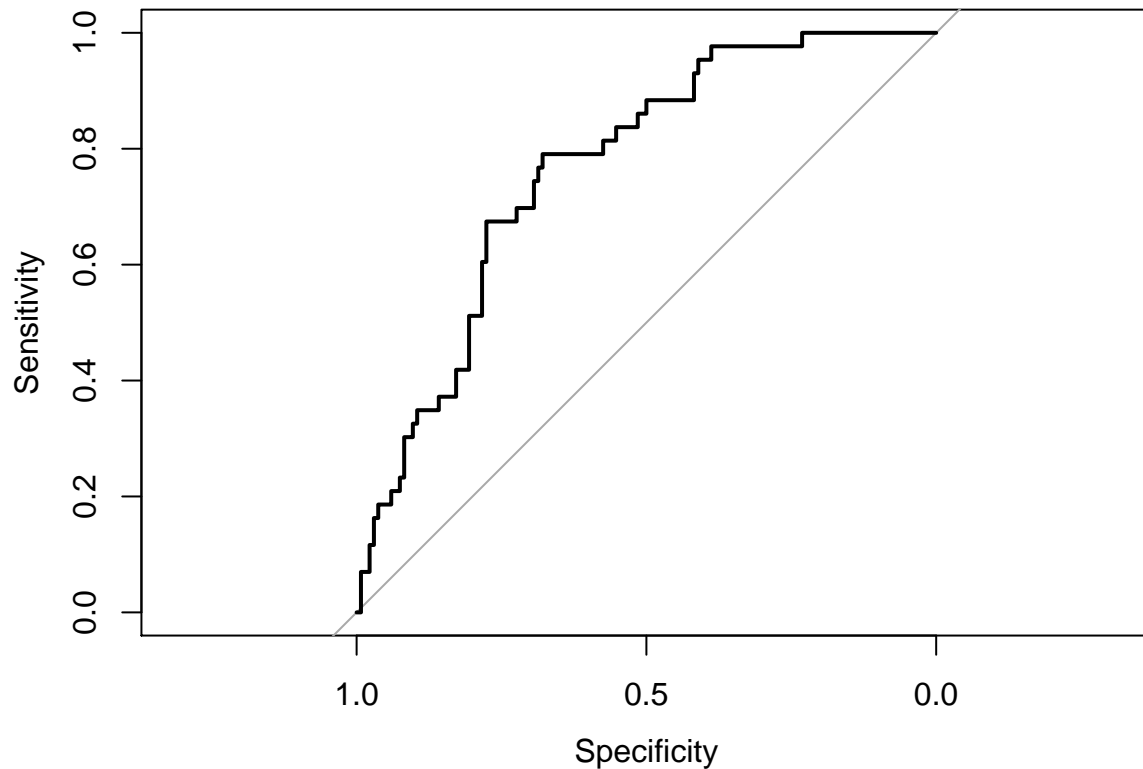
## TIF 5.622e-03 8.560e-02 0.066 0.947637
## OLDCLAIM -6.942e-06 2.056e-05 -0.338 0.735652
## CLM_FREQ 2.556e-01 9.373e-01 0.273 0.785105
## MVR_PTS 8.165e-02 1.611e-01 0.507 0.612269
## CAR_AGE -1.459e-02 6.757e-02 -0.216 0.829053
## PARENT1_Yes 6.071e-01 4.571e-01 1.328 0.184193
## MSTATUS_Yes -5.877e-01 3.270e-01 -1.797 0.072339 .
## SEX_z_F -6.916e-01 4.488e-01 -1.541 0.123286
## EDUCATION_.High.School -1.212e+00 7.759e-01 -1.562 0.118203
## EDUCATION_Bachelors -6.560e-01 6.350e-01 -1.033 0.301627
## EDUCATION_Masters 4.129e-01 5.489e-01 0.752 0.451843
## EDUCATION_z_High.School -6.321e-01 6.984e-01 -0.905 0.365418
## JOB_ -1.260e+00 6.639e-01 -1.899 0.057622 .
## JOB_Clerical 1.142e-02 4.813e-01 0.024 0.981063
## JOB_Doctor -5.175e-01 9.043e-01 -0.572 0.567134
## JOB_Home.Maker -6.183e-01 6.573e-01 -0.941 0.346832
## JOB_Lawyer -1.288e+00 6.311e-01 -2.041 0.041292 *
## JOB_Manager -9.294e-01 4.591e-01 -2.024 0.042924 *
## JOB_Student -1.166e+00 7.237e-01 -1.611 0.107142
## JOB_z_Blue.Collar -1.686e-01 4.478e-01 -0.376 0.706580
## CAR_USE_Commercial 9.046e-01 3.538e-01 2.557 0.010554 *
## CAR_TYPE_Panel.Truck 1.206e-01 5.991e-01 0.201 0.840526
## CAR_TYPE_Pickup 8.360e-01 3.985e-01 2.098 0.035898 *
## CAR_TYPE_Sports.Car 1.813e+00 5.236e-01 3.463 0.000535 ***
## CAR_TYPE_Van 3.539e-01 4.670e-01 0.758 0.448663
## CAR_TYPE_z_SUV 1.540e+00 4.640e-01 3.320 0.000901 ***
## RED_CAR_no -8.181e-02 3.384e-01 -0.242 0.808971
## REVOKED_Yes 7.037e-01 3.972e-01 1.772 0.076454 .
## URBANICITY_z_Highly.Rural..Rural -2.206e+00 4.146e-01 -5.320 1.04e-07 ***
## YOJ_NA -4.035e-01 4.345e-01 -0.929 0.353069
## INCOME_NA 3.544e-02 4.904e-01 0.072 0.942389
## CAR_AGE_NA -4.229e-01 4.521e-01 -0.936 0.349514
## HOME_VAL_NA 3.520e-03 2.798e-01 0.013 0.989965
## ageSquared 5.817e-04 1.140e-03 0.510 0.609833
## yojSquared 1.056e-02 6.451e-03 1.636 0.101782
## income_log 1.378e-01 3.125e-01 0.441 0.659196
## homeval_log -5.558e-01 1.379e+00 -0.403 0.687006
## travtime_log -5.278e-01 6.170e-01 -0.855 0.392348
## bluebook_log -3.923e-01 4.373e-01 -0.897 0.369759
## carage_log -2.470e-01 4.532e-01 -0.545 0.585697
## oldclaim_log 9.914e-02 1.618e-01 0.613 0.540011
## clm_freq_log -7.894e-01 2.920e+00 -0.270 0.786919
## mvr_pts_log -1.980e-01 4.790e-01 -0.413 0.679362
## tif_log -2.879e-01 4.953e-01 -0.581 0.561094
## kidsdriv_log 2.071e+00 2.272e+00 0.911 0.362059
## homekids_log 1.478e-01 1.448e+00 0.102 0.918679
## inter 6.260e-02 3.694e-02 1.695 0.090096 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 744.35 on 638 degrees of freedom
## Residual deviance: 557.53 on 583 degrees of freedom

```

```

## AIC: 669.53
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  28
##           1  16  15
##
##           Accuracy : 0.7514
##           95% CI : (0.681, 0.8132)
##           No Information Rate : 0.7571
##           P-Value [Acc > NIR] : 0.60902
##
##           Kappa : 0.2535
##
## Mcnemar's Test P-Value : 0.09725
##
##           Sensitivity : 0.8806
##           Specificity : 0.3488
##           Pos Pred Value : 0.8082
##           Neg Pred Value : 0.4839
##           Prevalence : 0.7571
##           Detection Rate : 0.6667
##           Detection Prevalence : 0.8249
##           Balanced Accuracy : 0.6147
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.766053453661923"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7661
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1022  -0.6842  -0.3307   0.3992   2.6406
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.808e+01  1.512e+01   1.196  0.231781
## KIDSDRIV      -3.972e+00  2.128e+00  -1.867  0.061966 .
## AGE           -1.419e-01  1.074e-01  -1.322  0.186297
## HOMEKIDS      -4.642e-01  6.777e-01  -0.685  0.493393
## YOJ           -1.902e-01  1.398e-01  -1.361  0.173481
## INCOME        -1.838e-05  1.435e-05  -1.281  0.200337
## HOME_VAL       1.037e-05  9.501e-06   1.091  0.275168
## TRAVTIME       4.982e-02  2.119e-02   2.351  0.018724 *
## BLUEBOOK       4.514e-05  3.824e-05   1.180  0.237831
```



```

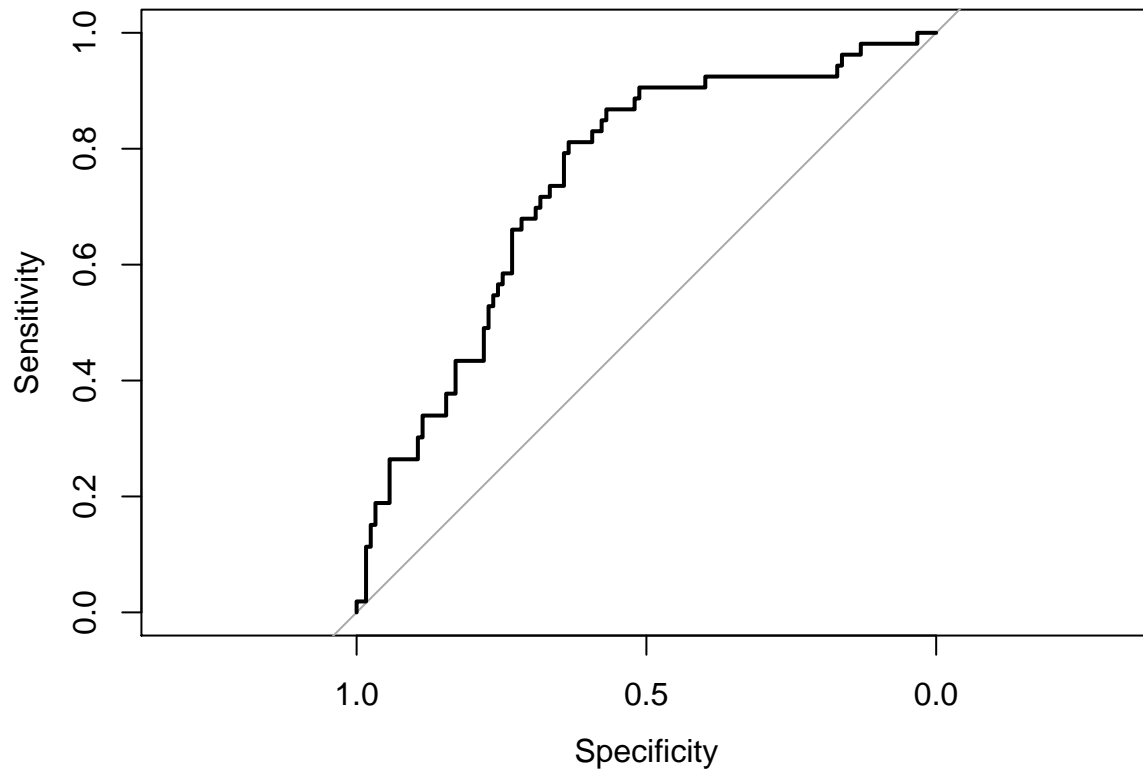
## TIF -4.334e-02 9.147e-02 -0.474 0.635595
## OLDCLAIM 1.256e-07 2.246e-05 0.006 0.995539
## CLM_FREQ -4.670e-02 9.361e-01 -0.050 0.960208
## MVR_PTS -1.525e-01 1.776e-01 -0.859 0.390461
## CAR_AGE -4.037e-02 7.050e-02 -0.573 0.566894
## PARENT1_Yes 4.565e-01 4.753e-01 0.961 0.336769
## MSTATUS_Yes -7.153e-01 3.391e-01 -2.110 0.034879 *
## SEX_z_F -2.443e-01 4.746e-01 -0.515 0.606744
## EDUCATION_.High.School 1.691e-01 8.122e-01 0.208 0.835063
## EDUCATION_Bachelors 6.462e-02 6.838e-01 0.095 0.924703
## EDUCATION_Masters 5.314e-01 5.728e-01 0.928 0.353536
## EDUCATION_z_High.School 7.037e-01 7.385e-01 0.953 0.340636
## JOB_ -9.834e-01 6.964e-01 -1.412 0.157901
## JOB_Clerical -7.530e-01 4.873e-01 -1.545 0.122313
## JOB_Doctor -1.942e-01 9.233e-01 -0.210 0.833443
## JOB_Home.Maker -9.416e-01 7.011e-01 -1.343 0.179225
## JOB_Lawyer -8.024e-01 6.879e-01 -1.166 0.243432
## JOB_Manager -1.109e+00 4.976e-01 -2.229 0.025839 *
## JOB_Student -1.411e+00 7.583e-01 -1.861 0.062753 .
## JOB_z_Blue.Collar -4.206e-01 4.410e-01 -0.954 0.340258
## CAR_USE_Commercial 7.190e-01 3.550e-01 2.025 0.042847 *
## CAR_TYPE_Panel.Truck 1.695e-01 6.473e-01 0.262 0.793418
## CAR_TYPE_Pickup 1.087e+00 4.128e-01 2.633 0.008466 **
## CAR_TYPE_Sports.Car 1.888e+00 5.284e-01 3.573 0.000352 ***
## CAR_TYPE_Van 7.068e-01 4.705e-01 1.502 0.133025
## CAR_TYPE_z_SUV 1.741e+00 4.672e-01 3.727 0.000194 ***
## RED_CAR_no -5.157e-01 3.559e-01 -1.449 0.147362
## REVOKED_Yes 3.731e-01 4.217e-01 0.885 0.376253
## URBANICITY_z_Highly.Rural..Rural -2.773e+00 5.228e-01 -5.303 1.14e-07 ***
## YOJ_NA -3.613e-01 4.404e-01 -0.820 0.412090
## INCOME_NA 1.502e-01 5.616e-01 0.268 0.789084
## CAR_AGE_NA -1.969e-01 4.641e-01 -0.424 0.671355
## HOME_VAL_NA 1.299e-01 2.975e-01 0.437 0.662447
## ageSquared 1.572e-03 1.162e-03 1.353 0.176104
## yojSquared 9.267e-03 7.257e-03 1.277 0.201580
## income_log 3.140e-03 3.145e-01 0.010 0.992034
## homeval_log -1.090e+00 1.443e+00 -0.756 0.449938
## travtime_log -7.748e-01 5.900e-01 -1.313 0.189129
## bluebook_log -2.870e-01 4.552e-01 -0.630 0.528458
## carage_log -7.822e-02 4.658e-01 -0.168 0.866629
## oldclaim_log 4.647e-03 1.615e-01 0.029 0.977052
## clm_freq_log 5.762e-01 2.910e+00 0.198 0.843037
## mvr_pts_log 5.476e-01 5.141e-01 1.065 0.286806
## tif_log 2.019e-01 5.226e-01 0.386 0.699234
## kidsdriv_log 3.649e+00 2.311e+00 1.579 0.114370
## homekids_log 1.051e+00 1.462e+00 0.718 0.472494
## inter 5.990e-02 4.011e-02 1.493 0.135344
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 724.15 on 639 degrees of freedom
## Residual deviance: 528.56 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 640.56
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  37
##           1  14  16
##
##           Accuracy : 0.7102
##           95% CI : (0.6372, 0.776)
##           No Information Rate : 0.6989
##           P-Value [Acc > NIR] : 0.406691
##
##           Kappa : 0.2146
##
## Mcnemar's Test P-Value : 0.002066
##
##           Sensitivity : 0.8862
##           Specificity : 0.3019
##           Pos Pred Value : 0.7466
##           Neg Pred Value : 0.5333
##           Prevalence : 0.6989
##           Detection Rate : 0.6193
##           Detection Prevalence : 0.8295
##           Balanced Accuracy : 0.5940
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.74305875134223"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 123 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.7431
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3948  -0.6686  -0.3603   0.5146   2.9198
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    5.482e+00  1.469e+01   0.373 0.709105
## KIDSDRIV       1.329e-01  1.731e+00   0.077 0.938807
## AGE           -1.415e-01  1.046e-01  -1.353 0.176213
## HOMEKIDS       -7.444e-01  6.906e-01  -1.078 0.281054
## YOJ            -2.442e-01  1.288e-01  -1.896 0.057959
## INCOME         -1.372e-05  1.350e-05  -1.016 0.309589
## HOME_VAL        6.924e-06  8.926e-06   0.776 0.437950
## TRAVTIME        3.232e-02  2.166e-02   1.492 0.135696
## BLUEBOOK       -1.395e-05  4.087e-05  -0.341 0.732770
```

```

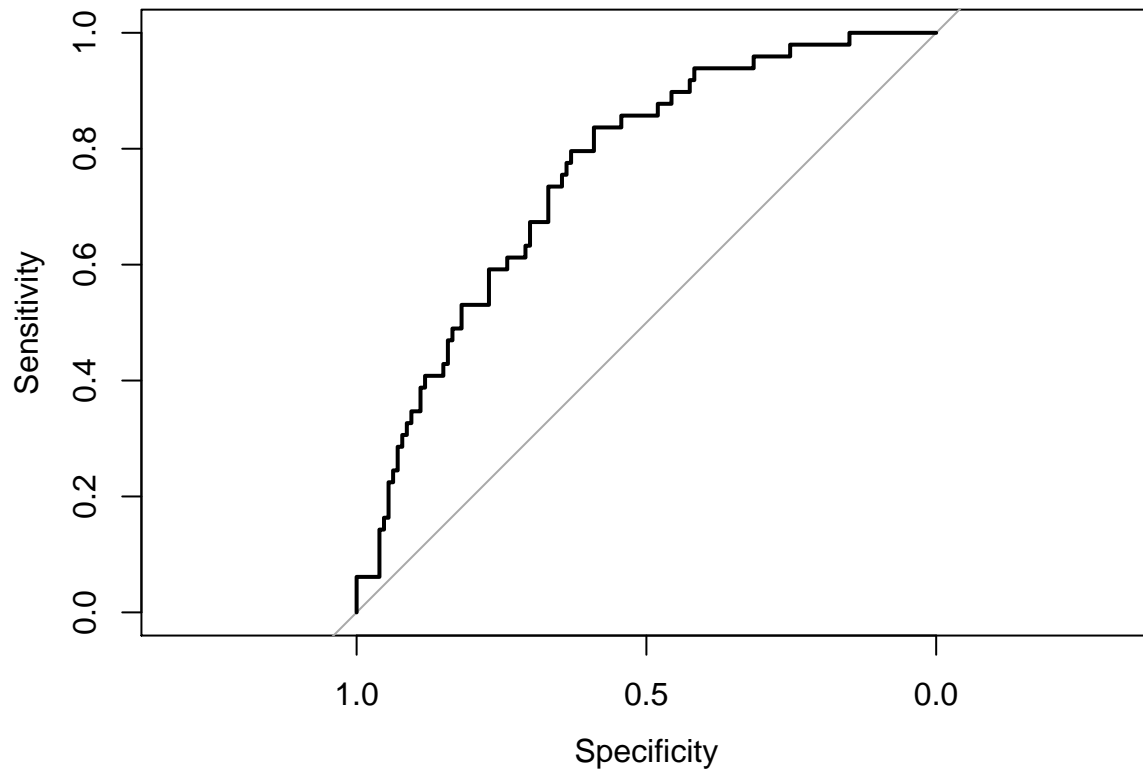
## TIF -5.463e-03 9.171e-02 -0.060 0.952500
## OLDCLAIM 9.193e-06 2.174e-05 0.423 0.672320
## CLM_FREQ -3.819e-01 9.910e-01 -0.385 0.699967
## MVR_PTS -5.990e-02 1.696e-01 -0.353 0.723885
## CAR_AGE -2.200e-02 6.930e-02 -0.317 0.750865
## PARENT1_Yes 7.967e-01 4.778e-01 1.667 0.095432 .
## MSTATUS_Yes -5.305e-01 3.368e-01 -1.575 0.115228
## SEX_z_F -7.438e-01 4.818e-01 -1.544 0.122660
## EDUCATION_.High.School 4.312e-01 8.020e-01 0.538 0.590855
## EDUCATION_Bachelors -3.429e-01 6.804e-01 -0.504 0.614328
## EDUCATION_Masters 6.225e-01 5.926e-01 1.050 0.293524
## EDUCATION_z_High.School 4.208e-01 7.287e-01 0.577 0.563653
## JOB_ -1.412e+00 6.945e-01 -2.034 0.041994 *
## JOB_Clerical -7.956e-01 4.928e-01 -1.615 0.106398
## JOB_Doctor -7.879e-01 9.858e-01 -0.799 0.424172
## JOB_Home.Maker -1.727e-01 6.554e-01 -0.263 0.792181
## JOB_Lawyer -1.128e+00 6.334e-01 -1.781 0.074932 .
## JOB_Manager -9.529e-01 4.764e-01 -2.000 0.045462 *
## JOB_Student -1.170e+00 7.484e-01 -1.563 0.118127
## JOB_z_Blue.Collar -7.997e-01 4.516e-01 -1.771 0.076577 .
## CAR_USE_Commercial 9.509e-01 3.574e-01 2.661 0.007799 **
## CAR_TYPE_Panel.Truck -2.944e-01 6.465e-01 -0.455 0.648835
## CAR_TYPE_Pickup 8.907e-01 4.088e-01 2.179 0.029356 *
## CAR_TYPE_Sports.Car 2.001e+00 5.376e-01 3.723 0.000197 ***
## CAR_TYPE_Van 2.972e-01 4.797e-01 0.620 0.535520
## CAR_TYPE_z_SUV 1.919e+00 4.761e-01 4.031 5.57e-05 ***
## RED_CAR_no -2.303e-01 3.415e-01 -0.674 0.500043
## REVOKED_Yes 5.091e-01 4.047e-01 1.258 0.208336
## URBANICITY_z_Highly.Rural..Rural -2.191e+00 4.280e-01 -5.119 3.06e-07 ***
## YOJ_NA -9.585e-02 4.334e-01 -0.221 0.824953
## INCOME_NA -1.246e-01 5.283e-01 -0.236 0.813604
## CAR_AGE_NA -2.892e-02 4.752e-01 -0.061 0.951478
## HOME_VAL_NA -1.911e-01 2.905e-01 -0.658 0.510540
## ageSquared 1.503e-03 1.141e-03 1.317 0.187774
## yojSquared 1.414e-02 6.736e-03 2.099 0.035799 *
## income_log 7.326e-02 2.757e-01 0.266 0.790471
## homeval_log -6.471e-01 1.362e+00 -0.475 0.634728
## travtime_log -2.477e-01 6.178e-01 -0.401 0.688525
## bluebook_log 4.709e-01 5.140e-01 0.916 0.359590
## carage_log -7.416e-02 4.603e-01 -0.161 0.872015
## oldclaim_log -9.137e-03 1.646e-01 -0.056 0.955720
## clm_freq_log 1.409e+00 3.041e+00 0.463 0.643231
## mvr_pts_log 2.783e-01 4.987e-01 0.558 0.576723
## tif_log -2.443e-01 5.224e-01 -0.468 0.639949
## kidsdriv_log 6.423e-01 2.170e+00 0.296 0.767249
## homekids_log 1.263e+00 1.477e+00 0.855 0.392613
## inter 7.449e-03 3.363e-02 0.222 0.824703
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 732.67 on 639 degrees of freedom
## Residual deviance: 539.53 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 651.53
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 115  32
##           1  12  17
##
##           Accuracy : 0.75
##           95% CI : (0.6793, 0.8121)
##           No Information Rate : 0.7216
##           P-Value [Acc > NIR] : 0.226228
##
##           Kappa : 0.2886
##
## Mcnemar's Test P-Value : 0.004179
##
##           Sensitivity : 0.9055
##           Specificity : 0.3469
##           Pos Pred Value : 0.7823
##           Neg Pred Value : 0.5862
##           Prevalence : 0.7216
##           Detection Rate : 0.6534
##           Detection Prevalence : 0.8352
##           Balanced Accuracy : 0.6262
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.757994536397236"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 127 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.758
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9917  -0.6565  -0.3504  -0.1106   3.0592
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    7.255e+00  1.556e+01   0.466  0.641062
## KIDSDRIV       3.440e-01  1.916e+00   0.179  0.857554
## AGE          -1.327e-01  1.105e-01  -1.201  0.229752
## HOMEKIDS     -1.326e+00  7.073e-01  -1.874  0.060912
## YOJ          -2.192e-01  1.348e-01  -1.626  0.103873
## INCOME       -1.961e-05  1.384e-05  -1.417  0.156561
## HOME_VAL      8.555e-06  9.126e-06   0.937  0.348574
## TRAVTIME      2.792e-02  2.171e-02   1.286  0.198395
## BLUEBOOK      5.583e-05  4.013e-05   1.391  0.164177
```

```

## TIF -6.936e-02 9.280e-02 -0.747 0.454841
## OLDCLAIM 2.853e-05 2.243e-05 1.272 0.203380
## CLM_FREQ -4.690e-01 1.095e+00 -0.428 0.668504
## MVRPTS 1.774e-01 1.756e-01 1.010 0.312474
## CAR_AGE 1.785e-02 7.384e-02 0.242 0.808971
## PARENT1_Yes 3.624e-01 4.793e-01 0.756 0.449571
## MSTATUS_Yes -8.257e-01 3.453e-01 -2.392 0.016779 *
## SEX_z_F -7.362e-01 5.080e-01 -1.449 0.147293
## EDUCATION_.High.School 1.173e+00 8.357e-01 1.403 0.160624
## EDUCATION_Bachelors 5.157e-01 7.074e-01 0.729 0.466039
## EDUCATION_Masters 1.177e+00 5.985e-01 1.966 0.049331 *
## EDUCATION_z_High.School 1.311e+00 7.575e-01 1.731 0.083458 .
## JOB_ -4.907e-01 7.168e-01 -0.685 0.493580
## JOB_Clerical -6.082e-01 5.174e-01 -1.175 0.239796
## JOB_Doctor -9.394e-01 1.290e+00 -0.728 0.466467
## JOB_Home.Maker 1.359e-01 6.865e-01 0.198 0.843108
## JOB_Lawyer -8.116e-01 6.946e-01 -1.168 0.242636
## JOB_Manager -7.515e-01 4.799e-01 -1.566 0.117369
## JOB_Student -5.563e-01 7.389e-01 -0.753 0.451520
## JOB_z_Blue.Collar -3.693e-01 4.707e-01 -0.785 0.432717
## CAR_USE_Commercial 7.124e-01 3.615e-01 1.970 0.048790 *
## CAR_TYPE_Panel.Truck -3.203e-01 6.388e-01 -0.501 0.616109
## CAR_TYPE_Pickup 5.697e-01 4.276e-01 1.332 0.182747
## CAR_TYPE_Sports.Car 1.917e+00 5.662e-01 3.386 0.000710 ***
## CAR_TYPE_Van 3.208e-01 4.924e-01 0.651 0.514733
## CAR_TYPE_z_SUV 1.822e+00 5.071e-01 3.592 0.000328 ***
## RED_CAR_no -3.295e-01 3.514e-01 -0.938 0.348365
## REVOKED_Yes 7.597e-02 4.363e-01 0.174 0.861765
## URBANICITY_z_Highly.Rural..Rural -2.475e+00 4.512e-01 -5.484 4.16e-08 ***
## YOJ_NA -4.403e-01 4.262e-01 -1.033 0.301593
## INCOME_NA 4.236e-01 5.890e-01 0.719 0.472035
## CAR_AGE_NA -3.077e-01 4.765e-01 -0.646 0.518488
## HOME_VAL_NA -2.934e-01 2.965e-01 -0.990 0.322296
## ageSquared 1.329e-03 1.187e-03 1.120 0.262856
## yojSquared 1.208e-02 6.973e-03 1.732 0.083279 .
## income_log -6.625e-02 3.072e-01 -0.216 0.829286
## homeval_log -1.290e-01 1.468e+00 -0.088 0.930001
## travtime_log -4.258e-01 6.036e-01 -0.705 0.480544
## bluebook_log -2.938e-01 4.964e-01 -0.592 0.554016
## carage_log -3.673e-01 4.782e-01 -0.768 0.442446
## oldclaim_log -4.904e-02 1.839e-01 -0.267 0.789718
## clm_freq_log 1.380e+00 3.384e+00 0.408 0.683491
## mvr_pts_log -4.991e-01 5.145e-01 -0.970 0.331958
## tif_log 1.947e-01 5.339e-01 0.365 0.715339
## kidsdriv_log -1.956e+00 2.394e+00 -0.817 0.414027
## homekids_log 2.794e+00 1.497e+00 1.867 0.061937 .
## inter 2.965e-02 3.554e-02 0.834 0.404058
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 692.73 on 640 degrees of freedom
## Residual deviance: 516.47 on 585 degrees of freedom

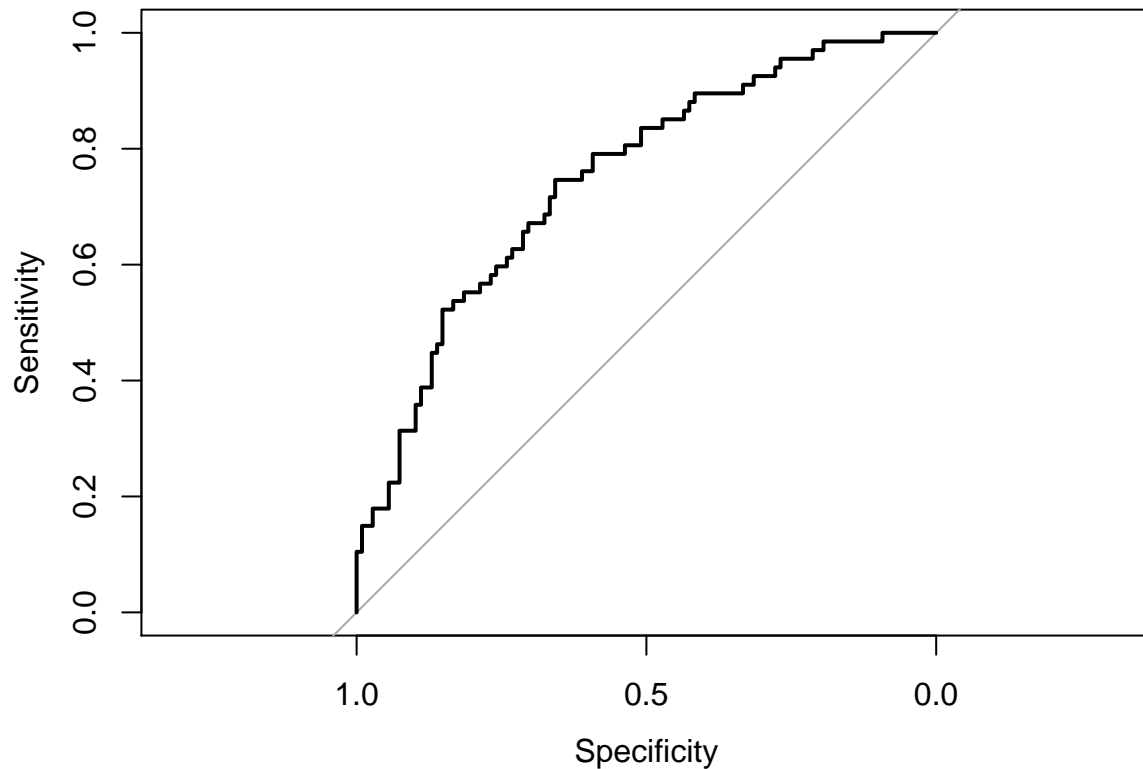
```

```

## (1 observation deleted due to missingness)
## AIC: 628.47
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction 0  1
##           0 96 43
##           1 12 24
##
##           Accuracy : 0.6857
##           95% CI : (0.6113, 0.7537)
##           No Information Rate : 0.6171
##           P-Value [Acc > NIR] : 0.03559
##
##           Kappa : 0.2709
##
## Mcnemar's Test P-Value : 5.228e-05
##
##           Sensitivity : 0.8889
##           Specificity : 0.3582
##           Pos Pred Value : 0.6906
##           Neg Pred Value : 0.6667
##           Prevalence : 0.6171
##           Detection Rate : 0.5486
##           Detection Prevalence : 0.7943
##           Balanced Accuracy : 0.6235
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.749447208402432"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 108 controls (dfPred_raw$class 0) < 67 cases (dfPred_raw$class 1).
## Area under the curve: 0.7494
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3887  -0.6542  -0.3500   0.4808   3.0308
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.795e+01  1.479e+01   1.214 0.224850
## KIDSDRIV      -2.980e+00  1.825e+00  -1.633 0.102573
## AGE           -1.346e-01  1.165e-01  -1.155 0.248037
## HOMEKIDS      -9.066e-02  6.710e-01  -0.135 0.892530
## YOJ           -1.281e-01  1.257e-01  -1.019 0.308169
## INCOME        -2.875e-05  1.338e-05  -2.149 0.031673 *
## HOME_VAL       1.711e-05  9.049e-06   1.891 0.058617 .
## TRAVTIME       2.852e-02  2.187e-02   1.304 0.192205
## BLUEBOOK       1.424e-05  4.155e-05   0.343 0.731801
```

```

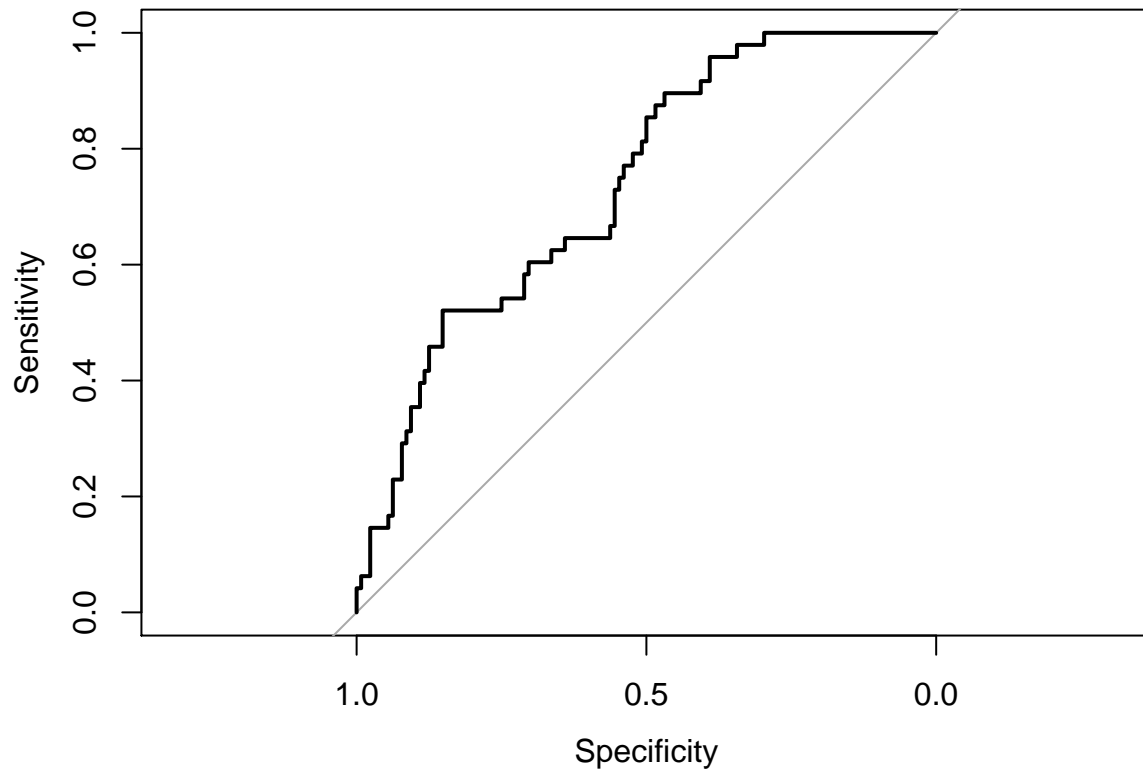
## TIF -3.148e-02 9.103e-02 -0.346 0.729477
## OLDCLAIM 2.675e-06 2.282e-05 0.117 0.906693
## CLM_FREQ -1.203e-01 1.049e+00 -0.115 0.908720
## MVRPTS 1.748e-01 1.725e-01 1.013 0.310953
## CAR_AGE 2.922e-02 7.009e-02 0.417 0.676770
## PARENT1_Yes 9.074e-01 4.798e-01 1.891 0.058586 .
## MSTATUS_Yes -5.038e-01 3.355e-01 -1.501 0.133239
## SEX_z_F -8.520e-01 4.588e-01 -1.857 0.063349 .
## EDUCATION_.High.School 5.232e-01 8.485e-01 0.617 0.537483
## EDUCATION_Bachelors 3.116e-01 7.050e-01 0.442 0.658476
## EDUCATION_Masters 1.245e+00 6.473e-01 1.923 0.054516 .
## EDUCATION_z_High.School 8.553e-01 7.628e-01 1.121 0.262227
## JOB_ -1.369e+00 6.784e-01 -2.019 0.043520 *
## JOB_Clerical -4.765e-01 4.816e-01 -0.989 0.322530
## JOB_Doctor 9.802e-02 9.436e-01 0.104 0.917273
## JOB_Home.Maker -4.068e-01 6.365e-01 -0.639 0.522777
## JOB_Lawyer -1.179e+00 6.530e-01 -1.805 0.071089 .
## JOB_Manager -1.006e+00 5.136e-01 -1.959 0.050067 .
## JOB_Student -7.955e-01 7.244e-01 -1.098 0.272115
## JOB_z_Blue.Collar -3.472e-01 4.533e-01 -0.766 0.443707
## CAR_USE_Commercial 6.453e-01 3.529e-01 1.828 0.067488 .
## CAR_TYPE_Panel.Truck 3.079e-01 6.267e-01 0.491 0.623263
## CAR_TYPE_Pickup 1.255e+00 4.121e-01 3.044 0.002335 **
## CAR_TYPE_Sports.Car 2.164e+00 5.686e-01 3.806 0.000141 ***
## CAR_TYPE_Van 5.837e-01 4.715e-01 1.238 0.215716
## CAR_TYPE_z_SUV 2.173e+00 4.806e-01 4.521 6.17e-06 ***
## RED_CAR_no -3.735e-02 3.427e-01 -0.109 0.913198
## REVOKED_Yes 4.090e-01 4.461e-01 0.917 0.359256
## URBANICITY_z_Highly.Rural..Rural -2.577e+00 4.336e-01 -5.944 2.79e-09 ***
## YOJ_NA 6.736e-03 4.534e-01 0.015 0.988145
## INCOME_NA 3.074e-01 5.284e-01 0.582 0.560705
## CAR_AGE_NA -3.653e-01 4.892e-01 -0.747 0.455199
## HOME_VAL_NA 1.936e-01 2.903e-01 0.667 0.504954
## ageSquared 1.455e-03 1.289e-03 1.129 0.258863
## yojSquared 6.202e-03 6.537e-03 0.949 0.342720
## income_log 1.816e-01 2.814e-01 0.645 0.518748
## homeval_log -1.826e+00 1.379e+00 -1.324 0.185518
## travtime_log -2.630e-01 6.248e-01 -0.421 0.673777
## bluebook_log 2.272e-01 5.142e-01 0.442 0.658668
## carage_log -4.030e-01 4.631e-01 -0.870 0.384195
## oldclaim_log 2.056e-02 1.718e-01 0.120 0.904750
## clm_freq_log 4.493e-01 3.206e+00 0.140 0.888559
## mvr_pts_log -2.579e-01 5.049e-01 -0.511 0.609577
## tif_log -2.159e-03 5.168e-01 -0.004 0.996667
## kidsdriv_log 3.751e+00 2.207e+00 1.699 0.089237 .
## homekids_log 1.584e-01 1.448e+00 0.109 0.912867
## inter 3.742e-02 3.326e-02 1.125 0.260626
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 734.77 on 639 degrees of freedom
## Residual deviance: 535.45 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 647.45
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  30
##           1  14  18
##
##           Accuracy : 0.75
##           95% CI : (0.6793, 0.8121)
##           No Information Rate : 0.7273
##           P-Value [Acc > NIR] : 0.27967
##
##           Kappa : 0.2965
##
## Mcnemar's Test P-Value : 0.02374
##
##           Sensitivity : 0.8906
##           Specificity : 0.3750
##           Pos Pred Value : 0.7917
##           Neg Pred Value : 0.5625
##           Prevalence : 0.7273
##           Detection Rate : 0.6477
##           Detection Prevalence : 0.8182
##           Balanced Accuracy : 0.6328
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.737955729166667"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 128 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.738
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1516  -0.6683  -0.3381   0.5316   2.5454
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.555e+01  1.459e+01   1.751 0.079921 .
## KIDSDRIV      -1.712e+00  1.747e+00  -0.980 0.327043
## AGE           -8.693e-02  1.154e-01  -0.753 0.451368
## HOMEKIDS      -8.350e-01  6.674e-01  -1.251 0.210891
## YOJ           -8.649e-02  1.311e-01  -0.660 0.509323
## INCOME        -9.568e-06  1.347e-05  -0.710 0.477467
## HOME_VAL       8.842e-06  8.950e-06   0.988 0.323203
## TRAVTIME       4.297e-02  2.218e-02   1.937 0.052716 .
## BLUEBOOK       3.982e-05  3.883e-05   1.025 0.305205
```

```

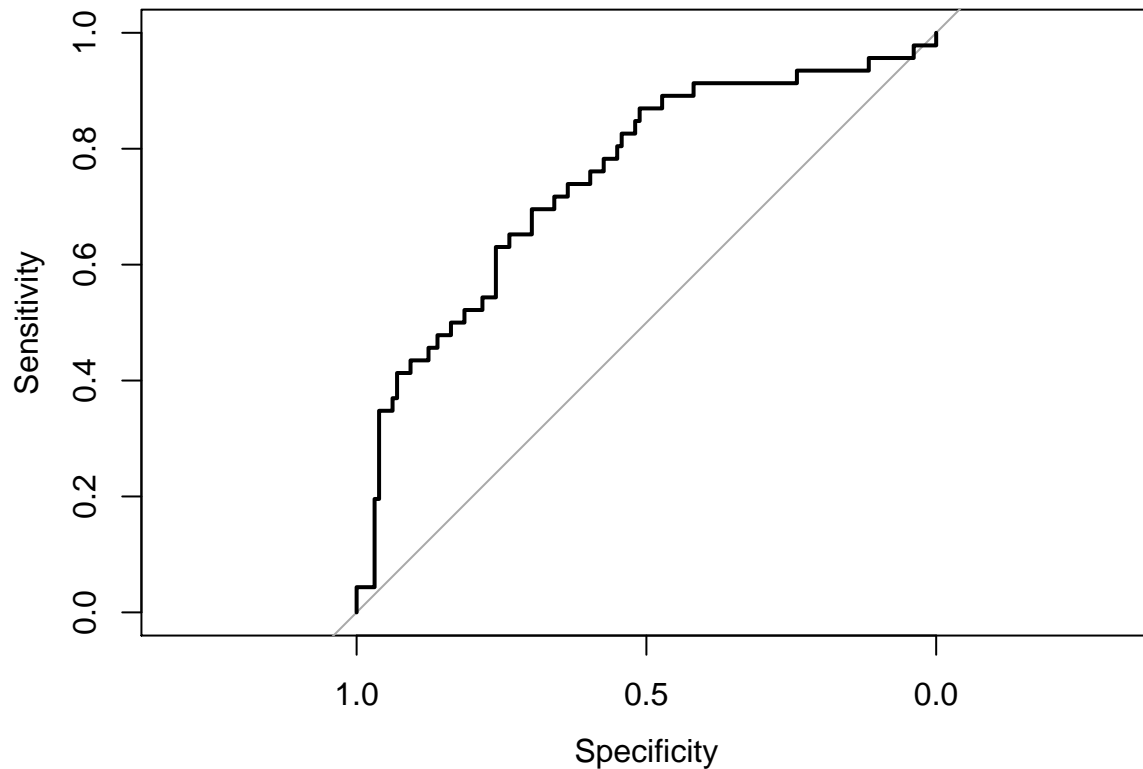
## TIF -7.112e-02 9.257e-02 -0.768 0.442270
## OLDCLAIM 6.483e-06 2.111e-05 0.307 0.758816
## CLM_FREQ 5.805e-01 9.756e-01 0.595 0.551814
## MVR_PTS 4.195e-02 1.862e-01 0.225 0.821745
## CAR_AGE -4.220e-02 7.123e-02 -0.592 0.553577
## PARENT1_Yes 5.827e-01 4.737e-01 1.230 0.218652
## MSTATUS_Yes -6.150e-01 3.430e-01 -1.793 0.072949 .
## SEX_z_F -9.467e-01 4.825e-01 -1.962 0.049775 *
## EDUCATION_.High.School -6.866e-01 7.640e-01 -0.899 0.368810
## EDUCATION_Bachelors -9.319e-01 6.348e-01 -1.468 0.142150
## EDUCATION_Masters 3.656e-01 5.401e-01 0.677 0.498374
## EDUCATION_z_High.School -2.440e-01 6.903e-01 -0.353 0.723781
## JOB_ -1.095e+00 6.925e-01 -1.582 0.113763
## JOB_Clerical -3.238e-02 4.902e-01 -0.066 0.947346
## JOB_Doctor -7.511e-01 9.636e-01 -0.780 0.435663
## JOB_Home.Maker -3.009e-01 6.870e-01 -0.438 0.661391
## JOB_Lawyer -9.720e-01 6.914e-01 -1.406 0.159732
## JOB_Manager -8.512e-01 4.993e-01 -1.705 0.088216 .
## JOB_Student -9.775e-01 7.418e-01 -1.318 0.187550
## JOB_z_Blue.Collar 2.811e-01 4.410e-01 0.638 0.523795
## CAR_USE_Commercial 5.950e-01 3.421e-01 1.739 0.081974 .
## CAR_TYPE_Panel.Truck 3.910e-01 6.406e-01 0.610 0.541597
## CAR_TYPE_Pickup 1.477e+00 4.127e-01 3.578 0.000346 ***
## CAR_TYPE_Sports.Car 1.918e+00 5.425e-01 3.535 0.000407 ***
## CAR_TYPE_Van 6.168e-01 4.771e-01 1.293 0.196057
## CAR_TYPE_z_SUV 2.199e+00 4.880e-01 4.507 6.57e-06 ***
## RED_CAR_no -1.012e-01 3.526e-01 -0.287 0.774030
## REVOKED_Yes 5.270e-01 4.080e-01 1.292 0.196474
## URBANICITY_z_Highly.Rural..Rural -2.606e+00 4.375e-01 -5.958 2.55e-09 ***
## YOJ_NA -4.196e-01 4.235e-01 -0.991 0.321726
## INCOME_NA 1.378e-01 5.186e-01 0.266 0.790507
## CAR_AGE_NA -1.227e-01 4.857e-01 -0.253 0.800589
## HOME_VAL_NA -3.824e-02 2.882e-01 -0.133 0.894467
## ageSquared 8.538e-04 1.269e-03 0.673 0.501107
## yojSquared 3.235e-03 6.883e-03 0.470 0.638335
## income_log 8.417e-02 2.967e-01 0.284 0.776647
## homeval_log -1.909e+00 1.356e+00 -1.407 0.159344
## travtime_log -5.756e-01 6.067e-01 -0.949 0.342750
## bluebook_log -2.590e-01 4.746e-01 -0.546 0.585270
## carage_log 2.441e-02 4.613e-01 0.053 0.957808
## oldclaim_log 6.726e-02 1.622e-01 0.415 0.678320
## clm_freq_log -1.257e+00 3.009e+00 -0.418 0.676042
## mvr_pts_log -1.372e-01 5.445e-01 -0.252 0.801123
## tif_log 1.781e-01 5.206e-01 0.342 0.732300
## kidsdriv_log 6.954e-01 2.191e+00 0.317 0.750970
## homekids_log 1.989e+00 1.439e+00 1.383 0.166715
## inter 4.284e-02 3.235e-02 1.324 0.185493
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 739.51 on 640 degrees of freedom
## Residual deviance: 534.08 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 646.08
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 113  25
##           1  16  21
##
##           Accuracy : 0.7657
##           95% CI : (0.6959, 0.8263)
##           No Information Rate : 0.7371
##           P-Value [Acc > NIR] : 0.2215
##
##           Kappa : 0.3548
##
## Mcnemar's Test P-Value : 0.2115
##
##           Sensitivity : 0.8760
##           Specificity : 0.4565
##           Pos Pred Value : 0.8188
##           Neg Pred Value : 0.5676
##           Prevalence : 0.7371
##           Detection Rate : 0.6457
##           Detection Prevalence : 0.7886
##           Balanced Accuracy : 0.6662
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.748567576676778"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 129 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7486
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1868  -0.6326  -0.3381   0.5029   3.0581
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.573e+01  1.516e+01   2.357  0.01842 *
## KIDSDRIV      -2.229e+00  1.906e+00  -1.169  0.24229
## AGE           -1.765e-01  1.096e-01  -1.610  0.10738
## HOMEKIDS      -4.049e-01  7.373e-01  -0.549  0.58285
## YOJ           -2.454e-01  1.357e-01  -1.808  0.07065 .
## INCOME        -1.127e-05  1.357e-05  -0.831  0.40620
## HOME_VAL       9.300e-06  9.398e-06   0.990  0.32239
## TRAVTIME      1.673e-02  2.235e-02   0.749  0.45396
## BLUEBOOK      9.210e-05  4.016e-05   2.294  0.02181 *
```

```

## TIF -6.485e-02 9.189e-02 -0.706 0.48036
## OLDCLAIM -1.094e-05 2.264e-05 -0.483 0.62897
## CLM_FREQ 7.245e-01 9.397e-01 0.771 0.44070
## MVR_PTS -5.792e-02 1.758e-01 -0.330 0.74177
## CAR_AGE -5.508e-02 7.391e-02 -0.745 0.45616
## PARENT1_Yes 1.129e+00 4.919e-01 2.294 0.02177 *
## MSTATUS_Yes -4.527e-01 3.337e-01 -1.356 0.17500
## SEX_z_F -9.817e-01 4.813e-01 -2.040 0.04137 *
## EDUCATION_.High.School -6.505e-01 8.252e-01 -0.788 0.43056
## EDUCATION_Bachelors -4.835e-01 7.151e-01 -0.676 0.49898
## EDUCATION_Masters 6.153e-01 5.993e-01 1.027 0.30458
## EDUCATION_z_High.School -1.845e-01 7.653e-01 -0.241 0.80944
## JOB_ -8.485e-01 7.085e-01 -1.198 0.23110
## JOB_Clerical -1.182e-01 4.743e-01 -0.249 0.80319
## JOB_Doctor 3.162e-02 9.409e-01 0.034 0.97319
## JOB_Home.Maker -5.692e-01 6.961e-01 -0.818 0.41355
## JOB_Lawyer -6.177e-01 6.884e-01 -0.897 0.36956
## JOB_Manager -7.293e-01 4.973e-01 -1.467 0.14251
## JOB_Student -1.315e+00 7.368e-01 -1.785 0.07418 .
## JOB_z_Blue.Collar 6.085e-02 4.451e-01 0.137 0.89125
## CAR_USE_Commercial 5.513e-01 3.553e-01 1.552 0.12071
## CAR_TYPE_Panel.Truck 3.976e-03 6.290e-01 0.006 0.99496
## CAR_TYPE_Pickup 1.109e+00 4.108e-01 2.699 0.00695 **
## CAR_TYPE_Sports.Car 1.467e+00 5.531e-01 2.653 0.00798 **
## CAR_TYPE_Van 2.716e-01 4.708e-01 0.577 0.56407
## CAR_TYPE_z_SUV 1.919e+00 4.815e-01 3.986 6.73e-05 ***
## RED_CAR_no -1.403e-01 3.453e-01 -0.406 0.68452
## REVOKED_Yes 7.744e-01 4.070e-01 1.903 0.05705 .
## URBANICITY_z_Highly.Rural..Rural -2.580e+00 4.212e-01 -6.125 9.07e-10 ***
## YOJ_NA -3.622e-01 4.342e-01 -0.834 0.40419
## INCOME_NA 6.350e-01 5.953e-01 1.067 0.28607
## CAR_AGE_NA 3.086e-01 5.281e-01 0.584 0.55900
## HOME_VAL_NA -6.772e-02 2.944e-01 -0.230 0.81806
## ageSquared 1.996e-03 1.191e-03 1.676 0.09380 .
## yojSquared 1.100e-02 6.881e-03 1.599 0.10975
## income_log 1.356e-01 2.962e-01 0.458 0.64707
## homeval_log -2.511e+00 1.452e+00 -1.730 0.08370 .
## travtime_log 5.291e-01 6.472e-01 0.818 0.41356
## bluebook_log -7.917e-01 4.796e-01 -1.651 0.09878 .
## carage_log -1.729e-01 4.681e-01 -0.369 0.71189
## oldclaim_log 1.830e-01 1.618e-01 1.131 0.25802
## clm_freq_log -2.181e+00 2.923e+00 -0.746 0.45562
## mvr_pts_log 2.635e-01 5.143e-01 0.512 0.60843
## tif_log 2.101e-02 5.216e-01 0.040 0.96787
## kidsdriv_log 2.841e+00 2.325e+00 1.222 0.22185
## homekids_log 8.744e-01 1.572e+00 0.556 0.57814
## inter 2.720e-02 3.342e-02 0.814 0.41578
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 754.79 on 639 degrees of freedom
## Residual deviance: 527.47 on 584 degrees of freedom

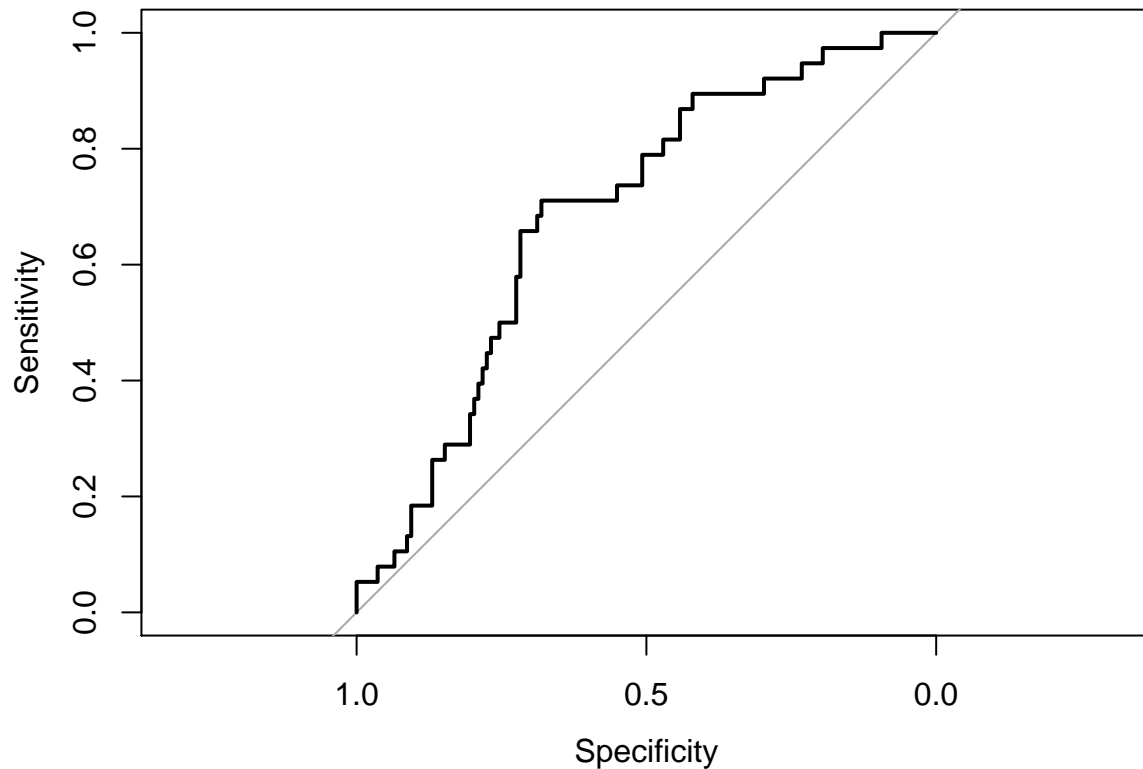
```



```

## (1 observation deleted due to missingness)
## AIC: 639.47
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 112  27
##           1  26  11
##
##           Accuracy : 0.6989
##           95% CI : (0.6253, 0.7656)
##           No Information Rate : 0.7841
##           P-Value [Acc > NIR] : 0.9969
##
##           Kappa : 0.102
##
## Mcnemar's Test P-Value : 1.0000
##
##           Sensitivity : 0.8116
##           Specificity : 0.2895
##           Pos Pred Value : 0.8058
##           Neg Pred Value : 0.2973
##           Prevalence : 0.7841
##           Detection Rate : 0.6364
##           Detection Prevalence : 0.7898
##           Balanced Accuracy : 0.5505
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.689740655987796"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 38 cases (dfPred_raw$class 1).
## Area under the curve: 0.6897
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2215  -0.6835  -0.3418   0.5448   2.6775
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.190e+01  1.438e+01   1.523 0.127724
## KIDSDRIV      -2.535e+00  1.725e+00  -1.469 0.141769
## AGE           -1.671e-01  1.252e-01  -1.334 0.182148
## HOMEKIDS      -8.198e-01  6.589e-01  -1.244 0.213453
## YOJ           -3.622e-03  1.364e-01  -0.027 0.978813
## INCOME        -3.470e-06  1.304e-05  -0.266 0.790161
## HOME_VAL       4.856e-06  8.625e-06   0.563 0.573419
## TRAVTIME       3.168e-02  2.248e-02   1.409 0.158840
## BLUEBOOK       6.459e-05  3.787e-05   1.705 0.088110 .
```

```

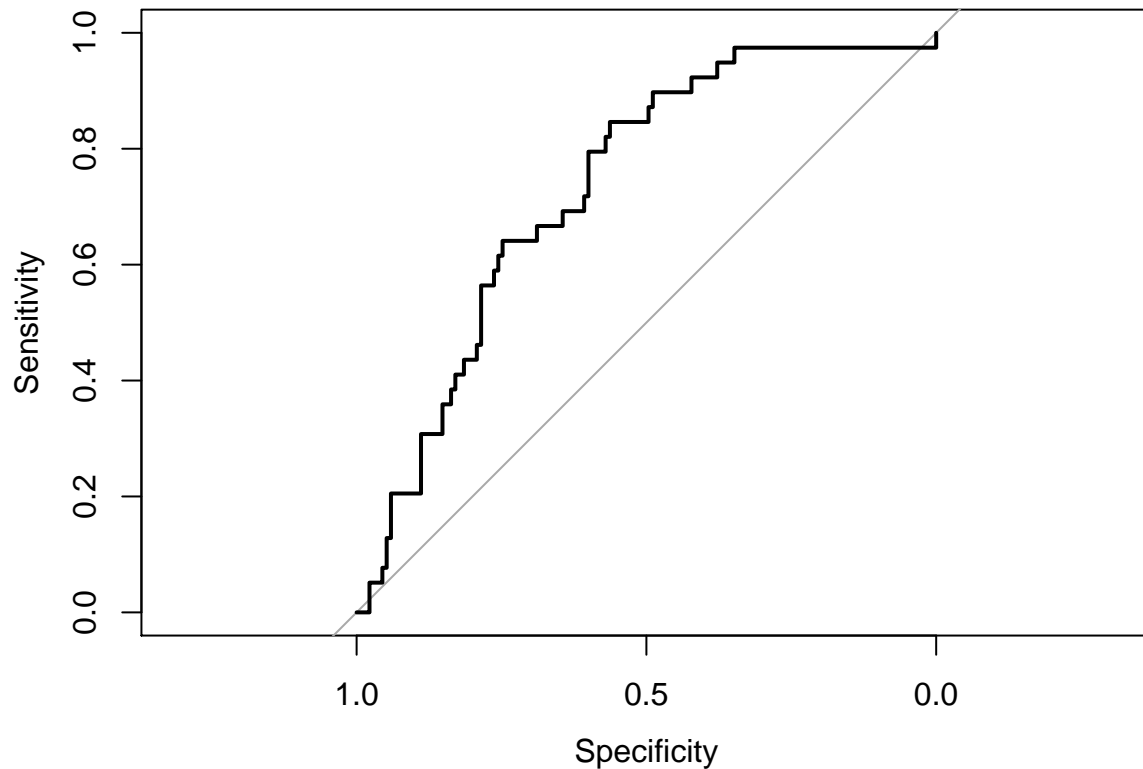
## TIF -9.966e-02 9.443e-02 -1.055 0.291263
## OLDCLAIM -2.574e-06 2.060e-05 -0.125 0.900556
## CLM_FREQ 7.355e-01 9.902e-01 0.743 0.457651
## MVR_PTS -1.754e-01 1.771e-01 -0.991 0.321891
## CAR_AGE -1.001e-02 7.413e-02 -0.135 0.892537
## PARENT1_Yes 3.059e-01 4.760e-01 0.643 0.520441
## MSTATUS_Yes -6.706e-01 3.284e-01 -2.042 0.041154 *
## SEX_z_F -7.552e-01 4.595e-01 -1.643 0.100312
## EDUCATION_.High.School -3.037e-01 7.535e-01 -0.403 0.686885
## EDUCATION_Bachelors -4.417e-01 6.264e-01 -0.705 0.480756
## EDUCATION_Masters 5.914e-01 5.453e-01 1.084 0.278157
## EDUCATION_z_High.School 8.777e-02 6.804e-01 0.129 0.897361
## JOB_ -1.157e+00 6.646e-01 -1.741 0.081651 .
## JOB_Clerical -9.937e-02 4.836e-01 -0.205 0.837205
## JOB_Doctor -6.267e-01 9.026e-01 -0.694 0.487468
## JOB_Home.Maker 1.784e-01 6.737e-01 0.265 0.791153
## JOB_Lawyer -9.158e-01 6.441e-01 -1.422 0.155052
## JOB_Manager -9.153e-01 4.949e-01 -1.849 0.064402 .
## JOB_Student -3.884e-01 7.516e-01 -0.517 0.605295
## JOB_z_Blue.Collar 8.907e-03 4.529e-01 0.020 0.984311
## CAR_USE_Commercial 7.592e-01 3.587e-01 2.117 0.034302 *
## CAR_TYPE_Panel.Truck -1.575e-01 6.286e-01 -0.250 0.802218
## CAR_TYPE_Pickup 8.909e-01 4.106e-01 2.170 0.030005 *
## CAR_TYPE_Sports.Car 1.739e+00 5.424e-01 3.207 0.001343 **
## CAR_TYPE_Van 4.638e-01 4.481e-01 1.035 0.300631
## CAR_TYPE_z_SUV 1.694e+00 4.745e-01 3.570 0.000358 ***
## RED_CAR_no -2.349e-01 3.307e-01 -0.710 0.477447
## REVOKED_Yes 3.404e-01 3.948e-01 0.862 0.388618
## URBANICITY_z_Highly.Rural..Rural -2.732e+00 4.555e-01 -5.997 2.01e-09 ***
## YOJ_NA -4.209e-01 4.273e-01 -0.985 0.324606
## INCOME_NA 1.856e-01 5.650e-01 0.329 0.742532
## CAR_AGE_NA -4.245e-01 4.653e-01 -0.912 0.361618
## HOME_VAL_NA -4.311e-02 2.851e-01 -0.151 0.879786
## ageSquared 1.735e-03 1.368e-03 1.269 0.204580
## yojSquared -6.820e-04 7.162e-03 -0.095 0.924133
## income_log 1.204e-01 2.912e-01 0.414 0.679173
## homeval_log -1.233e+00 1.358e+00 -0.908 0.364080
## travtime_log -2.557e-01 6.288e-01 -0.407 0.684311
## bluebook_log -6.404e-01 4.606e-01 -1.390 0.164410
## carage_log -4.170e-01 4.821e-01 -0.865 0.387073
## oldclaim_log 1.913e-01 1.639e-01 1.167 0.243131
## clm_freq_log -2.391e+00 3.057e+00 -0.782 0.433985
## mvr_pts_log 6.188e-01 5.126e-01 1.207 0.227397
## tif_log 3.165e-01 5.332e-01 0.594 0.552753
## kidsdriv_log 2.379e+00 2.093e+00 1.136 0.255865
## homekids_log 1.876e+00 1.438e+00 1.305 0.192017
## inter 4.686e-02 3.315e-02 1.414 0.157429
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 754.14 on 641 degrees of freedom
## Residual deviance: 542.66 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 654.66
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 117  27
##           1  18  12
##
##           Accuracy : 0.7414
##           95% CI : (0.6697, 0.8047)
##           No Information Rate : 0.7759
##           P-Value [Acc > NIR] : 0.8802
##
##           Kappa : 0.1899
##
## Mcnemar's Test P-Value : 0.2330
##
##           Sensitivity : 0.8667
##           Specificity : 0.3077
##           Pos Pred Value : 0.8125
##           Neg Pred Value : 0.4000
##           Prevalence : 0.7759
##           Detection Rate : 0.6724
##           Detection Prevalence : 0.8276
##           Balanced Accuracy : 0.5872
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.732763532763533"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 135 controls (dfPred_raw$class 0) < 39 cases (dfPred_raw$class 1).
## Area under the curve: 0.7328
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9171  -0.7033  -0.3513   0.6289   2.8854
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.150e+01  1.528e+01   2.061  0.03931 *
## KIDSDRIV       1.038e+00  1.740e+00   0.596  0.55092
## AGE          -3.015e-01  1.048e-01  -2.876  0.00403 **
## HOMEKIDS      -1.545e-01  6.597e-01  -0.234  0.81479
## YOJ           -1.589e-01  1.318e-01  -1.206  0.22779
## INCOME        -2.659e-05  1.319e-05  -2.016  0.04382 *
## HOME_VAL       1.836e-05  8.944e-06   2.052  0.04015 *
## TRAVTIME       4.566e-02  2.181e-02   2.094  0.03627 *
## BLUEBOOK       4.291e-05  4.182e-05   1.026  0.30481
```

```

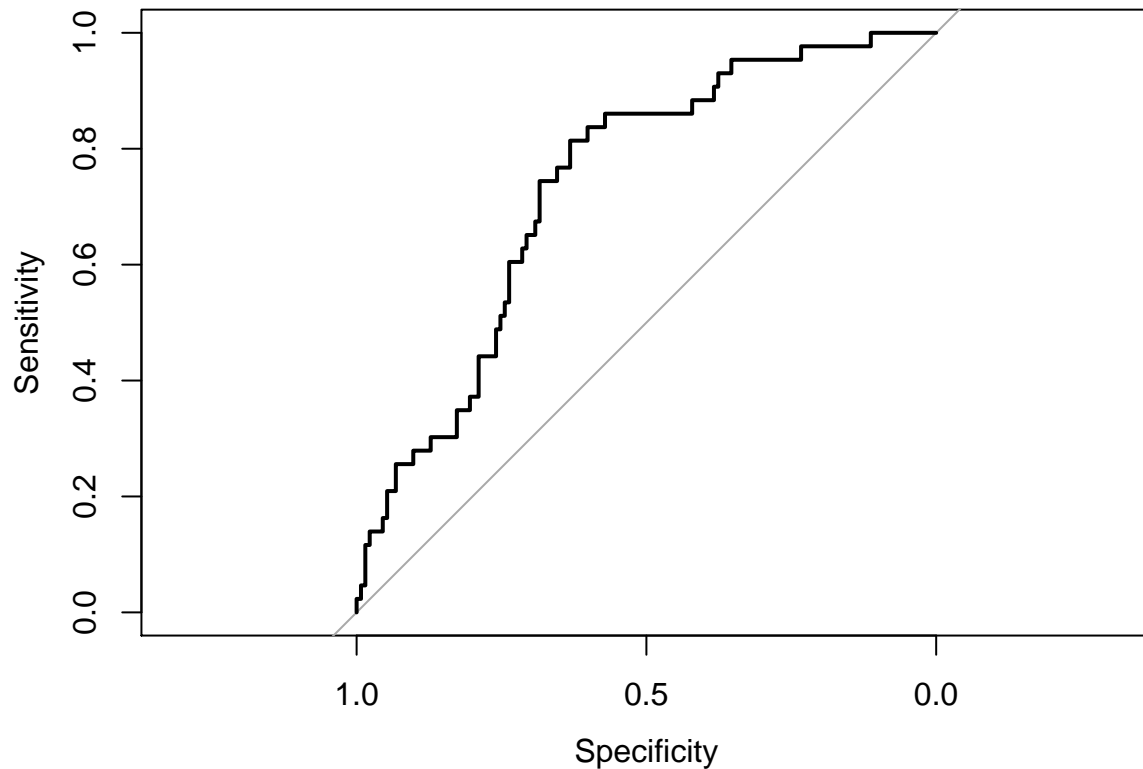
## TIF -3.270e-02 9.200e-02 -0.355 0.72227
## OLDCLAIM 1.006e-05 2.101e-05 0.479 0.63227
## CLM_FREQ 1.161e-01 1.034e+00 0.112 0.91059
## MVR_PTS 6.117e-02 1.691e-01 0.362 0.71760
## CAR_AGE 1.146e-03 6.814e-02 0.017 0.98658
## PARENT1_Yes 5.579e-01 4.596e-01 1.214 0.22479
## MSTATUS_Yes -6.819e-01 3.331e-01 -2.047 0.04064 *
## SEX_z_F -7.818e-01 4.607e-01 -1.697 0.08969 .
## EDUCATION_.High.School 1.095e-01 7.824e-01 0.140 0.88874
## EDUCATION_Bachelors 4.283e-02 6.621e-01 0.065 0.94842
## EDUCATION_Masters 1.174e+00 5.632e-01 2.085 0.03707 *
## EDUCATION_z_High.School 4.432e-01 7.178e-01 0.617 0.53695
## JOB_ -1.052e+00 6.624e-01 -1.588 0.11238
## JOB_Clerical -1.585e-01 4.811e-01 -0.329 0.74187
## JOB_Doctor -9.569e-01 1.061e+00 -0.902 0.36726
## JOB_Home.Maker 2.487e-02 6.343e-01 0.039 0.96873
## JOB_Lawyer -9.861e-01 6.219e-01 -1.586 0.11280
## JOB_Manager -1.045e+00 4.967e-01 -2.103 0.03543 *
## JOB_Student -1.207e+00 7.362e-01 -1.640 0.10104
## JOB_z_Blue.Collar -1.453e-01 4.563e-01 -0.318 0.75015
## CAR_USE_Commercial 5.599e-01 3.498e-01 1.600 0.10950
## CAR_TYPE_Panel.Truck 3.163e-01 6.137e-01 0.515 0.60629
## CAR_TYPE_Pickup 1.143e+00 4.147e-01 2.757 0.00583 **
## CAR_TYPE_Sports.Car 1.927e+00 5.452e-01 3.534 0.00041 ***
## CAR_TYPE_Van 7.418e-01 4.696e-01 1.580 0.11418
## CAR_TYPE_z_SUV 2.097e+00 4.823e-01 4.348 1.37e-05 ***
## RED_CAR_no -2.045e-01 3.273e-01 -0.625 0.53219
## REVOKED_Yes 2.246e-01 4.025e-01 0.558 0.57684
## URBANICITY_z_Highly.Rural..Rural -2.262e+00 4.123e-01 -5.485 4.14e-08 ***
## YOJ_NA -2.782e-02 4.312e-01 -0.065 0.94855
## INCOME_NA 5.396e-01 5.947e-01 0.907 0.36419
## CAR_AGE_NA -4.838e-02 4.986e-01 -0.097 0.92270
## HOME_VAL_NA -1.148e-01 2.835e-01 -0.405 0.68551
## ageSquared 3.205e-03 1.139e-03 2.813 0.00490 **
## yojSquared 9.751e-03 6.729e-03 1.449 0.14732
## income_log 2.811e-01 2.968e-01 0.947 0.34360
## homeval_log -2.346e+00 1.441e+00 -1.628 0.10346
## travtime_log -7.490e-01 6.164e-01 -1.215 0.22433
## bluebook_log -1.589e-01 5.266e-01 -0.302 0.76275
## carage_log -3.837e-01 4.490e-01 -0.855 0.39276
## oldclaim_log 5.302e-02 1.678e-01 0.316 0.75197
## clm_freq_log -1.149e-01 3.166e+00 -0.036 0.97106
## mvr_pts_log -1.582e-01 4.984e-01 -0.317 0.75087
## tif_log -1.735e-01 5.187e-01 -0.334 0.73803
## kidsdriv_log -5.573e-01 2.232e+00 -0.250 0.80282
## homekids_log 3.365e-01 1.431e+00 0.235 0.81408
## inter -2.419e-04 3.011e-02 -0.008 0.99359
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 744.97 on 639 degrees of freedom
## Residual deviance: 549.46 on 584 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 661.46
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 120  32
##           1  13  11
##
##           Accuracy : 0.7443
##           95% CI : (0.6732, 0.807)
##           No Information Rate : 0.7557
##           P-Value [Acc > NIR] : 0.67382
##
##           Kappa : 0.1859
##
## Mcnemar's Test P-Value : 0.00729
##
##           Sensitivity : 0.9023
##           Specificity : 0.2558
##           Pos Pred Value : 0.7895
##           Neg Pred Value : 0.4583
##           Prevalence : 0.7557
##           Detection Rate : 0.6818
##           Detection Prevalence : 0.8636
##           Balanced Accuracy : 0.5790
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.734918692079035"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7349
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1906  -0.6707  -0.3641   0.4023   3.0929
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    4.594e+00  1.501e+01   0.306  0.75949
## KIDSDRIV      -1.567e+00  1.949e+00  -0.804  0.42159
## AGE           -1.822e-01  1.080e-01  -1.687  0.09155 .
## HOMEKIDS      -1.310e+00  6.826e-01  -1.919  0.05498 .
## YOJ           -1.985e-01  1.354e-01  -1.466  0.14269
## INCOME        -1.198e-05  1.307e-05  -0.917  0.35915
## HOME_VAL       3.655e-06  8.847e-06   0.413  0.67948
## TRAVTIME       3.555e-02  2.169e-02   1.639  0.10120
## BLUEBOOK       5.362e-05  3.915e-05   1.369  0.17087
```



```

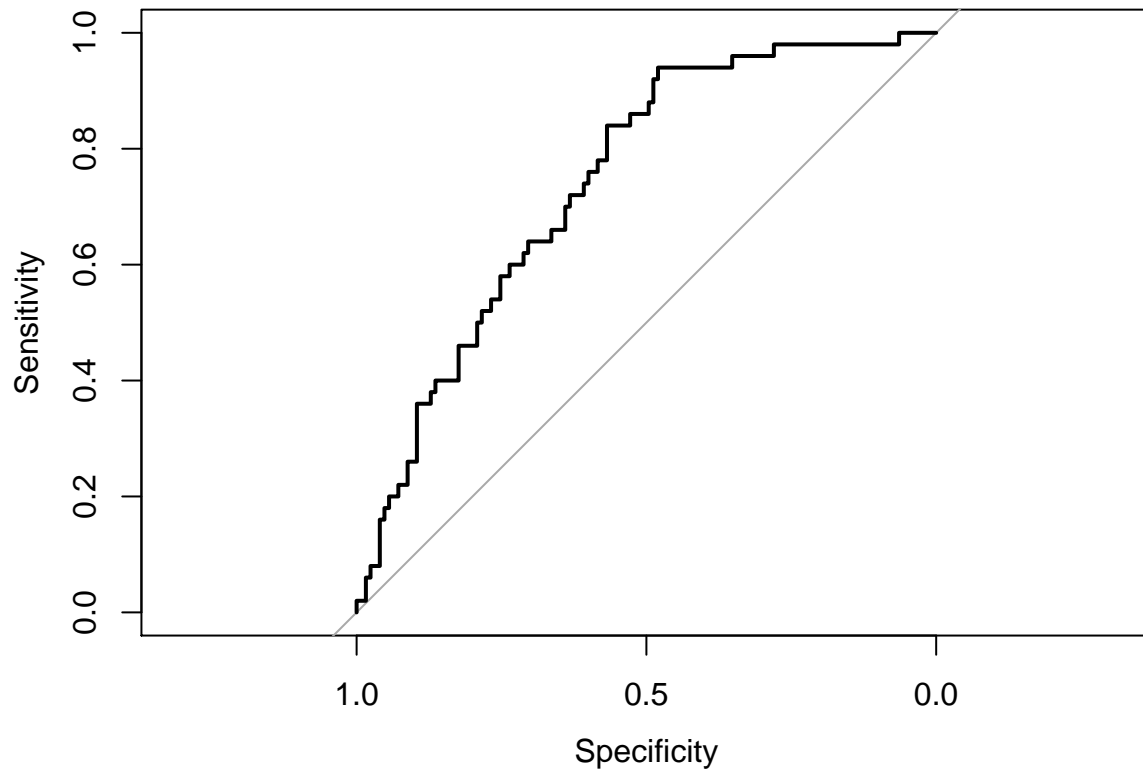
## TIF -5.531e-02 9.578e-02 -0.578 0.56358
## OLDCLAIM 3.036e-05 2.365e-05 1.284 0.19918
## CLM_FREQ -7.096e-01 1.039e+00 -0.683 0.49481
## MVR_PTS 1.850e-01 1.725e-01 1.072 0.28352
## CAR_AGE -8.520e-02 7.180e-02 -1.187 0.23534
## PARENT1_Yes 2.537e-01 4.751e-01 0.534 0.59336
## MSTATUS_Yes -7.211e-01 3.378e-01 -2.135 0.03277 *
## SEX_z_F -1.275e+00 4.795e-01 -2.660 0.00782 **
## EDUCATION_.High.School 1.068e+00 8.228e-01 1.299 0.19407
## EDUCATION_Bachelors 2.721e-01 7.056e-01 0.386 0.69974
## EDUCATION_Masters 7.688e-01 5.996e-01 1.282 0.19976
## EDUCATION_z_High.School 1.090e+00 7.472e-01 1.458 0.14475
## JOB_ -1.653e-01 7.202e-01 -0.230 0.81848
## JOB_Clerical 1.256e-02 5.039e-01 0.025 0.98011
## JOB_Doctor 8.735e-01 9.785e-01 0.893 0.37204
## JOB_Home.Maker -1.570e-01 7.334e-01 -0.214 0.83052
## JOB_Lawyer -1.833e-02 6.933e-01 -0.026 0.97891
## JOB_Manager -3.517e-01 4.774e-01 -0.737 0.46128
## JOB_Student -7.411e-01 7.520e-01 -0.985 0.32440
## JOB_z_Blue.Collar 2.646e-02 4.637e-01 0.057 0.95450
## CAR_USE_Commercial 4.311e-01 3.472e-01 1.242 0.21432
## CAR_TYPE_Panel.Truck -3.669e-01 6.184e-01 -0.593 0.55300
## CAR_TYPE_Pickup 1.182e+00 3.987e-01 2.964 0.00303 **
## CAR_TYPE_Sports.Car 2.323e+00 5.511e-01 4.216 2.48e-05 ***
## CAR_TYPE_Van 5.587e-01 4.547e-01 1.229 0.21916
## CAR_TYPE_z_SUV 2.091e+00 4.931e-01 4.242 2.22e-05 ***
## RED_CAR_no -1.707e-01 3.307e-01 -0.516 0.60559
## REVOKED_Yes -9.175e-03 4.427e-01 -0.021 0.98347
## URBANICITY_z_Highly.Rural..Rural -2.685e+00 4.426e-01 -6.066 1.31e-09 ***
## YOJ_NA -4.537e-01 4.130e-01 -1.099 0.27196
## INCOME_NA 1.999e-01 5.336e-01 0.375 0.70800
## CAR_AGE_NA 1.791e-01 5.233e-01 0.342 0.73213
## HOME_VAL_NA -1.841e-01 2.840e-01 -0.648 0.51685
## ageSquared 1.874e-03 1.165e-03 1.608 0.10790
## yojSquared 9.336e-03 6.851e-03 1.363 0.17292
## income_log -6.245e-02 2.953e-01 -0.211 0.83252
## homeval_log -9.469e-03 1.406e+00 -0.007 0.99462
## travtime_log -2.353e-01 6.080e-01 -0.387 0.69870
## bluebook_log -1.692e-01 4.722e-01 -0.358 0.72008
## carage_log 4.281e-01 4.670e-01 0.917 0.35936
## oldclaim_log -7.246e-02 1.721e-01 -0.421 0.67367
## clm_freq_log 1.912e+00 3.200e+00 0.598 0.55013
## mvr_pts_log -4.384e-01 5.049e-01 -0.868 0.38524
## tif_log 4.674e-02 5.385e-01 0.087 0.93084
## kidsdriv_log 5.937e-01 2.197e+00 0.270 0.78699
## homekids_log 2.933e+00 1.465e+00 2.003 0.04522 *
## inter 4.173e-02 4.010e-02 1.041 0.29802
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 731.16 on 640 degrees of freedom
## Residual deviance: 538.72 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 650.72
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 108  31
##           1  17  19
##
##           Accuracy : 0.7257
##           95% CI : (0.6533, 0.7903)
##           No Information Rate : 0.7143
##           P-Value [Acc > NIR] : 0.4053
##
##           Kappa : 0.2664
##
## Mcnemar's Test P-Value : 0.0606
##
##           Sensitivity : 0.8640
##           Specificity : 0.3800
##           Pos Pred Value : 0.7770
##           Neg Pred Value : 0.5278
##           Prevalence : 0.7143
##           Detection Rate : 0.6171
##           Detection Prevalence : 0.7943
##           Balanced Accuracy : 0.6220
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.7432"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 125 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7432
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3583  -0.6770  -0.3402   0.5308   3.0297
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.338e+01  1.481e+01   1.579 0.114437
## KIDSDRIV      -2.480e+00  1.919e+00  -1.292 0.196238
## AGE           -1.419e-01  1.059e-01  -1.340 0.180403
## HOMEKIDS       3.762e-01  6.712e-01   0.561 0.575100
## YOJ            -2.395e-01  1.354e-01  -1.769 0.076870 .
## INCOME         -2.641e-05  1.341e-05  -1.969 0.048900 *
## HOME_VAL       1.535e-05  9.162e-06   1.675 0.093913 .
## TRAVTIME       3.089e-02  2.108e-02   1.465 0.142828
## BLUEBOOK       4.100e-05  3.955e-05   1.037 0.299902
```

```

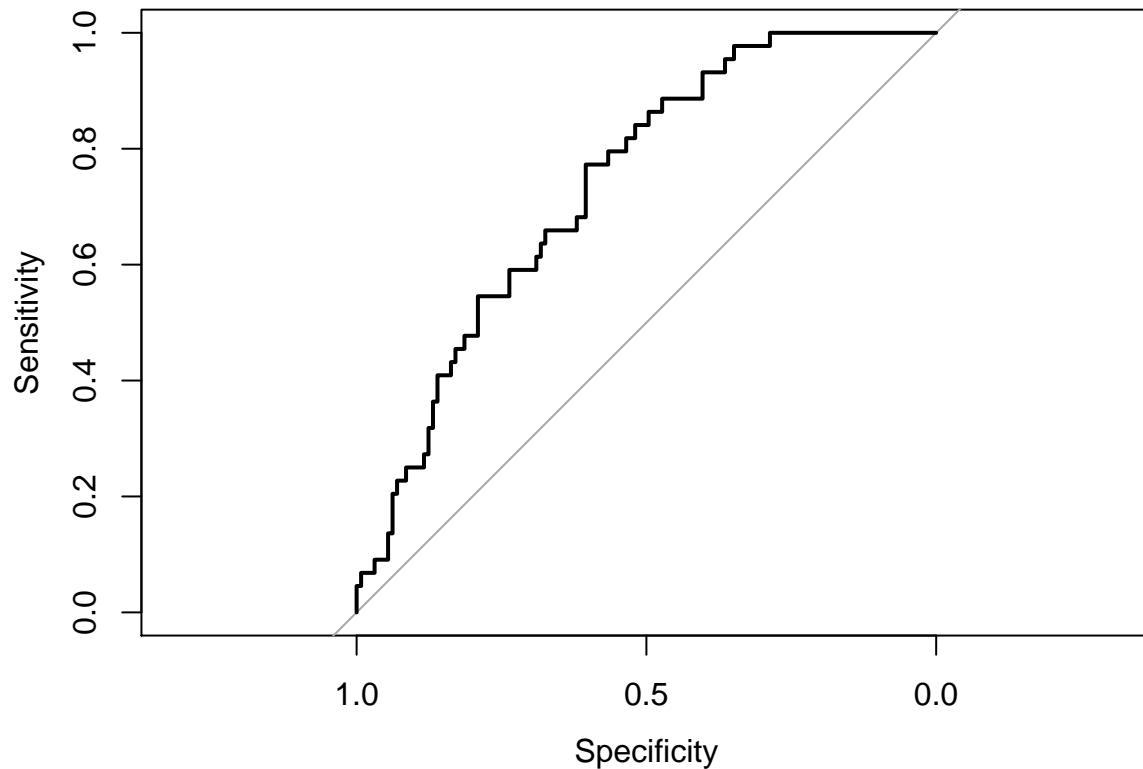
## TIF -3.768e-02 9.231e-02 -0.408 0.683146
## OLDCLAIM -2.308e-06 2.187e-05 -0.106 0.915956
## CLM_FREQ 2.076e-01 9.462e-01 0.219 0.826376
## MVRPTS -5.696e-02 1.769e-01 -0.322 0.747479
## CAR_AGE 6.650e-02 7.035e-02 0.945 0.344505
## PARENT1_Yes 7.408e-01 4.702e-01 1.576 0.115139
## MSTATUS_Yes -7.110e-01 3.339e-01 -2.129 0.033230 *
## SEX_z_F -7.463e-01 4.629e-01 -1.612 0.106918
## EDUCATION_High.School -3.820e-01 8.152e-01 -0.469 0.639359
## EDUCATION_Bachelors -4.736e-01 6.949e-01 -0.682 0.495485
## EDUCATION_Masters 5.721e-01 5.969e-01 0.958 0.337881
## EDUCATION_z_High.School -1.366e-01 7.444e-01 -0.184 0.854397
## JOB_ -1.734e+00 7.219e-01 -2.402 0.016308 *
## JOB_Clerical -9.863e-02 4.867e-01 -0.203 0.839414
## JOB_Doctor -7.505e-01 9.405e-01 -0.798 0.424899
## JOB_Home.Maker -9.942e-02 6.622e-01 -0.150 0.880655
## JOB_Lawyer -1.332e+00 6.512e-01 -2.045 0.040856 *
## JOB_Manager -1.189e+00 5.087e-01 -2.337 0.019441 *
## JOB_Student -8.588e-01 7.204e-01 -1.192 0.233208
## JOB_z_Blue.Collar -1.315e-01 4.530e-01 -0.290 0.771589
## CAR_USE_Commercial 7.128e-01 3.492e-01 2.041 0.041202 *
## CAR_TYPE_Panel.Truck 6.512e-02 6.305e-01 0.103 0.917743
## CAR_TYPE_Pickup 9.232e-01 4.035e-01 2.288 0.022144 *
## CAR_TYPE_Sports.Car 1.934e+00 5.339e-01 3.623 0.000291 ***
## CAR_TYPE_Van 3.316e-01 4.675e-01 0.709 0.478050
## CAR_TYPE_z_SUV 1.753e+00 4.663e-01 3.760 0.000170 ***
## RED_CAR_no -3.166e-01 3.373e-01 -0.939 0.347965
## REVOKED_Yes 8.008e-01 4.063e-01 1.971 0.048746 *
## URBANICITY_z_Highly.Rural..Rural -2.548e+00 4.256e-01 -5.988 2.13e-09 ***
## YOJ_NA -1.469e-01 4.318e-01 -0.340 0.733703
## INCOME_NA 6.138e-01 5.832e-01 1.052 0.292605
## CAR_AGE_NA 5.891e-03 4.670e-01 0.013 0.989936
## HOME_VAL_NA -1.342e-02 2.897e-01 -0.046 0.963045
## ageSquared 1.540e-03 1.148e-03 1.342 0.179606
## yojSquared 1.183e-02 6.929e-03 1.707 0.087836 .
## income_log 2.505e-01 2.767e-01 0.905 0.365247
## homeval_log -1.947e+00 1.411e+00 -1.380 0.167669
## travtime_log -1.224e-01 5.984e-01 -0.204 0.837972
## bluebook_log -1.580e-01 4.729e-01 -0.334 0.738245
## carage_log -7.212e-01 4.589e-01 -1.572 0.116043
## oldclaim_log 3.919e-02 1.652e-01 0.237 0.812451
## clm_freq_log -2.967e-01 2.955e+00 -0.100 0.920035
## mvr_pts_log 3.988e-01 5.145e-01 0.775 0.438224
## tif_log -6.047e-02 5.179e-01 -0.117 0.907059
## kidsdriv_log 3.036e+00 2.216e+00 1.370 0.170633
## homekids_log -6.647e-01 1.465e+00 -0.454 0.650011
## inter 3.282e-02 3.495e-02 0.939 0.347715
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 744.83 on 642 degrees of freedom
## Residual deviance: 538.24 on 587 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 650.24
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 112  30
##           1  17  14
##
##           Accuracy : 0.7283
##           95% CI : (0.6556, 0.7931)
##           No Information Rate : 0.7457
##           P-Value [Acc > NIR] : 0.73232
##
##           Kappa : 0.2065
##
## Mcnemar's Test P-Value : 0.08005
##
##           Sensitivity : 0.8682
##           Specificity : 0.3182
##           Pos Pred Value : 0.7887
##           Neg Pred Value : 0.4516
##           Prevalence : 0.7457
##           Detection Rate : 0.6474
##           Detection Prevalence : 0.8208
##           Balanced Accuracy : 0.5932
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.73661028893587"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 129 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7366
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3589  -0.6781  -0.3514   0.6093   2.9875
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.269e+01  1.436e+01   1.580 0.114219
## KIDSDRIV      -1.783e+00  1.912e+00  -0.933 0.350932
## AGE           -2.434e-01  1.024e-01  -2.376 0.017485 *
## HOMEKIDS      -4.109e-01  7.023e-01  -0.585 0.558540
## YOJ           -1.919e-01  1.286e-01  -1.493 0.135484
## INCOME        -1.263e-05  1.346e-05  -0.939 0.347904
## HOME_VAL       1.015e-05  8.961e-06   1.133 0.257127
## TRAVTIME       4.080e-02  2.167e-02   1.883 0.059707 .
## BLUEBOOK       3.472e-05  3.880e-05   0.895 0.370866
```

```

## TIF -3.301e-02 9.176e-02 -0.360 0.719003
## OLDCLAIM -2.318e-05 2.182e-05 -1.062 0.288080
## CLM_FREQ 7.141e-01 9.891e-01 0.722 0.470306
## MVR_PTS 3.269e-02 1.742e-01 0.188 0.851149
## CAR_AGE -3.677e-02 6.786e-02 -0.542 0.587946
## PARENT1_Yes 9.467e-01 4.766e-01 1.987 0.046971 *
## MSTATUS_Yes -6.515e-01 3.258e-01 -2.000 0.045536 *
## SEX_z_F -4.921e-01 4.492e-01 -1.096 0.273285
## EDUCATION_.High.School -7.512e-01 8.199e-01 -0.916 0.359549
## EDUCATION_Bachelors -6.497e-01 6.651e-01 -0.977 0.328635
## EDUCATION_Masters 6.015e-01 5.487e-01 1.096 0.272908
## EDUCATION_z_High.School -2.375e-02 7.268e-01 -0.033 0.973935
## JOB_ -1.298e+00 6.700e-01 -1.937 0.052754 .
## JOB_Clerical -4.603e-01 4.706e-01 -0.978 0.328019
## JOB_Doctor -9.635e-01 8.775e-01 -1.098 0.272190
## JOB_Home.Maker -3.898e-01 6.290e-01 -0.620 0.535488
## JOB_Lawyer -1.433e+00 6.431e-01 -2.229 0.025846 *
## JOB_Manager -1.051e+00 5.056e-01 -2.078 0.037666 *
## JOB_Student -8.340e-01 7.257e-01 -1.149 0.250457
## JOB_z_Blue.Collar -7.108e-02 4.564e-01 -0.156 0.876226
## CAR_USE_Commercial 5.070e-01 3.558e-01 1.425 0.154140
## CAR_TYPE_Panel.Truck 4.238e-01 6.184e-01 0.685 0.493166
## CAR_TYPE_Pickup 1.139e+00 4.089e-01 2.785 0.005359 **
## CAR_TYPE_Sports.Car 1.914e+00 5.225e-01 3.664 0.000248 ***
## CAR_TYPE_Van 7.017e-01 4.530e-01 1.549 0.121364
## CAR_TYPE_z_SUV 2.067e+00 4.598e-01 4.494 6.98e-06 ***
## RED_CAR_no -3.427e-01 3.383e-01 -1.013 0.311054
## REVOKED_Yes 4.485e-01 4.118e-01 1.089 0.276051
## URBANICITY_z_Highly.Rural..Rural -2.456e+00 4.358e-01 -5.635 1.75e-08 ***
## YOJ_NA -1.485e-01 4.037e-01 -0.368 0.713096
## INCOME_NA 3.114e-01 5.212e-01 0.597 0.550191
## CAR_AGE_NA -1.805e-01 4.944e-01 -0.365 0.715011
## HOME_VAL_NA 1.299e-01 2.858e-01 0.454 0.649602
## ageSquared 2.615e-03 1.119e-03 2.338 0.019376 *
## yojSquared 1.120e-02 6.659e-03 1.681 0.092733 .
## income_log -1.114e-02 2.996e-01 -0.037 0.970347
## homeval_log -1.416e+00 1.384e+00 -1.023 0.306133
## travtime_log -3.874e-01 6.178e-01 -0.627 0.530629
## bluebook_log -1.996e-01 4.714e-01 -0.424 0.671893
## carage_log -1.320e-02 4.476e-01 -0.029 0.976474
## oldclaim_log 1.847e-01 1.600e-01 1.155 0.248193
## clm_freq_log -2.100e+00 3.005e+00 -0.699 0.484538
## mvr_pts_log -9.956e-03 5.088e-01 -0.020 0.984389
## tif_log 2.836e-02 5.180e-01 0.055 0.956343
## kidsdriv_log 1.588e+00 2.345e+00 0.677 0.498284
## homekids_log 5.288e-01 1.509e+00 0.351 0.725911
## inter 3.732e-02 3.314e-02 1.126 0.260022
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 750.91 on 639 degrees of freedom
## Residual deviance: 552.51 on 584 degrees of freedom

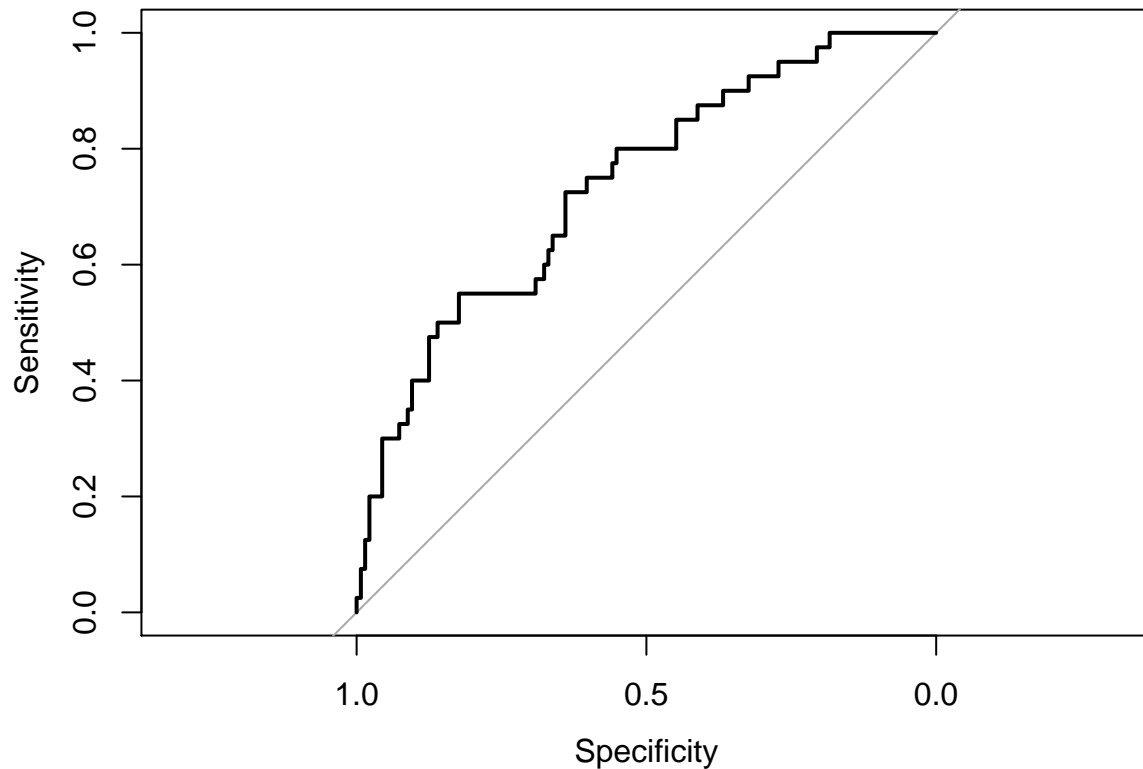
```

```

## (1 observation deleted due to missingness)
## AIC: 664.51
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 120  24
##           1  16  16
##
##           Accuracy : 0.7727
##           95% CI : (0.7036, 0.8324)
##           No Information Rate : 0.7727
##           P-Value [Acc > NIR] : 0.5423
##
##           Kappa : 0.3038
##
## Mcnemar's Test P-Value : 0.2684
##
##           Sensitivity : 0.8824
##           Specificity : 0.4000
##           Pos Pred Value : 0.8333
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7727
##           Detection Rate : 0.6818
##           Detection Prevalence : 0.8182
##           Balanced Accuracy : 0.6412
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.737132352941177"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 136 controls (dfPred_raw$class 0) < 40 cases (dfPred_raw$class 1).
## Area under the curve: 0.7371
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3469  -0.6638  -0.3826   0.4882   3.3219
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.850e+00  1.455e+01   0.265  0.79132
## KIDSDRIV     -1.534e+00  1.655e+00  -0.927  0.35395
## AGE          -4.641e-02  1.177e-01  -0.394  0.69343
## HOMEKIDS     -1.264e+00  6.683e-01  -1.891  0.05858
## YOJ          -1.864e-01  1.299e-01  -1.435  0.15141
## INCOME       -2.980e-06  1.299e-05  -0.229  0.81857
## HOME_VAL      7.317e-07  8.710e-06   0.084  0.93305
## TRAVTIME      1.784e-02  2.233e-02   0.799  0.42427
## BLUEBOOK      4.612e-05  3.904e-05   1.182  0.23740
```

```

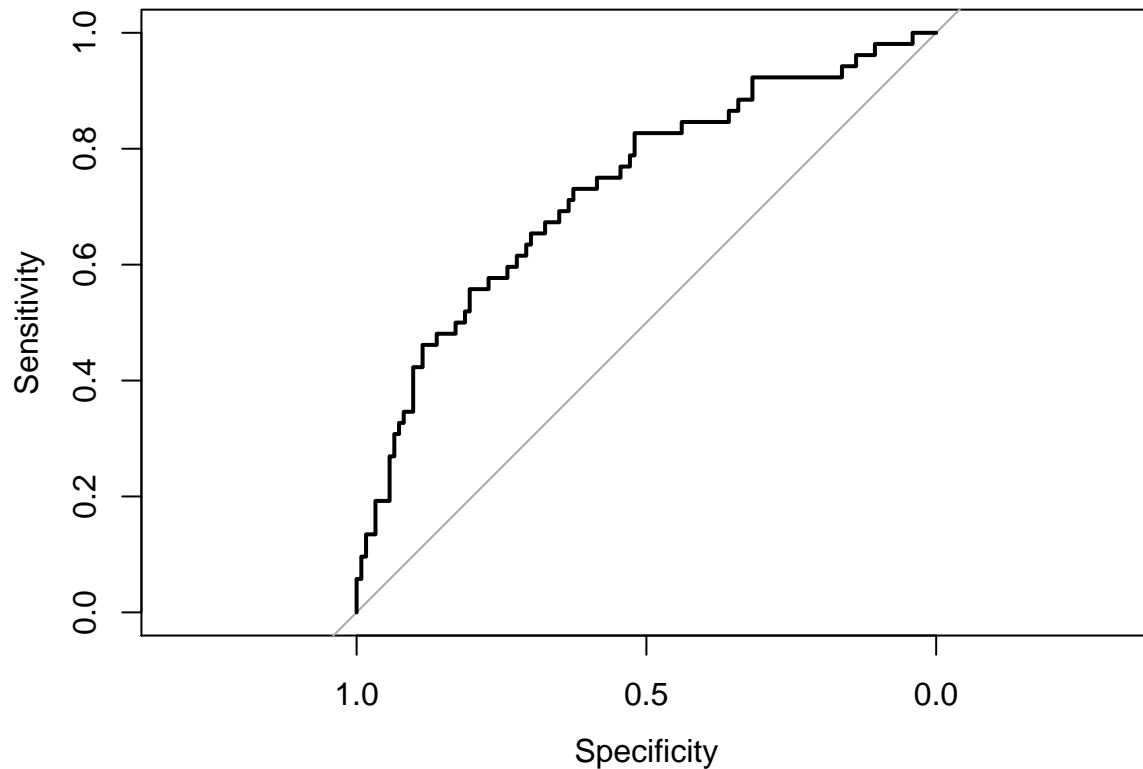
## TIF -9.229e-02 9.216e-02 -1.001 0.31664
## OLDCLAIM 9.157e-06 2.159e-05 0.424 0.67147
## CLM_FREQ -3.937e-01 1.114e+00 -0.353 0.72386
## MVR_PTS 1.017e-01 1.690e-01 0.602 0.54739
## CAR_AGE -3.092e-02 7.262e-02 -0.426 0.67028
## PARENT1_Yes 5.604e-01 4.730e-01 1.185 0.23605
## MSTATUS_Yes -4.821e-01 3.289e-01 -1.466 0.14276
## SEX_z_F -8.314e-01 4.765e-01 -1.745 0.08099 .
## EDUCATION_.High.School 3.244e-01 7.590e-01 0.427 0.66908
## EDUCATION_Bachelors -3.753e-01 6.135e-01 -0.612 0.54068
## EDUCATION_Masters 4.839e-01 5.241e-01 0.923 0.35579
## EDUCATION_z_High.School 5.593e-01 6.726e-01 0.832 0.40568
## JOB_ -3.423e-01 6.756e-01 -0.507 0.61240
## JOB_Clerical -3.954e-01 5.106e-01 -0.774 0.43871
## JOB_Doctor -3.654e-01 9.215e-01 -0.396 0.69175
## JOB_Home.Maker 1.634e-01 6.594e-01 0.248 0.80424
## JOB_Lawyer -4.832e-01 6.402e-01 -0.755 0.45036
## JOB_Manager -6.566e-01 4.932e-01 -1.331 0.18311
## JOB_Student -1.805e-01 7.481e-01 -0.241 0.80932
## JOB_z_Blue.Collar -2.145e-01 4.741e-01 -0.452 0.65105
## CAR_USE_Commercial 8.120e-01 3.705e-01 2.192 0.02841 *
## CAR_TYPE_Panel.Truck -7.092e-01 6.450e-01 -1.100 0.27154
## CAR_TYPE_Pickup 7.155e-01 4.077e-01 1.755 0.07924 .
## CAR_TYPE_Sports.Car 1.596e+00 5.431e-01 2.939 0.00329 **
## CAR_TYPE_Van 1.093e-02 4.672e-01 0.023 0.98133
## CAR_TYPE_z_SUV 1.849e+00 4.702e-01 3.932 8.41e-05 ***
## RED_CAR_no -7.132e-02 3.443e-01 -0.207 0.83591
## REVOKED_Yes 6.558e-03 4.238e-01 0.015 0.98765
## URBANICITY_z_Highly.Rural..Rural -2.782e+00 4.605e-01 -6.042 1.52e-09 ***
## YOJ_NA -2.453e-01 4.374e-01 -0.561 0.57492
## INCOME_NA -6.749e-02 5.444e-01 -0.124 0.90134
## CAR_AGE_NA -3.712e-01 4.643e-01 -0.799 0.42407
## HOME_VAL_NA -2.487e-01 2.866e-01 -0.868 0.38562
## ageSquared 2.606e-04 1.301e-03 0.200 0.84121
## yojSquared 1.156e-02 6.846e-03 1.689 0.09131 .
## income_log -5.658e-02 2.634e-01 -0.215 0.82989
## homeval_log -2.001e-01 1.340e+00 -0.149 0.88128
## travtime_log -7.190e-03 6.219e-01 -0.012 0.99078
## bluebook_log -7.367e-02 4.808e-01 -0.153 0.87821
## carage_log 1.484e-02 4.761e-01 0.031 0.97513
## oldclaim_log 3.664e-02 1.759e-01 0.208 0.83500
## clm_freq_log 6.237e-01 3.374e+00 0.185 0.85333
## mvr_pts_log -1.958e-01 5.003e-01 -0.391 0.69545
## tif_log 3.842e-01 5.257e-01 0.731 0.46494
## kidsdriv_log 1.984e+00 2.114e+00 0.939 0.34794
## homekids_log 2.372e+00 1.449e+00 1.637 0.10166
## inter 3.058e-02 3.129e-02 0.977 0.32837
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 726.89 on 640 degrees of freedom
## Residual deviance: 538.87 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 650.87
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 113  35
##           1  10  17
##
##           Accuracy : 0.7429
##           95% CI : (0.6715, 0.8058)
##           No Information Rate : 0.7029
##           P-Value [Acc > NIR] : 0.1406851
##
##           Kappa : 0.2852
##
## Mcnemar's Test P-Value : 0.0003466
##
##           Sensitivity : 0.9187
##           Specificity : 0.3269
##           Pos Pred Value : 0.7635
##           Neg Pred Value : 0.6296
##           Prevalence : 0.7029
##           Detection Rate : 0.6457
##           Detection Prevalence : 0.8457
##           Balanced Accuracy : 0.6228
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.730456535334584"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 123 controls (dfPred_raw$class 0) < 52 cases (dfPred_raw$class 1).
## Area under the curve: 0.7305
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9879  -0.6808  -0.3659   0.3677   2.8240
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.725e+01  1.452e+01   1.188 0.234893
## KIDSDRIV       3.421e-01  1.837e+00   0.186 0.852296
## AGE           -5.604e-02  1.029e-01  -0.544 0.586161
## HOMEKIDS      -8.641e-01  7.005e-01  -1.233 0.217404
## YOJ           -1.992e-01  1.239e-01  -1.608 0.107823
## INCOME        -3.741e-05  1.402e-05  -2.668 0.007641 **
## HOME_VAL       1.740e-05  8.935e-06   1.947 0.051521 .
## TRAVTIME       2.643e-02  2.168e-02   1.219 0.222795
## BLUEBOOK       1.126e-05  4.117e-05   0.273 0.784539
```

```

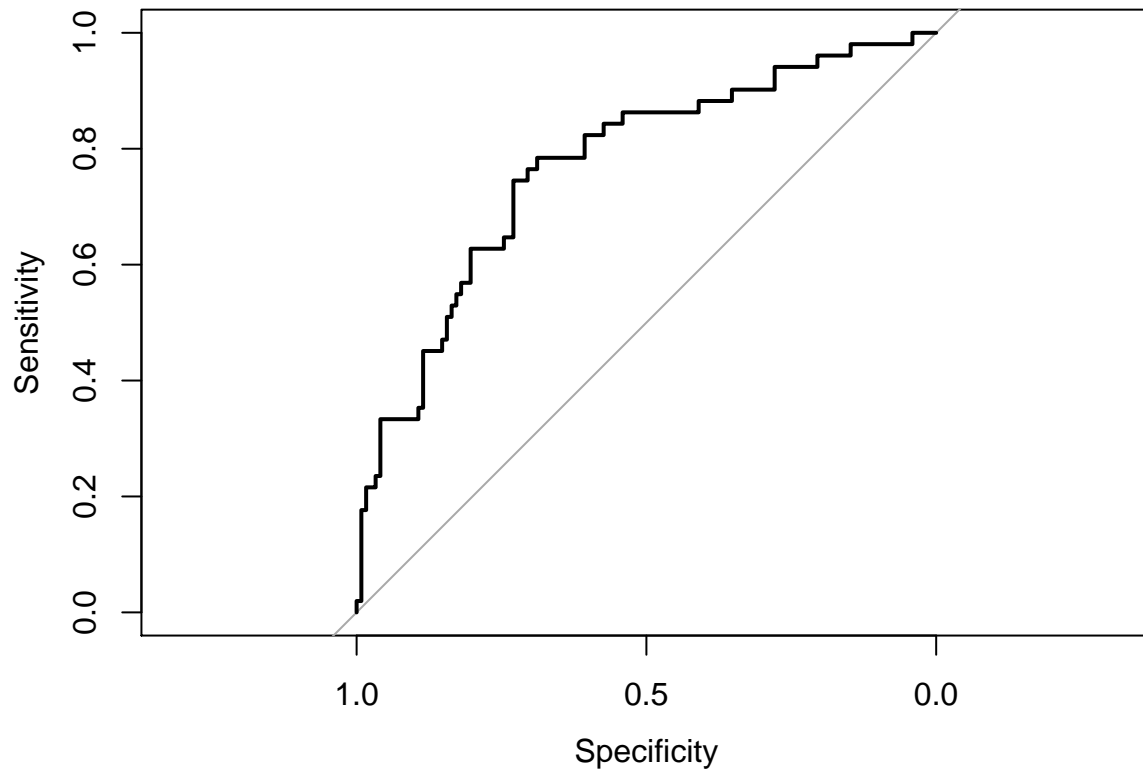
## TIF                1.347e-02  9.026e-02   0.149 0.881384
## OLDCLAIM           -1.067e-06  2.227e-05  -0.048 0.961798
## CLM_FREQ           -5.227e-01  1.007e+00  -0.519 0.603872
## MVR_PTS            1.614e-01  1.634e-01   0.988 0.323274
## CAR_AGE            -1.902e-02  7.166e-02  -0.265 0.790687
## PARENT1_Yes        7.074e-01  4.742e-01   1.492 0.135778
## MSTATUS_Yes        -5.162e-01  3.294e-01  -1.567 0.117138
## SEX_z_F            -9.335e-01  4.685e-01  -1.993 0.046295 *
## EDUCATION_.High.School -3.451e-01  7.848e-01  -0.440 0.660123
## EDUCATION_Bachelors -1.573e-01  6.536e-01  -0.241 0.809759
## EDUCATION_Masters   7.904e-01  5.713e-01   1.383 0.166520
## EDUCATION_z_High.School 1.188e-01  6.982e-01   0.170 0.864950
## JOB_               -1.050e+00  6.855e-01  -1.531 0.125786
## JOB_Clerical        8.528e-02  4.929e-01   0.173 0.862646
## JOB_Doctor          -3.739e-01  9.327e-01  -0.401 0.688472
## JOB_Home.Maker      -3.211e-01  6.780e-01  -0.474 0.635794
## JOB_Lawyer          -7.191e-01  6.304e-01  -1.141 0.253944
## JOB_Manager         -6.472e-01  4.564e-01  -1.418 0.156207
## JOB_Student         -6.439e-01  7.031e-01  -0.916 0.359811
## JOB_z_Blue.Collar  -3.450e-01  4.604e-01  -0.749 0.453644
## CAR_USE_Commercial  5.195e-01  3.416e-01   1.521 0.128318
## CAR_TYPE_Panel.Truck 4.219e-02  6.269e-01   0.067 0.946341
## CAR_TYPE_Pickup     9.298e-01  3.995e-01   2.327 0.019940 *
## CAR_TYPE_Sports.Car 1.780e+00  5.245e-01   3.393 0.000690 ***
## CAR_TYPE_Van        1.600e-01  4.781e-01   0.335 0.737848
## CAR_TYPE_z_SUV      1.626e+00  4.708e-01   3.453 0.000554 ***
## RED_CAR_no          -2.356e-01  3.371e-01  -0.699 0.484694
## REVOKED_Yes         8.555e-01  3.931e-01   2.177 0.029514 *
## URBANICITY_z_Highly.Rural..Rural -2.339e+00  4.438e-01  -5.271 1.36e-07 ***
## YOJ_NA              -2.898e-01  4.232e-01  -0.685 0.493493
## INCOME_NA           8.491e-02  4.936e-01   0.172 0.863418
## CAR_AGE_NA          -2.911e-01  5.176e-01  -0.562 0.573849
## HOME_VAL_NA         -1.882e-01  2.839e-01  -0.663 0.507452
## ageSquared          6.347e-04  1.115e-03   0.569 0.569329
## yojSquared          9.510e-03  6.439e-03   1.477 0.139701
## income_log          5.593e-01  3.209e-01   1.743 0.081396 .
## homeval_log         -1.980e+00  1.377e+00  -1.438 0.150306
## travtime_log        -2.795e-01  6.002e-01  -0.466 0.641439
## bluebook_log        1.818e-01  4.902e-01   0.371 0.710819
## carage_log          -1.821e-01  4.660e-01  -0.391 0.695966
## oldclaim_log        -1.116e-01  1.696e-01  -0.658 0.510577
## clm_freq_log        2.116e+00  3.110e+00   0.680 0.496267
## mvr_pts_log         -2.476e-01  4.882e-01  -0.507 0.611958
## tif_log             -3.758e-01  5.140e-01  -0.731 0.464667
## kidsdriv_log        -1.404e+00  2.236e+00  -0.628 0.529936
## homekids_log        1.854e+00  1.503e+00   1.233 0.217483
## inter              1.984e-02  3.526e-02   0.563 0.573716
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 730.22 on 642 degrees of freedom
## Residual deviance: 546.98 on 587 degrees of freedom

```

```

## AIC: 658.98
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  39
##           1   4  12
##
##           Accuracy : 0.7514
##           95% CI : (0.6802, 0.8139)
##           No Information Rate : 0.7052
##           P-Value [Acc > NIR] : 0.1042
##
##           Kappa : 0.253
##
## Mcnemar's Test P-Value : 2.161e-07
##
##           Sensitivity : 0.9672
##           Specificity : 0.2353
##           Pos Pred Value : 0.7516
##           Neg Pred Value : 0.7500
##           Prevalence : 0.7052
##           Detection Rate : 0.6821
##           Detection Prevalence : 0.9075
##           Balanced Accuracy : 0.6013
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.769527483124397"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 122 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.7695
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9704  -0.6492  -0.3533   0.4146   3.0805
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.834e+01  1.545e+01   1.188  0.23500
## KIDSDRIV      -4.992e+00  2.136e+00  -2.337  0.01945 *
## AGE           -8.714e-02  1.131e-01  -0.770  0.44111
## HOMEKIDS      -4.871e-01  6.825e-01  -0.714  0.47539
## YOJ           -1.656e-01  1.345e-01  -1.231  0.21834
## INCOME        -3.486e-05  1.484e-05  -2.348  0.01887 *
## HOME_VAL       1.660e-05  9.617e-06   1.726  0.08433 .
## TRAVTIME       4.241e-02  2.131e-02   1.990  0.04662 *
## BLUEBOOK       4.739e-05  3.844e-05   1.233  0.21759
```

```

## TIF -4.791e-02 9.638e-02 -0.497 0.61911
## OLDCLAIM -1.621e-05 2.244e-05 -0.722 0.47009
## CLM_FREQ 1.700e-01 9.685e-01 0.176 0.86065
## MVR_PTS 1.532e-01 1.752e-01 0.874 0.38214
## CAR_AGE -6.297e-02 7.573e-02 -0.832 0.40564
## PARENT1_Yes 4.692e-01 4.846e-01 0.968 0.33296
## MSTATUS_Yes -6.661e-01 3.428e-01 -1.943 0.05199 .
## SEX_z_F -5.189e-01 4.643e-01 -1.117 0.26383
## EDUCATION_.High.School -4.077e-01 8.216e-01 -0.496 0.61968
## EDUCATION_Bachelors -4.153e-01 6.705e-01 -0.619 0.53569
## EDUCATION_Masters 6.146e-01 5.945e-01 1.034 0.30123
## EDUCATION_z_High.School 2.192e-01 7.313e-01 0.300 0.76438
## JOB_ -1.546e+00 7.090e-01 -2.181 0.02922 *
## JOB_Clerical -4.288e-01 4.780e-01 -0.897 0.36959
## JOB_Doctor -1.406e-03 9.385e-01 -0.001 0.99880
## JOB_Home.Maker -1.201e-01 6.453e-01 -0.186 0.85231
## JOB_Lawyer -1.091e+00 6.721e-01 -1.623 0.10460
## JOB_Manager -9.760e-01 4.861e-01 -2.008 0.04466 *
## JOB_Student -6.741e-01 7.400e-01 -0.911 0.36230
## JOB_z_Blue.Collar -1.061e-01 4.427e-01 -0.240 0.81063
## CAR_USE_Commercial 5.542e-01 3.477e-01 1.594 0.11097
## CAR_TYPE_Panel.Truck 1.647e-01 6.542e-01 0.252 0.80128
## CAR_TYPE_Pickup 1.102e+00 4.059e-01 2.714 0.00664 **
## CAR_TYPE_Sports.Car 1.726e+00 5.352e-01 3.224 0.00126 **
## CAR_TYPE_Van 2.407e-01 4.763e-01 0.505 0.61326
## CAR_TYPE_z_SUV 1.852e+00 4.669e-01 3.967 7.27e-05 ***
## RED_CAR_no -3.255e-01 3.544e-01 -0.919 0.35835
## REVOKED_Yes 7.704e-01 4.177e-01 1.845 0.06510 .
## URBANICITY_z_Highly.Rural..Rural -2.384e+00 4.282e-01 -5.569 2.56e-08 ***
## YOJ_NA -2.648e-02 4.538e-01 -0.058 0.95347
## INCOME_NA 1.521e-01 5.428e-01 0.280 0.77935
## CAR_AGE_NA -3.257e-01 4.722e-01 -0.690 0.49036
## HOME_VAL_NA 1.126e-01 2.962e-01 0.380 0.70400
## ageSquared 7.659e-04 1.231e-03 0.622 0.53376
## yojSquared 7.456e-03 6.977e-03 1.069 0.28519
## income_log 3.642e-01 3.252e-01 1.120 0.26269
## homeval_log -1.582e+00 1.426e+00 -1.109 0.26722
## travtime_log -5.815e-01 5.855e-01 -0.993 0.32059
## bluebook_log -2.534e-01 4.620e-01 -0.549 0.58331
## carage_log 1.591e-01 4.829e-01 0.330 0.74171
## oldclaim_log 6.971e-02 1.673e-01 0.417 0.67685
## clm_freq_log -2.895e-01 3.019e+00 -0.096 0.92361
## mvr_pts_log -3.482e-01 5.154e-01 -0.676 0.49927
## tif_log 1.559e-01 5.365e-01 0.291 0.77141
## kidsdriv_log 1.676e+00 2.225e+00 0.753 0.45116
## homekids_log 1.142e+00 1.474e+00 0.775 0.43845
## inter 1.098e-01 4.253e-02 2.581 0.00985 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.87 on 641 degrees of freedom
## Residual deviance: 530.37 on 586 degrees of freedom

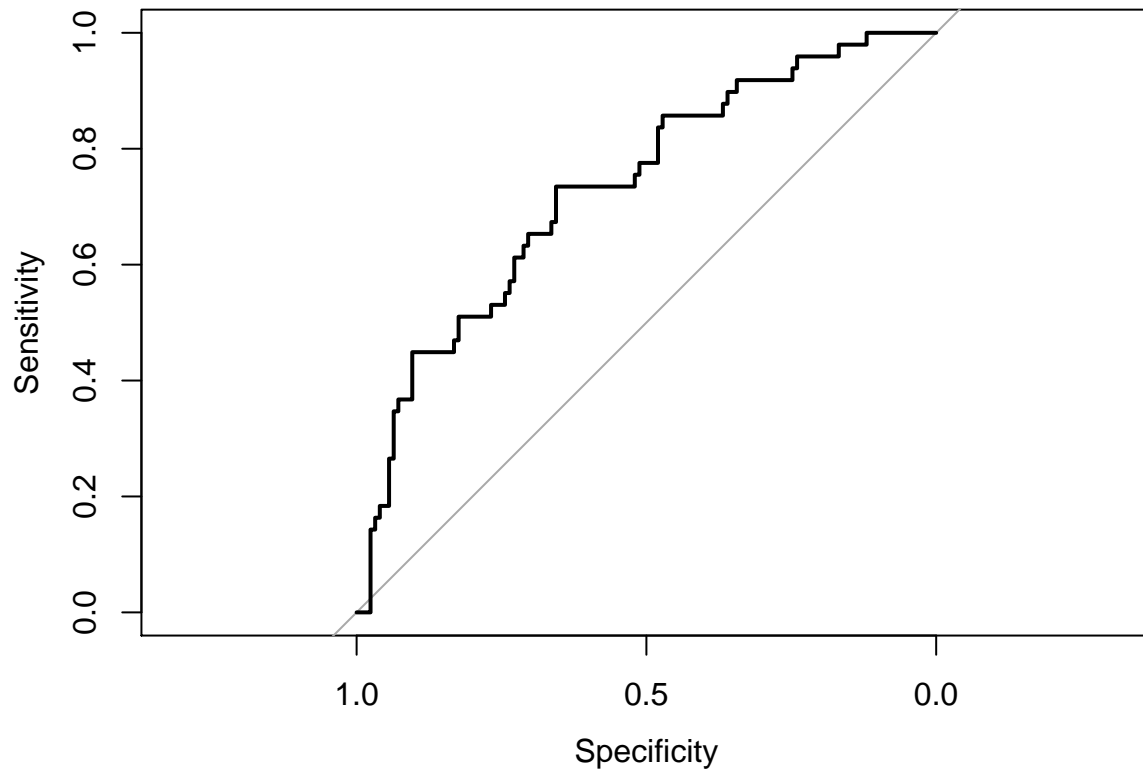
```



```

## (1 observation deleted due to missingness)
## AIC: 642.37
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 109  27
##           1  16  22
##
##           Accuracy : 0.7529
##           95% CI : (0.6819, 0.815)
##           No Information Rate : 0.7184
##           P-Value [Acc > NIR] : 0.1774
##
##           Kappa : 0.3445
##
## Mcnemar's Test P-Value : 0.1273
##
##           Sensitivity : 0.8720
##           Specificity : 0.4490
##           Pos Pred Value : 0.8015
##           Neg Pred Value : 0.5789
##           Prevalence : 0.7184
##           Detection Rate : 0.6264
##           Detection Prevalence : 0.7816
##           Balanced Accuracy : 0.6605
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.731591836734694"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 125 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.7316
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0268  -0.6326  -0.3230  -0.0668   3.1929
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.922e+01  1.612e+01   1.813 0.069780 .
## KIDSDRIV      -7.364e-01  1.677e+00  -0.439 0.660585
## AGE          -1.853e-01  1.091e-01  -1.699 0.089279 .
## HOMEKIDS      -9.981e-01  6.813e-01  -1.465 0.142936
## YOJ           -9.886e-02  1.380e-01  -0.716 0.473761
## INCOME        -2.743e-05  1.471e-05  -1.865 0.062247 .
## HOME_VAL       1.724e-05  9.710e-06   1.775 0.075837 .
## TRAVTIME       2.893e-02  2.280e-02   1.269 0.204545
## BLUEBOOK       1.074e-04  3.943e-05   2.724 0.006440 **
```

```

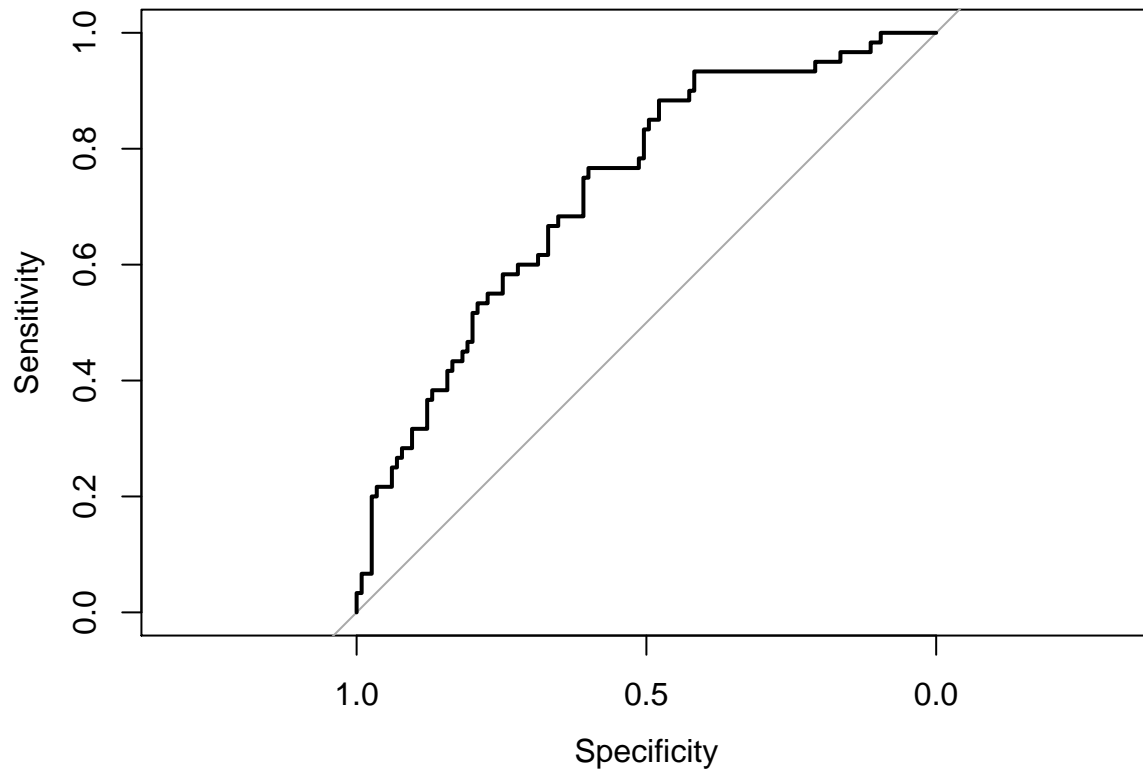
## TIF -3.284e-02 9.384e-02 -0.350 0.726383
## OLDCLAIM 2.001e-05 2.134e-05 0.938 0.348379
## CLM_FREQ 2.376e-01 1.025e+00 0.232 0.816720
## MVR_PTS 6.216e-02 1.820e-01 0.341 0.732761
## CAR_AGE -9.761e-02 7.512e-02 -1.299 0.193846
## PARENT1_Yes 2.416e-01 4.878e-01 0.495 0.620323
## MSTATUS_Yes -7.655e-01 3.564e-01 -2.148 0.031707 *
## SEX_z_F -1.212e+00 4.987e-01 -2.430 0.015113 *
## EDUCATION_.High.School 5.761e-01 8.092e-01 0.712 0.476513
## EDUCATION_Bachelors -1.302e-01 6.823e-01 -0.191 0.848667
## EDUCATION_Masters 9.512e-01 5.693e-01 1.671 0.094736 .
## EDUCATION_z_High.School 9.003e-01 7.334e-01 1.228 0.219609
## JOB_ -4.413e-01 7.291e-01 -0.605 0.544953
## JOB_Clerical -2.986e-01 5.279e-01 -0.566 0.571698
## JOB_Doctor -1.838e-01 1.013e+00 -0.181 0.856019
## JOB_Home.Maker 4.177e-01 7.418e-01 0.563 0.573369
## JOB_Lawyer -3.928e-01 7.303e-01 -0.538 0.590626
## JOB_Manager -9.939e-01 5.367e-01 -1.852 0.064046 .
## JOB_Student -4.278e-01 7.536e-01 -0.568 0.570285
## JOB_z_Blue.Collar -1.084e-01 4.834e-01 -0.224 0.822487
## CAR_USE_Commercial 5.647e-01 3.591e-01 1.573 0.115801
## CAR_TYPE_Panel.Truck -8.136e-01 6.679e-01 -1.218 0.223168
## CAR_TYPE_Pickup 1.020e+00 4.201e-01 2.427 0.015216 *
## CAR_TYPE_Sports.Car 2.215e+00 5.797e-01 3.821 0.000133 ***
## CAR_TYPE_Van 2.861e-01 4.790e-01 0.597 0.550288
## CAR_TYPE_z_SUV 2.135e+00 5.222e-01 4.089 4.33e-05 ***
## RED_CAR_no -5.126e-01 3.459e-01 -1.482 0.138345
## REVOKED_Yes 2.109e-01 4.098e-01 0.515 0.606791
## URBANICITY_z_Highly.Rural..Rural -2.802e+00 4.879e-01 -5.744 9.26e-09 ***
## YOJ_NA -5.371e-01 4.165e-01 -1.290 0.197214
## INCOME_NA 1.147e-02 5.595e-01 0.021 0.983642
## CAR_AGE_NA -4.397e-01 4.808e-01 -0.914 0.360492
## HOME_VAL_NA -3.221e-01 2.985e-01 -1.079 0.280598
## ageSquared 1.908e-03 1.182e-03 1.615 0.106315
## yojSquared 5.310e-03 7.138e-03 0.744 0.456924
## income_log 3.261e-01 3.395e-01 0.961 0.336795
## homeval_log -1.963e+00 1.558e+00 -1.259 0.207863
## travtime_log -2.563e-01 6.220e-01 -0.412 0.680327
## bluebook_log -8.364e-01 4.790e-01 -1.746 0.080767 .
## carage_log 3.288e-01 4.802e-01 0.685 0.493449
## oldclaim_log 3.732e-02 1.705e-01 0.219 0.826773
## clm_freq_log -5.516e-01 3.175e+00 -0.174 0.862083
## mvr_pts_log -1.585e-01 5.279e-01 -0.300 0.764074
## tif_log -6.862e-02 5.390e-01 -0.127 0.898681
## kidsdriv_log 2.041e-01 2.178e+00 0.094 0.925331
## homekids_log 2.321e+00 1.486e+00 1.563 0.118131
## inter 3.009e-02 3.064e-02 0.982 0.326014
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 709.15 on 640 degrees of freedom
## Residual deviance: 505.95 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 617.95
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 101  41
##           1  14  19
##
##           Accuracy : 0.6857
##           95% CI : (0.6113, 0.7537)
##           No Information Rate : 0.6571
##           P-Value [Acc > NIR] : 0.2381047
##
##           Kappa : 0.2184
##
## Mcnemar's Test P-Value : 0.0004552
##
##           Sensitivity : 0.8783
##           Specificity : 0.3167
##           Pos Pred Value : 0.7113
##           Neg Pred Value : 0.5758
##           Prevalence : 0.6571
##           Detection Rate : 0.5771
##           Detection Prevalence : 0.8114
##           Balanced Accuracy : 0.5975
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.730289855072464"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 115 controls (dfPred_raw$class 0) < 60 cases (dfPred_raw$class 1).
## Area under the curve: 0.7303
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4644  -0.6815  -0.3109   0.5519   2.3278
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.188e+01  1.502e+01   2.123 0.033768 *
## KIDSDRIV      -7.730e-01  1.731e+00  -0.446 0.655301
## AGE           -1.296e-01  1.086e-01  -1.192 0.233080
## HOMEKIDS      -9.217e-02  7.018e-01  -0.131 0.895513
## YOJ           -2.613e-01  1.332e-01  -1.962 0.049813 *
## INCOME        -3.832e-05  1.461e-05  -2.623 0.008720 **
## HOME_VAL       1.971e-05  9.411e-06   2.094 0.036247 *
## TRAVTIME       2.731e-02  2.206e-02   1.238 0.215874
## BLUEBOOK       3.075e-05  4.036e-05   0.762 0.446089
```

```

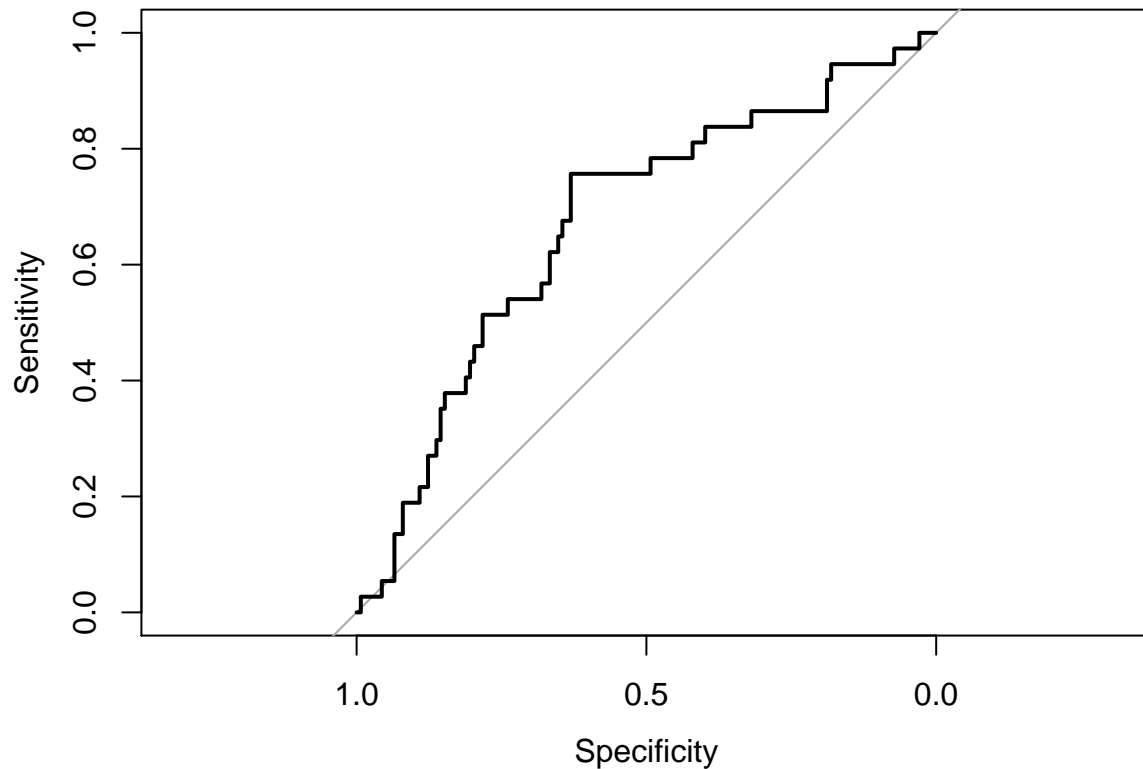
## TIF -8.208e-02 9.500e-02 -0.864 0.387594
## OLDCLAIM -3.217e-06 2.122e-05 -0.152 0.879507
## CLM_FREQ 2.384e-01 9.972e-01 0.239 0.811019
## MVR_PTS -1.559e-01 1.789e-01 -0.871 0.383514
## CAR_AGE -7.705e-02 7.468e-02 -1.032 0.302153
## PARENT1_Yes 5.325e-01 4.943e-01 1.077 0.281371
## MSTATUS_Yes -7.382e-01 3.378e-01 -2.185 0.028878 *
## SEX_z_F -8.921e-01 4.798e-01 -1.859 0.062998 .
## EDUCATION_.High.School -6.696e-01 8.107e-01 -0.826 0.408823
## EDUCATION_Bachelors -4.661e-01 6.913e-01 -0.674 0.500209
## EDUCATION_Masters 7.701e-01 6.185e-01 1.245 0.213057
## EDUCATION_z_High.School -2.096e-01 7.468e-01 -0.281 0.778969
## JOB_ -1.192e+00 7.044e-01 -1.692 0.090570 .
## JOB_Clerical -8.301e-02 4.992e-01 -0.166 0.867932
## JOB_Doctor 3.917e-01 9.228e-01 0.424 0.671229
## JOB_Home.Maker -6.296e-01 7.020e-01 -0.897 0.369824
## JOB_Lawyer -6.286e-01 6.775e-01 -0.928 0.353506
## JOB_Manager -5.923e-01 4.908e-01 -1.207 0.227529
## JOB_Student -8.673e-01 7.551e-01 -1.149 0.250737
## JOB_z_Blue.Collar -1.643e-02 4.630e-01 -0.035 0.971693
## CAR_USE_Commercial 6.980e-01 3.509e-01 1.989 0.046672 *
## CAR_TYPE_Panel.Truck 1.083e-01 6.523e-01 0.166 0.868078
## CAR_TYPE_Pickup 1.131e+00 4.177e-01 2.707 0.006795 **
## CAR_TYPE_Sports.Car 1.785e+00 5.389e-01 3.313 0.000925 ***
## CAR_TYPE_Van 4.078e-01 4.890e-01 0.834 0.404290
## CAR_TYPE_z_SUV 1.999e+00 4.851e-01 4.120 3.78e-05 ***
## RED_CAR_no -2.202e-01 3.463e-01 -0.636 0.524902
## REVOKED_Yes 7.078e-01 4.135e-01 1.712 0.086977 .
## URBANICITY_z_Highly.Rural..Rural -2.561e+00 4.505e-01 -5.685 1.31e-08 ***
## YOJ_NA -4.592e-01 4.383e-01 -1.048 0.294727
## INCOME_NA 1.934e-03 5.327e-01 0.004 0.997103
## CAR_AGE_NA -6.169e-01 5.060e-01 -1.219 0.222765
## HOME_VAL_NA 1.334e-01 2.944e-01 0.453 0.650590
## ageSquared 1.313e-03 1.182e-03 1.111 0.266639
## yojSquared 1.158e-02 7.007e-03 1.653 0.098298 .
## income_log 5.365e-01 3.160e-01 1.698 0.089547 .
## homeval_log -2.818e+00 1.427e+00 -1.974 0.048364 *
## travtime_log -1.045e-01 6.266e-01 -0.167 0.867510
## bluebook_log -1.849e-01 4.779e-01 -0.387 0.698788
## carage_log 9.648e-02 4.735e-01 0.204 0.838553
## oldclaim_log 3.880e-02 1.675e-01 0.232 0.816774
## clm_freq_log -2.417e-01 3.084e+00 -0.078 0.937519
## mvr_pts_log 6.648e-01 5.239e-01 1.269 0.204451
## tif_log 1.814e-01 5.296e-01 0.343 0.731928
## kidsdriv_log -8.517e-01 2.273e+00 -0.375 0.707880
## homekids_log 3.723e-01 1.531e+00 0.243 0.807889
## inter 4.372e-02 3.195e-02 1.368 0.171204
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 757.35 on 640 degrees of freedom
## Residual deviance: 528.80 on 585 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 640.8
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  25
##           1  20  12
##
##           Accuracy : 0.7429
##           95% CI : (0.6715, 0.8058)
##           No Information Rate : 0.7886
##           P-Value [Acc > NIR] : 0.9395
##
##           Kappa : 0.1887
##
## Mcnemar's Test P-Value : 0.5510
##
##           Sensitivity : 0.8551
##           Specificity : 0.3243
##           Pos Pred Value : 0.8252
##           Neg Pred Value : 0.3750
##           Prevalence : 0.7886
##           Detection Rate : 0.6743
##           Detection Prevalence : 0.8171
##           Balanced Accuracy : 0.5897
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.672150411280846"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 37 cases (dfPred_raw$class 1).
## Area under the curve: 0.6722
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0627  -0.6404  -0.3485   0.3511   2.5016
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.190e+01  1.579e+01   1.387 0.165441
## KIDSDRIV       1.763e+00  2.122e+00   0.831 0.406056
## AGE           -6.034e-02  1.049e-01  -0.575 0.565004
## HOMEKIDS      -8.397e-01  7.045e-01  -1.192 0.233289
## YOJ           -2.168e-01  1.369e-01  -1.584 0.113291
## INCOME        -2.584e-05  1.414e-05  -1.828 0.067532 .
## HOME_VAL       1.550e-05  9.552e-06   1.623 0.104566
## TRAVTIME       5.265e-02  2.159e-02   2.439 0.014726 *
## BLUEBOOK       4.407e-05  3.989e-05   1.105 0.269160
```



```

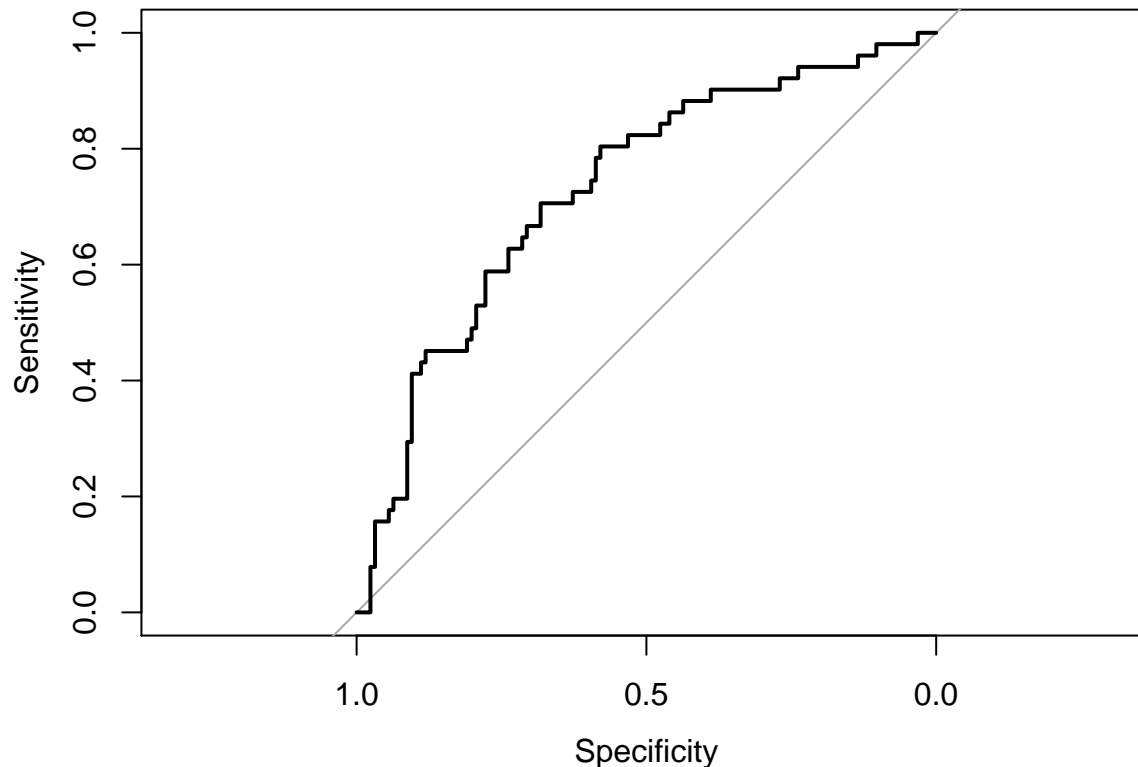
## TIF -3.402e-02 8.889e-02 -0.383 0.701904
## OLDCLAIM 7.864e-06 2.122e-05 0.371 0.710920
## CLM_FREQ -3.579e-02 9.595e-01 -0.037 0.970249
## MVRPTS 1.380e-02 1.725e-01 0.080 0.936267
## CAR_AGE 3.274e-02 7.143e-02 0.458 0.646678
## PARENT1_Yes 3.676e-01 4.803e-01 0.765 0.444000
## MSTATUS_Yes -6.902e-01 3.464e-01 -1.993 0.046306 *
## SEX_z_F -1.766e-01 4.707e-01 -0.375 0.707473
## EDUCATION_.High.School 8.389e-01 8.140e-01 1.031 0.302697
## EDUCATION_Bachelors 6.048e-01 6.933e-01 0.872 0.382948
## EDUCATION_Masters 1.382e+00 6.084e-01 2.272 0.023096 *
## EDUCATION_z_High.School 1.190e+00 7.464e-01 1.594 0.110865
## JOB_ -1.224e+00 7.067e-01 -1.732 0.083220 .
## JOB_Clerical -7.023e-01 5.041e-01 -1.393 0.163589
## JOB_Doctor 2.580e-01 9.347e-01 0.276 0.782539
## JOB_Home.Maker -8.868e-01 6.922e-01 -1.281 0.200160
## JOB_Lawyer -1.195e+00 6.746e-01 -1.772 0.076465 .
## JOB_Manager -1.076e+00 4.739e-01 -2.270 0.023229 *
## JOB_Student -1.594e+00 7.679e-01 -2.077 0.037847 *
## JOB_z_Blue.Collar -9.323e-01 4.605e-01 -2.024 0.042923 *
## CAR_USE_Commercial 1.156e+00 3.585e-01 3.226 0.001256 **
## CAR_TYPE_Panel.Truck -8.544e-02 6.411e-01 -0.133 0.893975
## CAR_TYPE_Pickup 7.317e-01 4.154e-01 1.762 0.078148 .
## CAR_TYPE_Sports.Car 1.162e+00 5.284e-01 2.199 0.027852 *
## CAR_TYPE_Van 3.784e-01 4.889e-01 0.774 0.438887
## CAR_TYPE_z_SUV 1.599e+00 4.756e-01 3.362 0.000775 ***
## RED_CAR_no -3.254e-01 3.467e-01 -0.939 0.347969
## REVOKED_Yes 3.859e-01 4.394e-01 0.878 0.379845
## URBANICITY_z_Highly.Rural..Rural -2.549e+00 4.813e-01 -5.297 1.18e-07 ***
## YOJ_NA -1.674e-01 4.645e-01 -0.360 0.718630
## INCOME_NA 1.819e-01 5.460e-01 0.333 0.739023
## CAR_AGE_NA -2.209e-01 5.056e-01 -0.437 0.662129
## HOME_VAL_NA -2.131e-01 2.958e-01 -0.721 0.471196
## ageSquared 6.530e-04 1.137e-03 0.574 0.565827
## yojSquared 1.069e-02 7.182e-03 1.488 0.136767
## income_log -4.063e-02 2.922e-01 -0.139 0.889385
## homeval_log -1.583e+00 1.491e+00 -1.061 0.288483
## travtime_log -1.053e+00 5.831e-01 -1.806 0.070981 .
## bluebook_log -1.763e-01 4.773e-01 -0.369 0.711854
## carage_log -4.097e-01 4.736e-01 -0.865 0.386990
## oldclaim_log 1.996e-02 1.650e-01 0.121 0.903665
## clm_freq_log 5.713e-01 2.992e+00 0.191 0.848552
## mvr_pts_log 4.526e-02 5.090e-01 0.089 0.929147
## tif_log 4.807e-03 5.123e-01 0.009 0.992513
## kidsdriv_log -3.419e+00 2.787e+00 -1.227 0.219930
## homekids_log 1.906e+00 1.503e+00 1.268 0.204630
## inter 2.352e-02 3.679e-02 0.639 0.522571
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 727.85 on 638 degrees of freedom
## Residual deviance: 525.73 on 583 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 637.73
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 107  28
##           1  19  23
##
##           Accuracy : 0.7345
##           95% CI : (0.663, 0.7979)
##           No Information Rate : 0.7119
##           P-Value [Acc > NIR] : 0.2834
##
##           Kappa : 0.3168
##
## Mcnemar's Test P-Value : 0.2432
##
##           Sensitivity : 0.8492
##           Specificity : 0.4510
##           Pos Pred Value : 0.7926
##           Neg Pred Value : 0.5476
##           Prevalence : 0.7119
##           Detection Rate : 0.6045
##           Detection Prevalence : 0.7627
##           Balanced Accuracy : 0.6501
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.73062558356676"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 126 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.7306
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2736  -0.6373  -0.3448   0.3637   3.1225
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.831e+01  1.539e+01   1.189  0.234336
## KIDSDRIV      -2.319e+00  1.775e+00  -1.306  0.191464
## AGE           -2.978e-01  1.108e-01  -2.688  0.007186 **
## HOMEKIDS      -8.445e-01  6.926e-01  -1.219  0.222724
## YOJ           -2.306e-01  1.449e-01  -1.591  0.111536
## INCOME        -1.734e-05  1.354e-05  -1.281  0.200310
## HOME_VAL       9.883e-06  9.012e-06   1.097  0.272755
## TRAVTIME       3.110e-02  2.255e-02   1.379  0.167877
## BLUEBOOK       6.758e-05  3.995e-05   1.692  0.090733 .
```

```

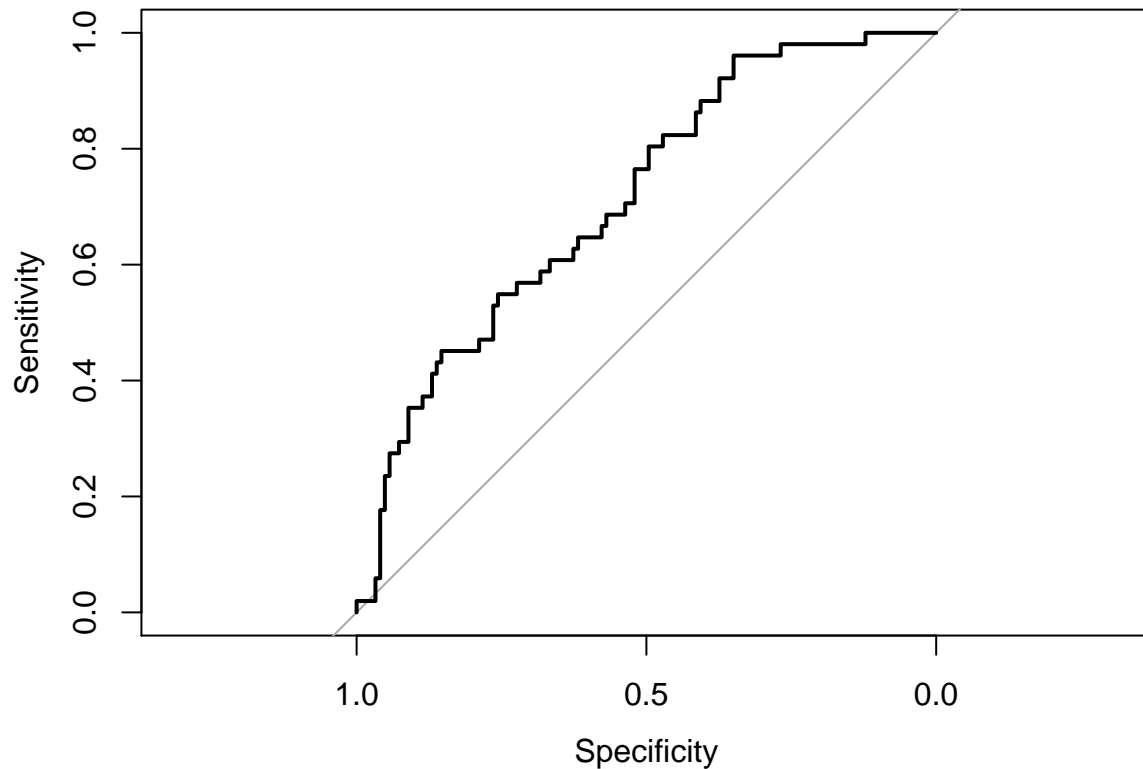
## TIF -1.012e-01 9.999e-02 -1.012 0.311359
## OLDCLAIM -1.600e-06 2.159e-05 -0.074 0.940924
## CLM_FREQ 1.562e-01 1.018e+00 0.153 0.878079
## MVR_PTS -2.671e-02 1.751e-01 -0.153 0.878777
## CAR_AGE -9.974e-03 7.624e-02 -0.131 0.895916
## PARENT1_Yes 2.980e-01 4.884e-01 0.610 0.541804
## MSTATUS_Yes -8.272e-01 3.444e-01 -2.401 0.016329 *
## SEX_z_F -5.744e-01 4.873e-01 -1.179 0.238520
## EDUCATION_.High.School -1.804e-01 7.783e-01 -0.232 0.816752
## EDUCATION_Bachelors -5.752e-01 6.601e-01 -0.871 0.383503
## EDUCATION_Masters 5.223e-01 5.596e-01 0.933 0.350669
## EDUCATION_z_High.School -1.909e-01 7.127e-01 -0.268 0.788774
## JOB_ -1.614e+00 7.057e-01 -2.287 0.022197 *
## JOB_Clerical -3.404e-01 4.934e-01 -0.690 0.490260
## JOB_Doctor -1.162e+00 9.874e-01 -1.176 0.239401
## JOB_Home.Maker 2.240e-01 6.501e-01 0.345 0.730452
## JOB_Lawyer -8.934e-01 6.418e-01 -1.392 0.163869
## JOB_Manager -1.038e+00 4.895e-01 -2.121 0.033901 *
## JOB_Student -2.101e-01 7.401e-01 -0.284 0.776554
## JOB_z_Blue.Collar 5.040e-02 4.697e-01 0.107 0.914548
## CAR_USE_Commercial 6.525e-01 3.628e-01 1.798 0.072118 .
## CAR_TYPE_Panel.Truck 1.425e-01 6.309e-01 0.226 0.821352
## CAR_TYPE_Pickup 1.096e+00 4.208e-01 2.605 0.009189 **
## CAR_TYPE_Sports.Car 1.638e+00 5.567e-01 2.943 0.003250 **
## CAR_TYPE_Van 3.818e-01 4.887e-01 0.781 0.434647
## CAR_TYPE_z_SUV 1.815e+00 4.941e-01 3.674 0.000239 ***
## RED_CAR_no -2.898e-01 3.509e-01 -0.826 0.408837
## REVOKED_Yes 6.342e-01 3.983e-01 1.592 0.111308
## URBANICITY_z_Highly.Rural..Rural -2.490e+00 4.224e-01 -5.896 3.72e-09 ***
## YOJ_NA -2.255e-01 4.487e-01 -0.503 0.615314
## INCOME_NA 6.135e-01 6.903e-01 0.889 0.374167
## CAR_AGE_NA 1.279e-01 5.023e-01 0.255 0.799029
## HOME_VAL_NA -9.176e-02 3.004e-01 -0.305 0.760002
## ageSquared 3.028e-03 1.186e-03 2.553 0.010688 *
## yojSquared 1.122e-02 7.294e-03 1.538 0.124043
## income_log 7.562e-02 2.934e-01 0.258 0.796583
## homeval_log -7.050e-01 1.443e+00 -0.489 0.625050
## travtime_log -1.709e-01 6.400e-01 -0.267 0.789482
## bluebook_log -6.124e-01 4.898e-01 -1.250 0.211159
## carage_log -3.097e-01 4.914e-01 -0.630 0.528566
## oldclaim_log 1.216e-01 1.726e-01 0.705 0.480986
## clm_freq_log -8.428e-01 3.158e+00 -0.267 0.789578
## mvr_pts_log 1.338e-01 5.123e-01 0.261 0.793936
## tif_log 2.951e-01 5.542e-01 0.532 0.594471
## kidsdriv_log 2.316e+00 2.211e+00 1.048 0.294701
## homekids_log 1.623e+00 1.507e+00 1.076 0.281731
## inter 4.348e-02 3.381e-02 1.286 0.198495
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 729.63 on 641 degrees of freedom
## Residual deviance: 524.24 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 636.24
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 107  30
##           1  16  21
##
##           Accuracy : 0.7356
##           95% CI : (0.6636, 0.7995)
##           No Information Rate : 0.7069
##           P-Value [Acc > NIR] : 0.22836
##
##           Kappa : 0.3063
##
## Mcnemar's Test P-Value : 0.05527
##
##           Sensitivity : 0.8699
##           Specificity : 0.4118
##           Pos Pred Value : 0.7810
##           Neg Pred Value : 0.5676
##           Prevalence : 0.7069
##           Detection Rate : 0.6149
##           Detection Prevalence : 0.7874
##           Balanced Accuracy : 0.6408
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.713055954088953"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 123 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.7131
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0526  -0.6711  -0.3475   0.5547   3.2624
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.496e+01  1.448e+01   1.724  0.08478 .
## KIDSDRIV      -3.259e+00  1.830e+00  -1.781  0.07498 .
## AGE           -1.339e-01  1.130e-01  -1.185  0.23592
## HOMEKIDS      -6.294e-01  6.641e-01  -0.948  0.34325
## YOJ           -1.385e-02  1.290e-01  -0.107  0.91452
## INCOME        -2.218e-05  1.321e-05  -1.679  0.09312 .
## HOME_VAL       1.387e-05  8.793e-06   1.578  0.11466
## TRAVTIME       4.223e-02  2.284e-02   1.849  0.06446 .
## BLUEBOOK       8.043e-05  3.743e-05   2.149  0.03163 *
```

```

## TIF 4.467e-02 9.302e-02 0.480 0.63110
## OLDCLAIM -1.119e-05 2.115e-05 -0.529 0.59661
## CLM_FREQ 5.446e-01 1.023e+00 0.533 0.59432
## MVR_PTS 1.427e-01 1.827e-01 0.781 0.43483
## CAR_AGE -7.425e-03 7.112e-02 -0.104 0.91685
## PARENT1_Yes 3.485e-01 4.702e-01 0.741 0.45850
## MSTATUS_Yes -7.253e-01 3.366e-01 -2.155 0.03119 *
## SEX_z_F -7.758e-01 4.787e-01 -1.621 0.10511
## EDUCATION_.High.School 1.727e-01 7.951e-01 0.217 0.82807
## EDUCATION_Bachelors 9.949e-03 6.553e-01 0.015 0.98789
## EDUCATION_Masters 6.527e-01 5.509e-01 1.185 0.23612
## EDUCATION_z_High.School 6.547e-01 7.093e-01 0.923 0.35604
## JOB_ -9.299e-01 6.929e-01 -1.342 0.17956
## JOB_Clerical -2.591e-01 4.819e-01 -0.538 0.59085
## JOB_Doctor -3.265e-01 9.842e-01 -0.332 0.74010
## JOB_Home.Maker -2.587e-01 6.725e-01 -0.385 0.70043
## JOB_Lawyer -4.303e-01 6.462e-01 -0.666 0.50545
## JOB_Manager -8.930e-01 4.956e-01 -1.802 0.07158 .
## JOB_Student -3.053e-01 7.147e-01 -0.427 0.66929
## JOB_z_Blue.Collar -1.329e-01 4.655e-01 -0.285 0.77530
## CAR_USE_Commercial 3.850e-01 3.545e-01 1.086 0.27741
## CAR_TYPE_Panel.Truck 6.778e-01 6.205e-01 1.092 0.27470
## CAR_TYPE_Pickup 1.260e+00 4.234e-01 2.975 0.00293 **
## CAR_TYPE_Sports.Car 2.452e+00 5.802e-01 4.226 2.37e-05 ***
## CAR_TYPE_Van 8.329e-01 4.679e-01 1.780 0.07508 .
## CAR_TYPE_z_SUV 2.246e+00 5.019e-01 4.476 7.61e-06 ***
## RED_CAR_no -3.164e-01 3.488e-01 -0.907 0.36444
## REVOKED_Yes 6.319e-01 4.075e-01 1.551 0.12099
## URBANICITY_z_Highly.Rural..Rural -2.822e+00 4.586e-01 -6.153 7.59e-10 ***
## YOJ_NA -2.109e-01 4.359e-01 -0.484 0.62855
## INCOME_NA 2.410e-01 5.218e-01 0.462 0.64410
## CAR_AGE_NA -3.345e-01 4.841e-01 -0.691 0.48951
## HOME_VAL_NA -3.822e-02 2.851e-01 -0.134 0.89336
## ageSquared 1.286e-03 1.229e-03 1.046 0.29550
## yojSquared -5.563e-04 6.734e-03 -0.083 0.93416
## income_log 1.459e-01 3.042e-01 0.480 0.63143
## homeval_log -1.597e+00 1.365e+00 -1.170 0.24181
## travtime_log -4.612e-01 6.269e-01 -0.736 0.46192
## bluebook_log -7.300e-01 4.495e-01 -1.624 0.10439
## carage_log -1.287e-01 4.650e-01 -0.277 0.78203
## oldclaim_log 1.048e-01 1.671e-01 0.627 0.53053
## clm_freq_log -1.586e+00 3.144e+00 -0.504 0.61395
## mvr_pts_log -2.136e-01 5.234e-01 -0.408 0.68318
## tif_log -4.216e-01 5.287e-01 -0.798 0.42514
## kidsdriv_log 1.748e+00 2.170e+00 0.806 0.42040
## homekids_log 1.625e+00 1.459e+00 1.114 0.26518
## inter 6.252e-02 3.360e-02 1.860 0.06282 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 740.94 on 639 degrees of freedom
## Residual deviance: 535.78 on 584 degrees of freedom

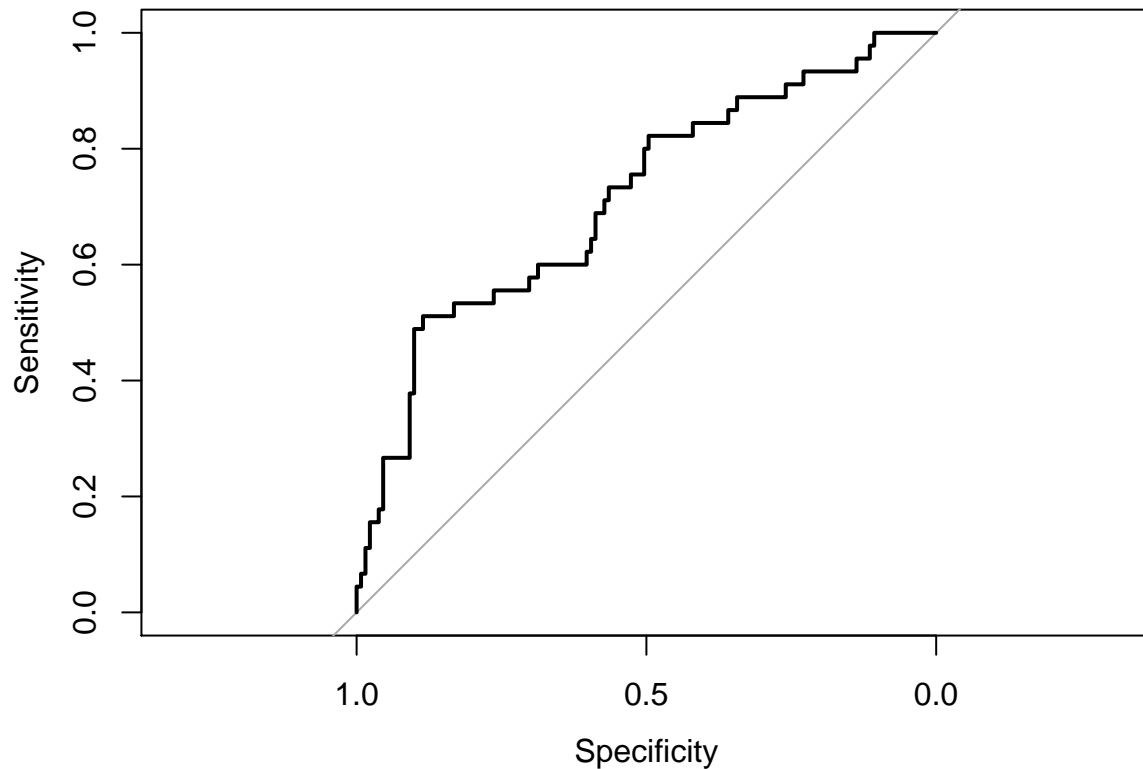
```

```

## (1 observation deleted due to missingness)
## AIC: 647.78
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  23
##           1  13  22
##
##           Accuracy : 0.7955
##           95% CI : (0.7282, 0.8524)
##           No Information Rate : 0.7443
##           P-Value [Acc > NIR] : 0.06847
##
##           Kappa : 0.4203
##
## Mcnemar's Test P-Value : 0.13361
##
##           Sensitivity : 0.9008
##           Specificity : 0.4889
##           Pos Pred Value : 0.8369
##           Neg Pred Value : 0.6286
##           Prevalence : 0.7443
##           Detection Rate : 0.6705
##           Detection Prevalence : 0.8011
##           Balanced Accuracy : 0.6948
##
##           'Positive' Class : 0
##

```





```
## [1] "AUC: 0.713825275657337"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7138
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3060  -0.6867  -0.3662   0.6517   3.1086
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.209e+01  1.430e+01   1.545  0.122324
## KIDSDRIV      -3.800e-01  1.642e+00  -0.231  0.816979
## AGE           -1.672e-01  1.039e-01  -1.609  0.107591
## HOMEKIDS      -7.897e-01  6.971e-01  -1.133  0.257240
## YOJ           -1.932e-01  1.260e-01  -1.534  0.125134
## INCOME        -4.387e-06  1.271e-05  -0.345  0.729930
## HOME_VAL       6.406e-06  8.453e-06   0.758  0.448550
## TRAVTIME      4.805e-02  2.327e-02   2.065  0.038908 *
## BLUEBOOK      4.396e-05  3.951e-05   1.113  0.265845
```

```

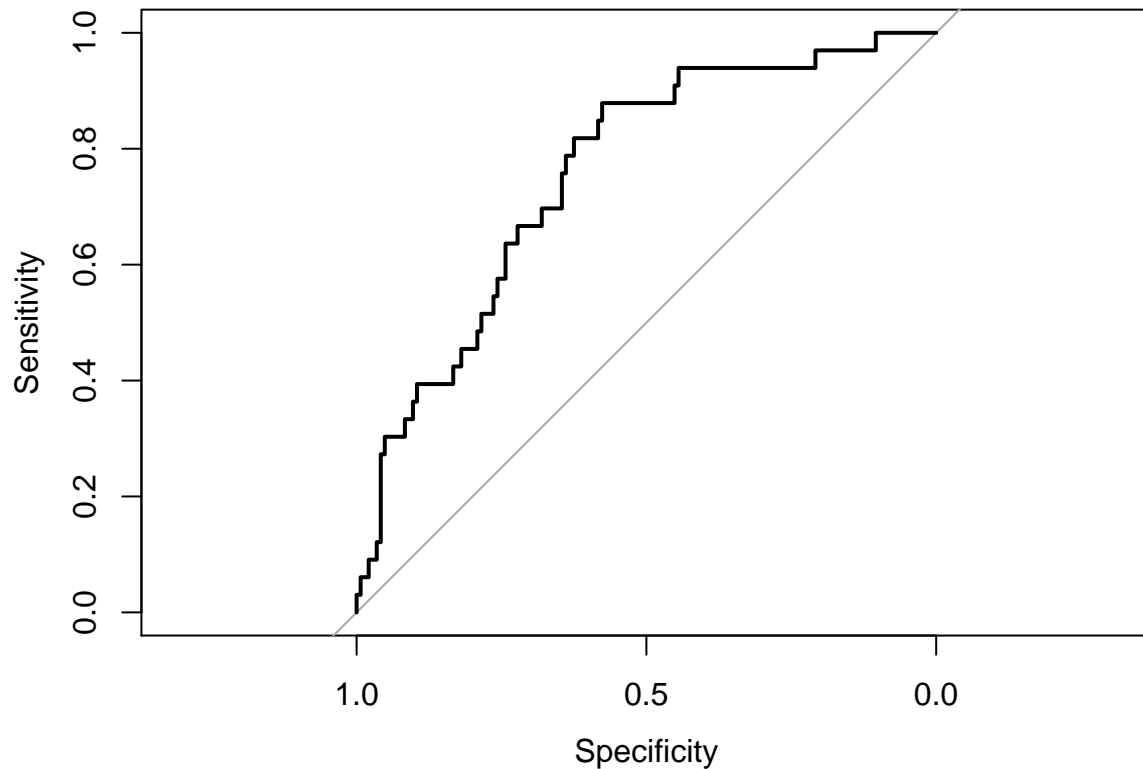
## TIF -5.726e-02 9.112e-02 -0.628 0.529759
## OLDCLAIM 3.853e-06 2.042e-05 0.189 0.850342
## CLM_FREQ 2.076e-01 9.588e-01 0.216 0.828625
## MVRPTS 2.927e-02 1.663e-01 0.176 0.860271
## CAR_AGE -3.941e-02 6.800e-02 -0.579 0.562264
## PARENT1_Yes 6.515e-01 4.695e-01 1.388 0.165252
## MSTATUS_Yes -5.005e-01 3.265e-01 -1.533 0.125289
## SEX_z_F -8.981e-01 4.585e-01 -1.959 0.050167
## EDUCATION_High.School -4.848e-01 7.524e-01 -0.644 0.519360
## EDUCATION_Bachelors -5.255e-01 6.312e-01 -0.833 0.405122
## EDUCATION_Masters 5.296e-01 5.336e-01 0.993 0.320947
## EDUCATION_z_High.School -8.677e-02 6.881e-01 -0.126 0.899644
## JOB_ -7.025e-01 6.624e-01 -1.060 0.288955
## JOB_Clerical 3.500e-01 4.856e-01 0.721 0.471095
## JOB_Doctor -5.790e-01 8.617e-01 -0.672 0.501585
## JOB_Home.Maker -4.743e-01 6.703e-01 -0.708 0.479249
## JOB_Lawyer -6.783e-01 6.370e-01 -1.065 0.286956
## JOB_Manager -7.717e-01 4.625e-01 -1.669 0.095189
## JOB_Student -8.846e-01 7.285e-01 -1.214 0.224649
## JOB_z_Blue.Collar 2.667e-01 4.486e-01 0.595 0.552094
## CAR_USE_Commercial 3.014e-01 3.469e-01 0.869 0.384994
## CAR_TYPE_Panel.Truck 2.085e-01 6.240e-01 0.334 0.738272
## CAR_TYPE_Pickup 1.536e+00 4.081e-01 3.765 0.000167 ***
## CAR_TYPE_Sports.Car 1.763e+00 5.251e-01 3.358 0.000786 ***
## CAR_TYPE_Van 6.817e-01 4.527e-01 1.506 0.132095
## CAR_TYPE_z_SUV 2.164e+00 4.709e-01 4.594 4.34e-06 ***
## RED_CAR_no -2.420e-01 3.360e-01 -0.720 0.471265
## REVOKED_Yes 3.293e-01 3.887e-01 0.847 0.396941
## URBANICITY_z_Highly.Rural..Rural -2.761e+00 4.381e-01 -6.302 2.94e-10 ***
## YOJ_NA -9.895e-02 4.380e-01 -0.226 0.821282
## INCOME_NA 1.793e-01 5.074e-01 0.353 0.723754
## CAR_AGE_NA 4.175e-01 5.309e-01 0.787 0.431574
## HOME_VAL_NA -2.335e-01 2.832e-01 -0.825 0.409652
## ageSquared 1.864e-03 1.135e-03 1.642 0.100563
## yojSquared 9.374e-03 6.488e-03 1.445 0.148543
## income_log -2.190e-02 2.912e-01 -0.075 0.940039
## homeval_log -1.520e+00 1.333e+00 -1.140 0.254300
## travtime_log -5.013e-01 6.366e-01 -0.787 0.431000
## bluebook_log -1.555e-01 4.869e-01 -0.319 0.749394
## carage_log -2.405e-02 4.494e-01 -0.054 0.957332
## oldclaim_log 1.102e-01 1.562e-01 0.705 0.480585
## clm_freq_log -7.920e-01 2.941e+00 -0.269 0.787730
## mvr_pts_log -1.443e-01 4.926e-01 -0.293 0.769619
## tif_log -1.427e-02 5.123e-01 -0.028 0.977783
## kidsdriv_log 6.912e-01 2.188e+00 0.316 0.752082
## homekids_log 1.729e+00 1.499e+00 1.153 0.248819
## inter 1.539e-02 3.026e-02 0.508 0.611123
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 763.54 on 638 degrees of freedom
## Residual deviance: 559.34 on 583 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 671.34
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 115  18
##           1  29  15
##
##           Accuracy : 0.7345
##           95% CI : (0.663, 0.7979)
##           No Information Rate : 0.8136
##           P-Value [Acc > NIR] : 0.9964
##
##           Kappa : 0.2243
##
## Mcnemar's Test P-Value : 0.1447
##
##           Sensitivity : 0.7986
##           Specificity : 0.4545
##           Pos Pred Value : 0.8647
##           Neg Pred Value : 0.3409
##           Prevalence : 0.8136
##           Detection Rate : 0.6497
##           Detection Prevalence : 0.7514
##           Balanced Accuracy : 0.6266
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.756313131313131"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 144 controls (dfPred_raw$class 0) < 33 cases (dfPred_raw$class 1).
## Area under the curve: 0.7563
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1614  -0.6646  -0.3494   0.3634   3.1551
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    4.916e+00  1.505e+01   0.327  0.74400
## KIDSDRIV      -3.587e+00  2.208e+00  -1.624  0.10429
## AGE           -1.681e-01  1.074e-01  -1.566  0.11744
## HOMEKIDS       -9.972e-01  7.267e-01  -1.372  0.16997
## YOJ            -2.083e-01  1.364e-01  -1.527  0.12667
## INCOME         -1.529e-05  1.402e-05  -1.090  0.27572
## HOME_VAL        3.662e-06  9.408e-06   0.389  0.69706
## TRAVTIME        3.180e-02  2.110e-02   1.507  0.13175
## BLUEBOOK        6.342e-05  3.830e-05   1.656  0.09773 .
```

```

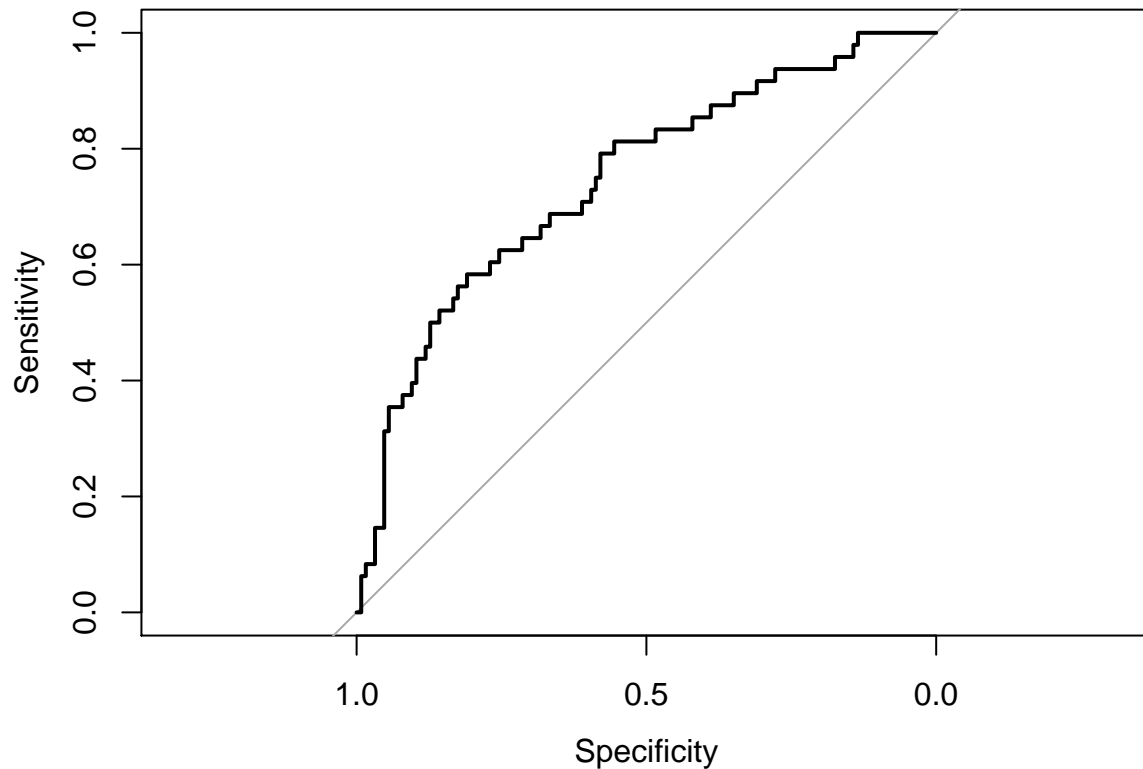
## TIF -1.096e-01 9.907e-02 -1.107 0.26849
## OLDCLAIM -4.864e-06 2.166e-05 -0.225 0.82234
## CLM_FREQ 8.378e-02 9.729e-01 0.086 0.93138
## MVRPTS 1.848e-02 1.629e-01 0.113 0.90971
## CAR_AGE -1.466e-02 7.214e-02 -0.203 0.83901
## PARENT1_Yes 6.615e-01 4.916e-01 1.345 0.17846
## MSTATUS_Yes -5.829e-01 3.317e-01 -1.757 0.07886
## SEX_z_F -2.558e-01 4.527e-01 -0.565 0.57200
## EDUCATION_.High.School 2.786e-01 8.055e-01 0.346 0.72943
## EDUCATION_Bachelors 2.935e-01 6.549e-01 0.448 0.65400
## EDUCATION_Masters 9.203e-01 5.856e-01 1.572 0.11604
## EDUCATION_z_High.School 6.167e-01 7.234e-01 0.853 0.39391
## JOB_ -5.945e-01 6.913e-01 -0.860 0.38980
## JOB_Clerical 1.004e-01 4.948e-01 0.203 0.83927
## JOB_Doctor 4.545e-01 9.034e-01 0.503 0.61489
## JOB_Home.Maker 2.974e-01 6.660e-01 0.447 0.65515
## JOB_Lawyer -5.239e-01 6.442e-01 -0.813 0.41605
## JOB_Manager -4.976e-01 4.719e-01 -1.054 0.29168
## JOB_Student 3.544e-02 7.386e-01 0.048 0.96173
## JOB_z_Blue.Collar -1.906e-01 4.762e-01 -0.400 0.68904
## CAR_USE_Commercial 8.680e-01 3.606e-01 2.407 0.01607 *
## CAR_TYPE_Panel.Truck -2.368e-01 6.230e-01 -0.380 0.70389
## CAR_TYPE_Pickup 6.171e-01 4.255e-01 1.450 0.14700
## CAR_TYPE_Sports.Car 1.466e+00 5.035e-01 2.912 0.00359 **
## CAR_TYPE_Van 1.441e-01 4.702e-01 0.307 0.75921
## CAR_TYPE_z_SUV 1.292e+00 4.524e-01 2.857 0.00428 **
## RED_CAR_no -4.708e-01 3.489e-01 -1.349 0.17719
## REVOKED_Yes 5.928e-01 4.182e-01 1.418 0.15629
## URBANICITY_z_Highly.Rural..Rural -2.843e+00 4.809e-01 -5.911 3.41e-09 ***
## YOJ_NA -5.571e-01 4.350e-01 -1.280 0.20037
## INCOME_NA 6.340e-02 5.440e-01 0.117 0.90723
## CAR_AGE_NA -5.233e-01 5.034e-01 -1.039 0.29858
## HOME_VAL_NA -4.687e-02 2.967e-01 -0.158 0.87448
## ageSquared 1.771e-03 1.161e-03 1.525 0.12735
## yojSquared 1.072e-02 7.089e-03 1.513 0.13033
## income_log 2.721e-02 2.967e-01 0.092 0.92692
## homeval_log 8.895e-02 1.434e+00 0.062 0.95053
## travtime_log -3.597e-01 5.871e-01 -0.613 0.54007
## bluebook_log -2.944e-01 4.518e-01 -0.651 0.51473
## carage_log -2.440e-01 4.677e-01 -0.522 0.60192
## oldclaim_log 2.337e-02 1.657e-01 0.141 0.88786
## clm_freq_log 5.381e-02 3.007e+00 0.018 0.98572
## mvr_pts_log 9.838e-02 4.892e-01 0.201 0.84063
## tif_log 4.550e-01 5.458e-01 0.834 0.40452
## kidsdriv_log 1.500e+00 2.310e+00 0.649 0.51603
## homekids_log 1.914e+00 1.538e+00 1.244 0.21346
## inter 8.243e-02 4.283e-02 1.925 0.05429
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 735.97 on 641 degrees of freedom
## Residual deviance: 537.26 on 586 degrees of freedom

```

```

## (1 observation deleted due to missingness)
## AIC: 649.26
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 111  27
##           1  15  21
##
##           Accuracy : 0.7586
##           95% CI : (0.6881, 0.8202)
##           No Information Rate : 0.7241
##           P-Value [Acc > NIR] : 0.17584
##
##           Kappa : 0.3452
##
## Mcnemar's Test P-Value : 0.08963
##
##           Sensitivity : 0.8810
##           Specificity : 0.4375
##           Pos Pred Value : 0.8043
##           Neg Pred Value : 0.5833
##           Prevalence : 0.7241
##           Detection Rate : 0.6379
##           Detection Prevalence : 0.7931
##           Balanced Accuracy : 0.6592
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.744212962962963"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 126 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.7442
## [1] "Accuracy: 0.743344577383705"
## [1] "AIC: 651.450894189602"
## [1] "AUC: 0.739864574020757"
```

The “kitchen sink” model offers a little improvement: Accuracy .752, AIC 654 and AUC .75.

**Create Model 3 - Use backward elimination to choose the best model:** We use backward elimination to achieve a better fit and lower AIC.

```
##
## Call:
## glm(formula = fla, family = "binomial", data = df)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9456  -0.7302  -0.4158   0.6675   2.8498
##
```

```

## Coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.967e+00  4.282e-01  -4.595 4.34e-06 ***
## INCOME          -1.003e-05  2.567e-06  -3.909 9.28e-05 ***
## TRAVTIME         2.081e-02  6.109e-03   3.406 0.000659 ***
## BLUEBOOK         3.122e-05  1.352e-05   2.310 0.020897 *
## TIF             -3.942e-02  2.308e-02  -1.708 0.087645 .
## OLDCLAIM         1.247e-05  1.231e-05   1.013 0.311218
## PARENT1_Yes      1.057e+00  2.652e-01   3.987 6.68e-05 ***
## SEX_z_F         -8.977e-01  3.053e-01  -2.941 0.003276 **
## JOB_Manager     -6.325e-01  3.123e-01  -2.025 0.042818 *
## CAR_USE_Commercial 4.699e-01  2.093e-01   2.245 0.024785 *
## CAR_TYPE_Pickup   9.750e-01  2.798e-01   3.485 0.000492 ***
## CAR_TYPE_Sports.Car 1.794e+00  4.165e-01   4.308 1.65e-05 ***
## CAR_TYPE_z_SUV    1.619e+00  3.704e-01   4.370 1.24e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.396e+00  3.571e-01  -6.709 1.95e-11 ***
## HOME_VAL_NA      -3.424e-01  1.941e-01  -1.763 0.077823 .
## oldclaim_log      5.281e-02  2.753e-02   1.919 0.055024 .
## inter            1.454e-02  3.655e-03   3.979 6.91e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 943.78  on 816  degrees of freedom
## Residual deviance: 743.16  on 800  degrees of freedom
## AIC: 777.16
##
## Number of Fisher Scoring iterations: 5

## [[1]]
##              (Intercept)              INCOME
##      -1.967428e+00      -1.003291e-05
##              TRAVTIME              BLUEBOOK
##      2.080659e-02      3.122284e-05
##              TIF              OLDCLAIM
##      -3.941844e-02      1.246735e-05
##      PARENT1_Yes              SEX_z_F
##      1.057252e+00      -8.977356e-01
##      JOB_Manager              CAR_USE_Commercial
##      -6.325399e-01      4.698847e-01
##      CAR_TYPE_Pickup              CAR_TYPE_Sports.Car
##      9.749544e-01      1.794413e+00
##      CAR_TYPE_z_SUV URBANICITY_z_Highly.Rural..Rural
##      1.618999e+00      -2.395647e+00
##      HOME_VAL_NA              oldclaim_log
##      -3.423721e-01      5.281420e-02
##              inter
##      1.454299e-02
##
## [[2]]
## [1] 0
##
## [[3]]

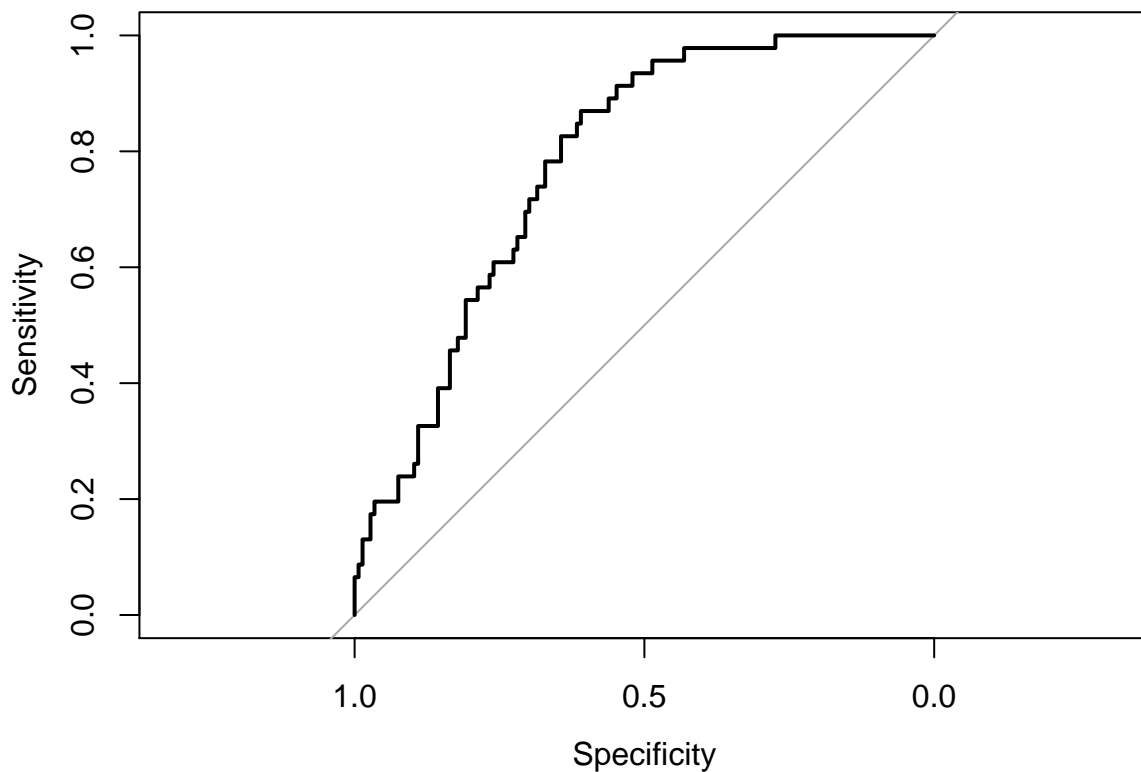
```



```
## [1] 0

##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1465  -0.7348  -0.4055   0.7038   2.7665
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.263e+00  5.052e-01  -4.480 7.48e-06 ***
## INCOME         -1.099e-05  2.929e-06  -3.752 0.000175 ***
## TRAVTIME        2.414e-02  7.191e-03   3.356 0.000790 ***
## BLUEBOOK        5.110e-05  1.604e-05   3.186 0.001441 **
## TIF            -6.506e-02  2.747e-02  -2.368 0.017873 *
## OLDCLAIM        1.766e-05  1.358e-05   1.301 0.193406
## PARENT1_Yes     9.767e-01  2.988e-01   3.269 0.001079 **
## SEX_z_F        -1.009e+00  3.510e-01  -2.874 0.004056 **
## JOB_Manager    -8.139e-01  3.745e-01  -2.173 0.029755 *
## CAR_USE_Commercial 2.396e-01  2.467e-01   0.971 0.331428
## CAR_TYPE_Pickup  1.273e+00  3.276e-01   3.886 0.000102 ***
## CAR_TYPE_Sports.Car 1.913e+00  4.927e-01   3.882 0.000104 ***
## CAR_TYPE_SUV     1.813e+00  4.281e-01   4.235 2.28e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.313e+00  3.960e-01  -5.840 5.22e-09 ***
## HOME_VAL_NA     -2.933e-01  2.227e-01  -1.317 0.187800
## oldclaim_log     6.007e-02  3.142e-02   1.912 0.055893 .
## inter           1.652e-02  4.311e-03   3.831 0.000128 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 731.55  on 624  degrees of freedom
## Residual deviance: 568.35  on 608  degrees of freedom
## AIC: 602.35
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 127  31
##              1  19  15
##
##              Accuracy : 0.7396
##              95% CI : (0.6715, 0.8001)
## No Information Rate : 0.7604
## P-Value [Acc > NIR] : 0.7785
##
##              Kappa : 0.2152
##
```

```
## McNemar's Test P-Value : 0.1198
##
##      Sensitivity : 0.8699
##      Specificity : 0.3261
##      Pos Pred Value : 0.8038
##      Neg Pred Value : 0.4412
##      Prevalence : 0.7604
##      Detection Rate : 0.6615
##      Detection Prevalence : 0.8229
##      Balanced Accuracy : 0.5980
##
##      'Positive' Class : 0
##
```



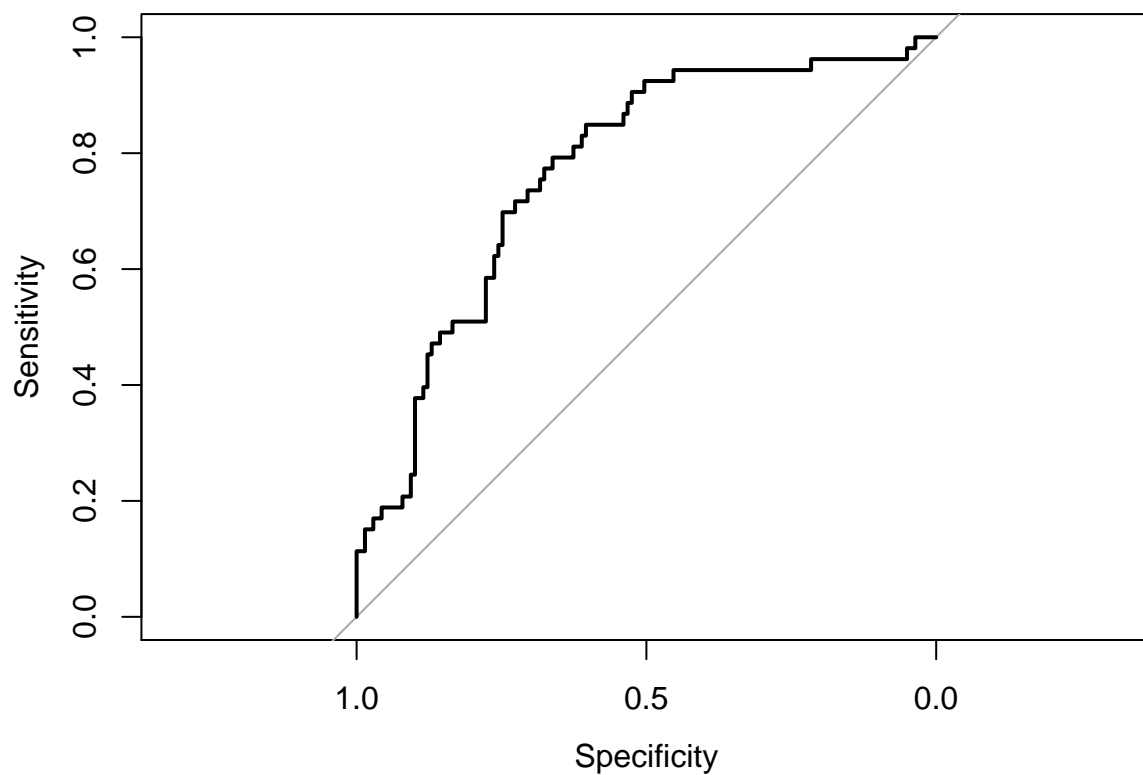
```
## [1] "AUC: 0.77933293627159"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7793
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
```

```

##      Min      1Q   Median      3Q      Max
## -1.9207  -0.7041  -0.3869   0.6167   2.8912
##
## Coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.621e+00  4.899e-01  -3.308 0.000939 ***
## INCOME         -9.651e-06  3.111e-06  -3.102 0.001921 **
## TRAVTIME        1.754e-02  6.964e-03   2.519 0.011774 *
## BLUEBOOK        1.034e-05  1.575e-05   0.657 0.511345
## TIF            -2.463e-02  2.747e-02  -0.897 0.369940
## OLDCLAIM        9.592e-06  1.502e-05   0.639 0.523139
## PARENT1_Yes     9.813e-01  3.066e-01   3.200 0.001373 ***
## SEX_z_F        -1.007e+00  3.598e-01  -2.797 0.005155 **
## JOB_Manager    -7.300e-01  3.834e-01  -1.904 0.056940 .
## CAR_USE_Commercial 4.832e-01  2.505e-01   1.929 0.053726 .
## CAR_TYPE_Pickup  9.774e-01  3.219e-01   3.037 0.002391 **
## CAR_TYPE_Sports.Car 1.776e+00  4.781e-01   3.715 0.000203 ***
## CAR_TYPE_z_SUV   1.684e+00  4.343e-01   3.877 0.000106 ***
## URBANICITY_z_Highly.Rural..Rural -2.262e+00  4.113e-01  -5.500 3.8e-08 ***
## HOME_VAL_NA     -4.795e-01  2.264e-01  -2.118 0.034134 *
## oldclaim_log     6.816e-02  3.134e-02   2.175 0.029611 *
## inter           1.626e-02  4.389e-03   3.706 0.000211 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 717.37  on 624  degrees of freedom
## Residual deviance: 556.43  on 608  degrees of freedom
## AIC: 590.43
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 126  40
##              1  13  13
##
##              Accuracy : 0.724
##              95% CI : (0.655, 0.7859)
##      No Information Rate : 0.724
##      P-Value [Acc > NIR] : 0.5369215
##
##              Kappa : 0.1801
##
## Mcnemar's Test P-Value : 0.0003551
##
##              Sensitivity : 0.9065
##              Specificity : 0.2453
##      Pos Pred Value : 0.7590
##      Neg Pred Value : 0.5000
##              Prevalence : 0.7240

```

```
##      Detection Rate : 0.6562
##      Detection Prevalence : 0.8646
##      Balanced Accuracy : 0.5759
##
##      'Positive' Class : 0
##
```

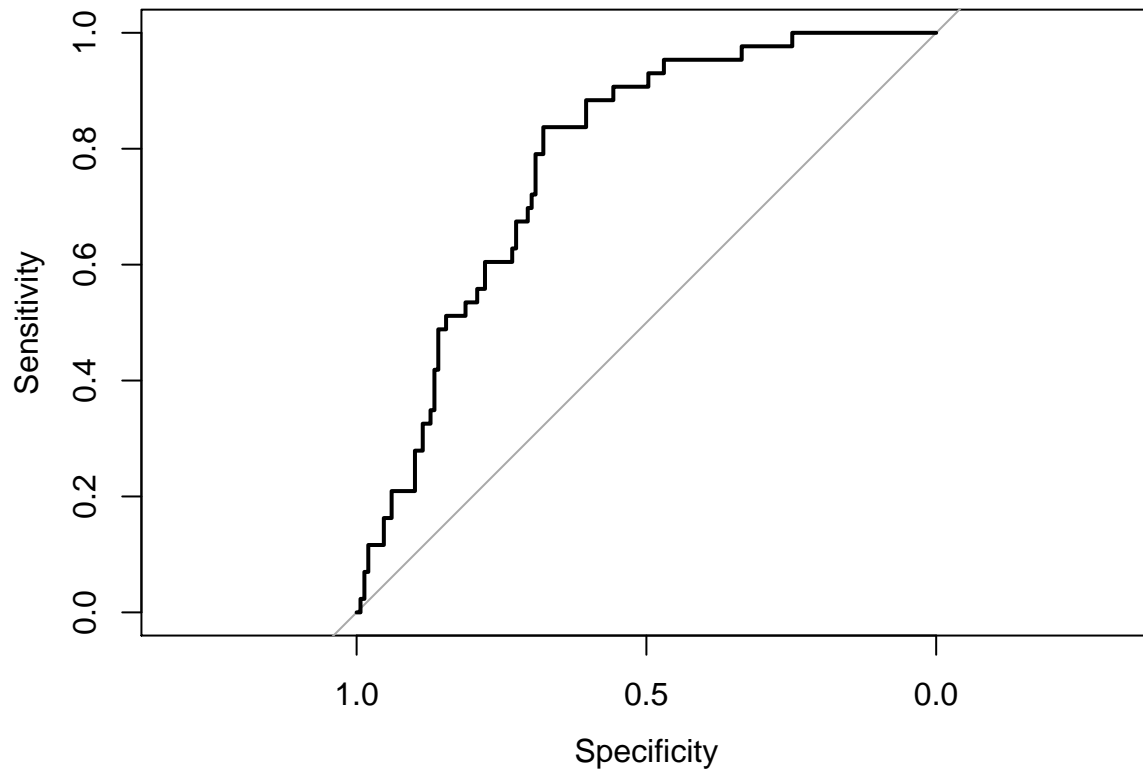


```
## [1] "AUC: 0.769648432197638"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 139 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.7696
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8099  -0.7447  -0.3989   0.6576   2.8704
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.844e+00  4.847e-01  -3.803 0.000143 ***
## INCOME        -9.875e-06  2.993e-06  -3.299 0.000969 ***
```

```

## TRAVTIME                2.149e-02  7.051e-03   3.048 0.002303 **
## BLUEBOOK                2.526e-05  1.514e-05   1.668 0.095302 .
## TIF                     -4.251e-02  2.733e-02  -1.555 0.119898
## OLDCLAIM                1.361e-05  1.407e-05   0.967 0.333414
## PARENT1_Yes             9.912e-01  3.159e-01   3.137 0.001704 **
## SEX_z_F                 -9.893e-01  3.597e-01  -2.750 0.005953 **
## JOB_Manager             -6.763e-01  3.526e-01  -1.918 0.055124 .
## CAR_USE_Commercial      5.550e-01  2.467e-01   2.250 0.024477 *
## CAR_TYPE_Pickup         8.725e-01  3.169e-01   2.753 0.005904 **
## CAR_TYPE_Sports.Car     1.872e+00  4.817e-01   3.887 0.000102 ***
## CAR_TYPE_z_SUV         1.753e+00  4.348e-01   4.031 5.56e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.426e+00  3.967e-01  -6.114 9.70e-10 ***
## HOME_VAL_NA             -4.159e-01  2.281e-01  -1.823 0.068321 .
## oldclaim_log            5.704e-02  3.142e-02   1.815 0.069470 .
## inter                   1.961e-02  4.244e-03   4.622 3.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 737.38  on 624  degrees of freedom
## Residual deviance: 568.24  on 608  degrees of freedom
## AIC: 602.24
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 128  25
##           1  21  18
##
##           Accuracy : 0.7604
##           95% CI : (0.6937, 0.8189)
##    No Information Rate : 0.776
##    P-Value [Acc > NIR] : 0.7310
##
##           Kappa : 0.2872
##
## Mcnemar's Test P-Value : 0.6583
##
##           Sensitivity : 0.8591
##           Specificity : 0.4186
##           Pos Pred Value : 0.8366
##           Neg Pred Value : 0.4615
##           Prevalence : 0.7760
##           Detection Rate : 0.6667
##           Detection Prevalence : 0.7969
##           Balanced Accuracy : 0.6388
##
##           'Positive' Class : 0
##

```

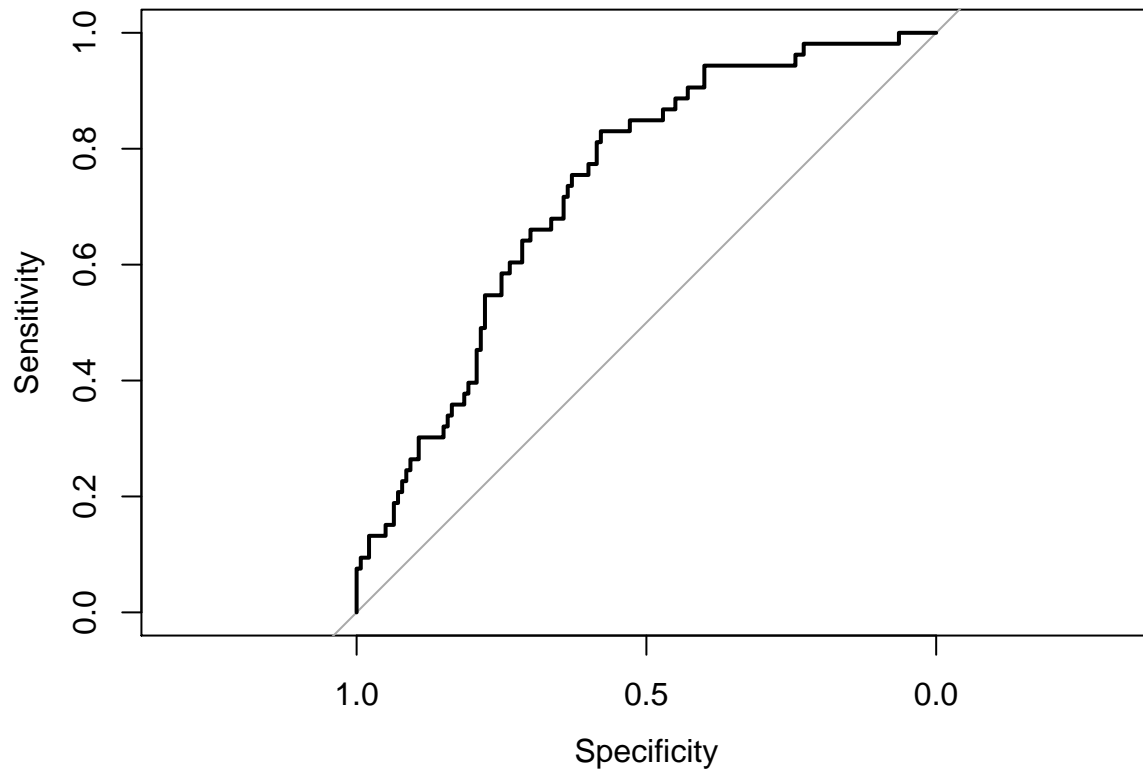


```
## [1] "AUC: 0.780552520680506"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 149 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7806
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9673  -0.7057  -0.3911   0.5850   2.7585
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.271e+00  5.058e-01  -4.489 7.16e-06 ***
## INCOME         -1.133e-05  2.954e-06  -3.836 0.000125 ***
## TRAVTIME        2.071e-02  7.137e-03   2.902 0.003710 **
## BLUEBOOK        5.591e-05  1.602e-05   3.489 0.000485 ***
## TIF            -4.197e-02  2.710e-02  -1.549 0.121434
## OLDCLAIM       3.272e-06  1.554e-05   0.211 0.833185
## PARENT1_Yes    1.423e+00  3.057e-01   4.654 3.25e-06 ***
## SEX_z_F       -9.441e-01  3.518e-01  -2.684 0.007279 **
## JOB_Manager   -6.777e-01  3.691e-01  -1.836 0.066327 .
```

```

## CAR_USE_Commercial          3.395e-01  2.435e-01  1.394 0.163327
## CAR_TYPE_Pickup             1.050e+00  3.303e-01  3.180 0.001472 **
## CAR_TYPE_Sports.Car         1.768e+00  4.838e-01  3.655 0.000257 ***
## CAR_TYPE_z_SUV              1.722e+00  4.295e-01  4.008 6.12e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.610e+00  4.281e-01 -6.097 1.08e-09 ***
## HOME_VAL_NA                 -4.483e-01  2.287e-01 -1.960 0.049979 *
## oldclaim_log                8.246e-02  3.239e-02  2.546 0.010897 *
## inter                       1.348e-02  4.241e-03  3.180 0.001475 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 716.76  on 623  degrees of freedom
## Residual deviance: 544.01  on 607  degrees of freedom
## AIC: 578.01
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 123  37
##           1  17  16
##
##           Accuracy : 0.7202
##           95% CI : (0.6512, 0.7823)
##      No Information Rate : 0.7254
##      P-Value [Acc > NIR] : 0.599955
##
##           Kappa : 0.2044
##
## Mcnemar's Test P-Value : 0.009722
##
##           Sensitivity : 0.8786
##           Specificity : 0.3019
##      Pos Pred Value : 0.7688
##      Neg Pred Value : 0.4848
##           Prevalence : 0.7254
##      Detection Rate : 0.6373
##      Detection Prevalence : 0.8290
##      Balanced Accuracy : 0.5902
##
##           'Positive' Class : 0
##

```



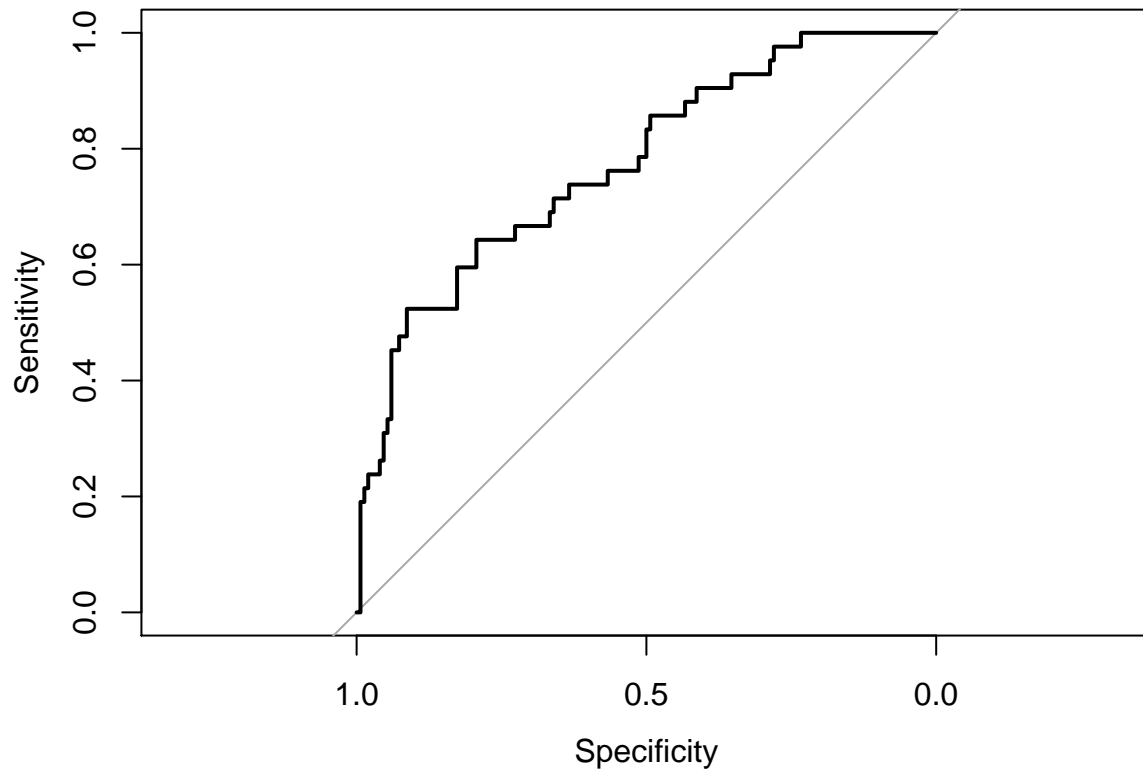
```
## [1] "AUC: 0.732210242587601"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 140 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.7322
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8945  -0.7502  -0.4196   0.7769   2.7885
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.705e+00  4.819e-01  -3.539 0.000402 ***
## INCOME        -9.542e-06  2.861e-06  -3.335 0.000854 ***
## TRAVTIME       2.039e-02  6.951e-03   2.933 0.003358 **
## BLUEBOOK       2.473e-05  1.523e-05   1.624 0.104444
## TIF           -5.315e-02  2.654e-02  -2.003 0.045216 *
## OLDCLAIM       1.051e-05  1.410e-05   0.745 0.455982
## PARENT1_Yes     8.482e-01  3.035e-01   2.795 0.005193 **
## SEX_z_F       -1.066e+00  3.705e-01  -2.878 0.004002 **
## JOB_Manager    -6.321e-01  3.476e-01  -1.819 0.068953 .
```



```

## CAR_USE_Commercial          7.839e-01  2.420e-01  3.240 0.001197 **
## CAR_TYPE_Pickup             7.793e-01  3.266e-01  2.386 0.017039 *
## CAR_TYPE_Sports.Car         2.039e+00  4.956e-01  4.115 3.87e-05 ***
## CAR_TYPE_z_SUV              1.952e+00  4.474e-01  4.364 1.28e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.325e+00  3.919e-01 -5.932 3.00e-09 ***
## HOME_VAL_NA                 -3.797e-01  2.196e-01 -1.729 0.083875 .
## oldclaim_log                1.647e-02  3.148e-02  0.523 0.600948
## inter                       1.704e-02  4.171e-03  4.086 4.38e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 739.29 on 624 degrees of freedom
## Residual deviance: 585.34 on 608 degrees of freedom
## AIC: 619.34
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 139  22
##           1  11  20
##
##           Accuracy : 0.8281
##           95% CI : (0.7672, 0.8786)
##           No Information Rate : 0.7812
##           P-Value [Acc > NIR] : 0.06598
##
##           Kappa : 0.4448
##
## Mcnemar's Test P-Value : 0.08172
##
##           Sensitivity : 0.9267
##           Specificity : 0.4762
##           Pos Pred Value : 0.8634
##           Neg Pred Value : 0.6452
##           Prevalence : 0.7812
##           Detection Rate : 0.7240
##           Detection Prevalence : 0.8385
##           Balanced Accuracy : 0.7014
##
##           'Positive' Class : 0
##

```

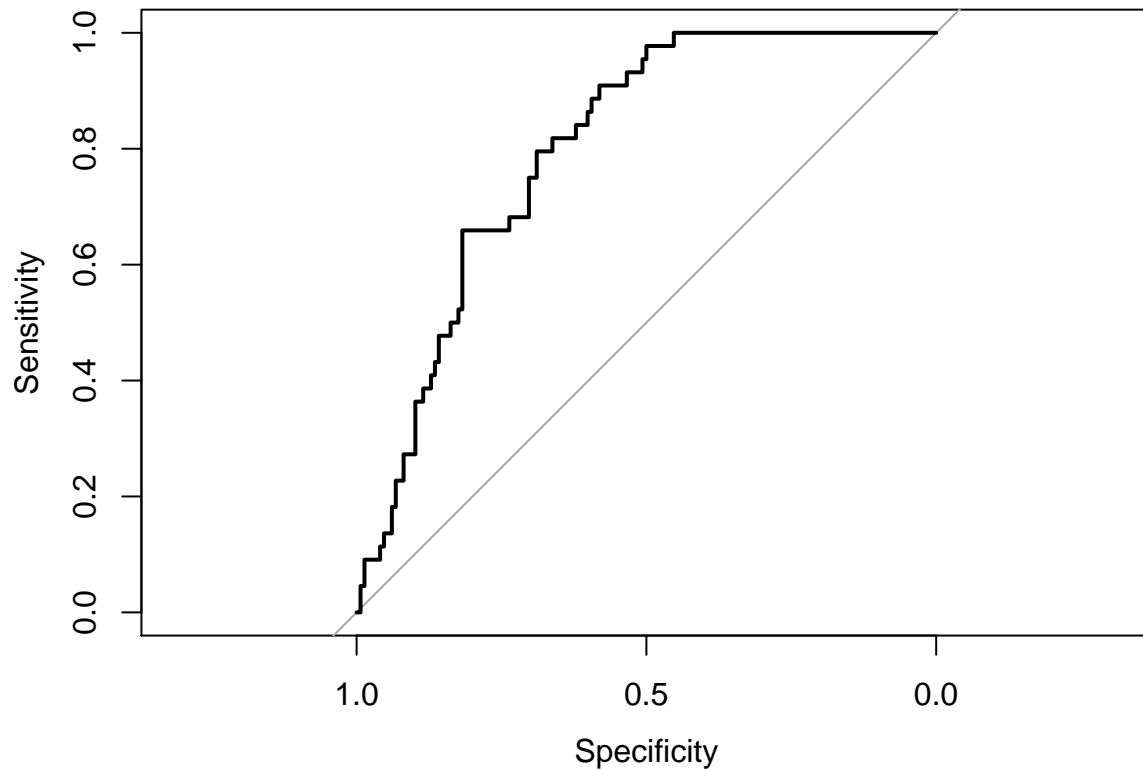


```
## [1] "AUC: 0.773968253968254"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 150 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.774
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8918  -0.7468  -0.4326   0.6752   2.7360
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.361e+00  4.871e-01  -4.846 1.26e-06 ***
## INCOME        -1.103e-05  2.952e-06  -3.737 0.000186 ***
## TRAVTIME       2.513e-02  7.020e-03   3.580 0.000343 ***
## BLUEBOOK       4.851e-05  1.554e-05   3.121 0.001803 **
## TIF           -3.618e-02  2.628e-02  -1.377 0.168580
## OLDCLAIM       1.307e-05  1.335e-05   0.979 0.327565
## PARENT1_Yes    1.310e+00  3.073e-01   4.262 2.02e-05 ***
## SEX_z_F       -9.074e-01  3.453e-01  -2.628 0.008597 **
## JOB_Manager    -7.327e-01  3.500e-01  -2.094 0.036301 *
```

```

## CAR_USE_Commercial          2.416e-01  2.373e-01  1.018 0.308527
## CAR_TYPE_Pickup             9.354e-01  3.168e-01  2.952 0.003157 **
## CAR_TYPE_Sports.Car        1.817e+00  4.764e-01  3.814 0.000137 ***
## CAR_TYPE_z_SUV             1.478e+00  4.176e-01  3.539 0.000402 ***
## URBANICITY_z_Highly.Rural..Rural -2.309e+00  3.876e-01 -5.958 2.55e-09 ***
## HOME_VAL_NA                -1.260e-01  2.216e-01 -0.568 0.569712
## oldclaim_log               7.018e-02  3.133e-02  2.240 0.025079 *
## inter                      1.228e-02  4.146e-03  2.962 0.003052 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 735.45 on 624 degrees of freedom
## Residual deviance: 578.73 on 608 degrees of freedom
## AIC: 612.73
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 127  25
##           1  21  19
##
##           Accuracy : 0.7604
##           95% CI : (0.6937, 0.8189)
##           No Information Rate : 0.7708
##           P-Value [Acc > NIR] : 0.6707
##
##           Kappa : 0.2995
##
## Mcnemar's Test P-Value : 0.6583
##
##           Sensitivity : 0.8581
##           Specificity : 0.4318
##           Pos Pred Value : 0.8355
##           Neg Pred Value : 0.4750
##           Prevalence : 0.7708
##           Detection Rate : 0.6615
##           Detection Prevalence : 0.7917
##           Balanced Accuracy : 0.6450
##
##           'Positive' Class : 0
##

```

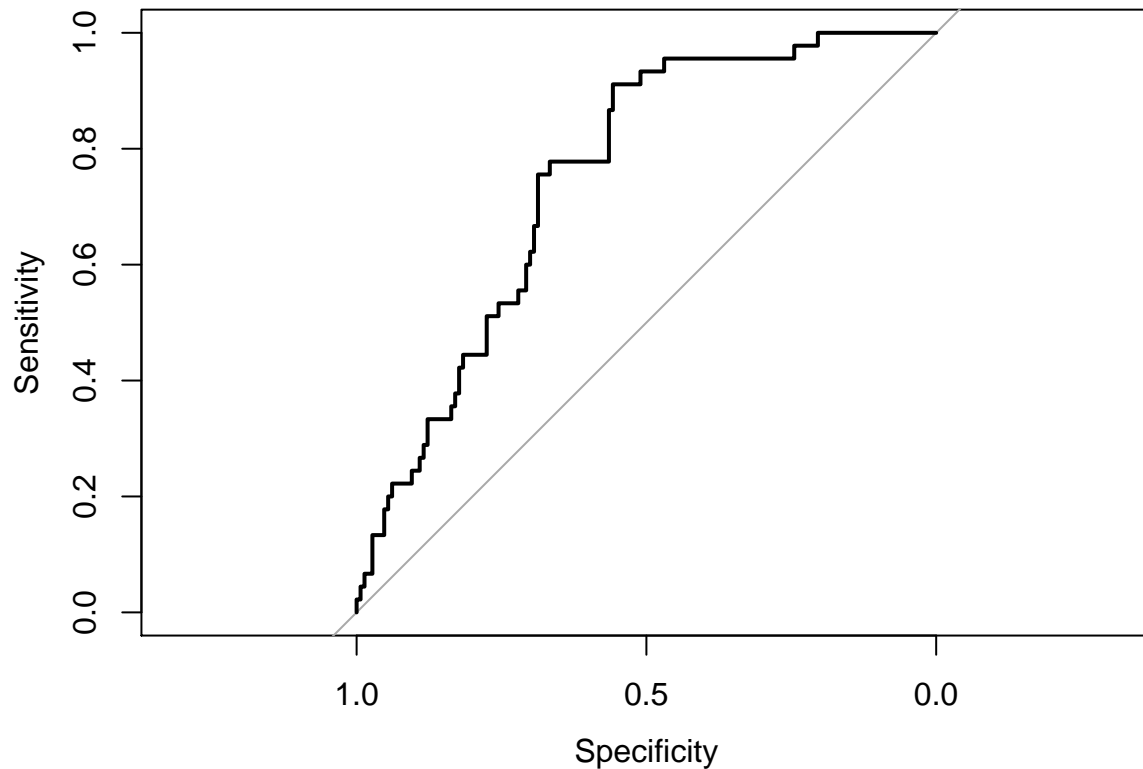


```
## [1] "AUC: 0.800675675675676"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 148 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.8007
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0642  -0.7188  -0.3919   0.6632   2.8775
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.088e+00  4.941e-01  -4.226 2.38e-05 ***
## INCOME        -1.112e-05  3.035e-06  -3.663 0.000249 ***
## TRAVTIME       2.163e-02  6.912e-03   3.130 0.001751 **
## BLUEBOOK       4.195e-05  1.563e-05   2.684 0.007275 **
## TIF           -3.238e-02  2.760e-02  -1.174 0.240590
## OLDCLAIM       6.692e-06  1.427e-05   0.469 0.639094
## PARENT1_Yes    1.022e+00  3.004e-01   3.402 0.000668 ***
## SEX_z_F       -9.624e-01  3.518e-01  -2.735 0.006232 **
## JOB_Manager   -8.617e-01  3.732e-01  -2.309 0.020928 *
```

```

## CAR_USE_Commercial          3.637e-01  2.474e-01  1.470 0.141600
## CAR_TYPE_Pickup             9.471e-01  3.232e-01  2.930 0.003386 **
## CAR_TYPE_Sports.Car        1.813e+00  4.795e-01  3.782 0.000156 ***
## CAR_TYPE_z_SUV             1.608e+00  4.256e-01  3.778 0.000158 ***
## URBANICITY_z_Highly.Rural..Rural -2.456e+00  4.144e-01 -5.928 3.07e-09 ***
## HOME_VAL_NA                -4.107e-01  2.305e-01 -1.782 0.074792 .
## oldclaim_log               8.127e-02  3.122e-02  2.603 0.009232 **
## inter                      1.786e-02  4.290e-03  4.162 3.15e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.51 on 624 degrees of freedom
## Residual deviance: 560.87 on 608 degrees of freedom
## AIC: 594.87
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 129  31
##           1  18  14
##
##           Accuracy : 0.7448
##           95% CI : (0.677, 0.8048)
##           No Information Rate : 0.7656
##           P-Value [Acc > NIR] : 0.78019
##
##           Kappa : 0.2097
##
## Mcnemar's Test P-Value : 0.08648
##
##           Sensitivity : 0.8776
##           Specificity : 0.3111
##           Pos Pred Value : 0.8062
##           Neg Pred Value : 0.4375
##           Prevalence : 0.7656
##           Detection Rate : 0.6719
##           Detection Prevalence : 0.8333
##           Balanced Accuracy : 0.5943
##
##           'Positive' Class : 0
##

```

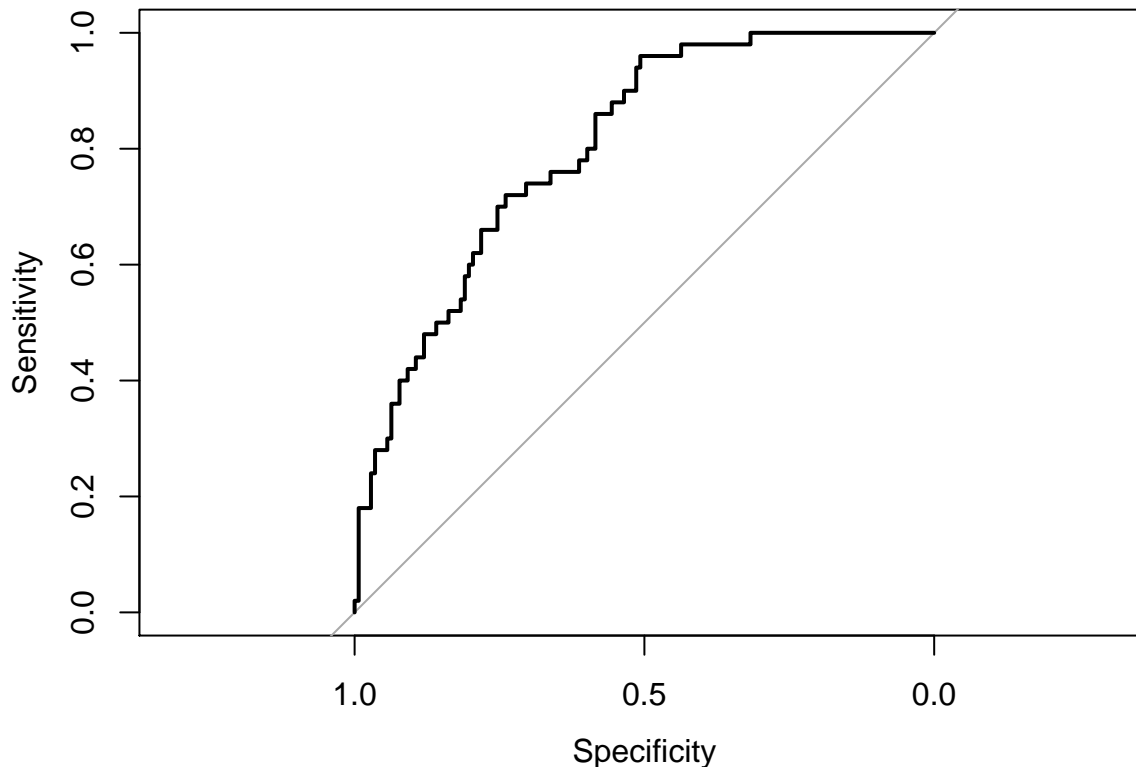


```
## [1] "AUC: 0.750566893424036"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 147 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7506
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7794  -0.7377  -0.4323   0.6873   2.6509
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.159e+00  5.026e-01  -4.295 1.74e-05 ***
## INCOME          -1.054e-05  2.937e-06  -3.589 0.000332 ***
## TRAVTIME         2.261e-02  6.917e-03   3.269 0.001081 **
## BLUEBOOK         3.918e-05  1.618e-05   2.423 0.015410 *
## TIF             -6.307e-02  2.786e-02  -2.264 0.023590 *
## OLDCLAIM         1.009e-05  1.350e-05   0.747 0.455115
## PARENT1_Yes      1.080e+00  3.072e-01   3.515 0.000439 ***
## SEX_z_F         -1.121e+00  3.718e-01  -3.014 0.002577 **
## JOB_Manager     -5.353e-01  3.565e-01  -1.502 0.133202
```

```

## CAR_USE_Commercial          3.327e-01  2.430e-01  1.369 0.170908
## CAR_TYPE_Pickup             1.134e+00  3.314e-01  3.423 0.000619 ***
## CAR_TYPE_Sports.Car         2.138e+00  4.944e-01  4.324 1.53e-05 ***
## CAR_TYPE_z_SUV              1.928e+00  4.448e-01  4.334 1.47e-05 ***
## URBANICITY_z_Highly.Rural..Rural -1.917e+00  3.700e-01 -5.180 2.22e-07 ***
## HOME_VAL_NA                 -2.264e-01  2.170e-01 -1.044 0.296714
## oldclaim_log                7.390e-02  3.135e-02  2.357 0.018408 *
## inter                       7.477e-03  4.278e-03  1.748 0.080500 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 723.54 on 624 degrees of freedom
## Residual deviance: 578.77 on 608 degrees of freedom
## AIC: 612.77
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 131  32
##           1  11  18
##
##           Accuracy : 0.776
##           95% CI : (0.7104, 0.8329)
##           No Information Rate : 0.7396
##           P-Value [Acc > NIR] : 0.142046
##
##           Kappa : 0.327
##
## Mcnemar's Test P-Value : 0.002289
##
##           Sensitivity : 0.9225
##           Specificity : 0.3600
##           Pos Pred Value : 0.8037
##           Neg Pred Value : 0.6207
##           Prevalence : 0.7396
##           Detection Rate : 0.6823
##           Detection Prevalence : 0.8490
##           Balanced Accuracy : 0.6413
##
##           'Positive' Class : 0
##

```



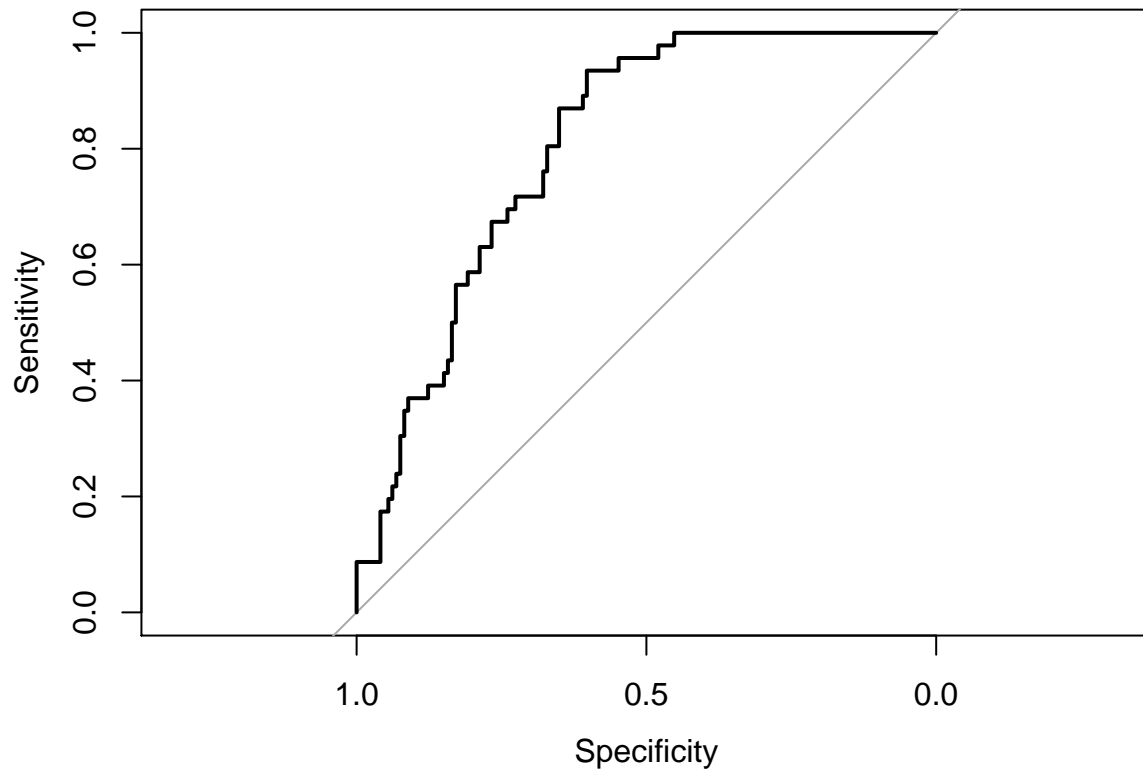
```
## [1] "AUC: 0.804084507042254"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 142 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.8041
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8921  -0.7468  -0.4307   0.7034   2.7429
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.186e+00  4.834e-01  -4.522 6.13e-06 ***
## INCOME        -1.001e-05  2.868e-06  -3.490 0.000483 ***
## TRAVTIME       2.362e-02  6.858e-03   3.444 0.000573 ***
## BLUEBOOK       4.218e-05  1.531e-05   2.755 0.005862 **
## TIF           -4.289e-02  2.645e-02  -1.621 0.104996
## OLDCLAIM       7.769e-06  1.461e-05   0.532 0.594931
## PARENT1_Yes    1.354e+00  3.085e-01   4.387 1.15e-05 ***
## SEX_z_F       -8.811e-01  3.463e-01  -2.544 0.010947 *
## JOB_Manager   -8.791e-01  3.780e-01  -2.326 0.020031 *
```



```

## CAR_USE_Commercial          3.912e-01  2.338e-01  1.673 0.094264 .
## CAR_TYPE_Pickup             9.275e-01  3.135e-01  2.959 0.003088 **
## CAR_TYPE_Sports.Car        1.870e+00  4.805e-01  3.893 9.91e-05 ***
## CAR_TYPE_z_SUV             1.558e+00  4.226e-01  3.688 0.000226 ***
## URBANICITY_z_Highly.Rural..Rural -2.310e+00  3.837e-01 -6.018 1.76e-09 ***
## HOME_VAL_NA                -2.046e-01  2.230e-01 -0.917 0.358945
## oldclaim_log               5.214e-02  3.186e-02  1.636 0.101758
## inter                      1.099e-02  4.030e-03  2.727 0.006395 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 731.55  on 624  degrees of freedom
## Residual deviance: 577.82  on 608  degrees of freedom
## AIC: 611.82
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 132  29
##           1  14  17
##
##           Accuracy : 0.776
##           95% CI : (0.7104, 0.8329)
##      No Information Rate : 0.7604
##      P-Value [Acc > NIR] : 0.34076
##
##           Kappa : 0.3081
##
## Mcnemar's Test P-Value : 0.03276
##
##           Sensitivity : 0.9041
##           Specificity : 0.3696
##      Pos Pred Value : 0.8199
##      Neg Pred Value : 0.5484
##           Prevalence : 0.7604
##      Detection Rate : 0.6875
##      Detection Prevalence : 0.8385
##      Balanced Accuracy : 0.6368
##
##           'Positive' Class : 0
##

```

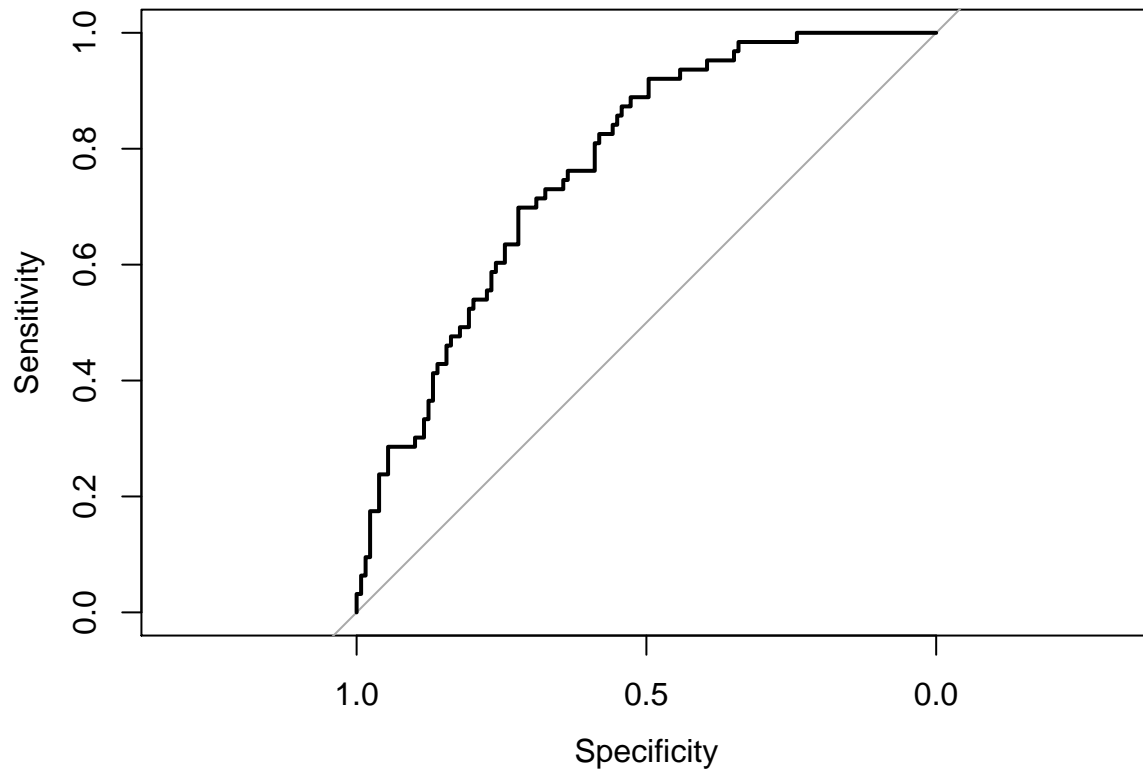


```
## [1] "AUC: 0.805687909469923"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.8057
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8787  -0.7045  -0.4089  -0.1203   2.8295
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.452e+00  5.250e-01  -4.670 3.01e-06 ***
## INCOME        -9.447e-06  2.983e-06  -3.166 0.001544 **
## TRAVTIME       2.236e-02  7.139e-03   3.132 0.001736 **
## BLUEBOOK       4.481e-05  1.616e-05   2.773 0.005553 **
## TIF          -5.876e-02  2.826e-02  -2.079 0.037584 *
## OLDCLAIM       1.387e-05  1.443e-05   0.961 0.336551
## PARENT1_Yes    1.320e+00  3.127e-01   4.222 2.42e-05 ***
## SEX_z_F       -8.572e-01  3.640e-01  -2.355 0.018528 *
## JOB_Manager   -4.090e-01  3.558e-01  -1.150 0.250303
```

```

## CAR_USE_Commercial          4.619e-01  2.457e-01  1.880 0.060128 .
## CAR_TYPE_Pickup             1.263e+00  3.353e-01  3.766 0.000166 ***
## CAR_TYPE_Sports.Car         1.981e+00  4.949e-01  4.003 6.26e-05 ***
## CAR_TYPE_z_SUV              1.665e+00  4.447e-01  3.744 0.000181 ***
## URBANICITY_z_Highly.Rural..Rural -2.179e+00  4.017e-01 -5.424 5.82e-08 ***
## HOME_VAL_NA                 -3.388e-01  2.254e-01 -1.503 0.132850
## oldclaim_log                6.183e-02  3.264e-02  1.894 0.058196 .
## inter                       7.986e-03  4.409e-03  1.811 0.070124 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 695.69  on 624  degrees of freedom
## Residual deviance: 548.63  on 608  degrees of freedom
## AIC: 582.63
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  44
##           1  13  19
##
##           Accuracy : 0.7031
##           95% CI : (0.6331, 0.7668)
##           No Information Rate : 0.6719
##           P-Value [Acc > NIR] : 0.1997
##
##           Kappa : 0.2297
##
## Mcnemar's Test P-Value : 7.08e-05
##
##           Sensitivity : 0.8992
##           Specificity : 0.3016
##           Pos Pred Value : 0.7250
##           Neg Pred Value : 0.5938
##           Prevalence : 0.6719
##           Detection Rate : 0.6042
##           Detection Prevalence : 0.8333
##           Balanced Accuracy : 0.6004
##
##           'Positive' Class : 0
##

```

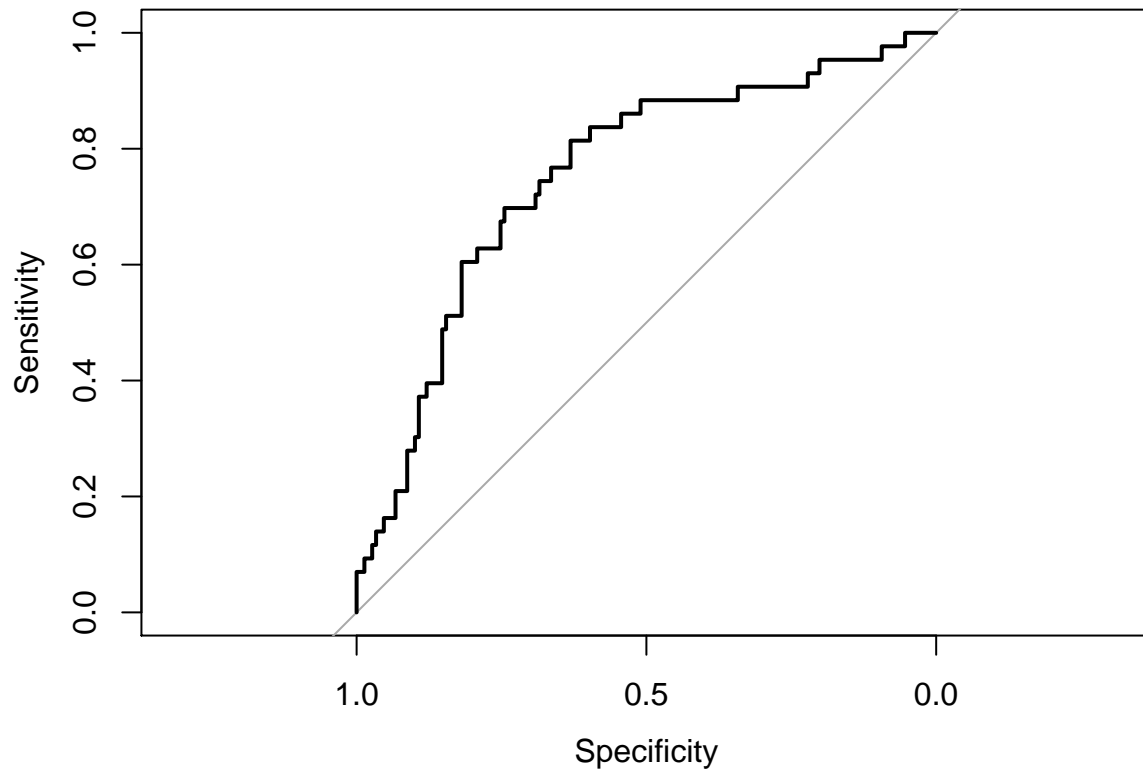


```
## [1] "AUC: 0.770394979697305"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 129 controls (dfPred_raw$class 0) < 63 cases (dfPred_raw$class 1).
## Area under the curve: 0.7704
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9573  -0.7398  -0.4032   0.6637   2.9741
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.651e+00  4.704e-01  -3.510 0.000448 ***
## INCOME        -1.166e-05  3.093e-06  -3.770 0.000163 ***
## TRAVTIME       1.907e-02  6.699e-03   2.847 0.004414 **
## BLUEBOOK       9.967e-06  1.533e-05   0.650 0.515524
## TIF           2.575e-03  2.588e-02   0.100 0.920726
## OLDCLAIM      -3.095e-06  1.487e-05  -0.208 0.835074
## PARENT1_Yes    1.144e+00  3.055e-01   3.745 0.000180 ***
## SEX_z_F       -6.877e-01  3.346e-01  -2.055 0.039833 *
## JOB_Manager   -5.607e-01  3.458e-01  -1.622 0.104879
```

```

## CAR_USE_Commercial          4.624e-01  2.378e-01  1.945 0.051819 .
## CAR_TYPE_Pickup             7.523e-01  3.092e-01  2.433 0.014981 *
## CAR_TYPE_Sports.Car        1.458e+00  4.518e-01  3.228 0.001248 **
## CAR_TYPE_z_SUV             1.284e+00  4.008e-01  3.204 0.001357 **
## URBANICITY_z_Highly.Rural..Rural -2.627e+00  4.214e-01 -6.234 4.55e-10 ***
## HOME_VAL_NA                -2.829e-01  2.255e-01 -1.255 0.209536
## oldclaim_log               7.830e-02  3.116e-02  2.513 0.011963 *
## inter                      1.558e-02  4.164e-03  3.742 0.000182 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 737.38 on 624 degrees of freedom
## Residual deviance: 570.87 on 608 degrees of freedom
## AIC: 604.87
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 134  31
##           1  15  12
##
##           Accuracy : 0.7604
##           95% CI : (0.6937, 0.8189)
##           No Information Rate : 0.776
##           P-Value [Acc > NIR] : 0.73100
##
##           Kappa : 0.2056
##
## Mcnemar's Test P-Value : 0.02699
##
##           Sensitivity : 0.8993
##           Specificity : 0.2791
##           Pos Pred Value : 0.8121
##           Neg Pred Value : 0.4444
##           Prevalence : 0.7760
##           Detection Rate : 0.6979
##           Detection Prevalence : 0.8594
##           Balanced Accuracy : 0.5892
##
##           'Positive' Class : 0
##

```

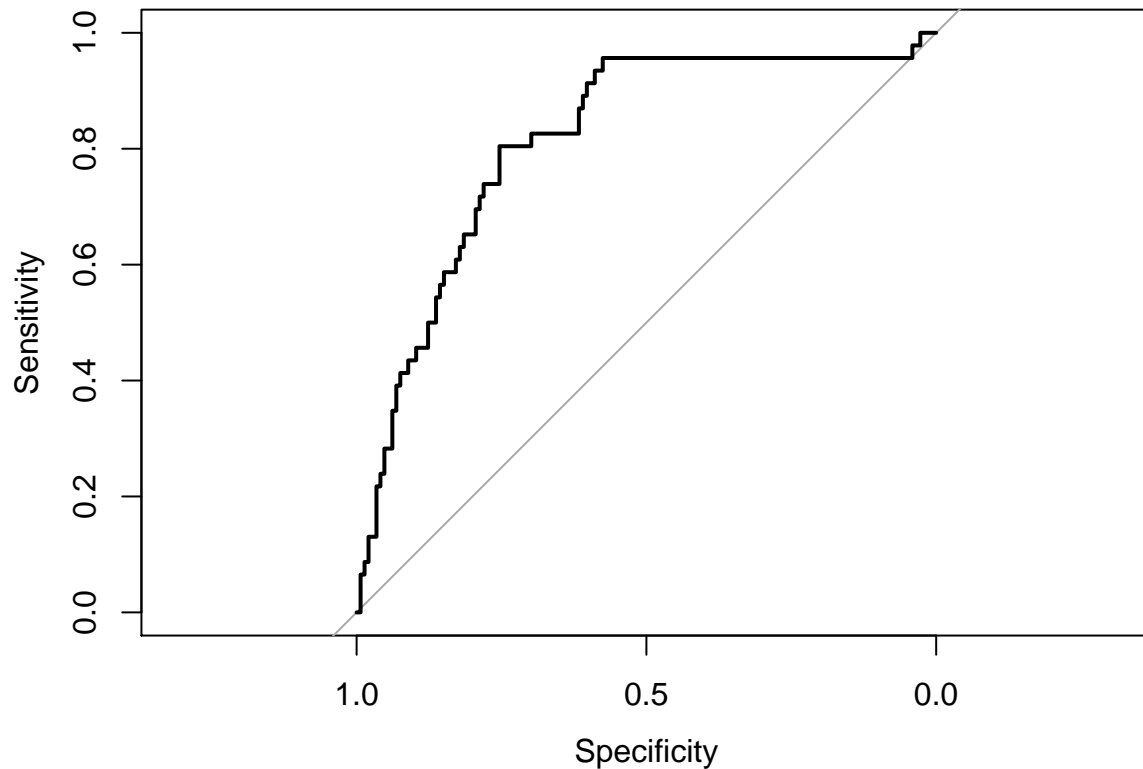


```
## [1] "AUC: 0.752926486655221"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 149 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7529
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8456  -0.7475  -0.4309   0.6919   2.8070
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.611e+00  4.755e-01  -3.389 0.000701 ***
## INCOME        -1.050e-05  2.928e-06  -3.584 0.000338 ***
## TRAVTIME       2.195e-02  6.662e-03   3.295 0.000985 ***
## BLUEBOOK       5.786e-06  1.531e-05   0.378 0.705425
## TIF          -3.175e-02  2.583e-02  -1.229 0.219101
## OLDCLAIM       6.700e-06  1.369e-05   0.490 0.624454
## PARENT1_Yes    9.944e-01  3.141e-01   3.166 0.001548 **
## SEX_z_F       -7.357e-01  3.407e-01  -2.160 0.030807 *
## JOB_Manager   -5.626e-01  3.465e-01  -1.623 0.104483
```

```

## CAR_USE_Commercial          4.316e-01  2.330e-01  1.853 0.063922 .
## CAR_TYPE_Pickup             8.260e-01  3.104e-01  2.661 0.007783 **
## CAR_TYPE_Sports.Car        1.681e+00  4.586e-01  3.665 0.000247 ***
## CAR_TYPE_z_SUV             1.431e+00  4.065e-01  3.519 0.000433 ***
## URBANICITY_z_Highly.Rural..Rural -2.324e+00  3.969e-01 -5.855 4.78e-09 ***
## HOME_VAL_NA                -1.947e-01  2.206e-01 -0.883 0.377394
## oldclaim_log               5.985e-02  3.139e-02  1.906 0.056604 .
## inter                      1.003e-02  4.221e-03  2.375 0.017535 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 731.55 on 624 degrees of freedom
## Residual deviance: 583.48 on 608 degrees of freedom
## AIC: 617.48
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 136  28
##           1  10  18
##
##           Accuracy : 0.8021
##           95% CI : (0.7386, 0.856)
##           No Information Rate : 0.7604
##           P-Value [Acc > NIR] : 0.10058
##
##           Kappa : 0.3728
##
## Mcnemar's Test P-Value : 0.00582
##
##           Sensitivity : 0.9315
##           Specificity : 0.3913
##           Pos Pred Value : 0.8293
##           Neg Pred Value : 0.6429
##           Prevalence : 0.7604
##           Detection Rate : 0.7083
##           Detection Prevalence : 0.8542
##           Balanced Accuracy : 0.6614
##
##           'Positive' Class : 0
##

```



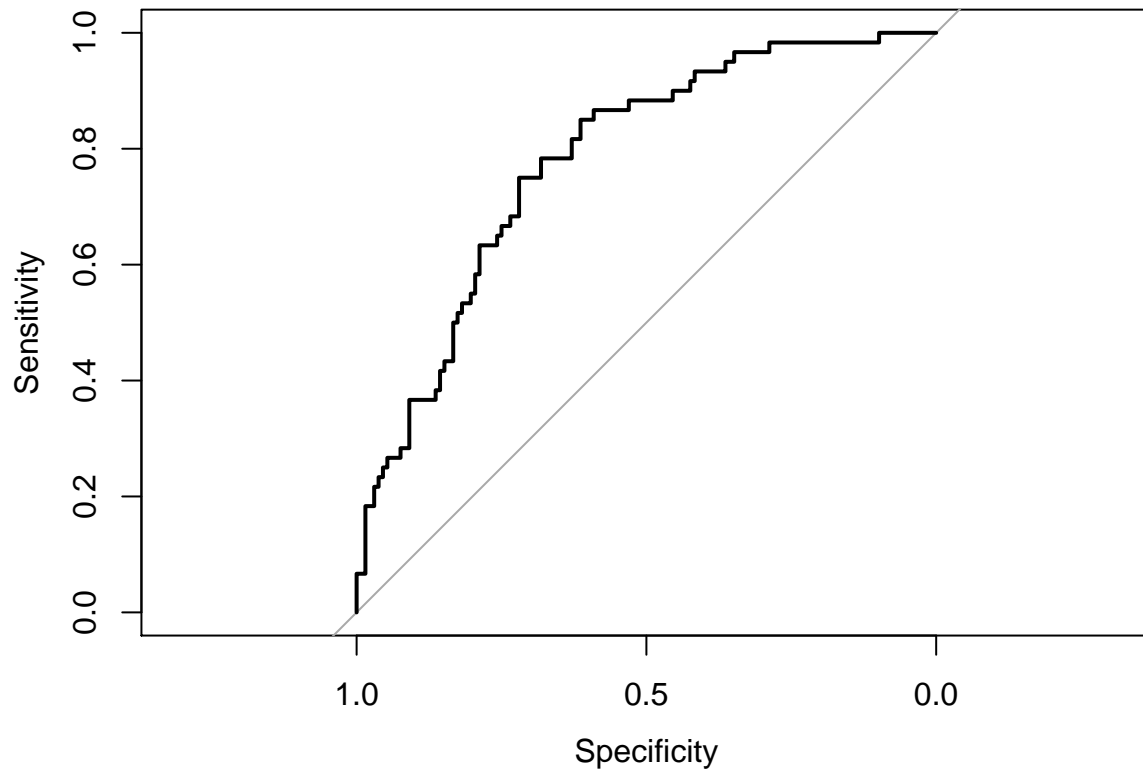
```
## [1] "AUC: 0.815366289458011"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.8154
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.00795  -0.71918  -0.42419  -0.07833   2.87101
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.006e+00  5.078e-01  -3.951 7.78e-05 ***
## INCOME        -8.365e-06  2.974e-06  -2.813 0.004911 **
## TRAVTIME       2.010e-02  7.470e-03   2.690 0.007140 **
## BLUEBOOK       3.435e-05  1.554e-05   2.210 0.027081 *
## TIF           -5.491e-02  2.686e-02  -2.045 0.040883 *
## OLDCLAIM       1.789e-05  1.459e-05   1.226 0.220329
## PARENT1_Yes    1.065e+00  3.084e-01   3.455 0.000551 ***
## SEX_z_F       -7.716e-01  3.515e-01  -2.195 0.028147 *
## JOB_Manager   -5.197e-01  3.624e-01  -1.434 0.151550
```



```

## CAR_USE_Commercial          6.732e-01  2.479e-01  2.715 0.006628 **
## CAR_TYPE_Pickup             9.327e-01  3.252e-01  2.868 0.004132 **
## CAR_TYPE_Sports.Car         1.606e+00  4.860e-01  3.304 0.000953 ***
## CAR_TYPE_z_SUV              1.604e+00  4.333e-01  3.702 0.000214 ***
## URBANICITY_z_Highly.Rural..Rural -2.305e+00  4.156e-01 -5.547 2.91e-08 ***
## HOME_VAL_NA                 -5.741e-01  2.281e-01 -2.517 0.011849 *
## oldclaim_log                2.068e-02  3.238e-02  0.639 0.523079
## inter                       1.697e-02  4.523e-03  3.753 0.000175 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 702.37  on 624  degrees of freedom
## Residual deviance: 552.79  on 608  degrees of freedom
## AIC: 586.79
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 121  43
##           1  11  17
##
##           Accuracy : 0.7188
##           95% CI : (0.6495, 0.7811)
##      No Information Rate : 0.6875
##      P-Value [Acc > NIR] : 0.1966
##
##           Kappa : 0.234
##
## Mcnemar's Test P-Value : 2.459e-05
##
##           Sensitivity : 0.9167
##           Specificity : 0.2833
##      Pos Pred Value : 0.7378
##      Neg Pred Value : 0.6071
##           Prevalence : 0.6875
##      Detection Rate : 0.6302
##      Detection Prevalence : 0.8542
##      Balanced Accuracy : 0.6000
##
##           'Positive' Class : 0
##

```

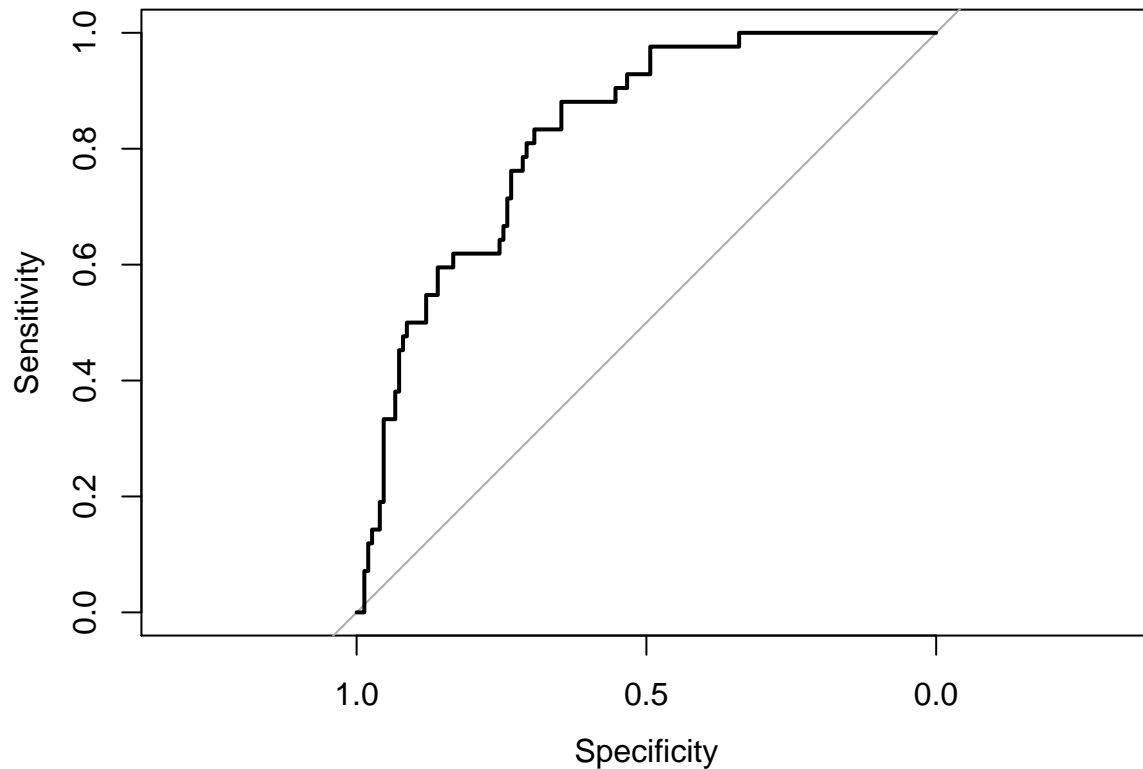


```
## [1] "AUC: 0.780176767676768"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 132 controls (dfPred_raw$class 0) < 60 cases (dfPred_raw$class 1).
## Area under the curve: 0.7802
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7423  -0.7591  -0.4299   0.7578   2.7496
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.837e+00  4.819e-01  -3.812 0.000138 ***
## INCOME        -9.623e-06  2.823e-06  -3.408 0.000654 ***
## TRAVTIME       2.454e-02  6.818e-03   3.600 0.000318 ***
## BLUEBOOK       2.021e-05  1.507e-05   1.340 0.180110
## TIF           -5.507e-02  2.649e-02  -2.079 0.037605 *
## OLDCLAIM       1.178e-05  1.334e-05   0.883 0.377009
## PARENT1_Yes     8.745e-01  3.186e-01   2.744 0.006061 **
## SEX_z_F        -9.313e-01  3.606e-01  -2.583 0.009806 **
## JOB_Manager    -6.524e-01  3.432e-01  -1.901 0.057345 .
```

```

## CAR_USE_Commercial          5.410e-01  2.334e-01  2.318 0.020462 *
## CAR_TYPE_Pickup             8.671e-01  3.187e-01  2.721 0.006514 **
## CAR_TYPE_Sports.Car        1.993e+00  4.836e-01  4.122 3.76e-05 ***
## CAR_TYPE_z_SUV             1.730e+00  4.318e-01  4.005 6.20e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.321e+00  3.852e-01 -6.025 1.69e-09 ***
## HOME_VAL_NA                -1.962e-01  2.206e-01 -0.889 0.373772
## oldclaim_log               4.348e-02  3.157e-02  1.377 0.168466
## inter                      1.288e-02  4.095e-03  3.145 0.001661 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 739.29 on 624 degrees of freedom
## Residual deviance: 589.00 on 608 degrees of freedom
## AIC: 623
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 132  20
##           1  18  22
##
##           Accuracy : 0.8021
##           95% CI : (0.7386, 0.856)
##           No Information Rate : 0.7812
##           P-Value [Acc > NIR] : 0.2741
##
##           Kappa : 0.4109
##
## Mcnemar's Test P-Value : 0.8711
##
##           Sensitivity : 0.8800
##           Specificity : 0.5238
##           Pos Pred Value : 0.8684
##           Neg Pred Value : 0.5500
##           Prevalence : 0.7812
##           Detection Rate : 0.6875
##           Detection Prevalence : 0.7917
##           Balanced Accuracy : 0.7019
##
##           'Positive' Class : 0
##

```

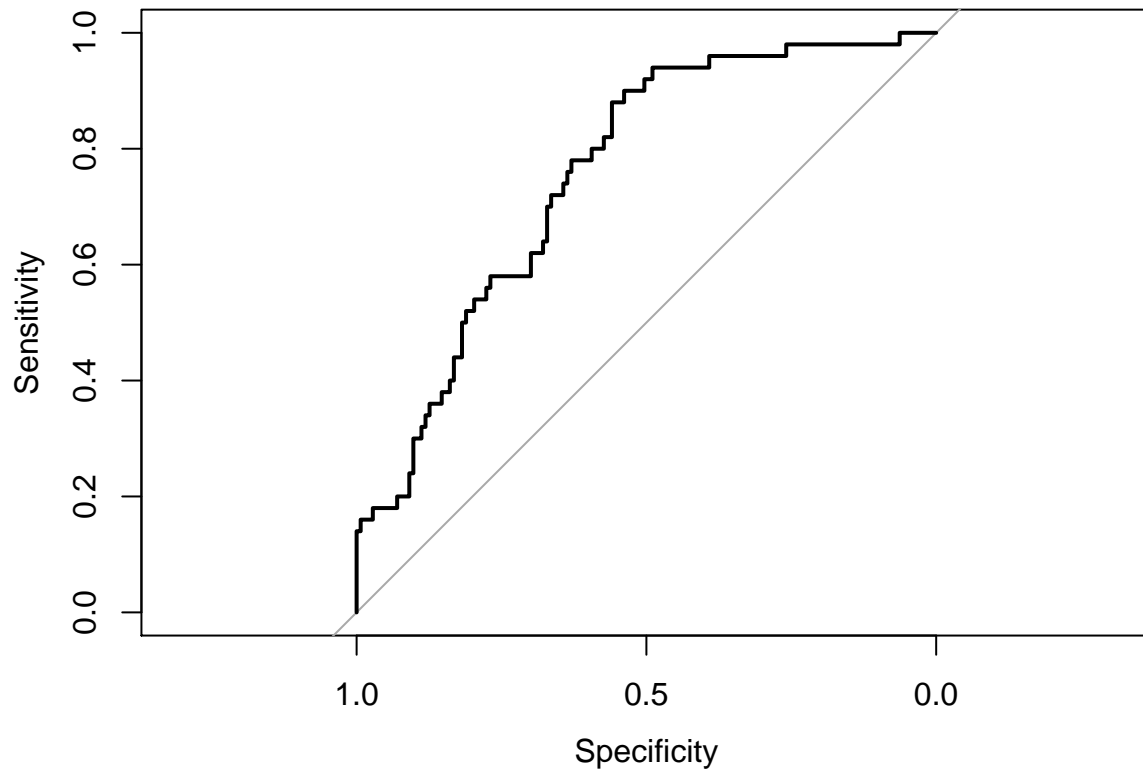


```
## [1] "AUC: 0.823650793650794"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 150 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.8237
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0051  -0.7002  -0.3930   0.5876   2.7058
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.089e+00  5.051e-01  -4.135 3.55e-05 ***
## INCOME        -1.126e-05  2.973e-06  -3.786 0.000153 ***
## TRAVTIME       1.998e-02  7.072e-03   2.826 0.004715 **
## BLUEBOOK       4.273e-05  1.565e-05   2.730 0.006331 **
## TIF           -4.882e-02  2.781e-02  -1.755 0.079199 .
## OLDCLAIM       1.524e-05  1.480e-05   1.030 0.303142
## PARENT1_Yes    1.504e+00  3.223e-01   4.666 3.07e-06 ***
## SEX_z_F       -1.099e+00  3.631e-01  -3.026 0.002482 **
## JOB_Manager    -7.427e-01  3.614e-01  -2.055 0.039882 *
```

```

## CAR_USE_Commercial          4.368e-01  2.489e-01  1.755 0.079329 .
## CAR_TYPE_Pickup             8.924e-01  3.266e-01  2.732 0.006288 **
## CAR_TYPE_Sports.Car         2.004e+00  4.899e-01  4.090 4.31e-05 ***
## CAR_TYPE_z_SUV              1.807e+00  4.386e-01  4.119 3.81e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.444e+00  4.067e-01 -6.011 1.85e-09 ***
## HOME_VAL_NA                 -2.868e-01  2.258e-01 -1.270 0.204034
## oldclaim_log                7.627e-02  3.167e-02  2.409 0.016015 *
## inter                       1.238e-02  4.027e-03  3.075 0.002103 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 722.92 on 623 degrees of freedom
## Residual deviance: 552.14 on 607 degrees of freedom
## AIC: 586.14
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 124  32
##           1  19  18
##
##           Accuracy : 0.7358
##           95% CI : (0.6676, 0.7965)
##           No Information Rate : 0.7409
##           P-Value [Acc > NIR] : 0.60208
##
##           Kappa : 0.2481
##
## Mcnemar's Test P-Value : 0.09289
##
##           Sensitivity : 0.8671
##           Specificity : 0.3600
##           Pos Pred Value : 0.7949
##           Neg Pred Value : 0.4865
##           Prevalence : 0.7409
##           Detection Rate : 0.6425
##           Detection Prevalence : 0.8083
##           Balanced Accuracy : 0.6136
##
##           'Positive' Class : 0
##

```

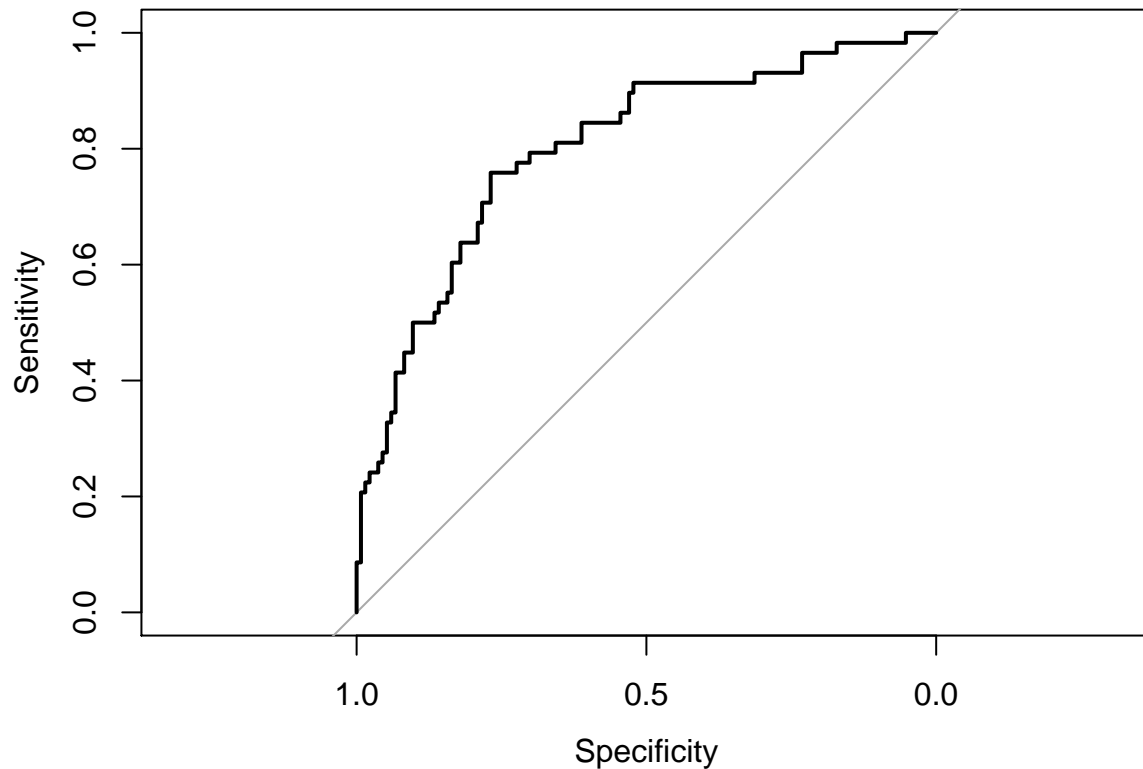


```
## [1] "AUC: 0.755664335664336"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 143 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7557
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8490  -0.7130  -0.4193   0.5575   2.9891
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.967e+00  4.906e-01  -4.010 6.06e-05 ***
## INCOME        -9.101e-06  2.993e-06  -3.040 0.00236 **
## TRAVTIME       2.220e-02  7.040e-03   3.153 0.00161 **
## BLUEBOOK       1.629e-05  1.576e-05   1.034 0.30124
## TIF           -2.020e-02  2.562e-02  -0.788 0.43047
## OLDCLAIM       1.376e-06  1.520e-05   0.091 0.92787
## PARENT1_Yes    1.167e+00  2.949e-01   3.958 7.56e-05 ***
## SEX_z_F       -5.584e-01  3.393e-01  -1.646 0.09981 .
## JOB_Manager   -3.702e-01  3.534e-01  -1.048 0.29482
```

```

## CAR_USE_Commercial          5.398e-01  2.372e-01  2.275  0.02289 *
## CAR_TYPE_Pickup             9.078e-01  3.169e-01  2.864  0.00418 **
## CAR_TYPE_Sports.Car         1.344e+00  4.650e-01  2.889  0.00386 **
## CAR_TYPE_z_SUV              1.308e+00  4.112e-01  3.180  0.00147 **
## URBANICITY_z_Highly.Rural..Rural -2.564e+00  4.432e-01 -5.786  7.23e-09 ***
## HOME_VAL_NA                 -4.333e-01  2.228e-01 -1.945  0.05180 .
## oldclaim_log                5.135e-02  3.203e-02  1.603  0.10896
## inter                       1.313e-02  4.359e-03  3.013  0.00259 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 706.74 on 624 degrees of freedom
## Residual deviance: 559.74 on 608 degrees of freedom
## AIC: 593.74
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 131  44
##           1   3  14
##
##           Accuracy : 0.7552
##           95% CI : (0.6881, 0.8143)
##           No Information Rate : 0.6979
##           P-Value [Acc > NIR] : 0.04739
##
##           Kappa : 0.2739
##
## Mcnemar's Test P-Value : 5.392e-09
##
##           Sensitivity : 0.9776
##           Specificity : 0.2414
##           Pos Pred Value : 0.7486
##           Neg Pred Value : 0.8235
##           Prevalence : 0.6979
##           Detection Rate : 0.6823
##           Detection Prevalence : 0.9115
##           Balanced Accuracy : 0.6095
##
##           'Positive' Class : 0
##

```



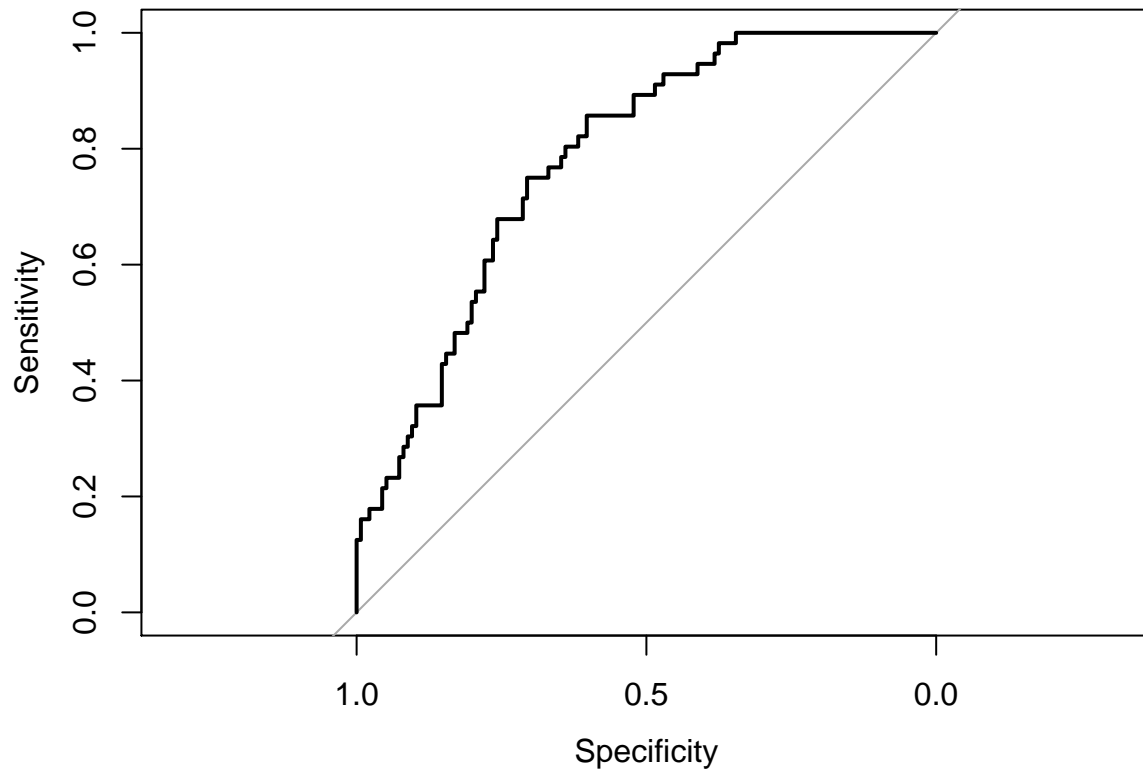
```
## [1] "AUC: 0.801595470921256"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 58 cases (dfPred_raw$class 1).
## Area under the curve: 0.8016
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8574  -0.7154  -0.4119   0.5964   2.7193
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.186e+00  5.038e-01  -4.338 1.44e-05 ***
## INCOME        -7.665e-06  2.908e-06  -2.635 0.008402 **
## TRAVTIME       2.116e-02  7.101e-03   2.980 0.002880 **
## BLUEBOOK       4.137e-05  1.578e-05   2.621 0.008764 **
## TIF           -6.668e-02  2.818e-02  -2.366 0.017963 *
## OLDCLAIM       1.023e-05  1.517e-05   0.674 0.500174
## PARENT1_Yes    1.306e+00  3.129e-01   4.173 3.01e-05 ***
## SEX_z_F       -1.182e+00  3.764e-01  -3.140 0.001688 **
## JOB_Manager   -9.011e-01  4.081e-01  -2.208 0.027233 *
```



```

## CAR_USE_Commercial          4.063e-01  2.504e-01  1.622 0.104726
## CAR_TYPE_Pickup             1.156e+00  3.298e-01  3.504 0.000459 ***
## CAR_TYPE_Sports.Car        2.026e+00  5.069e-01  3.996 6.43e-05 ***
## CAR_TYPE_z_SUV             1.948e+00  4.594e-01  4.240 2.24e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.038e+00  3.848e-01 -5.296 1.18e-07 ***
## HOME_VAL_NA                -4.275e-01  2.255e-01 -1.896 0.057967 .
## oldclaim_log               5.673e-02  3.199e-02  1.773 0.076209 .
## inter                      1.359e-02  4.391e-03  3.096 0.001963 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 711.04  on 624  degrees of freedom
## Residual deviance: 559.53  on 608  degrees of freedom
## AIC: 593.53
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 126  41
##           1  10  15
##
##           Accuracy : 0.7344
##           95% CI : (0.666, 0.7954)
##    No Information Rate : 0.7083
##    P-Value [Acc > NIR] : 0.2392
##
##           Kappa : 0.2321
##
## Mcnemar's Test P-Value : 2.659e-05
##
##           Sensitivity : 0.9265
##           Specificity : 0.2679
##    Pos Pred Value : 0.7545
##    Neg Pred Value : 0.6000
##           Prevalence : 0.7083
##    Detection Rate : 0.6562
##    Detection Prevalence : 0.8698
##    Balanced Accuracy : 0.5972
##
##           'Positive' Class : 0
##

```

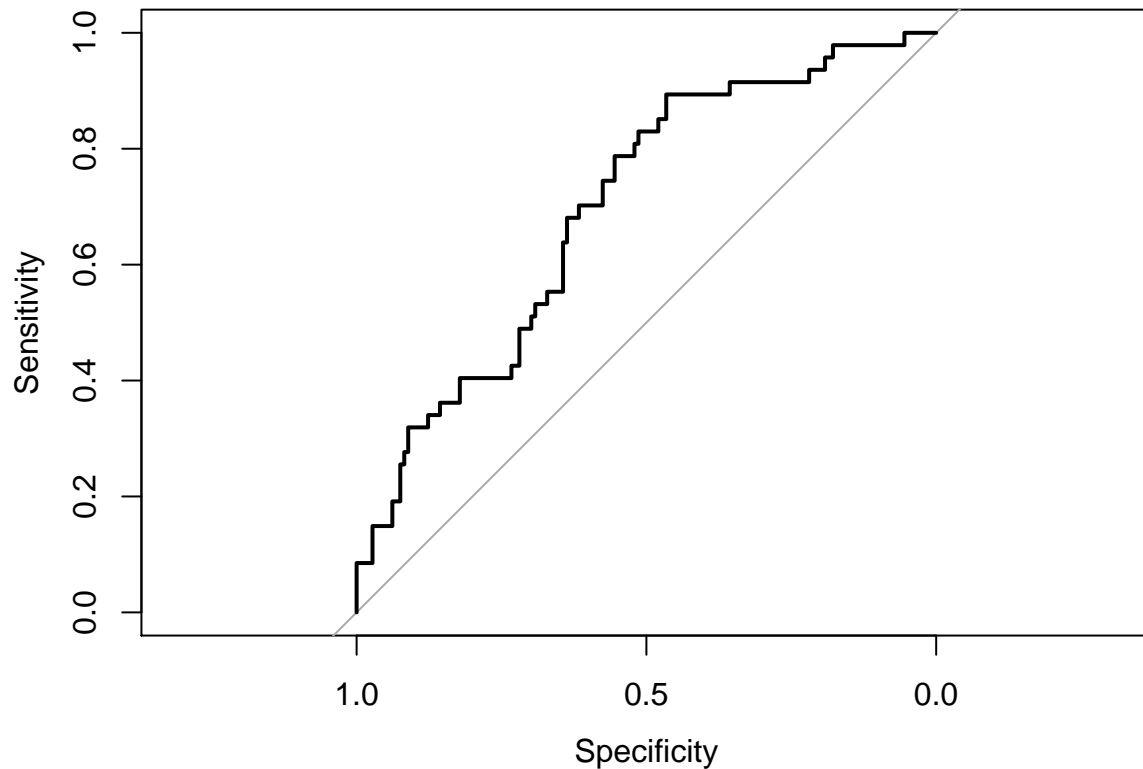


```
## [1] "AUC: 0.782956932773109"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 136 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.783
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9319  -0.7088  -0.3698   0.6552   2.9009
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.857e+00  4.898e-01  -3.791 0.000150 ***
## INCOME        -1.227e-05  3.042e-06  -4.033 5.50e-05 ***
## TRAVTIME       1.934e-02  6.993e-03   2.766 0.005683 **
## BLUEBOOK       3.484e-05  1.559e-05   2.234 0.025463 *
## TIF           -1.272e-02  2.623e-02  -0.485 0.627813
## OLDCLAIM       1.527e-06  1.490e-05   0.102 0.918366
## PARENT1_Yes    1.024e+00  2.981e-01   3.436 0.000591 ***
## SEX_z_F       -8.395e-01  3.411e-01  -2.461 0.013842 *
## JOB_Manager   -7.999e-01  3.701e-01  -2.161 0.030689 *
```

```

## CAR_USE_Commercial          4.193e-01  2.403e-01  1.745 0.080998 .
## CAR_TYPE_Pickup             7.544e-01  3.205e-01  2.354 0.018571 *
## CAR_TYPE_Sports.Car        1.677e+00  4.717e-01  3.556 0.000376 ***
## CAR_TYPE_z_SUV             1.450e+00  4.103e-01  3.534 0.000410 ***
## URBANICITY_z_Highly.Rural..Rural -2.893e+00  4.658e-01 -6.211 5.25e-10 ***
## HOME_VAL_NA                -3.933e-01  2.296e-01 -1.713 0.086692 .
## oldclaim_log               8.485e-02  3.164e-02  2.682 0.007326 **
## inter                      1.573e-02  4.064e-03  3.870 0.000109 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 728.94 on 623 degrees of freedom
## Residual deviance: 548.14 on 607 degrees of freedom
## AIC: 582.14
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 129  32
##           1  17  15
##
##           Accuracy : 0.7461
##           95% CI : (0.6786, 0.8059)
##           No Information Rate : 0.7565
##           P-Value [Acc > NIR] : 0.6667
##
##           Kappa : 0.2273
##
## Mcnemar's Test P-Value : 0.0455
##
##           Sensitivity : 0.8836
##           Specificity : 0.3191
##           Pos Pred Value : 0.8012
##           Neg Pred Value : 0.4687
##           Prevalence : 0.7565
##           Detection Rate : 0.6684
##           Detection Prevalence : 0.8342
##           Balanced Accuracy : 0.6014
##
##           'Positive' Class : 0
##

```

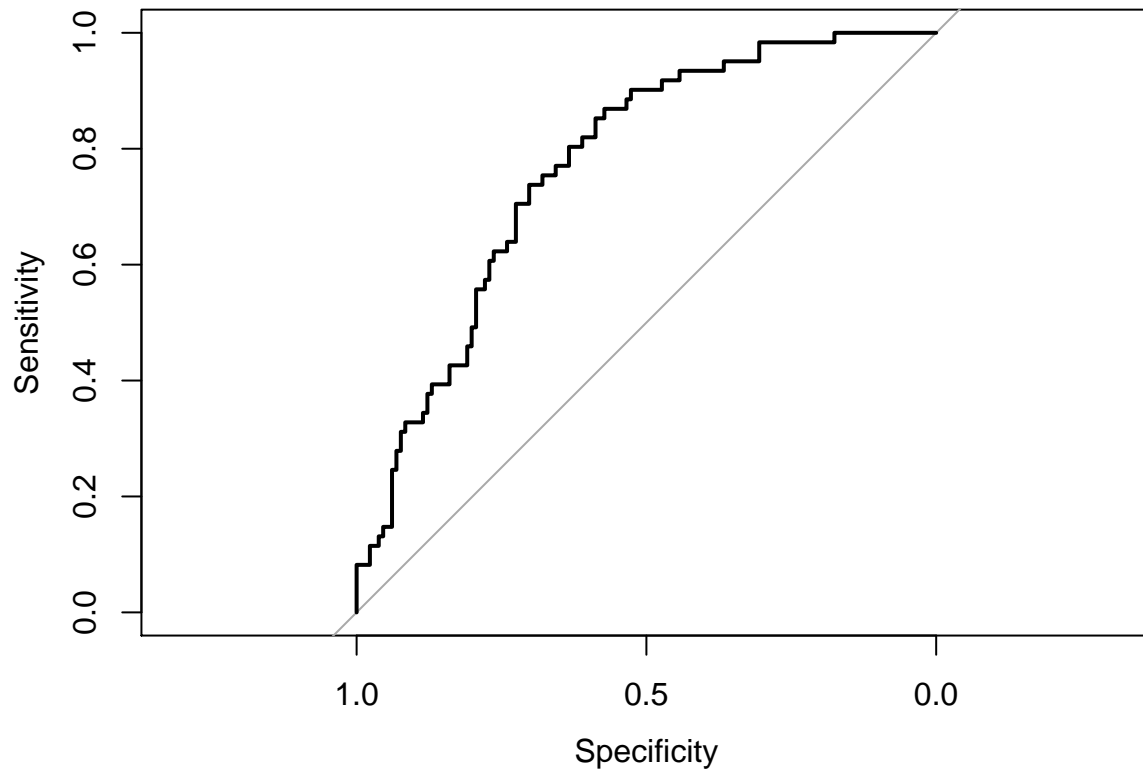


```
## [1] "AUC: 0.69804721655494"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.698
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.09637  -0.71025  -0.40193  -0.08668   2.84572
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.219e+00  5.004e-01  -4.434  9.27e-06 ***
## INCOME        -6.925e-06  2.907e-06  -2.382  0.017220 *
## TRAVTIME       2.125e-02  7.194e-03   2.954  0.003139 **
## BLUEBOOK       3.815e-05  1.582e-05   2.410  0.015931 *
## TIF           -4.386e-02  2.653e-02  -1.653  0.098252 .
## OLDCLAIM       1.684e-05  1.478e-05   1.139  0.254523
## PARENT1_Yes    1.153e+00  3.120e-01   3.696  0.000219 ***
## SEX_z_F       -1.064e+00  3.544e-01  -3.003  0.002674 **
## JOB_Manager   -1.022e+00  4.169e-01  -2.452  0.014225 *
```

```

## CAR_USE_Commercial          2.670e-01  2.462e-01  1.084 0.278214
## CAR_TYPE_Pickup             1.325e+00  3.292e-01  4.025 5.69e-05 ***
## CAR_TYPE_Sports.Car         1.780e+00  4.954e-01  3.593 0.000326 ***
## CAR_TYPE_z_SUV              1.647e+00  4.398e-01  3.745 0.000180 ***
## URBANICITY_z_Highly.Rural..Rural -2.398e+00  4.187e-01 -5.729 1.01e-08 ***
## HOME_VAL_NA                 -4.037e-01  2.291e-01 -1.762 0.078009 .
## oldclaim_log                4.804e-02  3.252e-02  1.477 0.139632
## inter                       1.961e-02  4.537e-03  4.321 1.55e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 700.16  on 624  degrees of freedom
## Residual deviance: 546.85  on 608  degrees of freedom
## AIC: 580.85
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 116  40
##           1  15  21
##
##               Accuracy : 0.7135
##               95% CI : (0.644, 0.7763)
##       No Information Rate : 0.6823
##       P-Value [Acc > NIR] : 0.197691
##
##               Kappa : 0.258
##
## Mcnemar's Test P-Value : 0.001211
##
##               Sensitivity : 0.8855
##               Specificity : 0.3443
##       Pos Pred Value : 0.7436
##       Neg Pred Value : 0.5833
##       Prevalence : 0.6823
##       Detection Rate : 0.6042
##       Detection Prevalence : 0.8125
##       Balanced Accuracy : 0.6149
##
##       'Positive' Class : 0
##

```

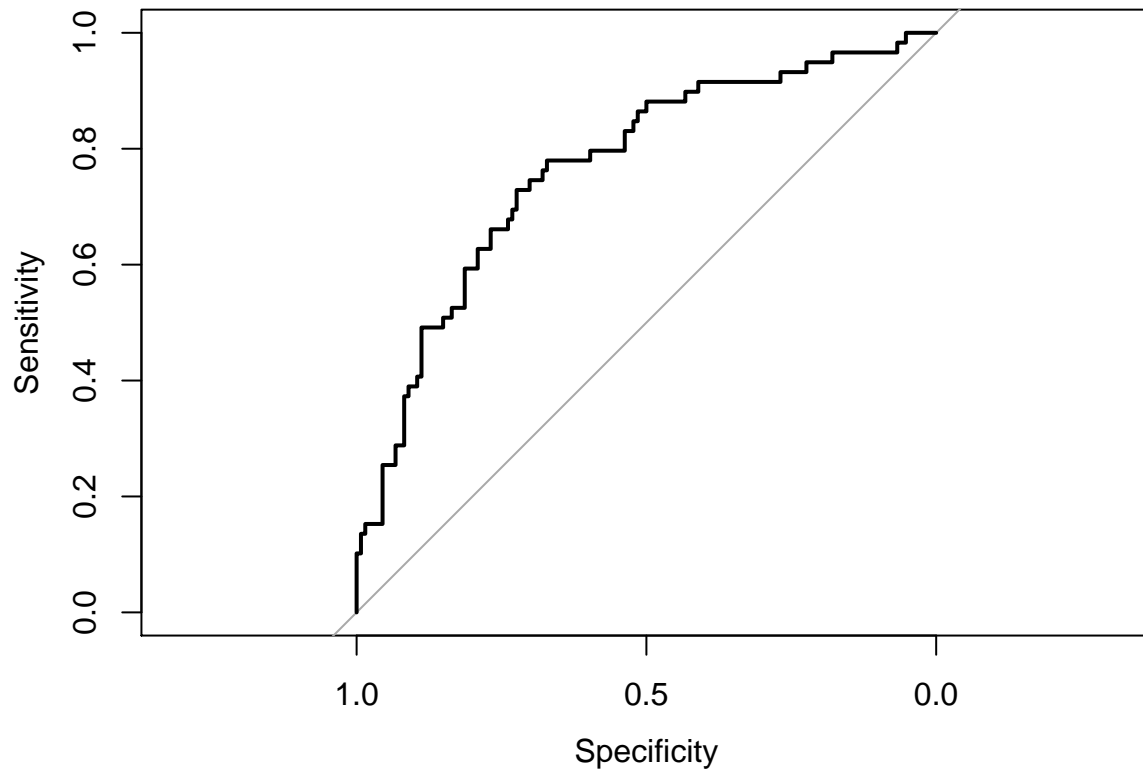


```
## [1] "AUC: 0.769490677011638"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 131 controls (dfPred_raw$class 0) < 61 cases (dfPred_raw$class 1).
## Area under the curve: 0.7695
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8573  -0.6933  -0.3879   0.3222   2.5515
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.837e+00  4.987e-01  -3.683 0.000230 ***
## INCOME          -9.806e-06  3.018e-06  -3.249 0.001158 **
## TRAVTIME         1.840e-02  7.097e-03   2.592 0.009531 **
## BLUEBOOK         2.138e-05  1.578e-05   1.355 0.175556
## TIF             -2.428e-02  2.663e-02  -0.912 0.362007
## OLDCLAIM         1.495e-05  1.489e-05   1.004 0.315422
## PARENT1_Yes      1.231e+00  3.102e-01   3.967 7.26e-05 ***
## SEX_z_F         -1.003e+00  3.573e-01  -2.806 0.005012 **
## JOB_Manager     -6.624e-01  3.685e-01  -1.798 0.072242 .
```

```

## CAR_USE_Commercial          4.025e-01  2.441e-01  1.649 0.099159 .
## CAR_TYPE_Pickup             9.732e-01  3.266e-01  2.980 0.002881 **
## CAR_TYPE_Sports.Car         1.756e+00  4.781e-01  3.674 0.000239 ***
## CAR_TYPE_z_SUV              1.572e+00  4.315e-01  3.643 0.000270 ***
## URBANICITY_z_Highly.Rural..Rural -2.555e+00  4.440e-01 -5.755 8.69e-09 ***
## HOME_VAL_NA                 -4.115e-01  2.261e-01 -1.820 0.068805 .
## oldclaim_log                7.147e-02  3.200e-02  2.233 0.025531 *
## inter                       1.426e-02  4.298e-03  3.318 0.000908 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 703.98  on 623  degrees of freedom
## Residual deviance: 542.28  on 607  degrees of freedom
## AIC: 576.28
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 123  39
##           1  11  20
##
##           Accuracy : 0.7409
##           95% CI : (0.6731, 0.8012)
##      No Information Rate : 0.6943
##      P-Value [Acc > NIR] : 0.0906524
##
##           Kappa : 0.2962
##
## Mcnemar's Test P-Value : 0.0001343
##
##           Sensitivity : 0.9179
##           Specificity : 0.3390
##      Pos Pred Value : 0.7593
##      Neg Pred Value : 0.6452
##           Prevalence : 0.6943
##      Detection Rate : 0.6373
##      Detection Prevalence : 0.8394
##      Balanced Accuracy : 0.6284
##
##           'Positive' Class : 0
##

```



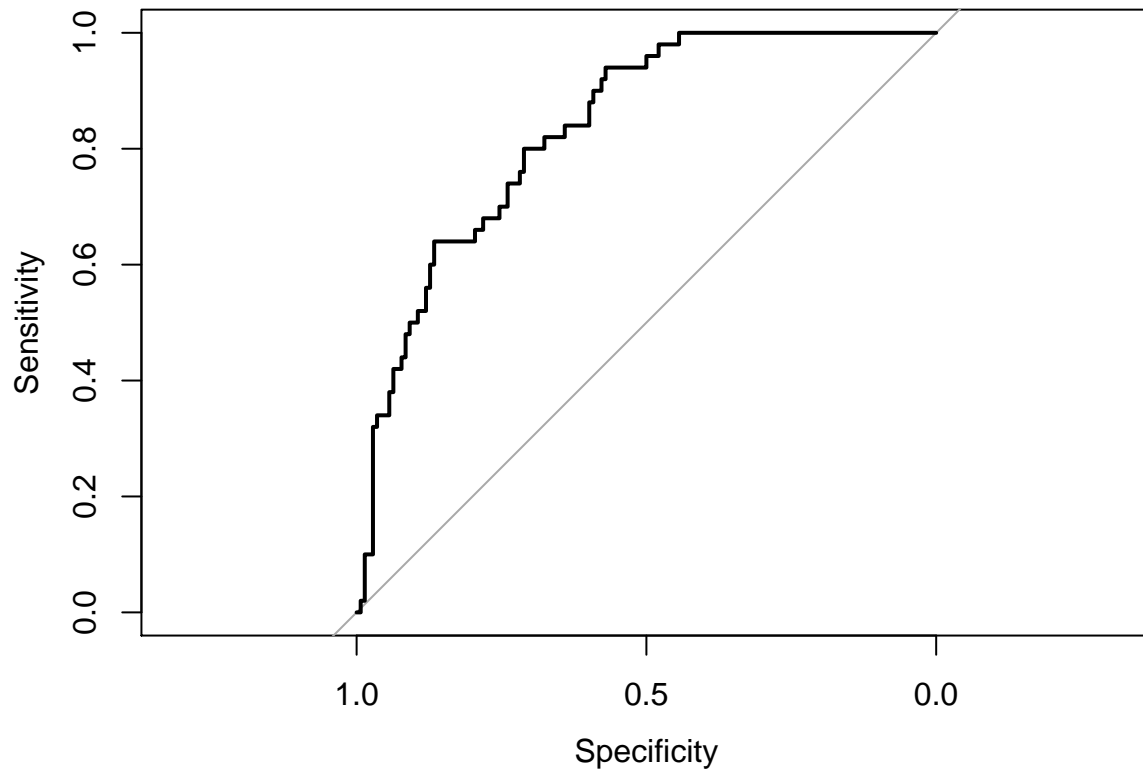
```
## [1] "AUC: 0.767391854287883"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 59 cases (dfPred_raw$class 1).
## Area under the curve: 0.7674
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8323  -0.7348  -0.4372   0.7331   2.7047
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.076e+00  4.862e-01  -4.270 1.95e-05 ***
## INCOME        -9.976e-06  2.900e-06  -3.440 0.000582 ***
## TRAVTIME       2.459e-02  6.900e-03   3.564 0.000366 ***
## BLUEBOOK       3.236e-05  1.565e-05   2.067 0.038710 *
## TIF           -3.937e-02  2.566e-02  -1.534 0.125005
## OLDCLAIM       1.180e-05  1.297e-05   0.909 0.363087
## PARENT1_Yes     8.681e-01  3.027e-01   2.868 0.004126 **
## SEX_z_F        -8.775e-01  3.466e-01  -2.532 0.011357 *
## JOB_Manager    -7.034e-01  3.617e-01  -1.945 0.051803 .
```



```

## CAR_USE_Commercial          2.720e-01  2.311e-01  1.177 0.239318
## CAR_TYPE_Pickup             9.438e-01  3.220e-01  2.931 0.003381 **
## CAR_TYPE_Sports.Car        1.787e+00  4.738e-01  3.771 0.000163 ***
## CAR_TYPE_z_SUV             1.495e+00  4.163e-01  3.592 0.000329 ***
## URBANICITY_z_Highly.Rural..Rural -2.194e+00  3.939e-01 -5.570 2.54e-08 ***
## HOME_VAL_NA                -1.804e-01  2.193e-01 -0.823 0.410720
## oldclaim_log               6.165e-02  3.147e-02  1.959 0.050065 .
## inter                      1.026e-02  4.344e-03  2.361 0.018220 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 723.54 on 624 degrees of freedom
## Residual deviance: 583.25 on 608 degrees of freedom
## AIC: 617.25
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 132  29
##           1  10  21
##
##           Accuracy : 0.7969
##           95% CI : (0.733, 0.8514)
##           No Information Rate : 0.7396
##           P-Value [Acc > NIR] : 0.039500
##
##           Kappa : 0.3987
##
## Mcnemar's Test P-Value : 0.003948
##
##           Sensitivity : 0.9296
##           Specificity : 0.4200
##           Pos Pred Value : 0.8199
##           Neg Pred Value : 0.6774
##           Prevalence : 0.7396
##           Detection Rate : 0.6875
##           Detection Prevalence : 0.8385
##           Balanced Accuracy : 0.6748
##
##           'Positive' Class : 0
##

```

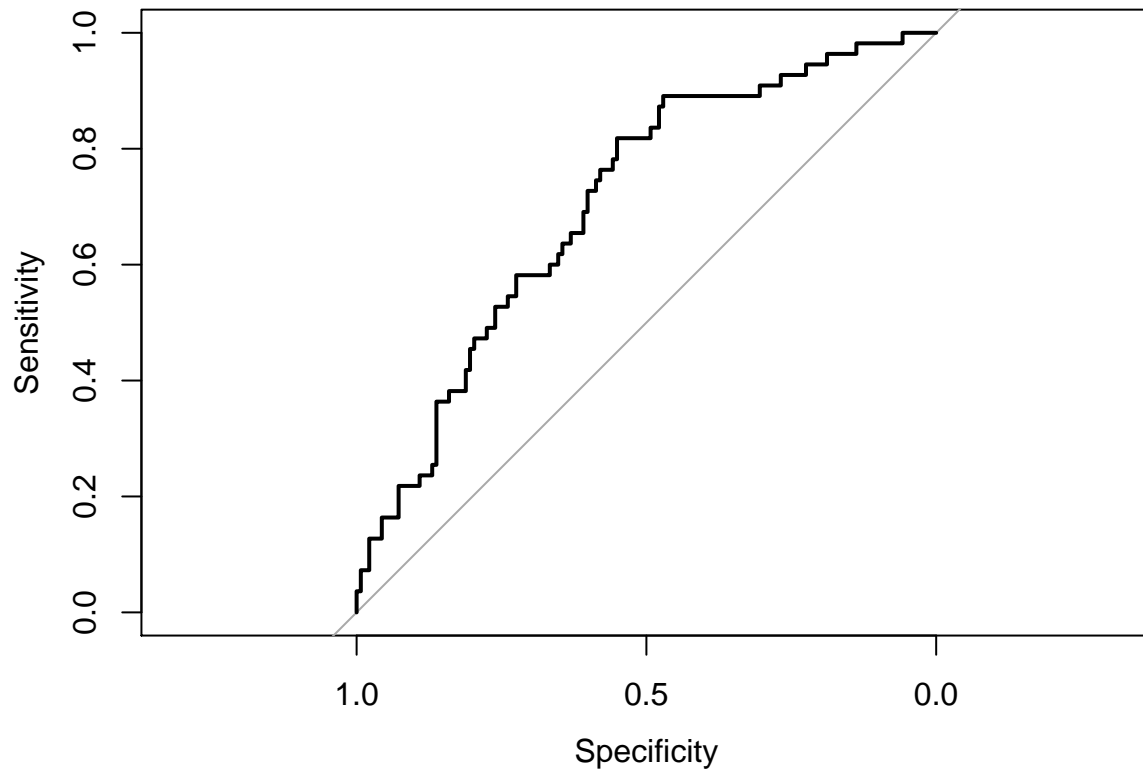


```
## [1] "AUC: 0.835492957746479"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 142 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.8355
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9128  -0.6912  -0.3603   0.5204   2.7333
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.272e+00  5.303e-01  -4.285 1.83e-05 ***
## INCOME        -1.290e-05  3.053e-06  -4.226 2.38e-05 ***
## TRAVTIME       1.979e-02  7.469e-03   2.650 0.008060 **
## BLUEBOOK       5.659e-05  1.583e-05   3.574 0.000351 ***
## TIF           -5.478e-02  2.804e-02  -1.953 0.050775 .
## OLDCLAIM       2.591e-05  1.445e-05   1.793 0.073018 .
## PARENT1_Yes    1.210e+00  3.078e-01   3.933 8.40e-05 ***
## SEX_z_F       -8.107e-01  3.525e-01  -2.300 0.021466 *
## JOB_Manager   -4.872e-01  3.433e-01  -1.419 0.155856
```

```

## CAR_USE_Commercial          6.604e-01  2.520e-01  2.621 0.008771 **
## CAR_TYPE_Pickup             9.752e-01  3.313e-01  2.944 0.003242 **
## CAR_TYPE_Sports.Car         1.944e+00  4.887e-01  3.978 6.96e-05 ***
## CAR_TYPE_z_SUV              1.683e+00  4.315e-01  3.901 9.58e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.676e+00  4.430e-01 -6.040 1.54e-09 ***
## HOME_VAL_NA                 -4.299e-01  2.296e-01 -1.873 0.061099 .
## oldclaim_log                4.181e-02  3.245e-02  1.288 0.197579
## inter                        1.334e-02  4.125e-03  3.233 0.001223 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 712.57 on 623 degrees of freedom
## Residual deviance: 533.87 on 607 degrees of freedom
## AIC: 567.87
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 123  42
##           1  15  13
##
##           Accuracy : 0.7047
##           95% CI : (0.6349, 0.768)
##           No Information Rate : 0.715
##           P-Value [Acc > NIR] : 0.6584089
##
##           Kappa : 0.1498
##
## Mcnemar's Test P-Value : 0.0005736
##
##           Sensitivity : 0.8913
##           Specificity : 0.2364
##           Pos Pred Value : 0.7455
##           Neg Pred Value : 0.4643
##           Prevalence : 0.7150
##           Detection Rate : 0.6373
##           Detection Prevalence : 0.8549
##           Balanced Accuracy : 0.5638
##
##           'Positive' Class : 0
##

```

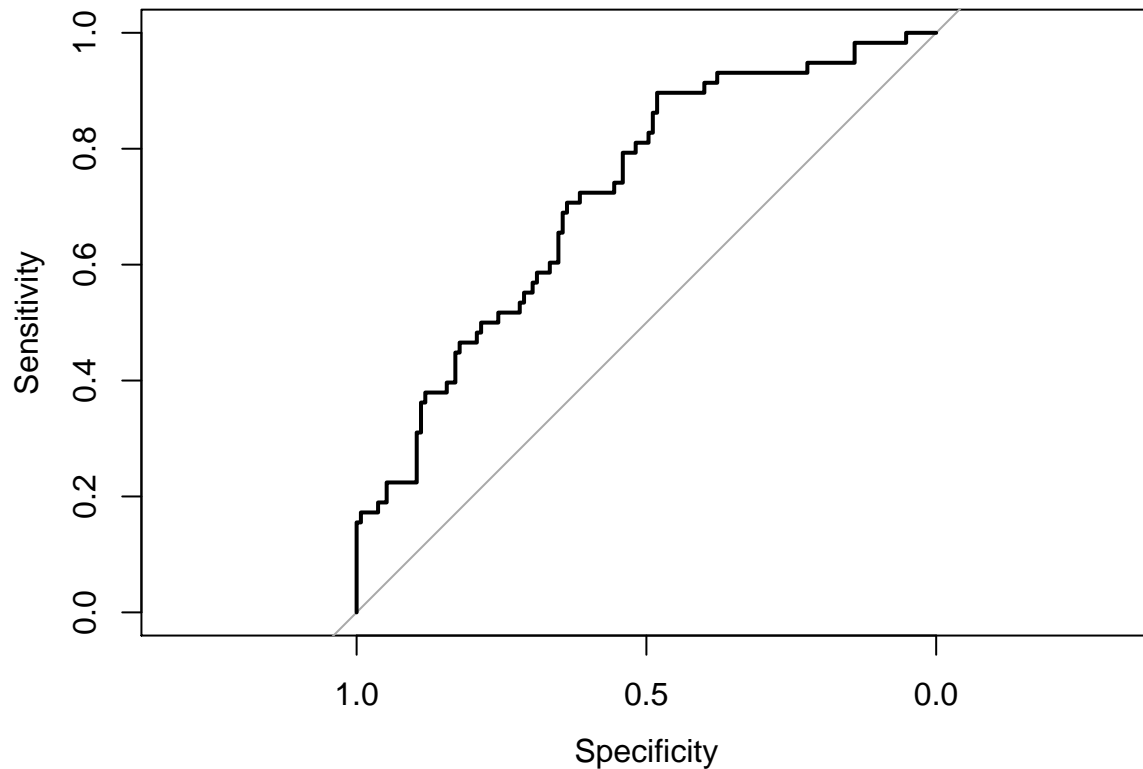


```
## [1] "AUC: 0.706324110671937"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 55 cases (dfPred_raw$class 1).
## Area under the curve: 0.7063
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8591  -0.6860  -0.3534   0.4293   2.8716
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.954e+00  5.104e-01  -3.829 0.000129 ***
## INCOME         -1.081e-05  3.089e-06  -3.501 0.000464 ***
## TRAVTIME        1.444e-02  7.171e-03   2.014 0.043992 *
## BLUEBOOK        3.528e-05  1.613e-05   2.187 0.028762 *
## TIF            -2.339e-02  2.760e-02  -0.847 0.396731
## OLDCLAIM        1.359e-05  1.588e-05   0.856 0.392136
## PARENT1_Yes     1.219e+00  3.126e-01   3.898 9.68e-05 ***
## SEX_z_F        -1.127e+00  3.608e-01  -3.125 0.001779 **
## JOB_Manager     -7.935e-01  3.916e-01  -2.026 0.042747 *
```

```

## CAR_USE_Commercial          4.464e-01  2.521e-01  1.771 0.076617 .
## CAR_TYPE_Pickup             1.188e+00  3.342e-01  3.555 0.000378 ***
## CAR_TYPE_Sports.Car         2.069e+00  4.966e-01  4.165 3.11e-05 ***
## CAR_TYPE_z_SUV              1.837e+00  4.423e-01  4.153 3.29e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.657e+00  4.424e-01 -6.007 1.89e-09 ***
## HOME_VAL_NA                 -4.251e-01  2.296e-01 -1.852 0.064086 .
## oldclaim_log                7.623e-02  3.217e-02  2.370 0.017804 *
## inter                       1.585e-02  4.068e-03  3.896 9.76e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 706.15 on 623 degrees of freedom
## Residual deviance: 531.64 on 607 degrees of freedom
## AIC: 565.64
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 120  37
##           1  15  21
##
##           Accuracy : 0.7306
##           95% CI : (0.6621, 0.7918)
##           No Information Rate : 0.6995
##           P-Value [Acc > NIR] : 0.194674
##
##           Kappa : 0.2814
##
## Mcnemar's Test P-Value : 0.003589
##
##           Sensitivity : 0.8889
##           Specificity : 0.3621
##           Pos Pred Value : 0.7643
##           Neg Pred Value : 0.5833
##           Prevalence : 0.6995
##           Detection Rate : 0.6218
##           Detection Prevalence : 0.8135
##           Balanced Accuracy : 0.6255
##
##           'Positive' Class : 0
##

```

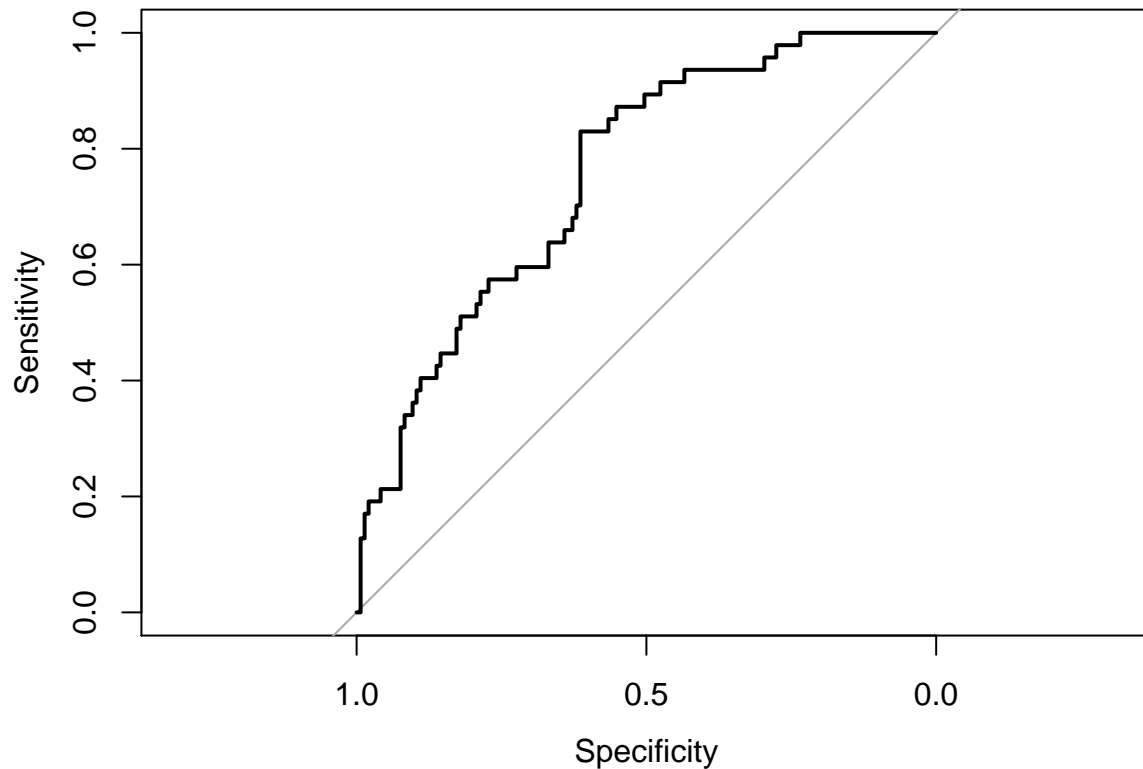


```
## [1] "AUC: 0.720945083014048"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 135 controls (dfPred_raw$class 0) < 58 cases (dfPred_raw$class 1).
## Area under the curve: 0.7209
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8526  -0.7409  -0.4024   0.6662   2.7778
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.284e+00  5.089e-01  -4.488 7.17e-06 ***
## INCOME        -9.173e-06  2.863e-06  -3.204 0.001356 **
## TRAVTIME       2.337e-02  7.092e-03   3.295 0.000984 ***
## BLUEBOOK       4.169e-05  1.584e-05   2.631 0.008511 **
## TIF           -7.308e-02  2.854e-02  -2.561 0.010449 *
## OLDCLAIM       2.022e-05  1.422e-05   1.422 0.155007
## PARENT1_Yes    1.061e+00  3.075e-01   3.451 0.000559 ***
## SEX_z_F       -1.298e+00  3.882e-01  -3.345 0.000822 ***
## JOB_Manager   -8.468e-01  3.786e-01  -2.236 0.025327 *
```

```

## CAR_USE_Commercial          5.562e-01  2.473e-01  2.249 0.024511 *
## CAR_TYPE_Pickup             1.342e+00  3.347e-01  4.008 6.11e-05 ***
## CAR_TYPE_Sports.Car        2.551e+00  5.267e-01  4.843 1.28e-06 ***
## CAR_TYPE_z_SUV             2.219e+00  4.736e-01  4.685 2.79e-06 ***
## URBANICITY_z_Highly.Rural..Rural -2.286e+00  3.900e-01 -5.863 4.55e-09 ***
## HOME_VAL_NA                -2.305e-01  2.222e-01 -1.037 0.299739
## oldclaim_log               2.985e-02  3.180e-02  0.939 0.347862
## inter                      1.541e-02  4.068e-03  3.787 0.000152 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 729.57 on 624 degrees of freedom
## Residual deviance: 567.06 on 608 degrees of freedom
## AIC: 601.06
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 131  30
##           1  14  17
##
##           Accuracy : 0.7708
##           95% CI : (0.7048, 0.8283)
##           No Information Rate : 0.7552
##           P-Value [Acc > NIR] : 0.34178
##
##           Kappa : 0.2996
##
## Mcnemar's Test P-Value : 0.02374
##
##           Sensitivity : 0.9034
##           Specificity : 0.3617
##           Pos Pred Value : 0.8137
##           Neg Pred Value : 0.5484
##           Prevalence : 0.7552
##           Detection Rate : 0.6823
##           Detection Prevalence : 0.8385
##           Balanced Accuracy : 0.6326
##
##           'Positive' Class : 0
##

```



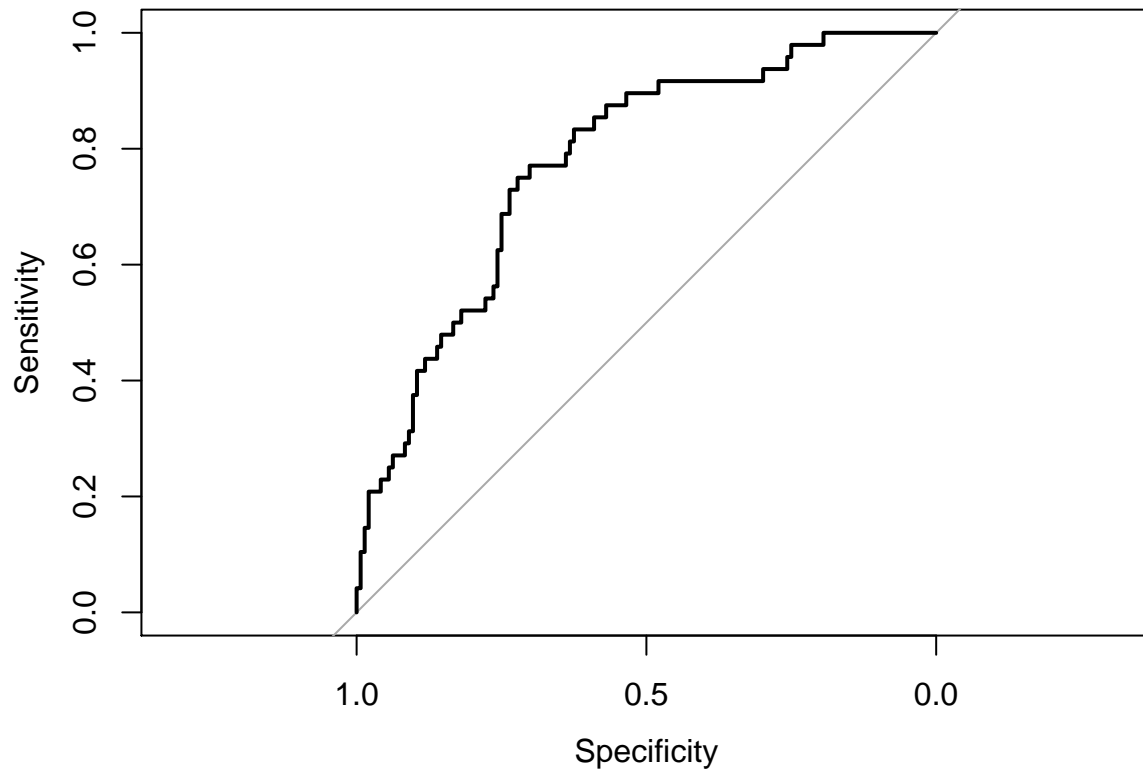
```
## [1] "AUC: 0.757740278796772"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 145 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.7577
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7846  -0.7393  -0.4081   0.6429   2.9095
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.009e+00  4.826e-01  -4.164 3.13e-05 ***
## INCOME        -9.706e-06  2.983e-06  -3.254 0.001137 **
## TRAVTIME       2.072e-02  6.968e-03   2.973 0.002947 **
## BLUEBOOK       3.329e-05  1.555e-05   2.141 0.032285 *
## TIF           -2.530e-02  2.649e-02  -0.955 0.339581
## OLDCLAIM      -2.015e-06  1.528e-05  -0.132 0.895054
## PARENT1_Yes    1.082e+00  2.967e-01   3.645 0.000267 ***
## SEX_z_F       -9.708e-01  3.578e-01  -2.713 0.006669 **
## JOB_Manager   -5.437e-01  3.626e-01  -1.499 0.133785
```



```

## CAR_USE_Commercial          5.353e-01  2.383e-01  2.246 0.024697 *
## CAR_TYPE_Pickup             1.007e+00  3.192e-01  3.154 0.001608 **
## CAR_TYPE_Sports.Car         1.854e+00  4.818e-01  3.849 0.000119 ***
## CAR_TYPE_z_SUV              1.744e+00  4.358e-01  4.002 6.28e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.482e+00  4.072e-01 -6.096 1.09e-09 ***
## HOME_VAL_NA                 -4.575e-01  2.249e-01 -2.034 0.041944 *
## oldclaim_log                5.714e-02  3.196e-02  1.788 0.073814 .
## inter                       1.700e-02  4.214e-03  4.033 5.50e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 727.58  on 624  degrees of freedom
## Residual deviance: 568.32  on 608  degrees of freedom
## AIC: 602.32
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 131  33
##           1  13  15
##
##           Accuracy : 0.7604
##           95% CI : (0.6937, 0.8189)
##      No Information Rate : 0.75
##      P-Value [Acc > NIR] : 0.406410
##
##           Kappa : 0.2581
##
## Mcnemar's Test P-Value : 0.005088
##
##           Sensitivity : 0.9097
##           Specificity : 0.3125
##      Pos Pred Value : 0.7988
##      Neg Pred Value : 0.5357
##           Prevalence : 0.7500
##      Detection Rate : 0.6823
##      Detection Prevalence : 0.8542
##      Balanced Accuracy : 0.6111
##
##           'Positive' Class : 0
##

```

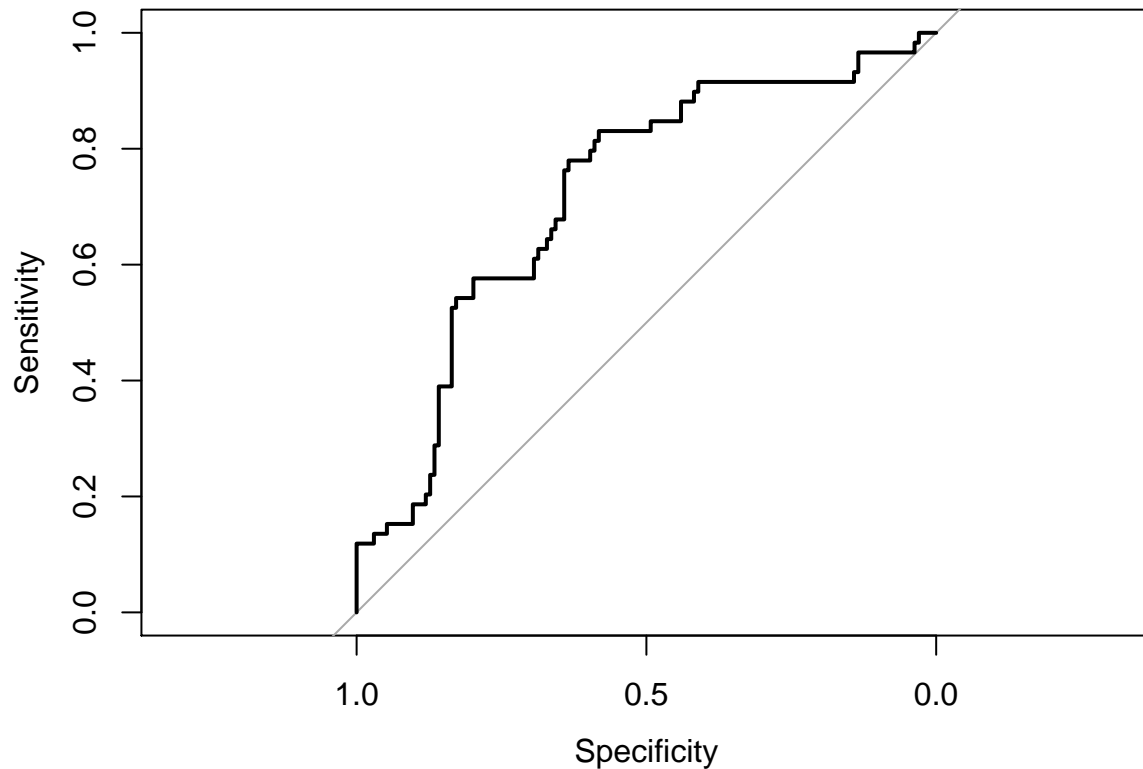


```
## [1] "AUC: 0.777777777777778"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 144 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.7778
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9380  -0.6787  -0.3615   0.4866   2.6503
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.881e+00  5.051e-01  -3.723 0.000196 ***
## INCOME        -1.209e-05  3.173e-06  -3.812 0.000138 ***
## TRAVTIME       1.533e-02  7.132e-03   2.150 0.031552 *
## BLUEBOOK       3.049e-05  1.615e-05   1.888 0.059060 .
## TIF           -1.107e-02  2.694e-02  -0.411 0.681094
## OLDCLAIM       1.343e-06  1.673e-05   0.080 0.936032
## PARENT1_Yes    1.387e+00  3.059e-01   4.535 5.76e-06 ***
## SEX_z_F       -8.897e-01  3.488e-01  -2.551 0.010743 *
## JOB_Manager   -4.195e-01  3.710e-01  -1.131 0.258229
```

```

## CAR_USE_Commercial          4.032e-01  2.462e-01  1.638 0.101438
## CAR_TYPE_Pickup             1.135e+00  3.246e-01  3.496 0.000472 ***
## CAR_TYPE_Sports.Car         1.694e+00  4.703e-01  3.603 0.000315 ***
## CAR_TYPE_z_SUV              1.582e+00  4.242e-01  3.728 0.000193 ***
## URBANICITY_z_Highly.Rural..Rural -2.669e+00  4.480e-01 -5.956 2.58e-09 ***
## HOME_VAL_NA                 -5.505e-01  2.297e-01 -2.396 0.016567 *
## oldclaim_log                8.978e-02  3.277e-02  2.740 0.006143 **
## inter                       1.430e-02  4.327e-03  3.304 0.000952 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 703.98 on 623 degrees of freedom
## Residual deviance: 527.07 on 607 degrees of freedom
## AIC: 561.07
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 120  48
##           1  14  11
##
##           Accuracy : 0.6788
##           95% CI : (0.6079, 0.744)
##           No Information Rate : 0.6943
##           P-Value [Acc > NIR] : 0.7101
##
##           Kappa : 0.0977
##
## Mcnemar's Test P-Value : 2.777e-05
##
##           Sensitivity : 0.8955
##           Specificity : 0.1864
##           Pos Pred Value : 0.7143
##           Neg Pred Value : 0.4400
##           Prevalence : 0.6943
##           Detection Rate : 0.6218
##           Detection Prevalence : 0.8705
##           Balanced Accuracy : 0.5410
##
##           'Positive' Class : 0
##

```

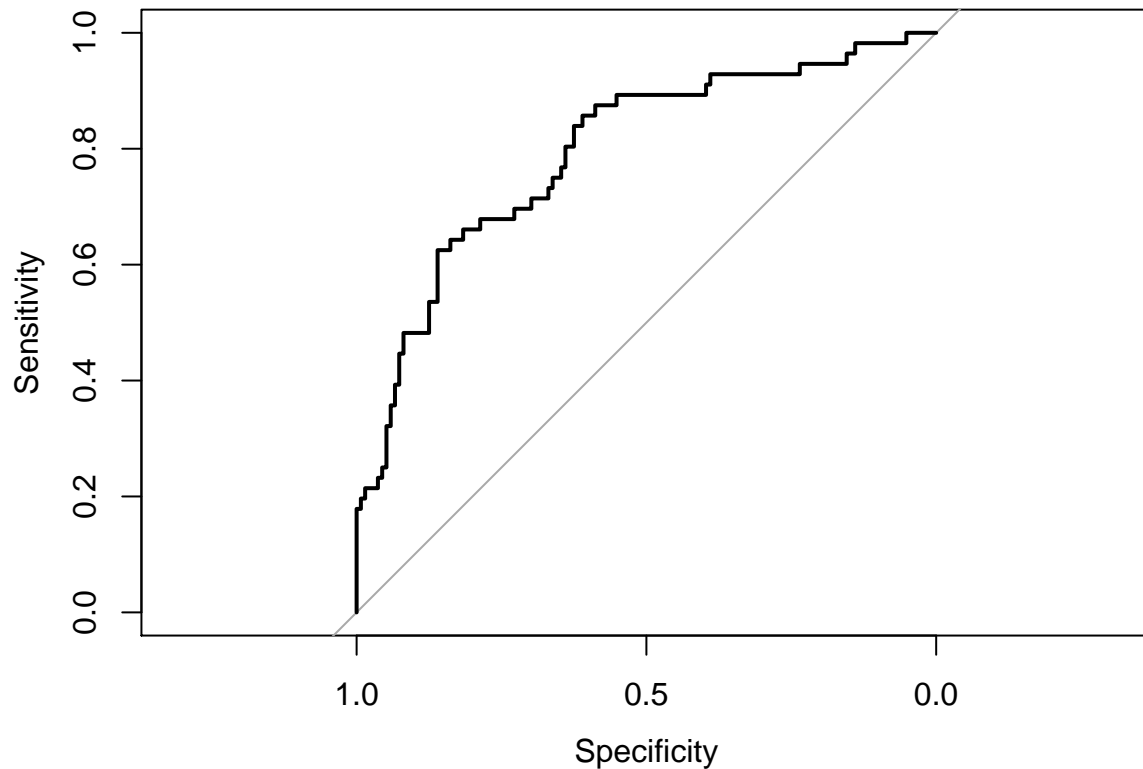


```
## [1] "AUC: 0.721477358967872"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 59 cases (dfPred_raw$class 1).
## Area under the curve: 0.7215
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7628  -0.7251  -0.4107   0.6085   2.9017
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.657e+00  4.975e-01  -3.331 0.000865 ***
## INCOME        -9.771e-06  3.036e-06  -3.218 0.001290 **
## TRAVTIME       1.700e-02  7.114e-03   2.389 0.016875 *
## BLUEBOOK       1.268e-05  1.571e-05   0.807 0.419460
## TIF           -3.689e-02  2.662e-02  -1.386 0.165813
## OLDCLAIM       1.783e-05  1.442e-05   1.236 0.216312
## PARENT1_Yes    1.001e+00  3.035e-01   3.297 0.000977 ***
## SEX_z_F       -7.731e-01  3.511e-01  -2.202 0.027670 *
## JOB_Manager   -4.091e-01  3.474e-01  -1.178 0.238887
```

```

## CAR_USE_Commercial          7.938e-01  2.453e-01  3.236 0.001210 **
## CAR_TYPE_Pickup             8.509e-01  3.217e-01  2.645 0.008164 **
## CAR_TYPE_Sports.Car        1.776e+00  4.754e-01  3.736 0.000187 ***
## CAR_TYPE_z_SUV             1.587e+00  4.247e-01  3.738 0.000186 ***
## URBANICITY_z_Highly.Rural..Rural -2.312e+00  4.176e-01 -5.538 3.07e-08 ***
## HOME_VAL_NA                -4.283e-01  2.209e-01 -1.939 0.052501 .
## oldclaim_log               1.941e-02  3.170e-02  0.612 0.540303
## inter                      1.233e-02  4.159e-03  2.965 0.003025 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 711.04 on 624 degrees of freedom
## Residual deviance: 563.68 on 608 degrees of freedom
## AIC: 597.68
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 129  41
##           1   7  15
##
##           Accuracy : 0.75
##           95% CI : (0.6826, 0.8096)
##           No Information Rate : 0.7083
##           P-Value [Acc > NIR] : 0.1158
##
##           Kappa : 0.2634
##
## Mcnemar's Test P-Value : 1.906e-06
##
##           Sensitivity : 0.9485
##           Specificity : 0.2679
##           Pos Pred Value : 0.7588
##           Neg Pred Value : 0.6818
##           Prevalence : 0.7083
##           Detection Rate : 0.6719
##           Detection Prevalence : 0.8854
##           Balanced Accuracy : 0.6082
##
##           'Positive' Class : 0
##

```

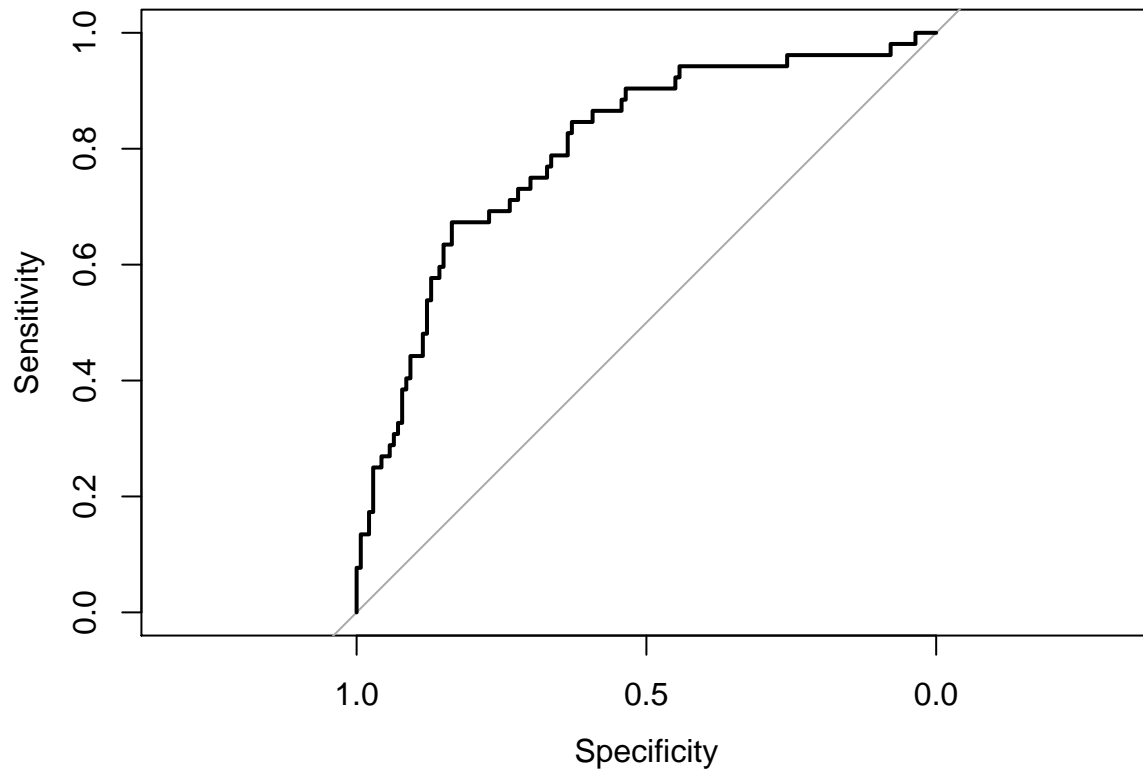


```
## [1] "AUC: 0.794248949579832"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 136 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7942
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9051  -0.7250  -0.4111   0.6332   2.8641
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.682e+00  4.958e-01  -3.393 0.000690 ***
## INCOME        -1.106e-05  2.985e-06  -3.704 0.000212 ***
## TRAVTIME       1.639e-02  6.820e-03   2.404 0.016218 *
## BLUEBOOK       1.936e-05  1.574e-05   1.229 0.218970
## TIF           -4.263e-02  2.638e-02  -1.616 0.106163
## OLDCLAIM       1.551e-05  1.384e-05   1.121 0.262423
## PARENT1_Yes     9.378e-01  3.113e-01   3.012 0.002595 **
## SEX_z_F        -9.001e-01  3.457e-01  -2.603 0.009229 **
## JOB_Manager    -4.286e-01  3.471e-01  -1.235 0.216839
```

```

## CAR_USE_Commercial          5.508e-01  2.420e-01  2.275 0.022883 *
## CAR_TYPE_Pickup             1.192e+00  3.245e-01  3.674 0.000239 ***
## CAR_TYPE_Sports.Car        1.960e+00  4.738e-01  4.136 3.53e-05 ***
## CAR_TYPE_z_SUV             1.756e+00  4.236e-01  4.146 3.38e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.323e+00  4.010e-01 -5.793 6.89e-09 ***
## HOME_VAL_NA                -3.012e-01  2.208e-01 -1.364 0.172507
## oldclaim_log               3.344e-02  3.183e-02  1.051 0.293463
## inter                      1.404e-02  4.363e-03  3.218 0.001291 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 719.44 on 624 degrees of freedom
## Residual deviance: 569.31 on 608 degrees of freedom
## AIC: 603.31
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 128  32
##           1  12  20
##
##           Accuracy : 0.7708
##           95% CI : (0.7048, 0.8283)
##           No Information Rate : 0.7292
##           P-Value [Acc > NIR] : 0.110326
##
##           Kappa : 0.34
##
## Mcnemar's Test P-Value : 0.004179
##
##           Sensitivity : 0.9143
##           Specificity : 0.3846
##           Pos Pred Value : 0.8000
##           Neg Pred Value : 0.6250
##           Prevalence : 0.7292
##           Detection Rate : 0.6667
##           Detection Prevalence : 0.8333
##           Balanced Accuracy : 0.6495
##
##           'Positive' Class : 0
##

```



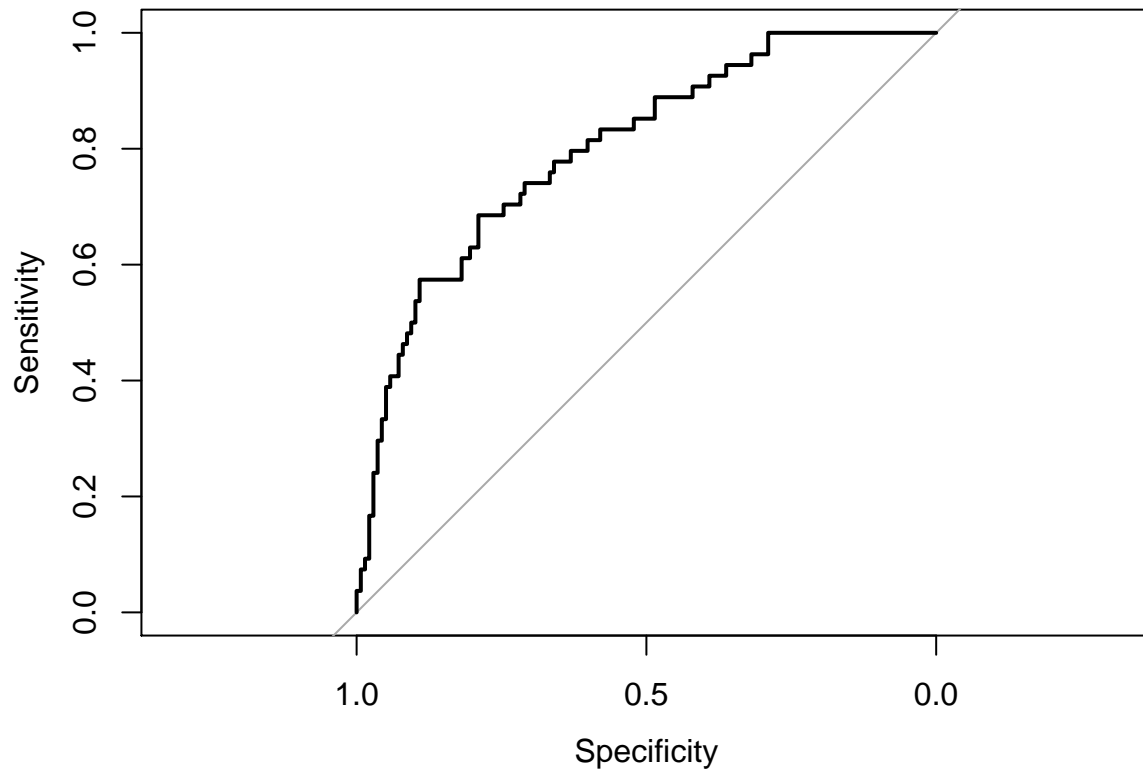
```
## [1] "AUC: 0.799175824175824"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 140 controls (dfPred_raw$class 0) < 52 cases (dfPred_raw$class 1).
## Area under the curve: 0.7992
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0316  -0.7344  -0.4005   0.5994   2.8354
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.011e+00  4.841e-01  -4.155 3.25e-05 ***
## INCOME         -7.580e-06  2.881e-06  -2.631 0.008514 **
## TRAVTIME        2.016e-02  6.867e-03   2.936 0.003327 **
## BLUEBOOK        2.494e-05  1.559e-05   1.600 0.109526
## TIF            -2.785e-02  2.594e-02  -1.073 0.283100
## OLDCLAIM        8.863e-06  1.438e-05   0.616 0.537607
## PARENT1_Yes     1.027e+00  3.073e-01   3.344 0.000826 ***
## SEX_z_F        -1.011e+00  3.531e-01  -2.864 0.004183 **
## JOB_Manager    -1.066e+00  4.019e-01  -2.654 0.007965 **
```



```

## CAR_USE_Commercial          3.899e-01  2.379e-01  1.639 0.101265
## CAR_TYPE_Pickup             1.064e+00  3.261e-01  3.263 0.001103 **
## CAR_TYPE_Sports.Car         1.832e+00  4.888e-01  3.748 0.000178 ***
## CAR_TYPE_z_SUV              1.559e+00  4.323e-01  3.605 0.000312 ***
## URBANICITY_z_Highly.Rural..Rural -2.462e+00  4.199e-01 -5.864 4.53e-09 ***
## HOME_VAL_NA                 -2.105e-01  2.240e-01 -0.940 0.347413
## oldclaim_log                5.210e-02  3.151e-02  1.653 0.098259 .
## inter                       1.711e-02  4.178e-03  4.095 4.22e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 715.27 on 624 degrees of freedom
## Residual deviance: 567.86 on 608 degrees of freedom
## AIC: 601.86
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 129  32
##           1   9  22
##
##           Accuracy : 0.7865
##           95% CI : (0.7217, 0.8422)
##           No Information Rate : 0.7188
##           P-Value [Acc > NIR] : 0.0202956
##
##           Kappa : 0.3932
##
## Mcnemar's Test P-Value : 0.0005908
##
##           Sensitivity : 0.9348
##           Specificity : 0.4074
##           Pos Pred Value : 0.8012
##           Neg Pred Value : 0.7097
##           Prevalence : 0.7188
##           Detection Rate : 0.6719
##           Detection Prevalence : 0.8385
##           Balanced Accuracy : 0.6711
##
##           'Positive' Class : 0
##

```

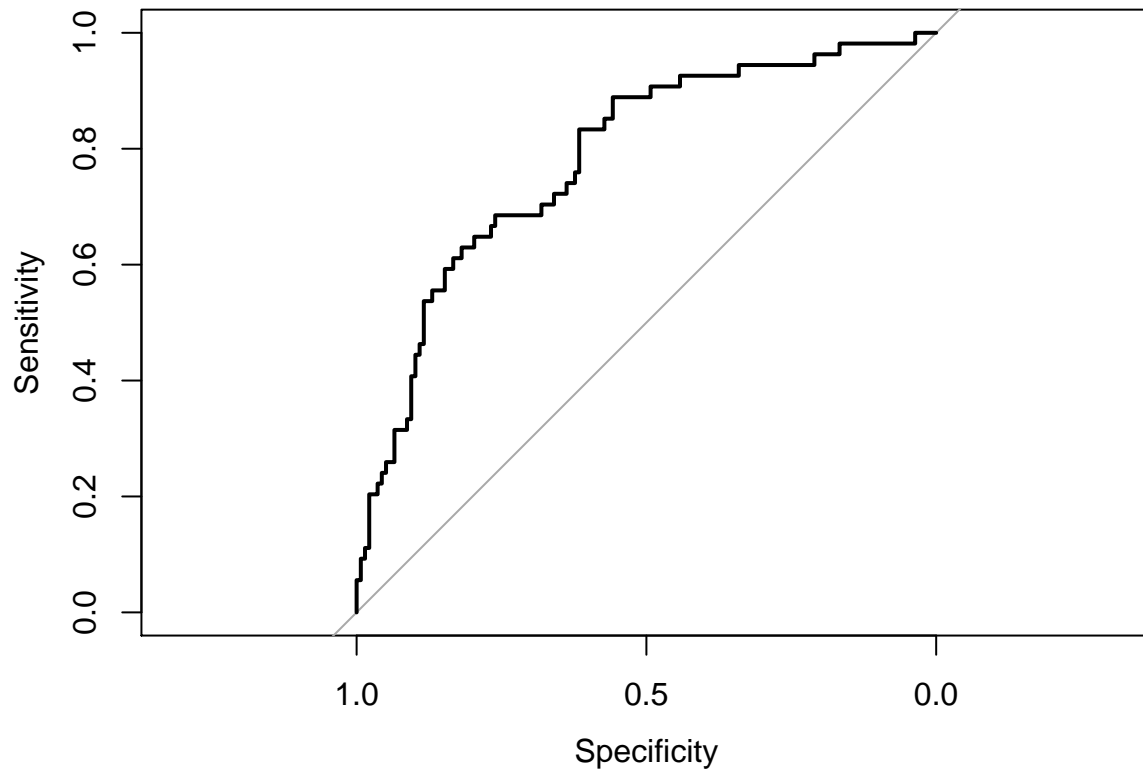


```
## [1] "AUC: 0.800456253354804"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 54 cases (dfPred_raw$class 1).
## Area under the curve: 0.8005
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9684  -0.7362  -0.3939   0.6472   2.9838
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.887e+00  4.935e-01  -3.823 0.000132 ***
## INCOME        -9.931e-06  2.958e-06  -3.357 0.000788 ***
## TRAVTIME       2.285e-02  7.057e-03   3.239 0.001201 **
## BLUEBOOK       1.543e-05  1.569e-05   0.983 0.325448
## TIF           -3.155e-02  2.608e-02  -1.210 0.226358
## OLDCLAIM       1.850e-05  1.427e-05   1.297 0.194800
## PARENT1_Yes     8.750e-01  2.954e-01   2.962 0.003054 **
## SEX_z_F       -6.482e-01  3.372e-01  -1.922 0.054564 .
## JOB_Manager    -7.317e-01  3.709e-01  -1.973 0.048484 *
```

```

## CAR_USE_Commercial          5.114e-01  2.368e-01  2.160 0.030795 *
## CAR_TYPE_Pickup             1.064e+00  3.151e-01  3.377 0.000733 ***
## CAR_TYPE_Sports.Car         1.683e+00  4.763e-01  3.534 0.000409 ***
## CAR_TYPE_z_SUV              1.398e+00  4.082e-01  3.426 0.000613 ***
## URBANICITY_z_Highly.Rural..Rural -2.743e+00  4.644e-01 -5.907 3.48e-09 ***
## HOME_VAL_NA                 -2.643e-01  2.236e-01 -1.182 0.237197
## oldclaim_log                2.821e-02  3.183e-02  0.886 0.375474
## inter                       1.326e-02  4.123e-03  3.216 0.001301 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 715.27 on 624 degrees of freedom
## Residual deviance: 559.33 on 608 degrees of freedom
## AIC: 593.33
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 130  40
##           1   8  14
##
##           Accuracy : 0.75
##           95% CI : (0.6826, 0.8096)
##           No Information Rate : 0.7188
##           P-Value [Acc > NIR] : 0.1894
##
##           Kappa : 0.2456
##
## Mcnemar's Test P-Value : 7.66e-06
##
##           Sensitivity : 0.9420
##           Specificity : 0.2593
##           Pos Pred Value : 0.7647
##           Neg Pred Value : 0.6364
##           Prevalence : 0.7188
##           Detection Rate : 0.6771
##           Detection Prevalence : 0.8854
##           Balanced Accuracy : 0.6006
##
##           'Positive' Class : 0
##

```

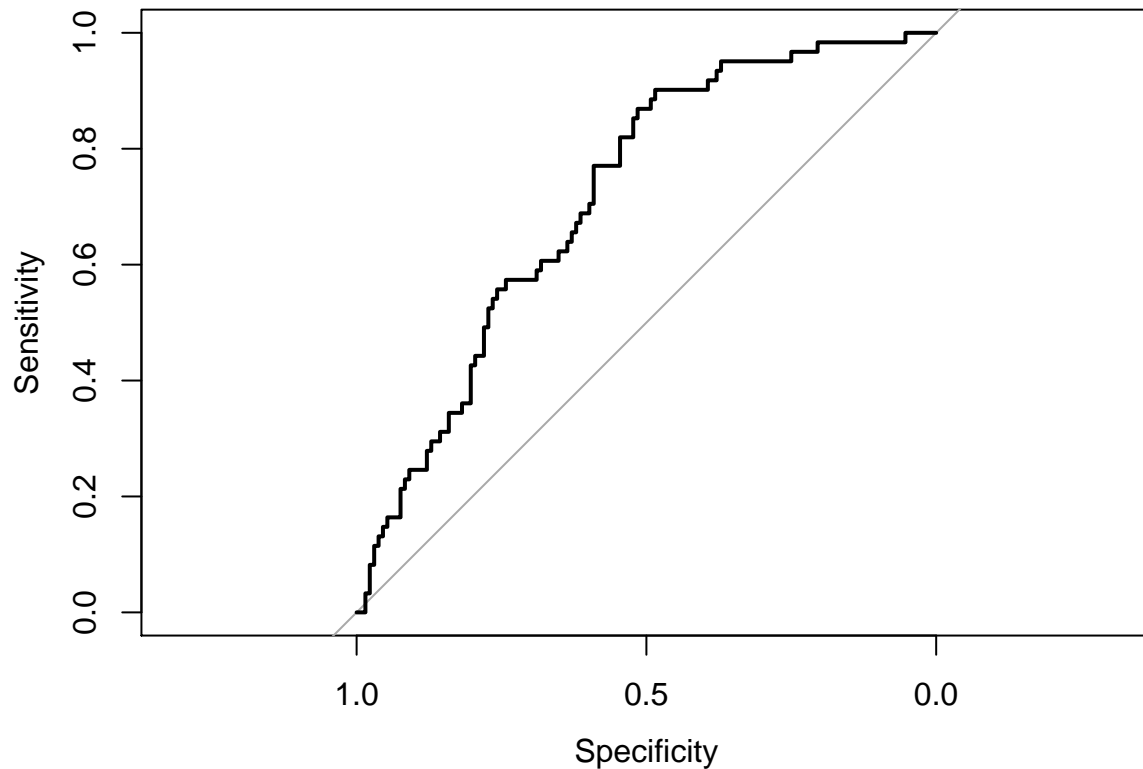


```
## [1] "AUC: 0.783011272141707"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 54 cases (dfPred_raw$class 1).
## Area under the curve: 0.783
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8984  -0.6771  -0.3741  -0.0816   2.8003
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.316e+00  5.425e-01  -4.270 1.95e-05 ***
## INCOME         -1.210e-05  3.050e-06  -3.967 7.27e-05 ***
## TRAVTIME        2.128e-02  7.327e-03   2.905 0.003675 **
## BLUEBOOK        4.960e-05  1.619e-05   3.063 0.002192 **
## TIF            -6.172e-02  2.914e-02  -2.118 0.034140 *
## OLDCLAIM        1.964e-05  1.416e-05   1.387 0.165436
## PARENT1_Yes     1.214e+00  3.183e-01   3.815 0.000136 ***
## SEX_z_F        -8.968e-01  3.646e-01  -2.459 0.013916 *
## JOB_Manager    -4.280e-01  3.437e-01  -1.245 0.212965
```

```

## CAR_USE_Commercial          4.310e-01  2.526e-01  1.706 0.087967 .
## CAR_TYPE_Pickup             1.128e+00  3.404e-01  3.314 0.000918 ***
## CAR_TYPE_Sports.Car         1.980e+00  4.932e-01  4.013 5.99e-05 ***
## CAR_TYPE_z_SUV              1.697e+00  4.408e-01  3.851 0.000118 ***
## URBANICITY_z_Highly.Rural..Rural -2.520e+00  4.413e-01 -5.711 1.12e-08 ***
## HOME_VAL_NA                 -3.720e-01  2.296e-01 -1.620 0.105210
## oldclaim_log                8.350e-02  3.252e-02  2.568 0.010237 *
## inter                       9.431e-03  4.356e-03  2.165 0.030366 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 699.59  on 623  degrees of freedom
## Residual deviance: 527.99  on 607  degrees of freedom
## AIC: 561.99
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 112  42
##           1  20  19
##
##           Accuracy : 0.6788
##           95% CI : (0.6079, 0.744)
##      No Information Rate : 0.6839
##      P-Value [Acc > NIR] : 0.595270
##
##           Kappa : 0.1771
##
## Mcnemar's Test P-Value : 0.007653
##
##           Sensitivity : 0.8485
##           Specificity : 0.3115
##      Pos Pred Value : 0.7273
##      Neg Pred Value : 0.4872
##           Prevalence : 0.6839
##      Detection Rate : 0.5803
##      Detection Prevalence : 0.7979
##      Balanced Accuracy : 0.5800
##
##      'Positive' Class : 0
##

```

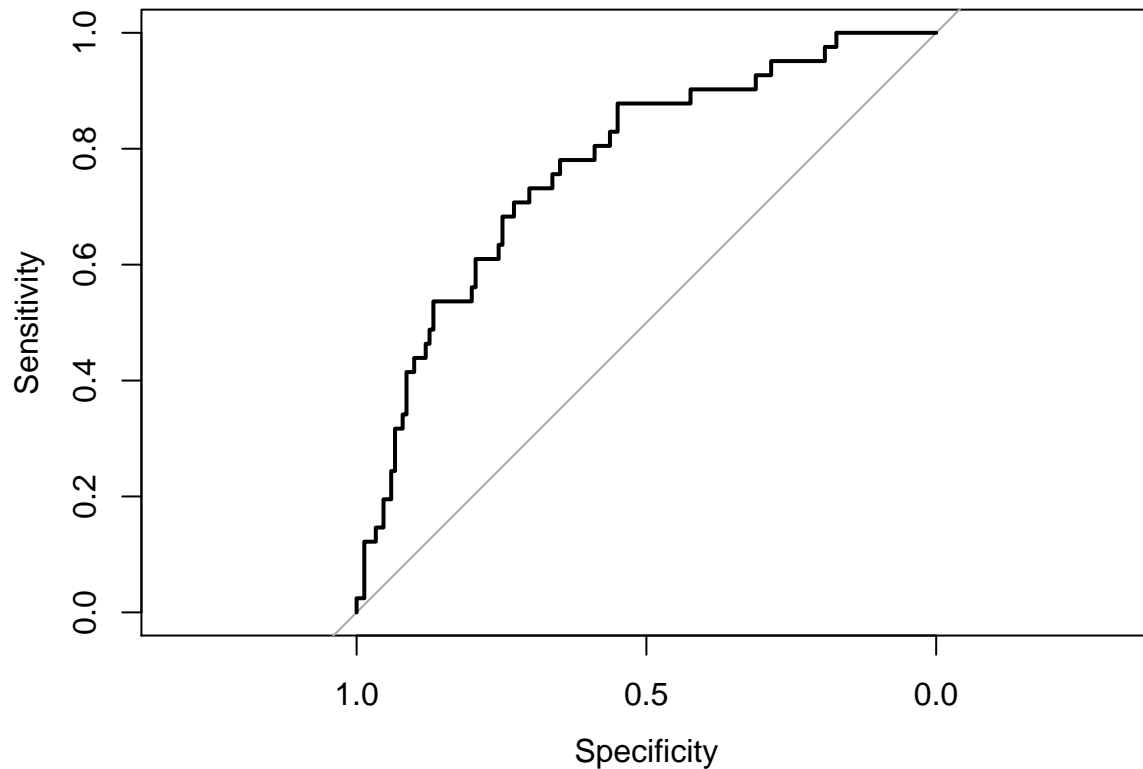


```
## [1] "AUC: 0.717337307501242"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 132 controls (dfPred_raw$class 0) < 61 cases (dfPred_raw$class 1).
## Area under the curve: 0.7173
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9853  -0.7491  -0.4181   0.7174   2.8958
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.702e+00  4.729e-01  -3.599 0.000320 ***
## INCOME        -1.165e-05  3.058e-06  -3.809 0.000140 ***
## TRAVTIME       1.642e-02  6.757e-03   2.430 0.015105 *
## BLUEBOOK       2.741e-05  1.514e-05   1.810 0.070293 .
## TIF           -1.257e-02  2.566e-02  -0.490 0.624094
## OLDCLAIM       1.922e-06  1.438e-05   0.134 0.893676
## PARENT1_Yes    1.097e+00  3.096e-01   3.541 0.000398 ***
## SEX_z_F       -7.962e-01  3.356e-01  -2.372 0.017680 *
## JOB_Manager   -6.132e-01  3.427e-01  -1.789 0.073537 .
```

```

## CAR_USE_Commercial          6.324e-01  2.393e-01  2.643 0.008223 **
## CAR_TYPE_Pickup             6.753e-01  3.132e-01  2.156 0.031055 *
## CAR_TYPE_Sports.Car         1.668e+00  4.616e-01  3.615 0.000301 ***
## CAR_TYPE_z_SUV              1.465e+00  4.080e-01  3.592 0.000328 ***
## URBANICITY_z_Highly.Rural..Rural -2.532e+00  4.008e-01 -6.317 2.67e-10 ***
## HOME_VAL_NA                 -3.107e-01  2.257e-01 -1.377 0.168639
## oldclaim_log                5.279e-02  3.110e-02  1.697 0.089602 .
## inter                       1.661e-02  4.065e-03  4.087 4.36e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 741.19  on 624  degrees of freedom
## Residual deviance: 578.38  on 608  degrees of freedom
## AIC: 612.38
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 137  24
##           1  14  17
##
##           Accuracy : 0.8021
##           95% CI : (0.7386, 0.856)
##           No Information Rate : 0.7865
##           P-Value [Acc > NIR] : 0.3349
##
##           Kappa : 0.3533
##
## Mcnemar's Test P-Value : 0.1443
##
##           Sensitivity : 0.9073
##           Specificity : 0.4146
##           Pos Pred Value : 0.8509
##           Neg Pred Value : 0.5484
##           Prevalence : 0.7865
##           Detection Rate : 0.7135
##           Detection Prevalence : 0.8385
##           Balanced Accuracy : 0.6610
##
##           'Positive' Class : 0
##

```



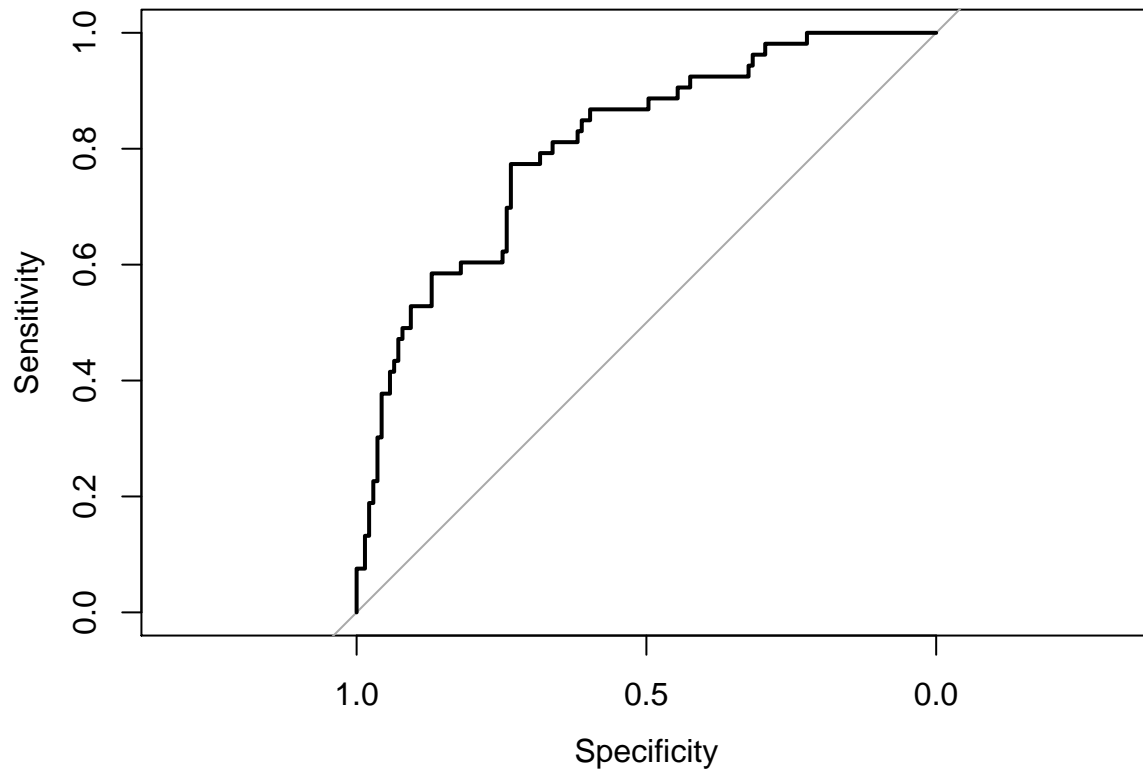
```
## [1] "AUC: 0.770150218058472"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 151 controls (dfPred_raw$class 0) < 41 cases (dfPred_raw$class 1).
## Area under the curve: 0.7702
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9851  -0.7354  -0.4298   0.5958   2.8302
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.035e+00  4.820e-01  -4.223 2.42e-05 ***
## INCOME         -8.984e-06  2.940e-06  -3.056 0.002245 **
## TRAVTIME        1.944e-02  6.952e-03   2.796 0.005172 **
## BLUEBOOK        3.777e-05  1.542e-05   2.450 0.014300 *
## TIF            -2.670e-02  2.527e-02  -1.057 0.290637
## OLDCLAIM        5.857e-06  1.454e-05   0.403 0.687148
## PARENT1_Yes     1.169e+00  3.039e-01   3.846 0.000120 ***
## SEX_z_F        -8.790e-01  3.442e-01  -2.554 0.010647 *
## JOB_Manager    -6.758e-01  3.665e-01  -1.844 0.065158 .
```



```

## CAR_USE_Commercial          5.187e-01  2.364e-01  2.195 0.028194 *
## CAR_TYPE_Pickup             9.034e-01  3.220e-01  2.806 0.005017 **
## CAR_TYPE_Sports.Car        1.671e+00  4.744e-01  3.523 0.000426 ***
## CAR_TYPE_z_SUV             1.502e+00  4.246e-01  3.537 0.000405 ***
## URBANICITY_z_Highly.Rural..Rural -2.391e+00  4.054e-01 -5.898 3.68e-09 ***
## HOME_VAL_NA                -3.718e-01  2.245e-01 -1.656 0.097632 .
## oldclaim_log               3.595e-02  3.212e-02  1.119 0.262955
## inter                      1.692e-02  4.365e-03  3.876 0.000106 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 717.37  on 624  degrees of freedom
## Residual deviance: 571.64  on 608  degrees of freedom
## AIC: 605.64
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 131  32
##           1   8  21
##
##           Accuracy : 0.7917
##           95% CI : (0.7273, 0.8468)
##    No Information Rate : 0.724
##    P-Value [Acc > NIR] : 0.0196445
##
##           Kappa : 0.3938
##
## Mcnemar's Test P-Value : 0.0002762
##
##           Sensitivity : 0.9424
##           Specificity : 0.3962
##    Pos Pred Value : 0.8037
##    Neg Pred Value : 0.7241
##           Prevalence : 0.7240
##    Detection Rate : 0.6823
##    Detection Prevalence : 0.8490
##    Balanced Accuracy : 0.6693
##
##           'Positive' Class : 0
##

```

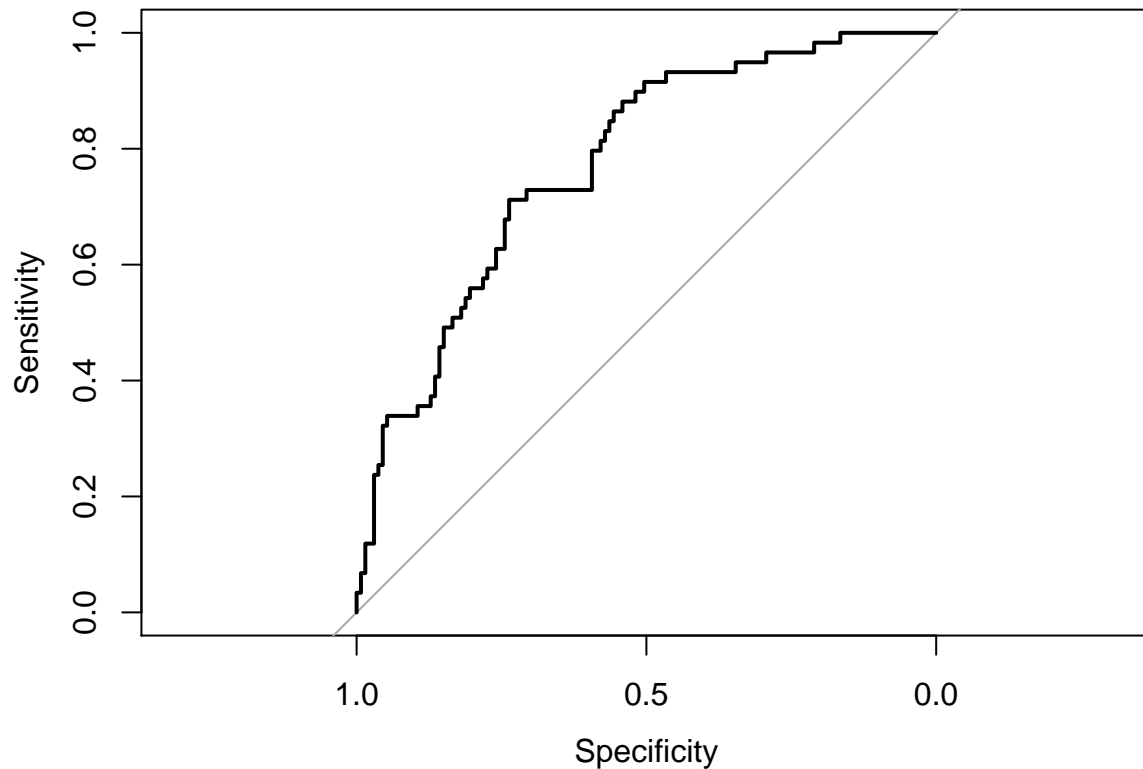


```
## [1] "AUC: 0.805755395683453"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 139 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.8058
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8304  -0.7211  -0.3972   0.4460   2.9410
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.317e+00  5.112e-01  -4.533 5.81e-06 ***
## INCOME          -7.833e-06  2.898e-06  -2.703 0.006868 **
## TRAVTIME         2.407e-02  7.427e-03   3.241 0.001193 **
## BLUEBOOK         3.488e-05  1.581e-05   2.206 0.027374 *
## TIF             -5.691e-02  2.719e-02  -2.093 0.036327 *
## OLDCLAIM         2.006e-05  1.495e-05   1.342 0.179496
## PARENT1_Yes      1.089e+00  3.031e-01   3.591 0.000329 ***
## SEX_z_F         -8.804e-01  3.623e-01  -2.430 0.015106 *
## JOB_Manager     -4.820e-01  3.650e-01  -1.321 0.186667
```

```

## CAR_USE_Commercial          6.154e-01  2.435e-01  2.527 0.011504 *
## CAR_TYPE_Pickup             1.316e+00  3.299e-01  3.988 6.67e-05 ***
## CAR_TYPE_Sports.Car        1.959e+00  5.029e-01  3.896 9.80e-05 ***
## CAR_TYPE_z_SUV             1.799e+00  4.479e-01  4.017 5.90e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.530e+00  4.274e-01 -5.920 3.22e-09 ***
## HOME_VAL_NA                -4.236e-01  2.245e-01 -1.887 0.059125 .
## oldclaim_log               1.645e-02  3.287e-02  0.500 0.616814
## inter                      1.652e-02  4.254e-03  3.884 0.000103 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 704.56  on 624  degrees of freedom
## Residual deviance: 551.66  on 608  degrees of freedom
## AIC: 585.66
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 120  39
##           1  13  20
##
##           Accuracy : 0.7292
##           95% CI : (0.6605, 0.7906)
##    No Information Rate : 0.6927
##    P-Value [Acc > NIR] : 0.1545173
##
##           Kappa : 0.2749
##
## Mcnemar's Test P-Value : 0.0005265
##
##           Sensitivity : 0.9023
##           Specificity : 0.3390
##    Pos Pred Value : 0.7547
##    Neg Pred Value : 0.6061
##           Prevalence : 0.6927
##    Detection Rate : 0.6250
##    Detection Prevalence : 0.8281
##    Balanced Accuracy : 0.6206
##
##           'Positive' Class : 0
##

```

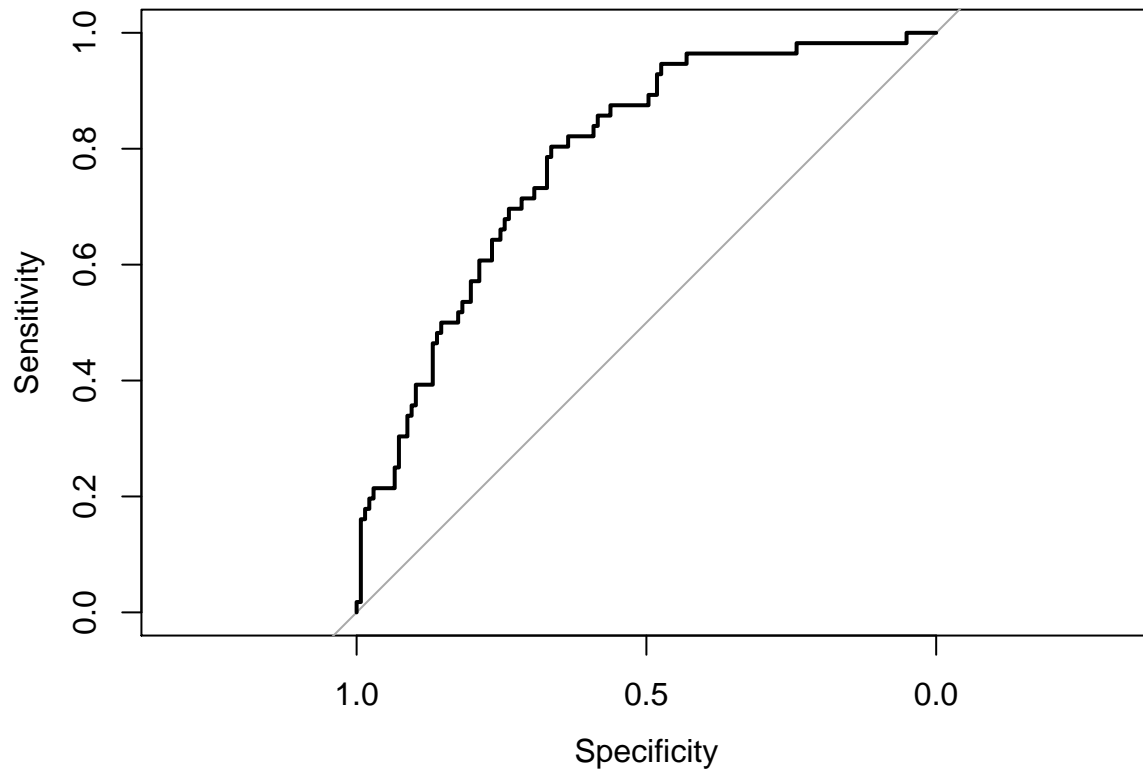


```
## [1] "AUC: 0.774181215751242"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 59 cases (dfPred_raw$class 1).
## Area under the curve: 0.7742
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8201  -0.7129  -0.4125   0.5272   2.7726
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.751e+00  5.021e-01  -3.487 0.000488 ***
## INCOME        -1.029e-05  2.972e-06  -3.460 0.000540 ***
## TRAVTIME       1.724e-02  6.974e-03   2.472 0.013420 *
## BLUEBOOK       2.301e-05  1.567e-05   1.469 0.141967
## TIF           -4.454e-02  2.716e-02  -1.640 0.101101
## OLDCLAIM       9.511e-06  1.439e-05   0.661 0.508669
## PARENT1_Yes    1.351e+00  3.348e-01   4.036 5.43e-05 ***
## SEX_z_F       -1.047e+00  3.632e-01  -2.882 0.003952 **
## JOB_Manager   -5.330e-01  3.500e-01  -1.523 0.127807
```

```

## CAR_USE_Commercial          4.627e-01  2.473e-01  1.871 0.061341 .
## CAR_TYPE_Pickup             8.603e-01  3.299e-01  2.607 0.009122 **
## CAR_TYPE_Sports.Car        1.861e+00  4.813e-01  3.866 0.000111 ***
## CAR_TYPE_z_SUV             1.760e+00  4.355e-01  4.041 5.33e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.345e+00  4.049e-01 -5.791 6.99e-09 ***
## HOME_VAL_NA                -3.352e-01  2.239e-01 -1.497 0.134381
## oldclaim_log               8.426e-02  3.161e-02  2.666 0.007673 **
## inter                      1.100e-02  4.192e-03  2.623 0.008706 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 710.45 on 623 degrees of freedom
## Residual deviance: 555.68 on 607 degrees of freedom
## AIC: 589.68
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 119  31
##           1  18  25
##
##           Accuracy : 0.7461
##           95% CI : (0.6786, 0.8059)
##           No Information Rate : 0.7098
##           P-Value [Acc > NIR] : 0.15107
##
##           Kappa : 0.3383
##
## Mcnemar's Test P-Value : 0.08648
##
##           Sensitivity : 0.8686
##           Specificity : 0.4464
##           Pos Pred Value : 0.7933
##           Neg Pred Value : 0.5814
##           Prevalence : 0.7098
##           Detection Rate : 0.6166
##           Detection Prevalence : 0.7772
##           Balanced Accuracy : 0.6575
##
##           'Positive' Class : 0
##

```

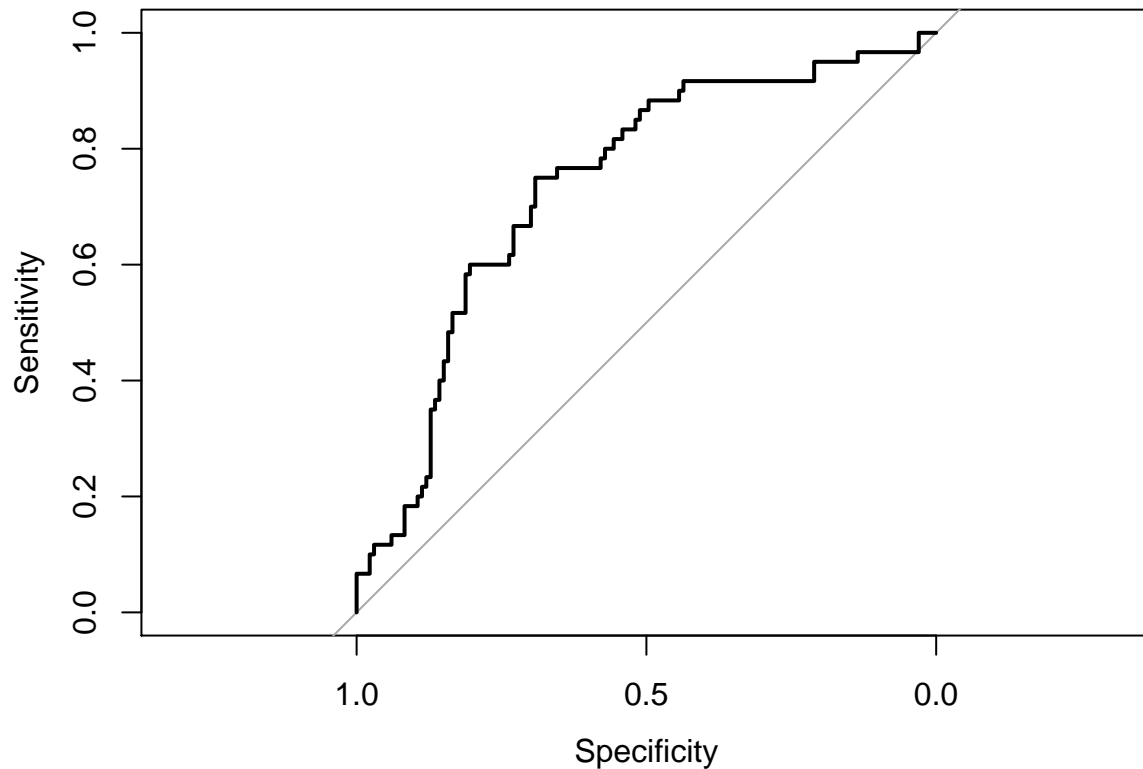


```
## [1] "AUC: 0.784801876955162"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 137 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7848
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.96167  -0.68678  -0.36298   0.00649   2.60181
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.786e+00  5.164e-01  -3.459 0.000543 ***
## INCOME         -1.069e-05  3.069e-06  -3.482 0.000498 ***
## TRAVTIME        2.017e-02  7.451e-03   2.707 0.006800 **
## BLUEBOOK        2.221e-05  1.579e-05   1.407 0.159457
## TIF            -4.222e-02  2.744e-02  -1.539 0.123874
## OLDCLAIM        1.968e-05  1.467e-05   1.341 0.179798
## PARENT1_Yes     1.025e+00  3.063e-01   3.345 0.000822 ***
## SEX_z_F        -7.015e-01  3.482e-01  -2.014 0.043958 *
## JOB_Manager    -4.847e-01  3.517e-01  -1.378 0.168130
```

```

## CAR_USE_Commercial          4.609e-01  2.523e-01  1.827 0.067762 .
## CAR_TYPE_Pickup             8.950e-01  3.284e-01  2.725 0.006422 **
## CAR_TYPE_Sports.Car        1.400e+00  4.728e-01  2.962 0.003060 **
## CAR_TYPE_z_SUV             1.486e+00  4.195e-01  3.543 0.000395 ***
## URBANICITY_z_Highly.Rural..Rural -2.824e+00  4.989e-01 -5.660 1.51e-08 ***
## HOME_VAL_NA                -6.009e-01  2.310e-01 -2.601 0.009296 **
## oldclaim_log               6.904e-02  3.214e-02  2.148 0.031678 *
## inter                      1.630e-02  4.455e-03  3.658 0.000254 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 701.79  on 623  degrees of freedom
## Residual deviance: 527.21  on 607  degrees of freedom
## AIC: 561.21
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  47
##           1  15  13
##
##           Accuracy : 0.6788
##           95% CI : (0.6079, 0.744)
##           No Information Rate : 0.6891
##           P-Value [Acc > NIR] : 0.6543
##
##           Kappa : 0.1217
##
## Mcnemar's Test P-Value : 8.251e-05
##
##           Sensitivity : 0.8872
##           Specificity : 0.2167
##           Pos Pred Value : 0.7152
##           Neg Pred Value : 0.4643
##           Prevalence : 0.6891
##           Detection Rate : 0.6114
##           Detection Prevalence : 0.8549
##           Balanced Accuracy : 0.5519
##
##           'Positive' Class : 0
##

```



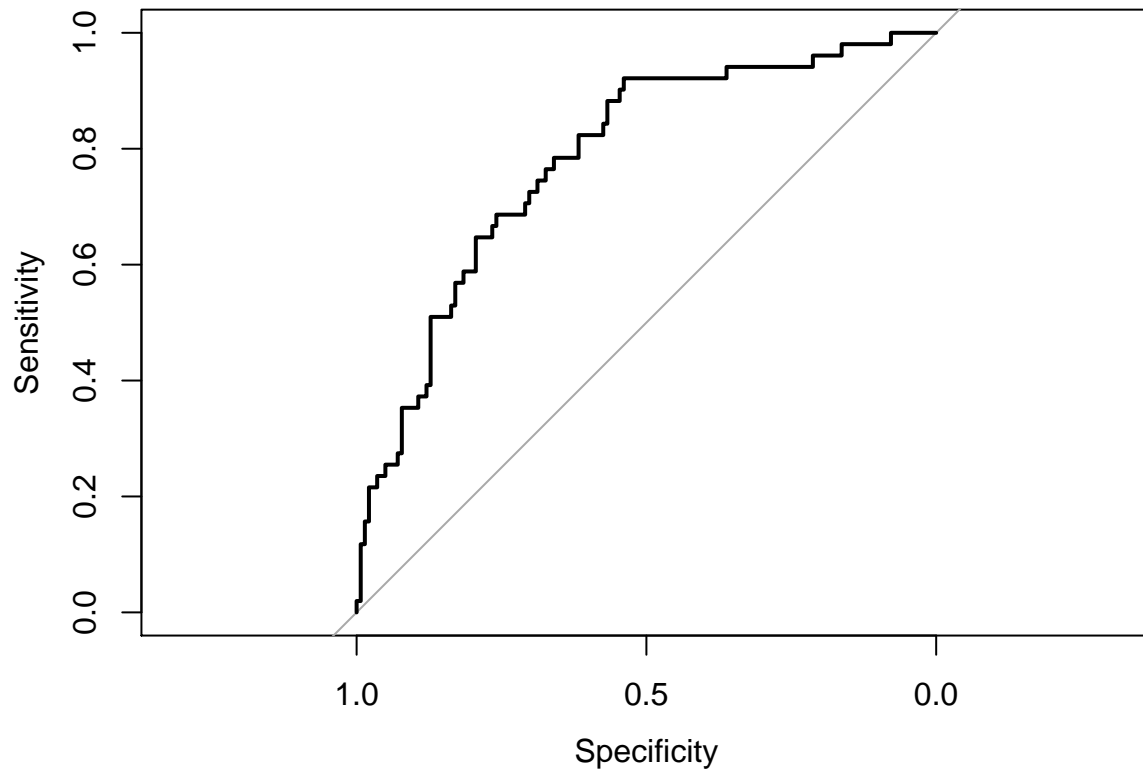
```
## [1] "AUC: 0.737218045112782"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 60 cases (dfPred_raw$class 1).
## Area under the curve: 0.7372
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7741  -0.7444  -0.4155   0.6441   2.9275
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.974e+00  4.951e-01  -3.988 6.66e-05 ***
## INCOME        -1.232e-05  3.137e-06  -3.929 8.54e-05 ***
## TRAVTIME       1.744e-02  7.221e-03   2.415 0.015754 *
## BLUEBOOK       3.841e-05  1.585e-05   2.424 0.015371 *
## TIF           -2.543e-02  2.612e-02  -0.973 0.330421
## OLDCLAIM       1.532e-05  1.388e-05   1.104 0.269731
## PARENT1_Yes     9.179e-01  2.932e-01   3.131 0.001744 **
## SEX_z_F        -6.907e-01  3.375e-01  -2.047 0.040684 *
## JOB_Manager    -3.260e-01  3.403e-01  -0.958 0.338023
```



```

## CAR_USE_Commercial          6.663e-01  2.424e-01  2.749 0.005976 **
## CAR_TYPE_Pickup             8.934e-01  3.174e-01  2.815 0.004881 **
## CAR_TYPE_Sports.Car        1.722e+00  4.667e-01  3.690 0.000225 ***
## CAR_TYPE_z_SUV             1.478e+00  4.098e-01  3.608 0.000309 ***
## URBANICITY_z_Highly.Rural..Rural -2.428e+00  4.147e-01 -5.855 4.77e-09 ***
## HOME_VAL_NA                -4.219e-01  2.219e-01 -1.901 0.057240 .
## oldclaim_log               4.067e-02  3.143e-02  1.294 0.195735
## inter                      1.357e-02  4.095e-03  3.313 0.000922 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 721.50  on 624  degrees of freedom
## Residual deviance: 565.64  on 608  degrees of freedom
## AIC: 599.64
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 130  36
##           1  11  15
##
##           Accuracy : 0.7552
##           95% CI : (0.6881, 0.8143)
##           No Information Rate : 0.7344
##           P-Value [Acc > NIR] : 0.2866885
##
##           Kappa : 0.2562
##
## Mcnemar's Test P-Value : 0.0004639
##
##           Sensitivity : 0.9220
##           Specificity : 0.2941
##           Pos Pred Value : 0.7831
##           Neg Pred Value : 0.5769
##           Prevalence : 0.7344
##           Detection Rate : 0.6771
##           Detection Prevalence : 0.8646
##           Balanced Accuracy : 0.6081
##
##           'Positive' Class : 0
##

```

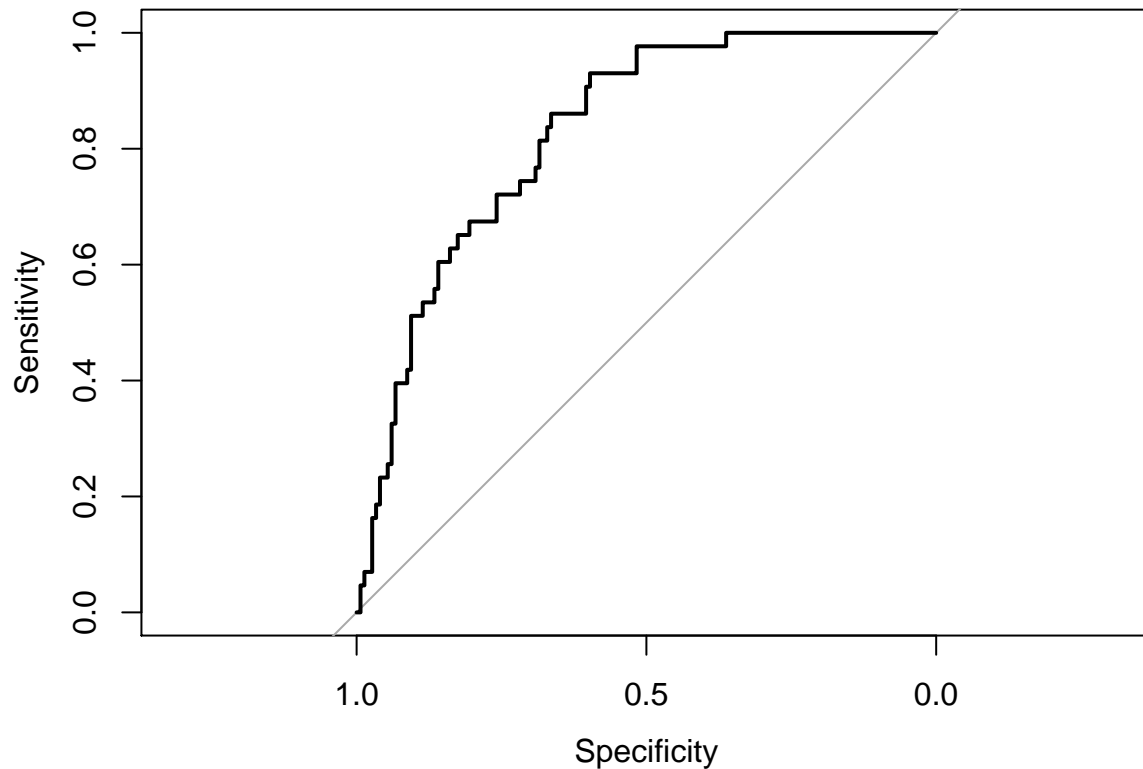


```
## [1] "AUC: 0.782505910165485"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 141 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.7825
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9971  -0.7637  -0.4424   0.7301   2.7428
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.039e+00  4.697e-01  -4.340 1.42e-05 ***
## INCOME          -9.739e-06  2.884e-06  -3.377 0.000732 ***
## TRAVTIME         2.486e-02  6.937e-03   3.584 0.000338 ***
## BLUEBOOK         3.107e-05  1.515e-05   2.051 0.040271 *
## TIF             -2.673e-02  2.502e-02  -1.069 0.285250
## OLDCLAIM         1.119e-05  1.326e-05   0.844 0.398612
## PARENT1_Yes      1.042e+00  3.025e-01   3.443 0.000576 ***
## SEX_z_F         -7.728e-01  3.350e-01  -2.307 0.021062 *
## JOB_Manager     -8.636e-01  3.618e-01  -2.387 0.017004 *
```

```

## CAR_USE_Commercial          3.083e-01  2.292e-01  1.345 0.178695
## CAR_TYPE_Pickup             7.290e-01  3.095e-01  2.355 0.018502 *
## CAR_TYPE_Sports.Car        1.596e+00  4.681e-01  3.408 0.000654 ***
## CAR_TYPE_z_SUV             1.315e+00  4.036e-01  3.258 0.001122 **
## URBANICITY_z_Highly.Rural..Rural -2.391e+00  3.959e-01 -6.039 1.56e-09 ***
## HOME_VAL_NA                -1.401e-01  2.201e-01 -0.637 0.524368
## oldclaim_log               5.449e-02  3.113e-02  1.750 0.080078 .
## inter                      1.312e-02  4.083e-03  3.214 0.001309 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 737.38  on 624  degrees of freedom
## Residual deviance: 589.86  on 608  degrees of freedom
## AIC: 623.86
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0   1
##           0 135  24
##           1  14  19
##
##           Accuracy : 0.8021
##           95% CI : (0.7386, 0.856)
##      No Information Rate : 0.776
##      P-Value [Acc > NIR] : 0.2199
##
##           Kappa : 0.3793
##
## Mcnemar's Test P-Value : 0.1443
##
##           Sensitivity : 0.9060
##           Specificity : 0.4419
##      Pos Pred Value : 0.8491
##      Neg Pred Value : 0.5758
##           Prevalence : 0.7760
##      Detection Rate : 0.7031
##      Detection Prevalence : 0.8281
##      Balanced Accuracy : 0.6740
##
##           'Positive' Class : 0
##

```

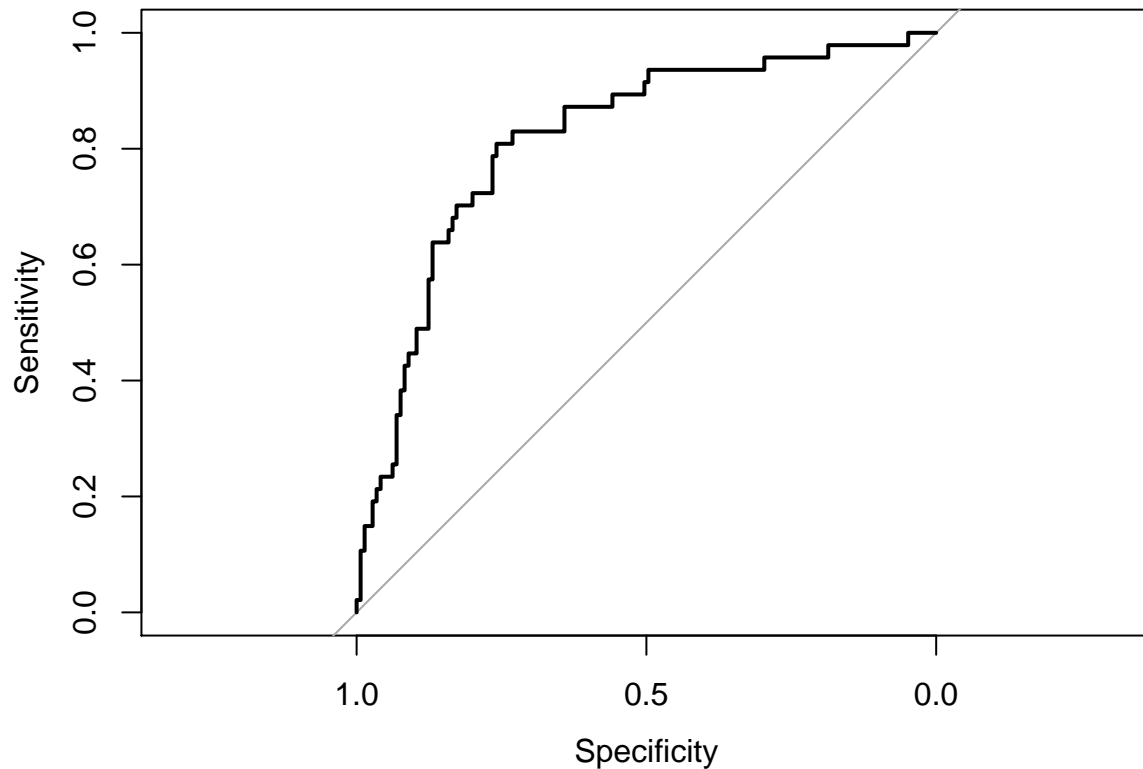


```
## [1] "AUC: 0.828468862181988"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 149 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.8285
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7275  -0.7541  -0.4122   0.6867   2.8282
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.837e+00  4.804e-01  -3.824 0.000131 ***
## INCOME        -9.320e-06  2.864e-06  -3.254 0.001136 **
## TRAVTIME       2.477e-02  6.971e-03   3.553 0.000380 ***
## BLUEBOOK       1.871e-05  1.532e-05   1.221 0.222125
## TIF           -4.435e-02  2.605e-02  -1.703 0.088608 .
## OLDCLAIM       9.769e-06  1.465e-05   0.667 0.505043
## PARENT1_Yes     9.505e-01  2.947e-01   3.226 0.001256 **
## SEX_z_F        -8.935e-01  3.558e-01  -2.511 0.012037 *
## JOB_Manager    -5.685e-01  3.663e-01  -1.552 0.120664
```

```

## CAR_USE_Commercial          5.368e-01  2.329e-01  2.305 0.021152 *
## CAR_TYPE_Pickup             1.014e+00  3.133e-01  3.235 0.001215 **
## CAR_TYPE_Sports.Car         1.834e+00  4.801e-01  3.820 0.000133 ***
## CAR_TYPE_z_SUV              1.723e+00  4.304e-01  4.004 6.23e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.376e+00  4.040e-01 -5.882 4.05e-09 ***
## HOME_VAL_NA                 -3.739e-01  2.191e-01 -1.707 0.087870 .
## oldclaim_log                2.602e-02  3.226e-02  0.807 0.419813
## inter                       1.419e-02  4.275e-03  3.320 0.000901 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 729.57 on 624 degrees of freedom
## Residual deviance: 577.53 on 608 degrees of freedom
## AIC: 611.53
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 136  35
##           1   9  12
##
##           Accuracy : 0.7708
##           95% CI : (0.7048, 0.8283)
##           No Information Rate : 0.7552
##           P-Value [Acc > NIR] : 0.341781
##
##           Kappa : 0.2377
##
## Mcnemar's Test P-Value : 0.000164
##
##           Sensitivity : 0.9379
##           Specificity : 0.2553
##           Pos Pred Value : 0.7953
##           Neg Pred Value : 0.5714
##           Prevalence : 0.7552
##           Detection Rate : 0.7083
##           Detection Prevalence : 0.8906
##           Balanced Accuracy : 0.5966
##
##           'Positive' Class : 0
##

```

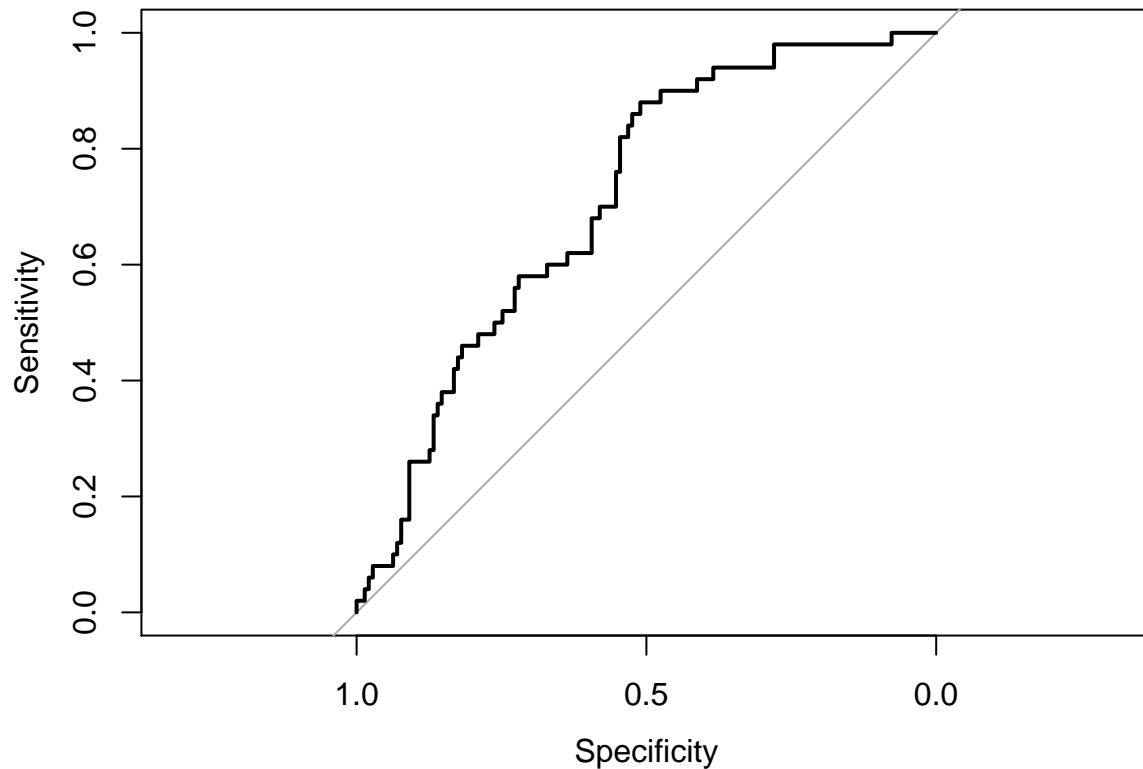


```
## [1] "AUC: 0.817754952311079"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 145 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.8178
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8761  -0.6908  -0.3769   0.5717   2.7192
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.457e+00  5.196e-01  -4.729 2.26e-06 ***
## INCOME        -1.114e-05  2.957e-06  -3.769 0.000164 ***
## TRAVTIME       2.503e-02  7.334e-03   3.413 0.000643 ***
## BLUEBOOK       5.747e-05  1.579e-05   3.640 0.000273 ***
## TIF           -5.808e-02  2.809e-02  -2.067 0.038697 *
## OLDCLAIM       2.375e-05  1.422e-05   1.670 0.094847 .
## PARENT1_Yes    1.247e+00  3.093e-01   4.031 5.56e-05 ***
## SEX_z_F       -1.158e+00  3.761e-01  -3.080 0.002071 **
## JOB_Manager   -7.229e-01  3.548e-01  -2.038 0.041593 *
```

```

## CAR_USE_Commercial          4.014e-01  2.468e-01  1.626 0.103859
## CAR_TYPE_Pickup             1.134e+00  3.349e-01  3.386 0.000710 ***
## CAR_TYPE_Sports.Car        2.214e+00  5.071e-01  4.366 1.27e-05 ***
## CAR_TYPE_z_SUV             1.861e+00  4.567e-01  4.076 4.58e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.617e+00  4.273e-01 -6.125 9.07e-10 ***
## HOME_VAL_NA                -3.425e-01  2.289e-01 -1.496 0.134542
## oldclaim_log               6.200e-02  3.258e-02  1.903 0.057001 .
## inter                      1.612e-02  4.217e-03  3.822 0.000132 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 722.92 on 623 degrees of freedom
## Residual deviance: 539.45 on 607 degrees of freedom
## AIC: 573.45
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 123  33
##           1  20  17
##
##           Accuracy : 0.7254
##           95% CI : (0.6567, 0.787)
##           No Information Rate : 0.7409
##           P-Value [Acc > NIR] : 0.72030
##
##           Kappa : 0.2186
##
## Mcnemar's Test P-Value : 0.09929
##
##           Sensitivity : 0.8601
##           Specificity : 0.3400
##           Pos Pred Value : 0.7885
##           Neg Pred Value : 0.4595
##           Prevalence : 0.7409
##           Detection Rate : 0.6373
##           Detection Prevalence : 0.8083
##           Balanced Accuracy : 0.6001
##
##           'Positive' Class : 0
##

```



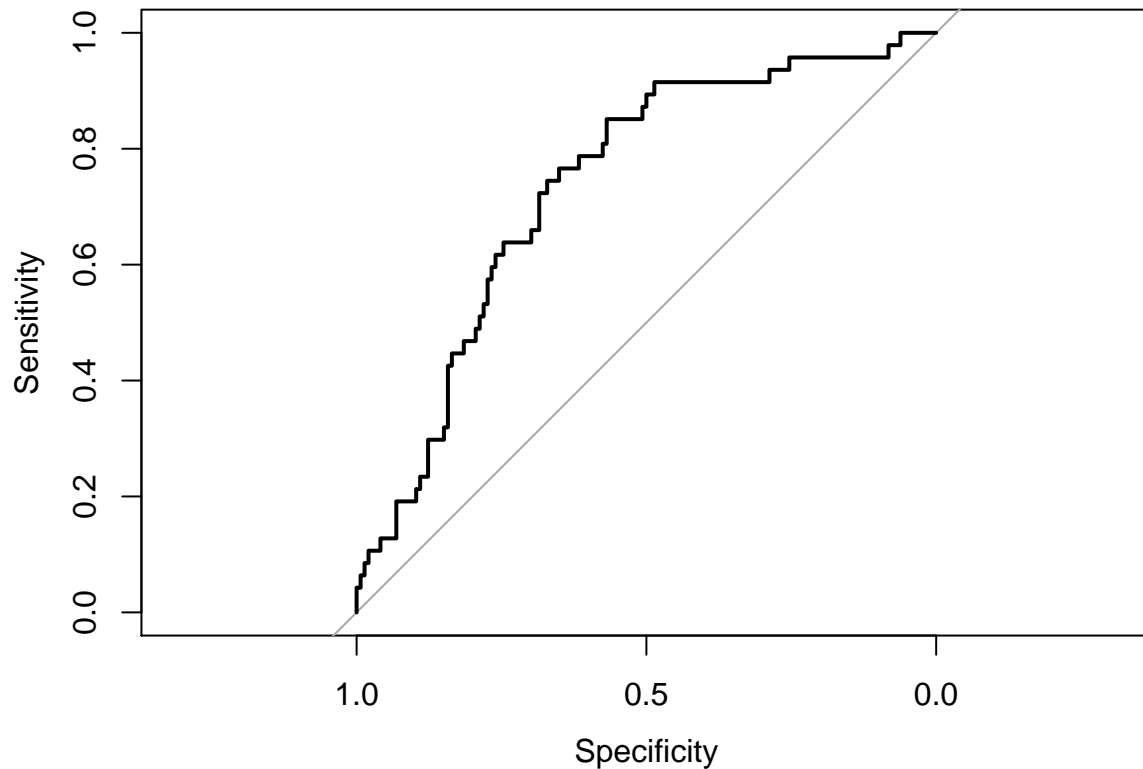
```
## [1] "AUC: 0.712167832167832"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 143 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7122
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9118  -0.7101  -0.3900   0.6310   2.5001
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.061e+00  5.043e-01  -4.088 4.36e-05 ***
## INCOME        -1.370e-05  3.030e-06  -4.520 6.18e-06 ***
## TRAVTIME       2.224e-02  7.174e-03   3.101 0.001930 **
## BLUEBOOK       4.449e-05  1.609e-05   2.766 0.005683 **
## TIF           -4.367e-02  2.713e-02  -1.610 0.107481
## OLDCLAIM       1.384e-05  1.419e-05   0.975 0.329343
## PARENT1_Yes    1.196e+00  2.966e-01   4.033 5.50e-05 ***
## SEX_z_F       -8.860e-01  3.455e-01  -2.564 0.010343 *
## JOB_Manager   -4.208e-01  3.442e-01  -1.223 0.221482
```



```

## CAR_USE_Commercial          2.710e-01  2.433e-01  1.114 0.265275
## CAR_TYPE_Pickup             1.050e+00  3.232e-01  3.247 0.001165 **
## CAR_TYPE_Sports.Car         1.736e+00  4.689e-01  3.701 0.000214 ***
## CAR_TYPE_z_SUV              1.685e+00  4.142e-01  4.068 4.75e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.509e+00  4.211e-01 -5.957 2.56e-09 ***
## HOME_VAL_NA                 -3.651e-01  2.225e-01 -1.640 0.100904
## oldclaim_log                8.144e-02  3.165e-02  2.573 0.010083 *
## inter                       1.150e-02  4.261e-03  2.698 0.006970 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 728.94  on 623  degrees of freedom
## Residual deviance: 553.37  on 607  degrees of freedom
## AIC: 587.37
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 129  36
##           1  17  11
##
##           Accuracy : 0.7254
##           95% CI : (0.6567, 0.787)
##    No Information Rate : 0.7565
##    P-Value [Acc > NIR] : 0.86168
##
##           Kappa : 0.1363
##
## Mcnemar's Test P-Value : 0.01342
##
##           Sensitivity : 0.8836
##           Specificity : 0.2340
##    Pos Pred Value : 0.7818
##    Neg Pred Value : 0.3929
##           Prevalence : 0.7565
##    Detection Rate : 0.6684
##    Detection Prevalence : 0.8549
##    Balanced Accuracy : 0.5588
##
##           'Positive' Class : 0
##

```

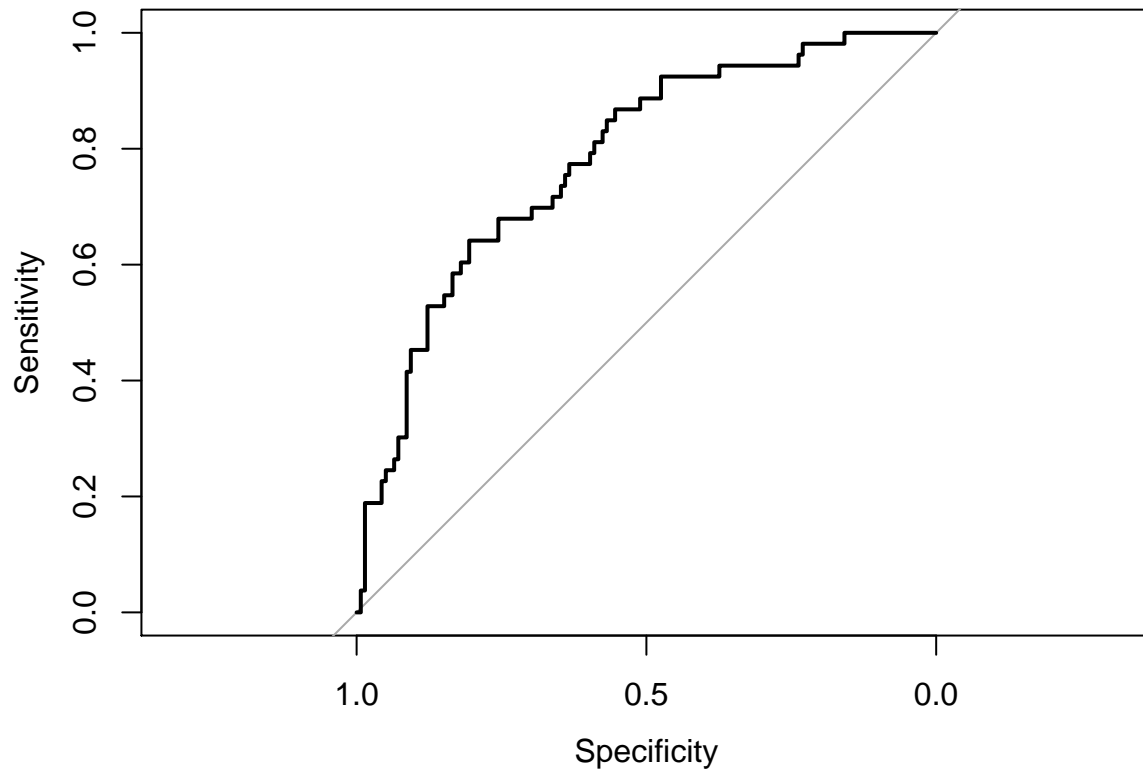


```
## [1] "AUC: 0.736374234916934"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.7364
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7846  -0.7347  -0.4176   0.5956   2.8603
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.946e+00  5.101e-01  -3.815 0.000136 ***
## INCOME        -1.014e-05  2.945e-06  -3.444 0.000573 ***
## TRAVTIME       2.097e-02  7.239e-03   2.897 0.003772 **
## BLUEBOOK       2.884e-05  1.551e-05   1.859 0.063080 .
## TIF           -6.591e-02  2.747e-02  -2.400 0.016412 *
## OLDCLAIM       2.289e-05  1.343e-05   1.704 0.088363 .
## PARENT1_Yes     8.124e-01  3.131e-01   2.595 0.009458 **
## SEX_z_F       -8.688e-01  3.690e-01  -2.354 0.018557 *
## JOB_Manager    -2.842e-01  3.271e-01  -0.869 0.384856
```

```

## CAR_USE_Commercial      8.191e-01  2.455e-01  3.337 0.000848 ***
## CAR_TYPE_Pickup         9.741e-01  3.310e-01  2.943 0.003249 **
## CAR_TYPE_Sports.Car     2.063e+00  4.936e-01  4.178 2.94e-05 ***
## CAR_TYPE_z_SUV          1.814e+00  4.456e-01  4.070 4.69e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.323e+00  4.048e-01 -5.739 9.54e-09 ***
## HOME_VAL_NA             -3.952e-01  2.222e-01 -1.778 0.075329 .
## oldclaim_log            1.558e-02  3.216e-02  0.485 0.628015
## inter                   1.365e-02  4.260e-03  3.204 0.001356 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 717.37 on 624 degrees of freedom
## Residual deviance: 565.69 on 608 degrees of freedom
## AIC: 599.69
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 126  30
##           1  13  23
##
##           Accuracy : 0.776
##           95% CI : (0.7104, 0.8329)
##           No Information Rate : 0.724
##           P-Value [Acc > NIR] : 0.06032
##
##           Kappa : 0.3779
##
## Mcnemar's Test P-Value : 0.01469
##
##           Sensitivity : 0.9065
##           Specificity : 0.4340
##           Pos Pred Value : 0.8077
##           Neg Pred Value : 0.6389
##           Prevalence : 0.7240
##           Detection Rate : 0.6562
##           Detection Prevalence : 0.8125
##           Balanced Accuracy : 0.6702
##
##           'Positive' Class : 0
##

```

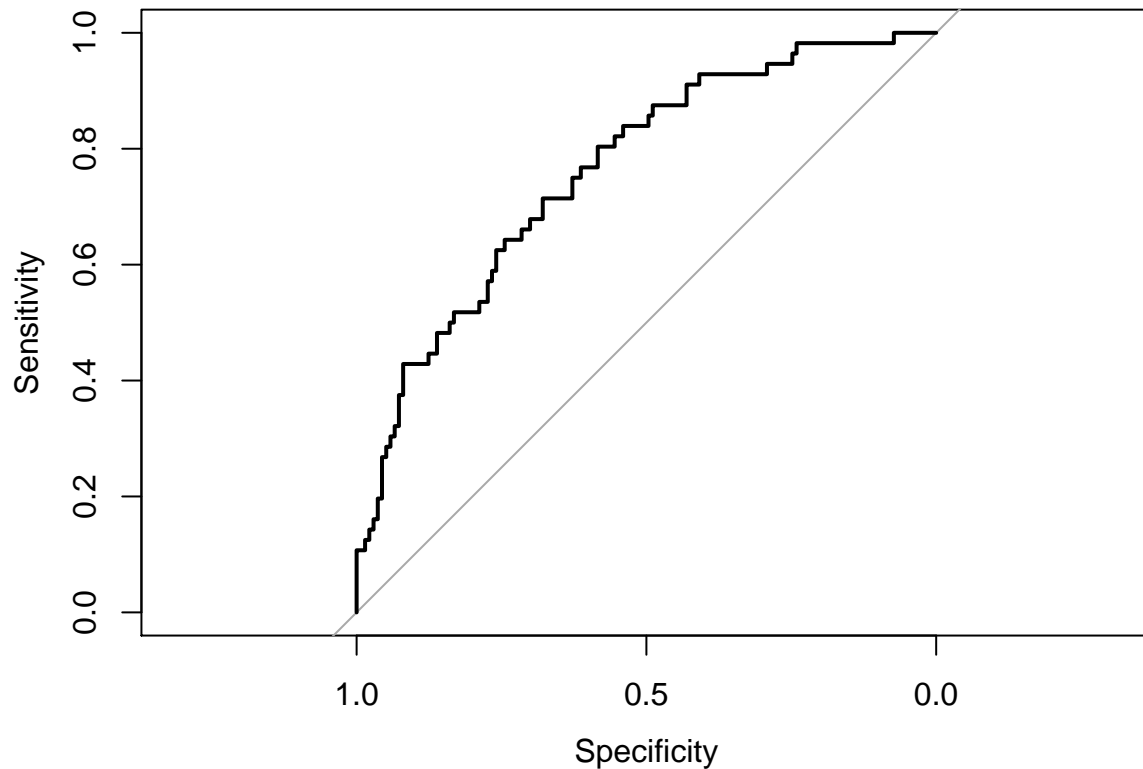


```
## [1] "AUC: 0.781457852585856"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 139 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.7815
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9057  -0.7077  -0.3949   0.5863   2.8050
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.975e+00  4.936e-01  -4.001 6.30e-05 ***
## INCOME        -9.304e-06  2.951e-06  -3.153 0.001614 **
## TRAVTIME       1.965e-02  7.058e-03   2.785 0.005359 **
## BLUEBOOK       3.167e-05  1.561e-05   2.029 0.042476 *
## TIF           -2.557e-02  2.623e-02  -0.975 0.329577
## OLDCLAIM       8.267e-06  1.469e-05   0.563 0.573642
## PARENT1_Yes    1.171e+00  3.095e-01   3.785 0.000154 ***
## SEX_z_F       -9.908e-01  3.565e-01  -2.779 0.005453 **
## JOB_Manager   -8.192e-01  3.737e-01  -2.192 0.028394 *
```

```

## CAR_USE_Commercial          3.928e-01  2.423e-01  1.621 0.104983
## CAR_TYPE_Pickup             8.144e-01  3.300e-01  2.468 0.013594 *
## CAR_TYPE_Sports.Car        1.661e+00  4.829e-01  3.439 0.000584 ***
## CAR_TYPE_z_SUV             1.529e+00  4.311e-01  3.546 0.000391 ***
## URBANICITY_z_Highly.Rural..Rural -2.621e+00  4.454e-01 -5.883 4.03e-09 ***
## HOME_VAL_NA                -3.711e-01  2.273e-01 -1.632 0.102609
## oldclaim_log               7.837e-02  3.173e-02  2.470 0.013504 *
## inter                      1.602e-02  4.216e-03  3.800 0.000145 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 710.45  on 623  degrees of freedom
## Residual deviance: 549.61  on 607  degrees of freedom
## AIC: 583.61
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 126  34
##           1  11  22
##
##           Accuracy : 0.7668
##           95% CI : (0.7007, 0.8246)
##      No Information Rate : 0.7098
##      P-Value [Acc > NIR] : 0.04572
##
##           Kappa : 0.3558
##
## Mcnemar's Test P-Value : 0.00104
##
##           Sensitivity : 0.9197
##           Specificity : 0.3929
##      Pos Pred Value : 0.7875
##      Neg Pred Value : 0.6667
##           Prevalence : 0.7098
##      Detection Rate : 0.6528
##      Detection Prevalence : 0.8290
##      Balanced Accuracy : 0.6563
##
##           'Positive' Class : 0
##

```

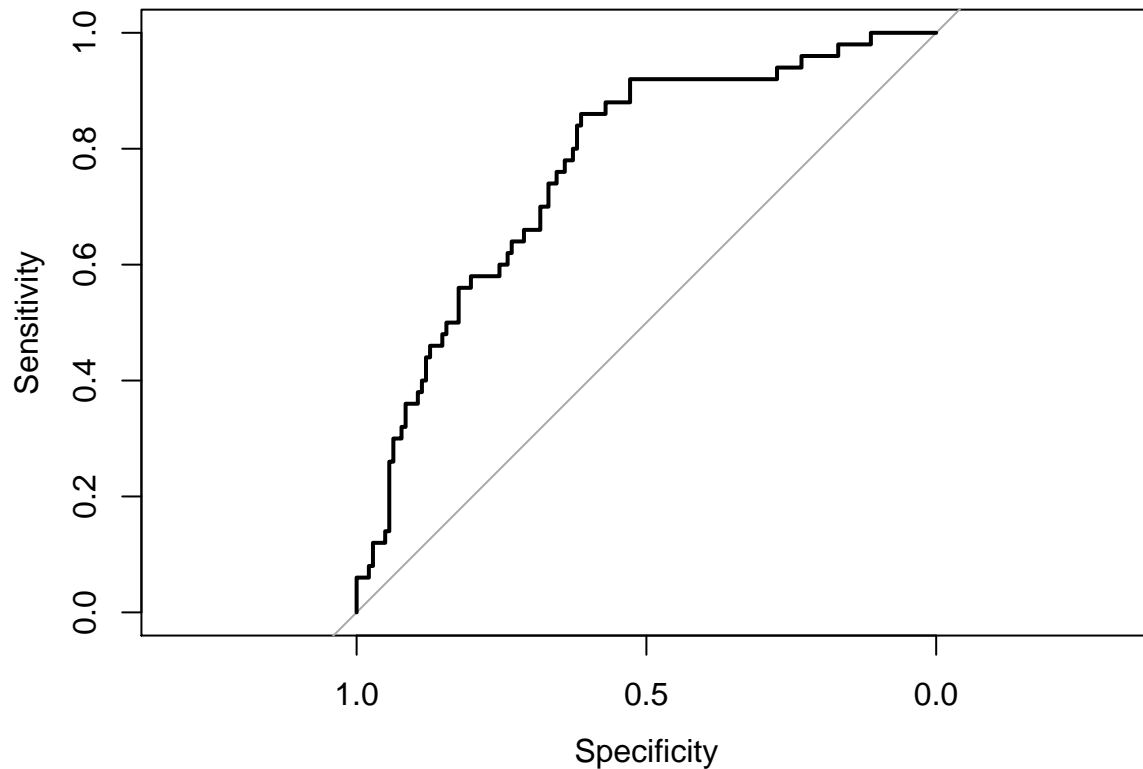


```
## [1] "AUC: 0.762513034410845"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 137 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7625
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8680  -0.7267  -0.4024   0.6410   3.0031
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.974e+00  4.835e-01  -4.082  4.46e-05 ***
## INCOME        -1.055e-05  3.084e-06  -3.420  0.000627 ***
## TRAVTIME       2.022e-02  6.931e-03   2.917  0.003529 **
## BLUEBOOK       2.207e-05  1.536e-05   1.437  0.150611
## TIF           -2.769e-03  2.605e-02  -0.106  0.915344
## OLDCLAIM       3.427e-06  1.452e-05   0.236  0.813389
## PARENT1_Yes    1.126e+00  3.052e-01   3.689  0.000225 ***
## SEX_z_F       -5.908e-01  3.375e-01  -1.750  0.080058 .
## JOB_Manager   -5.653e-01  3.442e-01  -1.642  0.100495
```

```

## CAR_USE_Commercial          5.641e-01  2.402e-01  2.349 0.018837 *
## CAR_TYPE_Pickup             6.885e-01  3.139e-01  2.193 0.028283 *
## CAR_TYPE_Sports.Car        1.458e+00  4.628e-01  3.150 0.001633 **
## CAR_TYPE_z_SUV             1.188e+00  4.063e-01  2.924 0.003454 **
## URBANICITY_z_Highly.Rural..Rural -2.693e+00  4.386e-01 -6.140 8.23e-10 ***
## HOME_VAL_NA                -3.083e-01  2.276e-01 -1.354 0.175583
## oldclaim_log               7.021e-02  3.122e-02  2.249 0.024534 *
## inter                      1.482e-02  4.002e-03  3.703 0.000213 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 723.54 on 624 degrees of freedom
## Residual deviance: 562.20 on 608 degrees of freedom
## AIC: 596.2
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 128  32
##           1  14  18
##
##           Accuracy : 0.7604
##           95% CI : (0.6937, 0.8189)
##           No Information Rate : 0.7396
##           P-Value [Acc > NIR] : 0.28551
##
##           Kappa : 0.2959
##
## Mcnemar's Test P-Value : 0.01219
##
##           Sensitivity : 0.9014
##           Specificity : 0.3600
##           Pos Pred Value : 0.8000
##           Neg Pred Value : 0.5625
##           Prevalence : 0.7396
##           Detection Rate : 0.6667
##           Detection Prevalence : 0.8333
##           Balanced Accuracy : 0.6307
##
##           'Positive' Class : 0
##

```



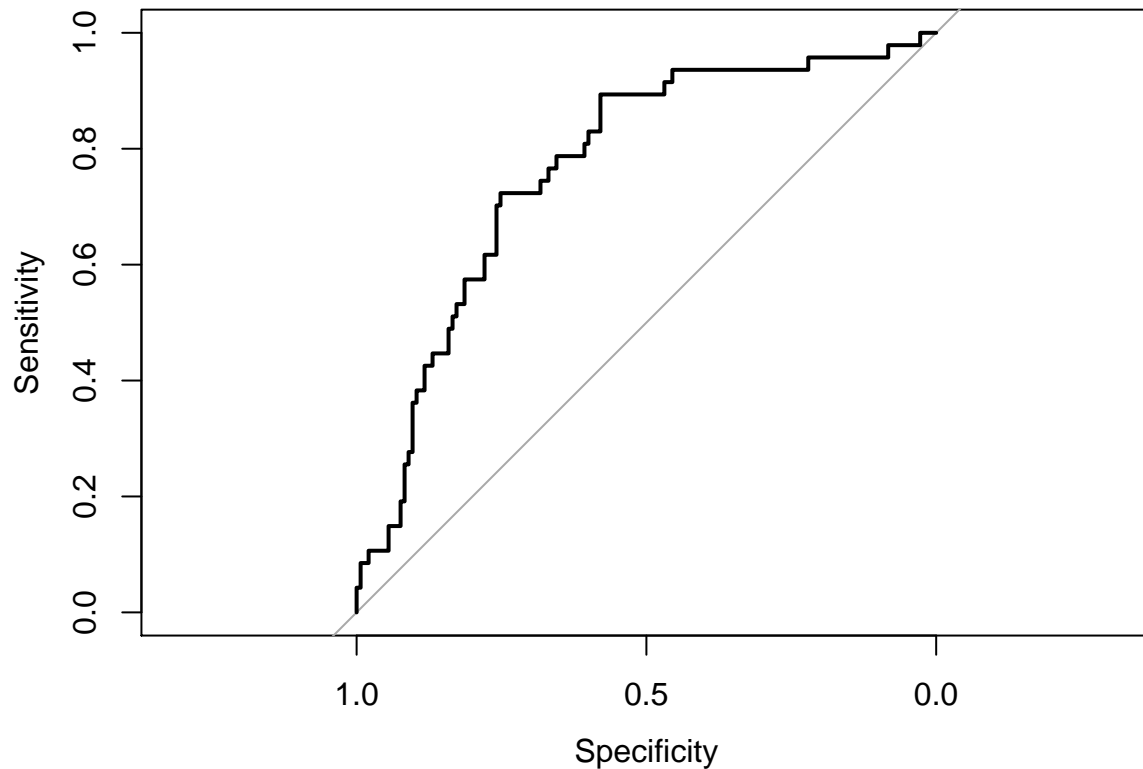
```
## [1] "AUC: 0.767605633802817"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 142 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7676
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0558  -0.7196  -0.3916   0.6867   2.9062
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.644e+00  4.814e-01  -3.415 0.000638 ***
## INCOME        -1.084e-05  3.057e-06  -3.544 0.000394 ***
## TRAVTIME       2.083e-02  6.899e-03   3.019 0.002535 **
## BLUEBOOK       1.520e-05  1.536e-05   0.990 0.322378
## TIF           -2.367e-02  2.657e-02  -0.891 0.372921
## OLDCLAIM       1.035e-05  1.437e-05   0.721 0.471069
## PARENT1_Yes    8.398e-01  2.978e-01   2.820 0.004795 **
## SEX_z_F       -8.348e-01  3.429e-01  -2.435 0.014908 *
## JOB_Manager   -8.370e-01  3.811e-01  -2.196 0.028064 *
```



```

## CAR_USE_Commercial          4.101e-01  2.412e-01  1.700 0.089145 .
## CAR_TYPE_Pickup             9.179e-01  3.129e-01  2.934 0.003348 **
## CAR_TYPE_Sports.Car        1.654e+00  4.653e-01  3.555 0.000378 ***
## CAR_TYPE_z_SUV             1.466e+00  4.121e-01  3.557 0.000375 ***
## URBANICITY_z_Highly.Rural..Rural -2.478e+00  4.285e-01 -5.783 7.33e-09 ***
## HOME_VAL_NA                -4.188e-01  2.264e-01 -1.850 0.064289 .
## oldclaim_log               6.125e-02  3.133e-02  1.955 0.050579 .
## inter                      1.671e-02  4.327e-03  3.862 0.000113 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 729.57  on 624  degrees of freedom
## Residual deviance: 562.75  on 608  degrees of freedom
## AIC: 596.75
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 134   39
##           1  11    8
##
##           Accuracy : 0.7396
##           95% CI : (0.6715, 0.8001)
##    No Information Rate : 0.7552
##    P-Value [Acc > NIR] : 0.7246395
##
##           Kappa : 0.1181
##
## Mcnemar's Test P-Value : 0.0001343
##
##           Sensitivity : 0.9241
##           Specificity : 0.1702
##    Pos Pred Value : 0.7746
##    Neg Pred Value : 0.4211
##           Prevalence : 0.7552
##    Detection Rate : 0.6979
##    Detection Prevalence : 0.9010
##    Balanced Accuracy : 0.5472
##
##           'Positive' Class : 0
##

```

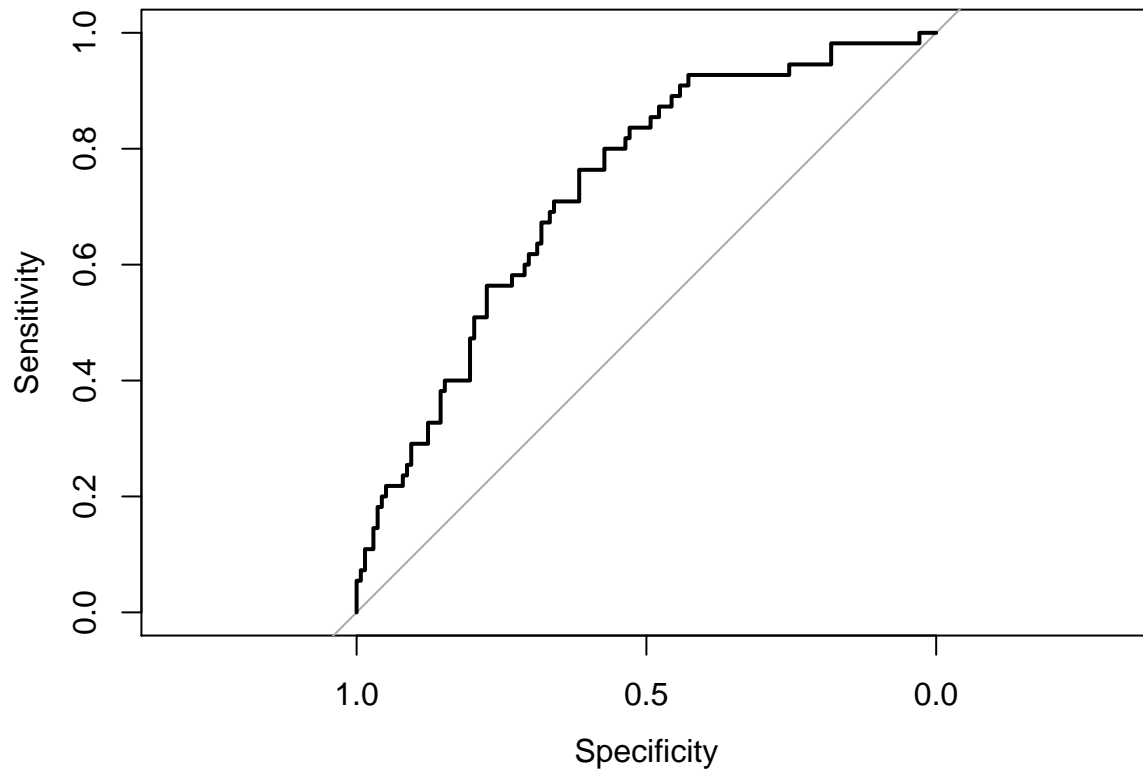


```
## [1] "AUC: 0.766691122523845"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 145 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.7667
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8298  -0.6907  -0.3782   0.5179   2.7956
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.748e+00  5.148e-01  -3.396 0.000684 ***
## INCOME        -1.357e-05  3.164e-06  -4.288 1.80e-05 ***
## TRAVTIME       1.248e-02  7.065e-03   1.766 0.077399 .
## BLUEBOOK       4.036e-05  1.593e-05   2.533 0.011308 *
## TIF           -3.377e-02  2.723e-02  -1.240 0.214978
## OLDCLAIM       1.189e-05  1.449e-05   0.820 0.411962
## PARENT1_Yes    1.135e+00  3.176e-01   3.572 0.000354 ***
## SEX_z_F       -1.035e+00  3.551e-01  -2.914 0.003569 **
## JOB_Manager   -3.744e-01  3.453e-01  -1.084 0.278216
```

```

## CAR_USE_Commercial          5.586e-01  2.498e-01  2.236 0.025353 *
## CAR_TYPE_Pickup             1.016e+00  3.324e-01  3.058 0.002232 **
## CAR_TYPE_Sports.Car         2.029e+00  4.787e-01  4.239 2.25e-05 ***
## CAR_TYPE_z_SUV              1.796e+00  4.325e-01  4.154 3.27e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.504e+00  4.201e-01 -5.959 2.53e-09 ***
## HOME_VAL_NA                 -4.848e-01  2.292e-01 -2.115 0.034417 *
## oldclaim_log                7.363e-02  3.209e-02  2.295 0.021751 *
## inter                       1.389e-02  4.241e-03  3.275 0.001058 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 712.57 on 623 degrees of freedom
## Residual deviance: 540.15 on 607 degrees of freedom
## AIC: 574.15
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 118  37
##           1  20  18
##
##           Accuracy : 0.7047
##           95% CI : (0.6349, 0.768)
##           No Information Rate : 0.715
##           P-Value [Acc > NIR] : 0.65841
##
##           Kappa : 0.201
##
## Mcnemar's Test P-Value : 0.03407
##
##           Sensitivity : 0.8551
##           Specificity : 0.3273
##           Pos Pred Value : 0.7613
##           Neg Pred Value : 0.4737
##           Prevalence : 0.7150
##           Detection Rate : 0.6114
##           Detection Prevalence : 0.8031
##           Balanced Accuracy : 0.5912
##
##           'Positive' Class : 0
##

```

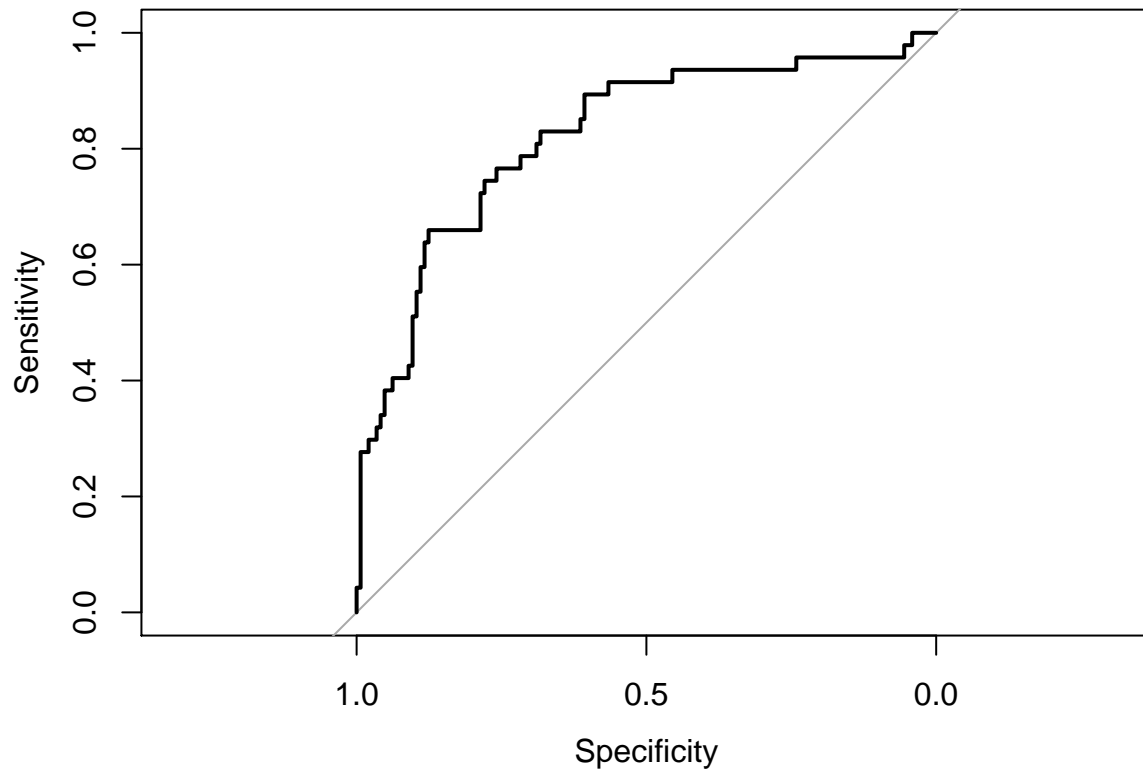


```
## [1] "AUC: 0.731093544137022"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 55 cases (dfPred_raw$class 1).
## Area under the curve: 0.7311
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7704  -0.7418  -0.4310   0.6999   2.7744
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.629e+00  4.849e-01  -3.359 0.000781 ***
## INCOME        -1.203e-05  3.002e-06  -4.008 6.13e-05 ***
## TRAVTIME       1.850e-02  6.871e-03   2.692 0.007104 **
## BLUEBOOK       1.949e-05  1.572e-05   1.240 0.215051
## TIF           -4.071e-02  2.584e-02  -1.576 0.115128
## OLDCLAIM       1.214e-05  1.329e-05   0.914 0.360972
## PARENT1_Yes     8.740e-01  3.004e-01   2.909 0.003621 **
## SEX_z_F       -8.128e-01  3.414e-01  -2.381 0.017273 *
## JOB_Manager    -1.975e-01  3.296e-01  -0.599 0.549043
```

```

## CAR_USE_Commercial          5.114e-01  2.368e-01  2.160 0.030796 *
## CAR_TYPE_Pickup             9.580e-01  3.174e-01  3.018 0.002542 **
## CAR_TYPE_Sports.Car        1.723e+00  4.586e-01  3.757 0.000172 ***
## CAR_TYPE_z_SUV             1.675e+00  4.108e-01  4.078 4.53e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.099e+00  3.830e-01 -5.481 4.23e-08 ***
## HOME_VAL_NA                 -3.405e-01  2.158e-01 -1.578 0.114599
## oldclaim_log                4.541e-02  3.134e-02  1.449 0.147322
## inter                       1.020e-02  4.396e-03  2.321 0.020310 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 729.57 on 624 degrees of freedom
## Residual deviance: 585.65 on 608 degrees of freedom
## AIC: 619.65
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 139  31
##           1   6  16
##
##           Accuracy : 0.8073
##           95% CI : (0.7443, 0.8605)
##           No Information Rate : 0.7552
##           P-Value [Acc > NIR] : 0.05265
##
##           Kappa : 0.3646
##
## Mcnemar's Test P-Value : 7.961e-05
##
##           Sensitivity : 0.9586
##           Specificity : 0.3404
##           Pos Pred Value : 0.8176
##           Neg Pred Value : 0.7273
##           Prevalence : 0.7552
##           Detection Rate : 0.7240
##           Detection Prevalence : 0.8854
##           Balanced Accuracy : 0.6495
##
##           'Positive' Class : 0
##

```

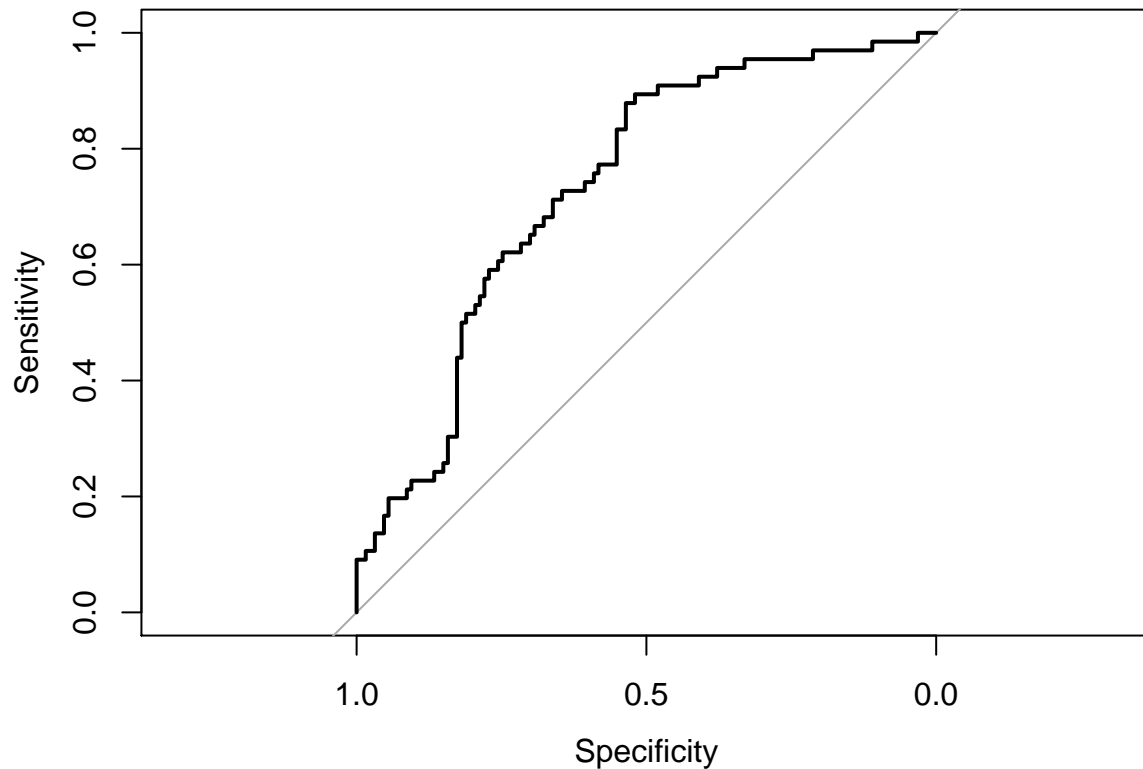


```
## [1] "AUC: 0.820836390315481"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 145 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.8208
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8737  -0.6782  -0.3810  -0.1054   2.8639
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.121e+00  5.248e-01  -4.042 5.30e-05 ***
## INCOME        -8.983e-06  3.014e-06  -2.980 0.002880 **
## TRAVTIME       2.061e-02  7.666e-03   2.688 0.007185 **
## BLUEBOOK       3.999e-05  1.594e-05   2.508 0.012135 *
## TIF           -5.135e-02  2.753e-02  -1.865 0.062171 .
## OLDCLAIM       1.724e-05  1.549e-05   1.113 0.265544
## PARENT1_Yes    1.242e+00  3.113e-01   3.989 6.63e-05 ***
## SEX_z_F       -8.108e-01  3.562e-01  -2.276 0.022841 *
## JOB_Manager   -5.650e-01  3.714e-01  -1.521 0.128163
```

```

## CAR_USE_Commercial          4.557e-01  2.521e-01  1.807 0.070718 .
## CAR_TYPE_Pickup             1.001e+00  3.331e-01  3.004 0.002661 **
## CAR_TYPE_Sports.Car         1.477e+00  4.899e-01  3.015 0.002568 **
## CAR_TYPE_z_SUV              1.574e+00  4.359e-01  3.611 0.000305 ***
## URBANICITY_z_Highly.Rural..Rural -2.645e+00  4.674e-01 -5.660 1.51e-08 ***
## HOME_VAL_NA                 -6.513e-01  2.328e-01 -2.798 0.005143 **
## oldclaim_log                6.452e-02  3.286e-02  1.963 0.049616 *
## inter                       1.594e-02  4.471e-03  3.565 0.000364 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 688.30 on 623 degrees of freedom
## Residual deviance: 522.19 on 607 degrees of freedom
## AIC: 556.19
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 111  51
##           1  16  15
##
##           Accuracy : 0.6528
##           95% CI : (0.5811, 0.7198)
##           No Information Rate : 0.658
##           P-Value [Acc > NIR] : 0.5929
##
##           Kappa : 0.1161
##
## Mcnemar's Test P-Value : 3.271e-05
##
##           Sensitivity : 0.8740
##           Specificity : 0.2273
##           Pos Pred Value : 0.6852
##           Neg Pred Value : 0.4839
##           Prevalence : 0.6580
##           Detection Rate : 0.5751
##           Detection Prevalence : 0.8394
##           Balanced Accuracy : 0.5506
##
##           'Positive' Class : 0
##

```



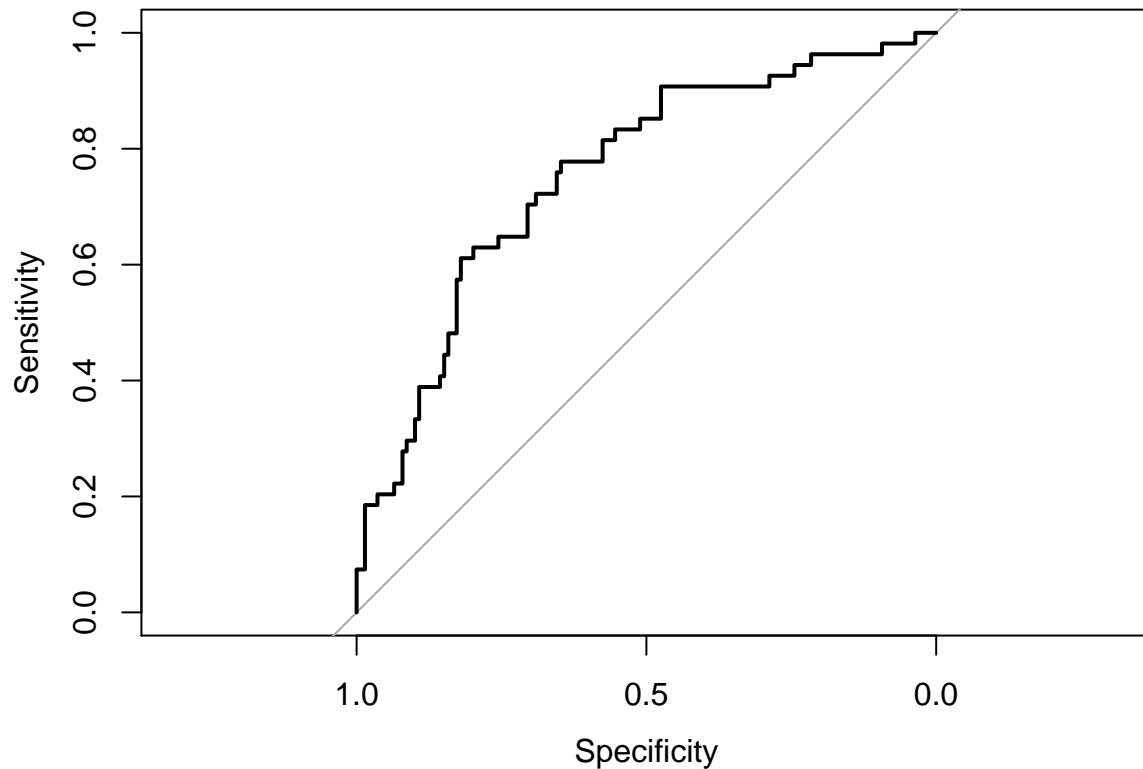
```
## [1] "AUC: 0.735504652827487"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 127 controls (dfPred_raw$class 0) < 66 cases (dfPred_raw$class 1).
## Area under the curve: 0.7355
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8406  -0.7213  -0.3948   0.5555   2.7724
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.997e+00  5.110e-01  -3.907  9.33e-05 ***
## INCOME         -1.263e-05  3.038e-06  -4.156  3.24e-05 ***
## TRAVTIME        1.889e-02  7.514e-03   2.514  0.011951 *
## BLUEBOOK        4.466e-05  1.593e-05   2.804  0.005044 **
## TIF            -4.397e-02  2.633e-02  -1.670  0.094891 .
## OLDCLAIM        1.984e-05  1.394e-05   1.423  0.154770
## PARENT1_Yes     9.999e-01  2.986e-01   3.349  0.000811 ***
## SEX_z_F        -6.724e-01  3.395e-01  -1.981  0.047611 *
## JOB_Manager    -3.064e-01  3.350e-01  -0.915  0.360363
```



```

## CAR_USE_Commercial          6.023e-01  2.441e-01  2.467 0.013626 *
## CAR_TYPE_Pickup             8.398e-01  3.252e-01  2.583 0.009807 **
## CAR_TYPE_Sports.Car        1.578e+00  4.713e-01  3.349 0.000811 ***
## CAR_TYPE_z_SUV             1.554e+00  4.117e-01  3.775 0.000160 ***
## URBANICITY_z_Highly.Rural..Rural -2.618e+00  4.403e-01 -5.944 2.78e-09 ***
## HOME_VAL_NA                -4.794e-01  2.235e-01 -2.145 0.031932 *
## oldclaim_log               4.615e-02  3.193e-02  1.446 0.148313
## inter                      1.347e-02  4.174e-03  3.228 0.001245 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 714.67 on 623 degrees of freedom
## Residual deviance: 550.25 on 607 degrees of freedom
## AIC: 584.25
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 127  38
##           1  12  16
##
##           Accuracy : 0.7409
##           95% CI : (0.6731, 0.8012)
##           No Information Rate : 0.7202
##           P-Value [Acc > NIR] : 0.290150
##
##           Kappa : 0.2462
##
## Mcnemar's Test P-Value : 0.000407
##
##           Sensitivity : 0.9137
##           Specificity : 0.2963
##           Pos Pred Value : 0.7697
##           Neg Pred Value : 0.5714
##           Prevalence : 0.7202
##           Detection Rate : 0.6580
##           Detection Prevalence : 0.8549
##           Balanced Accuracy : 0.6050
##
##           'Positive' Class : 0
##

```

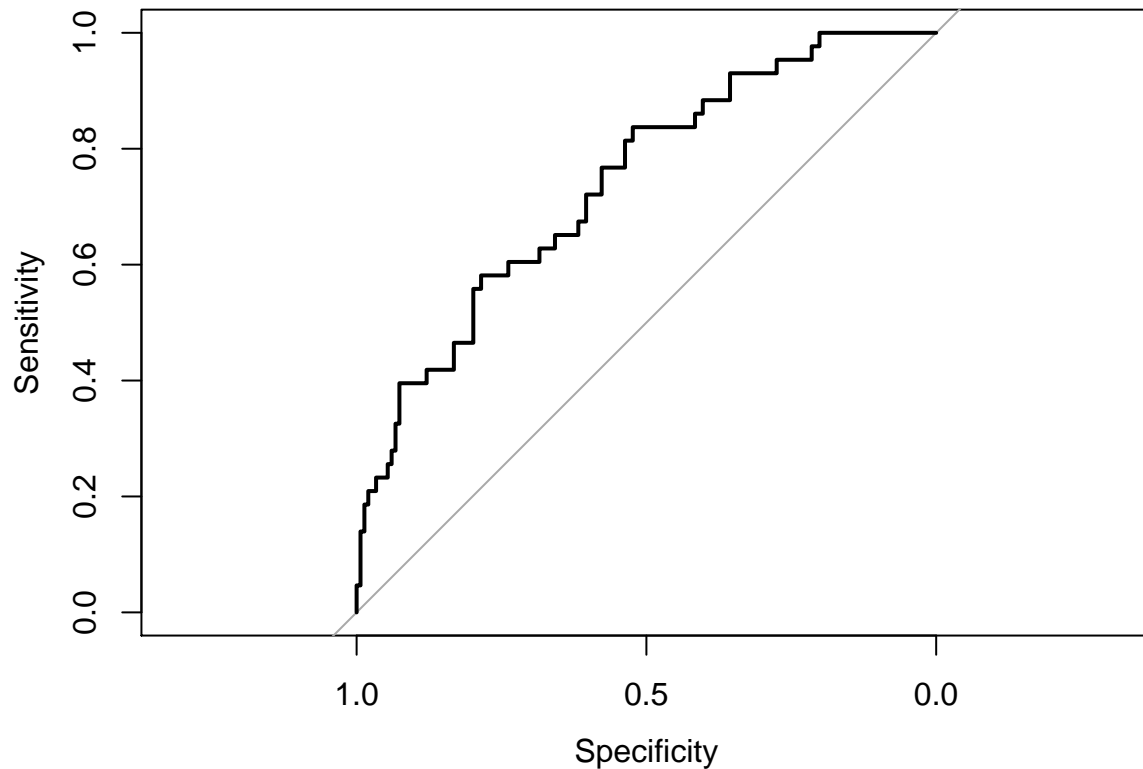


```
## [1] "AUC: 0.755795363709033"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 139 controls (dfPred_raw$class 0) < 54 cases (dfPred_raw$class 1).
## Area under the curve: 0.7558
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8592  -0.7407  -0.4086   0.7097   2.7412
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.067e+00  5.021e-01  -4.117 3.84e-05 ***
## INCOME        -1.019e-05  2.892e-06  -3.524 0.000425 ***
## TRAVTIME       2.118e-02  7.077e-03   2.993 0.002758 **
## BLUEBOOK       4.151e-05  1.548e-05   2.681 0.007337 **
## TIF           -7.511e-02  2.814e-02  -2.669 0.007599 **
## OLDCLAIM       2.115e-05  1.400e-05   1.511 0.130874
## PARENT1_Yes    1.044e+00  3.076e-01   3.393 0.000691 ***
## SEX_z_F       -1.253e+00  3.890e-01  -3.221 0.001275 **
## JOB_Manager   -6.401e-01  3.524e-01  -1.816 0.069322 .
```

```

## CAR_USE_Commercial          7.834e-01  2.492e-01  3.144 0.001668 **
## CAR_TYPE_Pickup             1.024e+00  3.307e-01  3.097 0.001958 **
## CAR_TYPE_Sports.Car         2.508e+00  5.188e-01  4.834 1.33e-06 ***
## CAR_TYPE_z_SUV              2.218e+00  4.718e-01  4.702 2.58e-06 ***
## URBANICITY_z_Highly.Rural..Rural -2.165e+00  3.783e-01 -5.722 1.05e-08 ***
## HOME_VAL_NA                 -3.228e-01  2.203e-01 -1.465 0.142900
## oldclaim_log                1.224e-02  3.184e-02  0.385 0.700582
## inter                       1.492e-02  4.102e-03  3.637 0.000276 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 737.38 on 624 degrees of freedom
## Residual deviance: 575.15 on 608 degrees of freedom
## AIC: 609.15
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 136  26
##           1  13  17
##
##           Accuracy : 0.7969
##           95% CI : (0.733, 0.8514)
##           No Information Rate : 0.776
##           P-Value [Acc > NIR] : 0.27573
##
##           Kappa : 0.3452
##
## Mcnemar's Test P-Value : 0.05466
##
##           Sensitivity : 0.9128
##           Specificity : 0.3953
##           Pos Pred Value : 0.8395
##           Neg Pred Value : 0.5667
##           Prevalence : 0.7760
##           Detection Rate : 0.7083
##           Detection Prevalence : 0.8438
##           Balanced Accuracy : 0.6541
##
##           'Positive' Class : 0
##

```

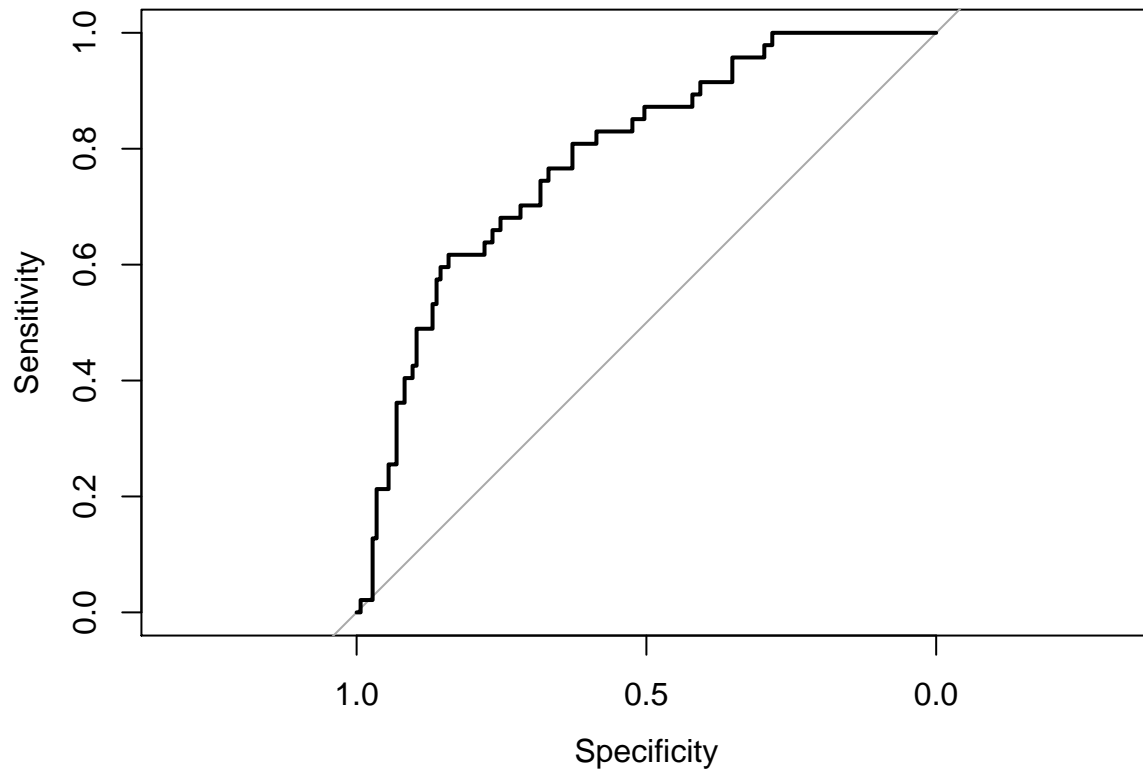


```
## [1] "AUC: 0.740127985016388"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 149 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7401
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9279  -0.7544  -0.4080   0.6614   2.8897
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.123e+00  4.862e-01  -4.367 1.26e-05 ***
## INCOME         -8.674e-06  2.861e-06  -3.032 0.002430 **
## TRAVTIME        2.471e-02  7.005e-03   3.527 0.000420 ***
## BLUEBOOK        3.158e-05  1.539e-05   2.052 0.040165 *
## TIF            -4.004e-02  2.622e-02  -1.527 0.126835
## OLDCLAIM        1.386e-05  1.380e-05   1.004 0.315222
## PARENT1_Yes     9.004e-01  3.028e-01   2.973 0.002947 **
## SEX_z_F        -9.923e-01  3.575e-01  -2.775 0.005512 **
## JOB_Manager    -8.809e-01  3.686e-01  -2.390 0.016848 *
```

```

## CAR_USE_Commercial          4.307e-01  2.366e-01  1.820 0.068761 .
## CAR_TYPE_Pickup             1.112e+00  3.257e-01  3.413 0.000643 ***
## CAR_TYPE_Sports.Car         1.961e+00  4.924e-01  3.984 6.78e-05 ***
## CAR_TYPE_z_SUV              1.666e+00  4.374e-01  3.809 0.000140 ***
## URBANICITY_z_Highly.Rural..Rural -2.679e+00  4.289e-01 -6.247 4.20e-10 ***
## HOME_VAL_NA                 -2.441e-01  2.239e-01 -1.090 0.275572
## oldclaim_log                3.812e-02  3.188e-02  1.195 0.231900
## inter                       2.044e-02  4.274e-03  4.782 1.74e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 729.57 on 624 degrees of freedom
## Residual deviance: 569.08 on 608 degrees of freedom
## AIC: 603.08
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 130  24
##           1  15  23
##
##           Accuracy : 0.7969
##           95% CI : (0.733, 0.8514)
##           No Information Rate : 0.7552
##           P-Value [Acc > NIR] : 0.1023
##
##           Kappa : 0.4126
##
## Mcnemar's Test P-Value : 0.2002
##
##           Sensitivity : 0.8966
##           Specificity : 0.4894
##           Pos Pred Value : 0.8442
##           Neg Pred Value : 0.6053
##           Prevalence : 0.7552
##           Detection Rate : 0.6771
##           Detection Prevalence : 0.8021
##           Balanced Accuracy : 0.6930
##
##           'Positive' Class : 0
##

```

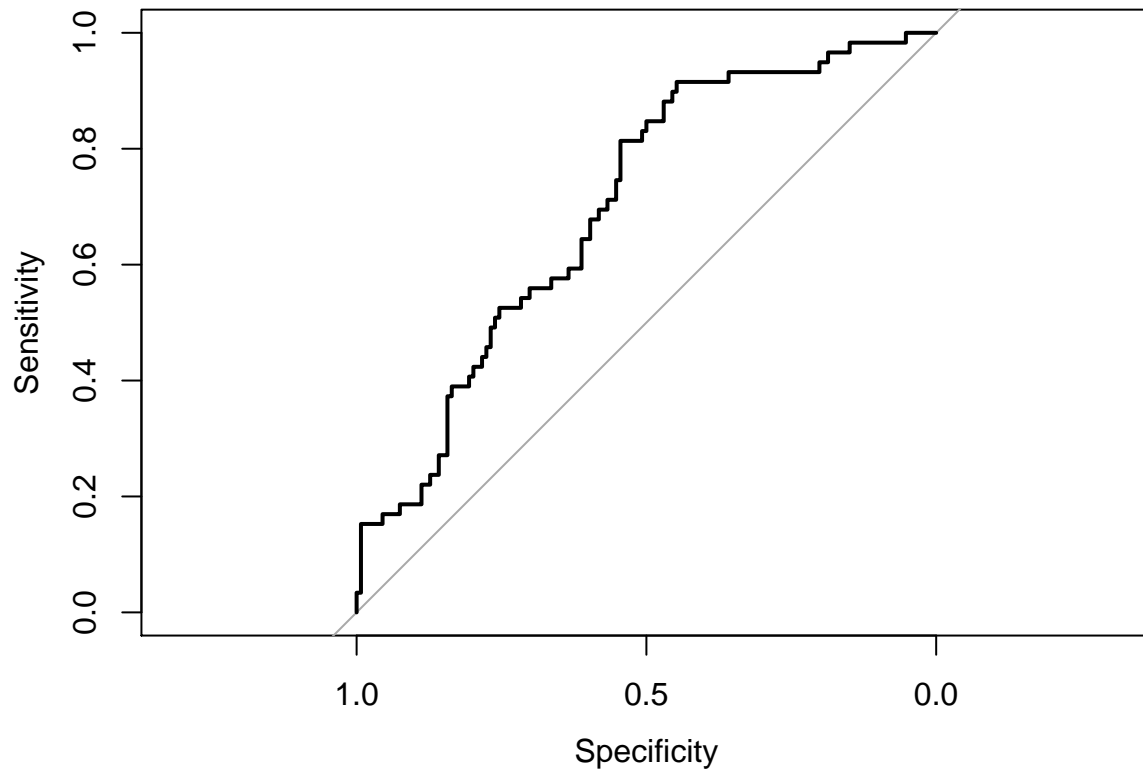


```
## [1] "AUC: 0.78459280997799"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 145 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.7846
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0018  -0.6866  -0.3577   0.2759   2.8618
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.488e+00  5.399e-01  -4.609 4.05e-06 ***
## INCOME         -1.198e-05  3.036e-06  -3.948 7.89e-05 ***
## TRAVTIME        2.197e-02  7.507e-03   2.926 0.003430 **
## BLUEBOOK        5.660e-05  1.618e-05   3.499 0.000467 ***
## TIF            -5.198e-02  2.843e-02  -1.828 0.067513 .
## OLDCLAIM       2.601e-05  1.478e-05   1.759 0.078532 .
## PARENT1_Yes    1.255e+00  3.091e-01   4.060 4.90e-05 ***
## SEX_z_F       -8.760e-01  3.548e-01  -2.469 0.013550 *
## JOB_Manager   -6.795e-01  3.640e-01  -1.867 0.061961 .
```

```

## CAR_USE_Commercial          4.420e-01  2.528e-01  1.749 0.080374 .
## CAR_TYPE_Pickup             1.271e+00  3.369e-01  3.771 0.000162 ***
## CAR_TYPE_Sports.Car         2.015e+00  5.014e-01  4.019 5.84e-05 ***
## CAR_TYPE_z_SUV              1.699e+00  4.357e-01  3.899 9.65e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.858e+00  4.699e-01 -6.083 1.18e-09 ***
## HOME_VAL_NA                 -3.423e-01  2.326e-01 -1.472 0.141082
## oldclaim_log                6.037e-02  3.256e-02  1.854 0.063732 .
## inter                       1.394e-02  4.142e-03  3.364 0.000767 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 703.98  on 623  degrees of freedom
## Residual deviance: 523.09  on 607  degrees of freedom
## AIC: 557.09
##
## Number of Fisher Scoring iterations: 6
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 115  44
##           1  19  15
##
##           Accuracy : 0.6736
##           95% CI : (0.6025, 0.7392)
##      No Information Rate : 0.6943
##      P-Value [Acc > NIR] : 0.760557
##
##           Kappa : 0.1276
##
## Mcnemar's Test P-Value : 0.002497
##
##           Sensitivity : 0.8582
##           Specificity : 0.2542
##      Pos Pred Value : 0.7233
##      Neg Pred Value : 0.4412
##           Prevalence : 0.6943
##      Detection Rate : 0.5959
##      Detection Prevalence : 0.8238
##      Balanced Accuracy : 0.5562
##
##           'Positive' Class : 0
##

```



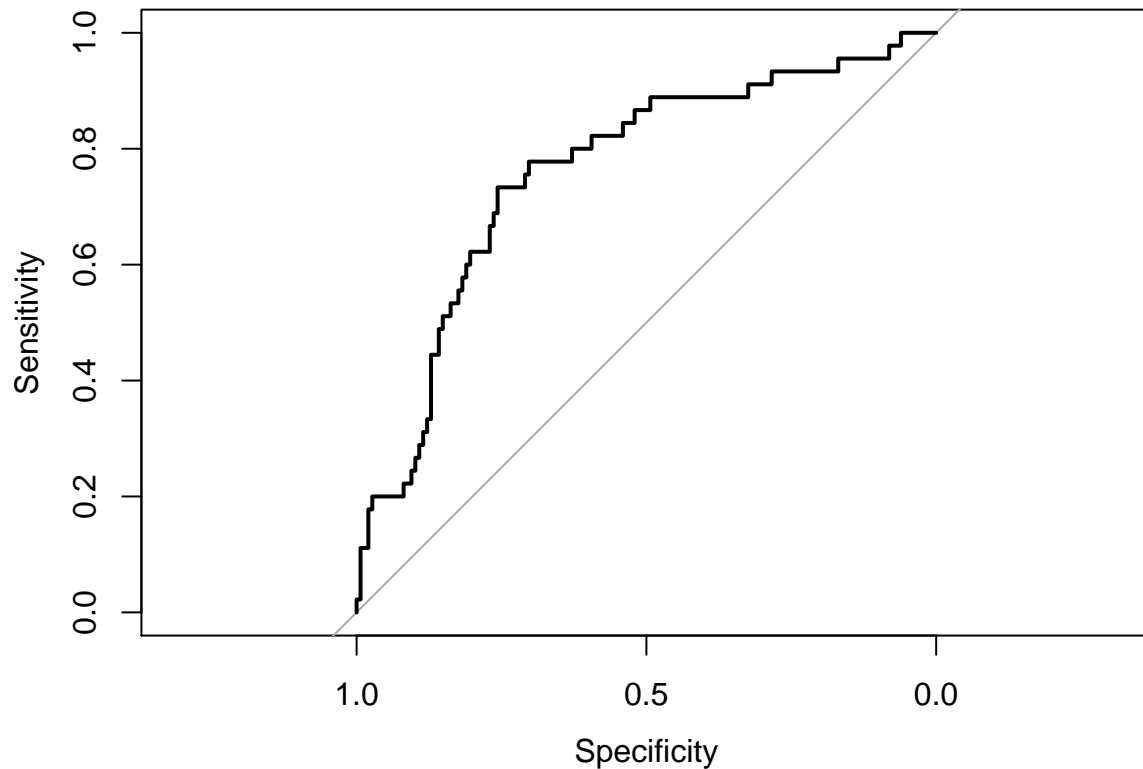
```
## [1] "AUC: 0.699721730331394"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 59 cases (dfPred_raw$class 1).
## Area under the curve: 0.6997
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8186  -0.7075  -0.3879   0.6927   2.5137
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.797e+00  4.918e-01  -3.653 0.000259 ***
## INCOME        -1.216e-05  2.959e-06  -4.110 3.96e-05 ***
## TRAVTIME       2.274e-02  7.272e-03   3.127 0.001769 **
## BLUEBOOK       2.789e-05  1.572e-05   1.775 0.075961 .
## TIF           -4.592e-02  2.656e-02  -1.729 0.083816 .
## OLDCLAIM       1.372e-05  1.401e-05   0.979 0.327484
## PARENT1_Yes     9.720e-01  2.943e-01   3.302 0.000959 ***
## SEX_z_F       -8.923e-01  3.515e-01  -2.539 0.011122 *
## JOB_Manager    -2.909e-01  3.341e-01  -0.871 0.383957
```



```

## CAR_USE_Commercial          4.320e-01  2.398e-01  1.801 0.071638 .
## CAR_TYPE_Pickup             9.533e-01  3.206e-01  2.974 0.002944 **
## CAR_TYPE_Sports.Car        1.708e+00  4.702e-01  3.633 0.000280 ***
## CAR_TYPE_z_SUV             1.791e+00  4.202e-01  4.262 2.02e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.566e+00  4.241e-01 -6.051 1.44e-09 ***
## HOME_VAL_NA                -4.596e-01  2.203e-01 -2.086 0.036948 *
## oldclaim_log               6.135e-02  3.175e-02  1.933 0.053290 .
## inter                      1.479e-02  4.249e-03  3.482 0.000498 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 732.87 on 623 degrees of freedom
## Residual deviance: 561.94 on 607 degrees of freedom
## AIC: 595.94
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 132  33
##           1  16  12
##
##           Accuracy : 0.7461
##           95% CI : (0.6786, 0.8059)
##           No Information Rate : 0.7668
##           P-Value [Acc > NIR] : 0.78003
##
##           Kappa : 0.1826
##
## Mcnemar's Test P-Value : 0.02227
##
##           Sensitivity : 0.8919
##           Specificity : 0.2667
##           Pos Pred Value : 0.8000
##           Neg Pred Value : 0.4286
##           Prevalence : 0.7668
##           Detection Rate : 0.6839
##           Detection Prevalence : 0.8549
##           Balanced Accuracy : 0.5793
##
##           'Positive' Class : 0
##

```

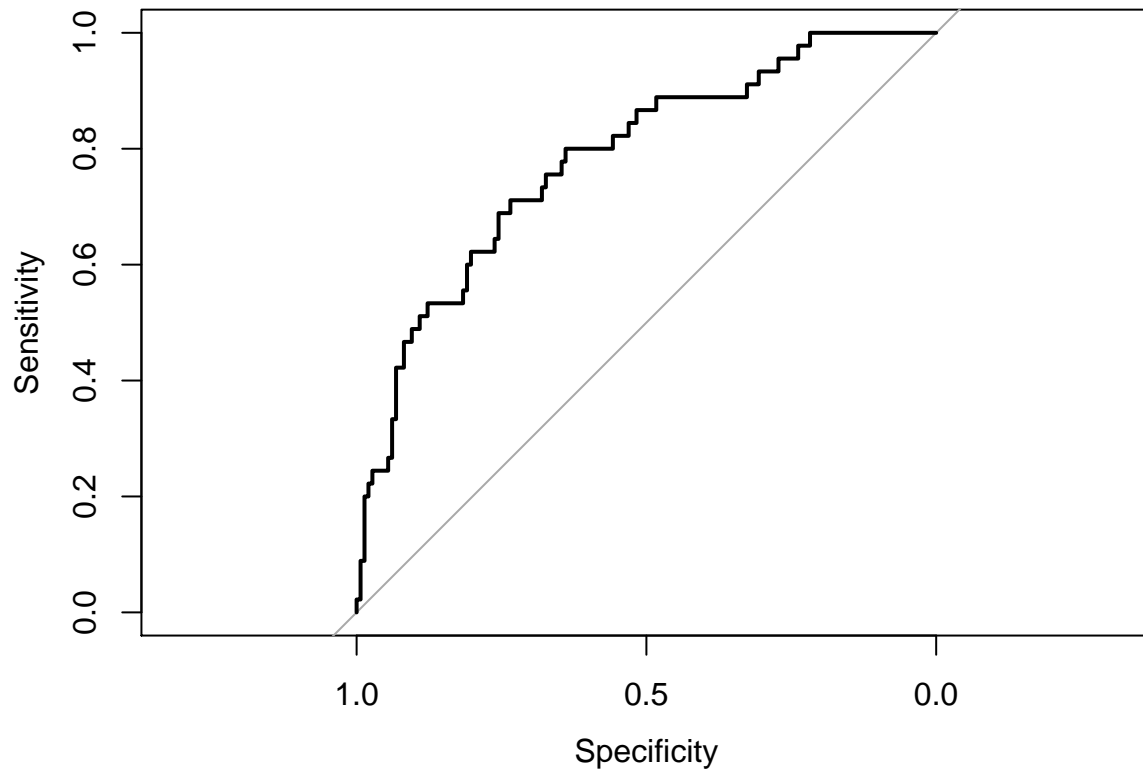


```
## [1] "AUC: 0.76021021021021"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 148 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7602
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8593  -0.7469  -0.4231   0.7090   2.8326
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.919e+00  4.828e-01  -3.974 7.07e-05 ***
## INCOME        -9.663e-06  2.872e-06  -3.365 0.000765 ***
## TRAVTIME       2.150e-02  7.050e-03   3.049 0.002292 **
## BLUEBOOK       3.516e-05  1.532e-05   2.296 0.021672 *
## TIF           -4.940e-02  2.598e-02  -1.901 0.057250 .
## OLDCLAIM       6.297e-06  1.462e-05   0.431 0.666569
## PARENT1_Yes    1.003e+00  3.010e-01   3.330 0.000867 ***
## SEX_z_F       -9.574e-01  3.580e-01  -2.674 0.007487 **
## JOB_Manager   -4.785e-01  3.476e-01  -1.377 0.168599
```

```

## CAR_USE_Commercial          6.924e-01  2.386e-01  2.903 0.003701 **
## CAR_TYPE_Pickup             9.126e-01  3.234e-01  2.822 0.004779 **
## CAR_TYPE_Sports.Car        1.897e+00  4.870e-01  3.896 9.78e-05 ***
## CAR_TYPE_z_SUV             1.884e+00  4.391e-01  4.292 1.77e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.406e+00  3.958e-01 -6.080 1.21e-09 ***
## HOME_VAL_NA                -4.406e-01  2.214e-01 -1.990 0.046610 *
## oldclaim_log               1.753e-02  3.221e-02  0.544 0.586317
## inter                      1.790e-02  4.342e-03  4.122 3.76e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.51 on 624 degrees of freedom
## Residual deviance: 579.43 on 608 degrees of freedom
## AIC: 613.43
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 135  26
##           1  12  19
##
##           Accuracy : 0.8021
##           95% CI : (0.7386, 0.856)
##           No Information Rate : 0.7656
##           P-Value [Acc > NIR] : 0.13319
##
##           Kappa : 0.3818
##
## Mcnemar's Test P-Value : 0.03496
##
##           Sensitivity : 0.9184
##           Specificity : 0.4222
##           Pos Pred Value : 0.8385
##           Neg Pred Value : 0.6129
##           Prevalence : 0.7656
##           Detection Rate : 0.7031
##           Detection Prevalence : 0.8385
##           Balanced Accuracy : 0.6703
##
##           'Positive' Class : 0
##

```

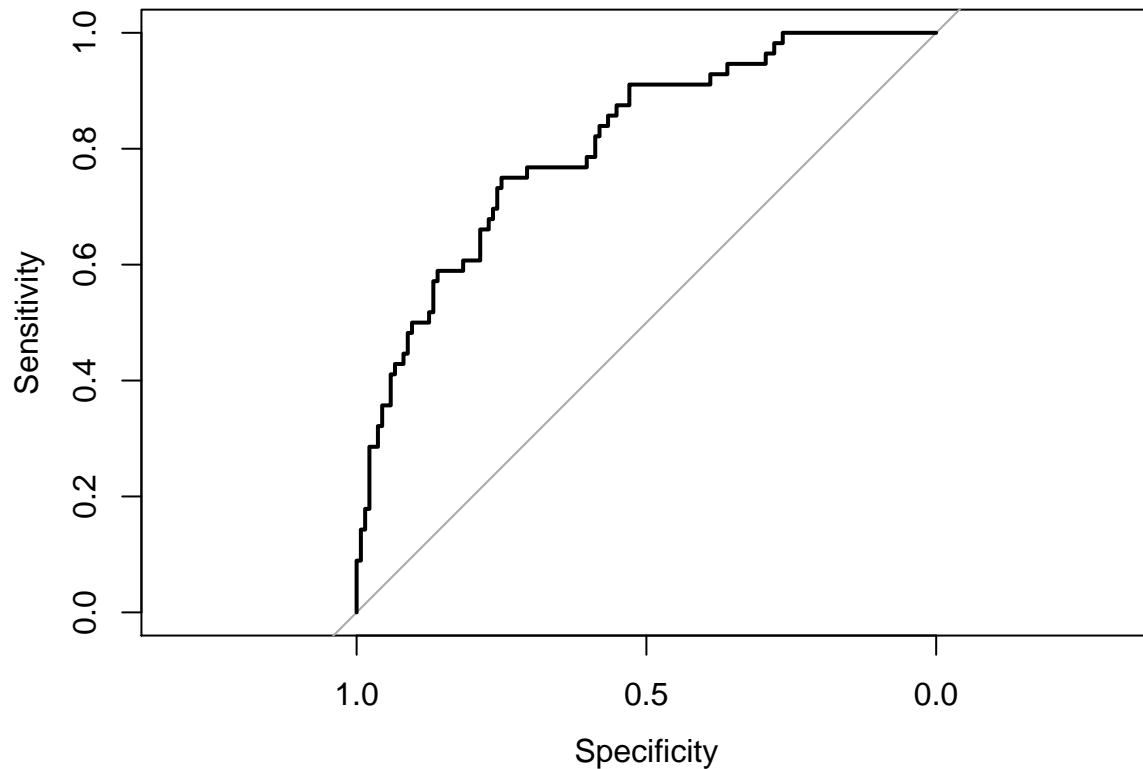


```
## [1] "AUC: 0.782161753590325"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 147 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7822
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9342  -0.7292  -0.4154   0.5984   2.8214
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.974e+00  4.846e-01  -4.073  4.64e-05 ***
## INCOME        -7.392e-06  2.853e-06  -2.591  0.009577 **
## TRAVTIME       1.885e-02  6.908e-03   2.728  0.006370 **
## BLUEBOOK       2.688e-05  1.577e-05   1.705  0.088228 .
## TIF           -3.586e-02  2.600e-02  -1.379  0.167833
## OLDCLAIM      -2.454e-06  1.610e-05  -0.152  0.878847
## PARENT1_Yes    1.179e+00  3.056e-01   3.858  0.000114 ***
## SEX_z_F       -1.013e+00  3.521e-01  -2.878  0.004004 **
## JOB_Manager   -8.631e-01  4.074e-01  -2.119  0.034108 *
```

```

## CAR_USE_Commercial          4.482e-01  2.390e-01  1.876 0.060703 .
## CAR_TYPE_Pickup             1.145e+00  3.248e-01  3.526 0.000422 ***
## CAR_TYPE_Sports.Car         1.778e+00  4.891e-01  3.635 0.000278 ***
## CAR_TYPE_z_SUV              1.767e+00  4.332e-01  4.080 4.51e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.327e+00  4.027e-01 -5.780 7.47e-09 ***
## HOME_VAL_NA                 -3.357e-01  2.219e-01 -1.513 0.130314
## oldclaim_log                5.042e-02  3.197e-02  1.577 0.114743
## inter                       1.496e-02  4.192e-03  3.570 0.000357 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 711.04  on 624  degrees of freedom
## Residual deviance: 569.22  on 608  degrees of freedom
## AIC: 603.22
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 130  36
##           1   6  20
##
##           Accuracy : 0.7812
##           95% CI : (0.716, 0.8376)
##      No Information Rate : 0.7083
##      P-Value [Acc > NIR] : 0.01426
##
##           Kappa : 0.3716
##
## Mcnemar's Test P-Value : 7.648e-06
##
##           Sensitivity : 0.9559
##           Specificity : 0.3571
##      Pos Pred Value : 0.7831
##      Neg Pred Value : 0.7692
##           Prevalence : 0.7083
##      Detection Rate : 0.6771
##      Detection Prevalence : 0.8646
##      Balanced Accuracy : 0.6565
##
##           'Positive' Class : 0
##

```

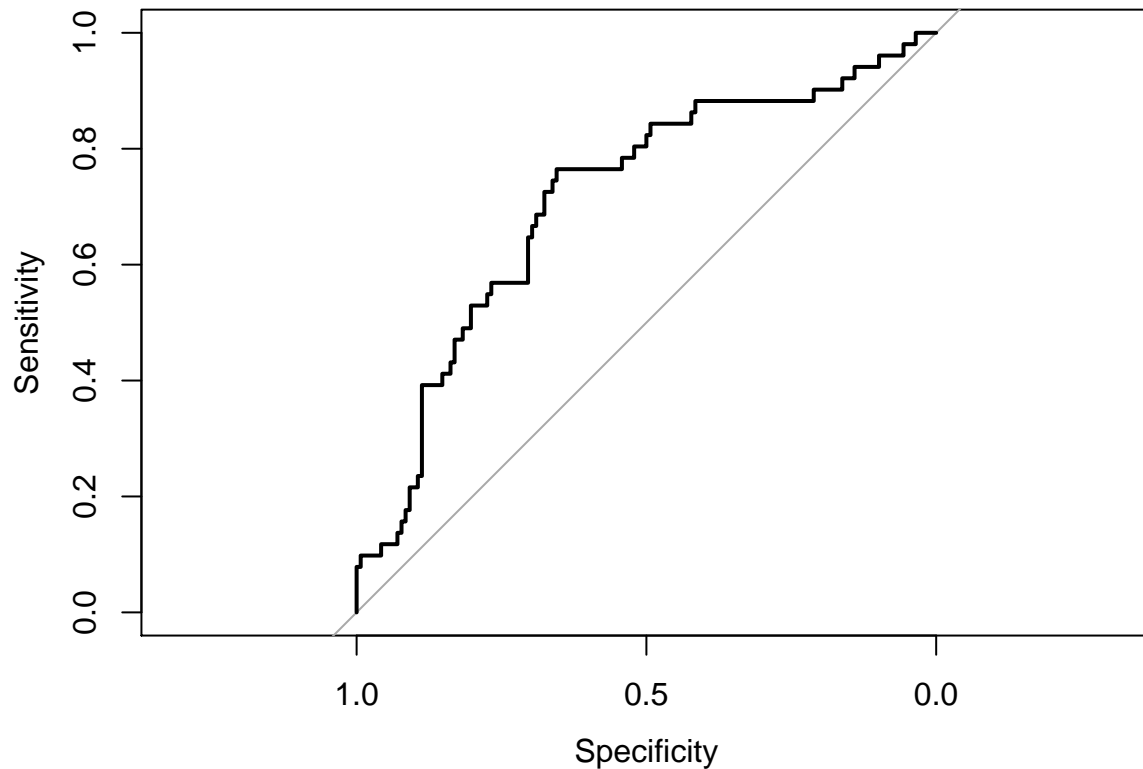


```
## [1] "AUC: 0.807247899159664"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 136 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.8072
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8914  -0.6853  -0.3707   0.5776   2.6905
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.605e+00  5.017e-01  -3.199  0.001379 **
## INCOME        -1.425e-05  3.210e-06  -4.439  9.04e-06 ***
## TRAVTIME       1.255e-02  7.086e-03   1.771  0.076641 .
## BLUEBOOK       3.125e-05  1.578e-05   1.980  0.047742 *
## TIF           -1.766e-02  2.677e-02  -0.660  0.509560
## OLDCLAIM       1.198e-05  1.527e-05   0.785  0.432505
## PARENT1_Yes    1.134e+00  3.050e-01   3.720  0.000199 ***
## SEX_z_F       -8.951e-01  3.441e-01  -2.601  0.009284 **
## JOB_Manager   -3.880e-01  3.479e-01  -1.115  0.264860
```

```

## CAR_USE_Commercial          5.941e-01  2.485e-01  2.390 0.016832 *
## CAR_TYPE_Pickup             9.579e-01  3.240e-01  2.956 0.003114 **
## CAR_TYPE_Sports.Car         1.835e+00  4.669e-01  3.931 8.47e-05 ***
## CAR_TYPE_z_SUV              1.668e+00  4.184e-01  3.987 6.68e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.709e+00  4.433e-01 -6.110 9.97e-10 ***
## HOME_VAL_NA                 -5.240e-01  2.285e-01 -2.293 0.021838 *
## oldclaim_log                6.257e-02  3.211e-02  1.948 0.051378 .
## inter                       1.617e-02  4.219e-03  3.833 0.000127 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 720.89 on 623 degrees of freedom
## Residual deviance: 538.30 on 607 degrees of freedom
## AIC: 572.3
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 127  39
##           1  15  12
##
##           Accuracy : 0.7202
##           95% CI : (0.6512, 0.7823)
##           No Information Rate : 0.7358
##           P-Value [Acc > NIR] : 0.718997
##
##           Kappa : 0.1527
##
## Mcnemar's Test P-Value : 0.001749
##
##           Sensitivity : 0.8944
##           Specificity : 0.2353
##           Pos Pred Value : 0.7651
##           Neg Pred Value : 0.4444
##           Prevalence : 0.7358
##           Detection Rate : 0.6580
##           Detection Prevalence : 0.8601
##           Balanced Accuracy : 0.5648
##
##           'Positive' Class : 0
##

```



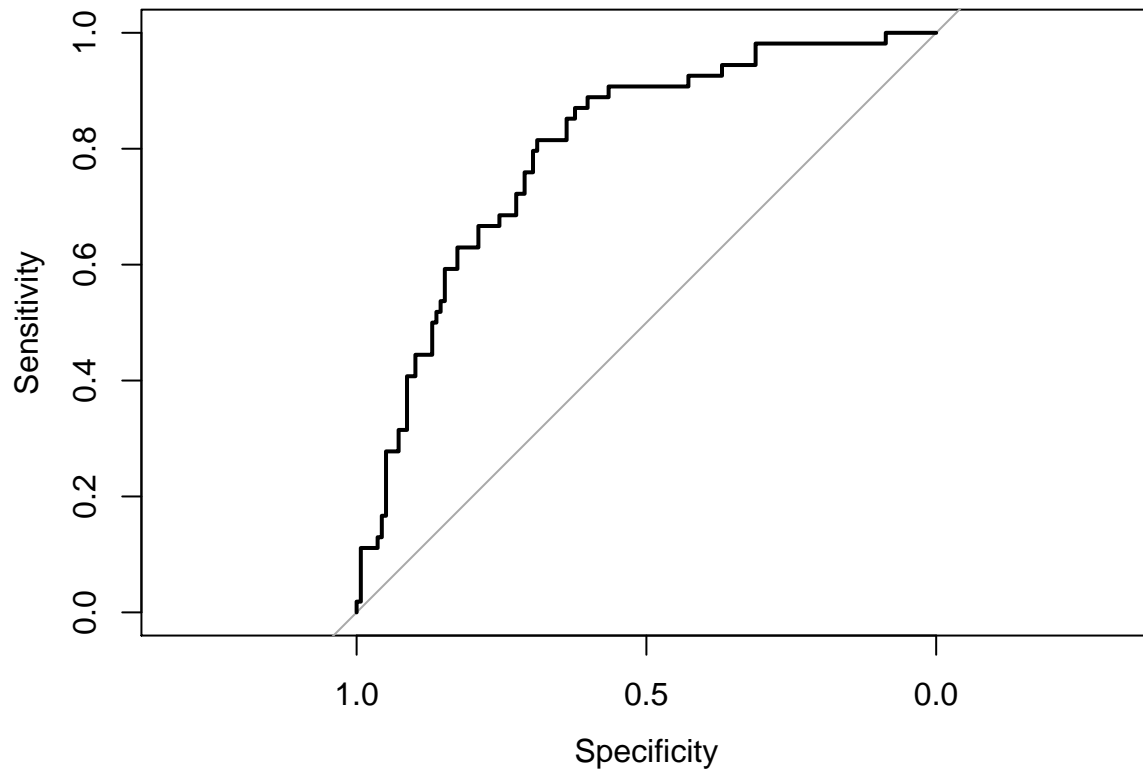
```
## [1] "AUC: 0.712096106048053"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 142 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.7121
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8927  -0.7437  -0.4200   0.5654   2.8844
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.047e+00  4.907e-01  -4.172 3.02e-05 ***
## INCOME        -9.411e-06  2.996e-06  -3.142 0.001680 **
## TRAVTIME       1.975e-02  7.142e-03   2.766 0.005674 **
## BLUEBOOK       3.064e-05  1.515e-05   2.022 0.043150 *
## TIF           -2.767e-02  2.575e-02  -1.074 0.282616
## OLDCLAIM       2.150e-05  1.363e-05   1.578 0.114646
## PARENT1_Yes    1.167e+00  3.146e-01   3.708 0.000209 ***
## SEX_z_F       -6.652e-01  3.413e-01  -1.949 0.051267 .
## JOB_Manager   -5.919e-01  3.418e-01  -1.732 0.083319 .
```



```

## CAR_USE_Commercial          6.707e-01  2.398e-01  2.796 0.005167 **
## CAR_TYPE_Pickup             7.367e-01  3.173e-01  2.322 0.020247 *
## CAR_TYPE_Sports.Car         1.632e+00  4.713e-01  3.464 0.000533 ***
## CAR_TYPE_z_SUV              1.244e+00  4.174e-01  2.981 0.002871 **
## URBANICITY_z_Highly.Rural..Rural -2.485e+00  4.174e-01 -5.954 2.61e-09 ***
## HOME_VAL_NA                 -2.755e-01  2.262e-01 -1.218 0.223287
## oldclaim_log                2.543e-02  3.183e-02  0.799 0.424385
## inter                       1.530e-02  4.104e-03  3.729 0.000192 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 715.27 on 624 degrees of freedom
## Residual deviance: 565.22 on 608 degrees of freedom
## AIC: 599.22
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 126  33
##           1  12  21
##
##           Accuracy : 0.7656
##           95% CI : (0.6992, 0.8236)
##           No Information Rate : 0.7188
##           P-Value [Acc > NIR] : 0.084452
##
##           Kappa : 0.3425
##
## Mcnemar's Test P-Value : 0.002869
##
##           Sensitivity : 0.9130
##           Specificity : 0.3889
##           Pos Pred Value : 0.7925
##           Neg Pred Value : 0.6364
##           Prevalence : 0.7188
##           Detection Rate : 0.6562
##           Detection Prevalence : 0.8281
##           Balanced Accuracy : 0.6510
##
##           'Positive' Class : 0
##

```

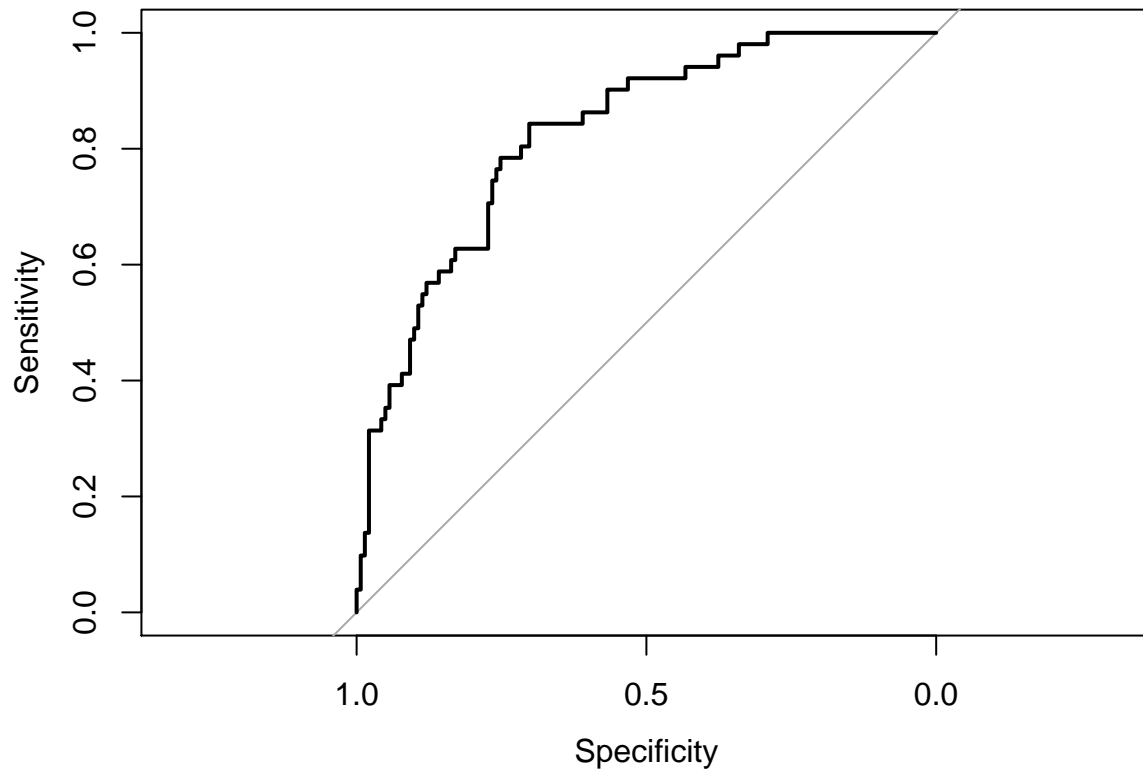


```
## [1] "AUC: 0.798711755233494"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 54 cases (dfPred_raw$class 1).
## Area under the curve: 0.7987
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7571  -0.7506  -0.4276   0.6613   2.8044
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.037e+00  4.915e-01  -4.144 3.42e-05 ***
## INCOME        -8.746e-06  2.857e-06  -3.061 0.002203 **
## TRAVTIME       2.310e-02  6.910e-03   3.344 0.000827 ***
## BLUEBOOK       2.871e-05  1.560e-05   1.840 0.065784 .
## TIF           -5.073e-02  2.647e-02  -1.917 0.055284 .
## OLDCLAIM       2.480e-06  1.430e-05   0.173 0.862334
## PARENT1_Yes    1.048e+00  3.086e-01   3.397 0.000682 ***
## SEX_z_F       -9.743e-01  3.651e-01  -2.668 0.007623 **
## JOB_Manager   -3.950e-01  3.475e-01  -1.137 0.255712
```

```

## CAR_USE_Commercial          4.962e-01  2.367e-01  2.096 0.036044 *
## CAR_TYPE_Pickup             1.071e+00  3.292e-01  3.254 0.001136 **
## CAR_TYPE_Sports.Car         1.881e+00  4.868e-01  3.864 0.000112 ***
## CAR_TYPE_z_SUV              1.835e+00  4.423e-01  4.149 3.34e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.270e+00  3.934e-01 -5.771 7.89e-09 ***
## HOME_VAL_NA                 -3.677e-01  2.206e-01 -1.667 0.095444 .
## oldclaim_log                5.341e-02  3.203e-02  1.668 0.095378 .
## inter                       1.444e-02  4.420e-03  3.267 0.001086 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 721.50  on 624  degrees of freedom
## Residual deviance: 578.68  on 608  degrees of freedom
## AIC: 612.68
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 128  28
##           1  13  23
##
##           Accuracy : 0.7865
##           95% CI : (0.7217, 0.8422)
##      No Information Rate : 0.7344
##      P-Value [Acc > NIR] : 0.05789
##
##           Kappa : 0.3959
##
## Mcnemar's Test P-Value : 0.02878
##
##           Sensitivity : 0.9078
##           Specificity : 0.4510
##      Pos Pred Value : 0.8205
##      Neg Pred Value : 0.6389
##           Prevalence : 0.7344
##      Detection Rate : 0.6667
##      Detection Prevalence : 0.8125
##      Balanced Accuracy : 0.6794
##
##           'Positive' Class : 0
##

```

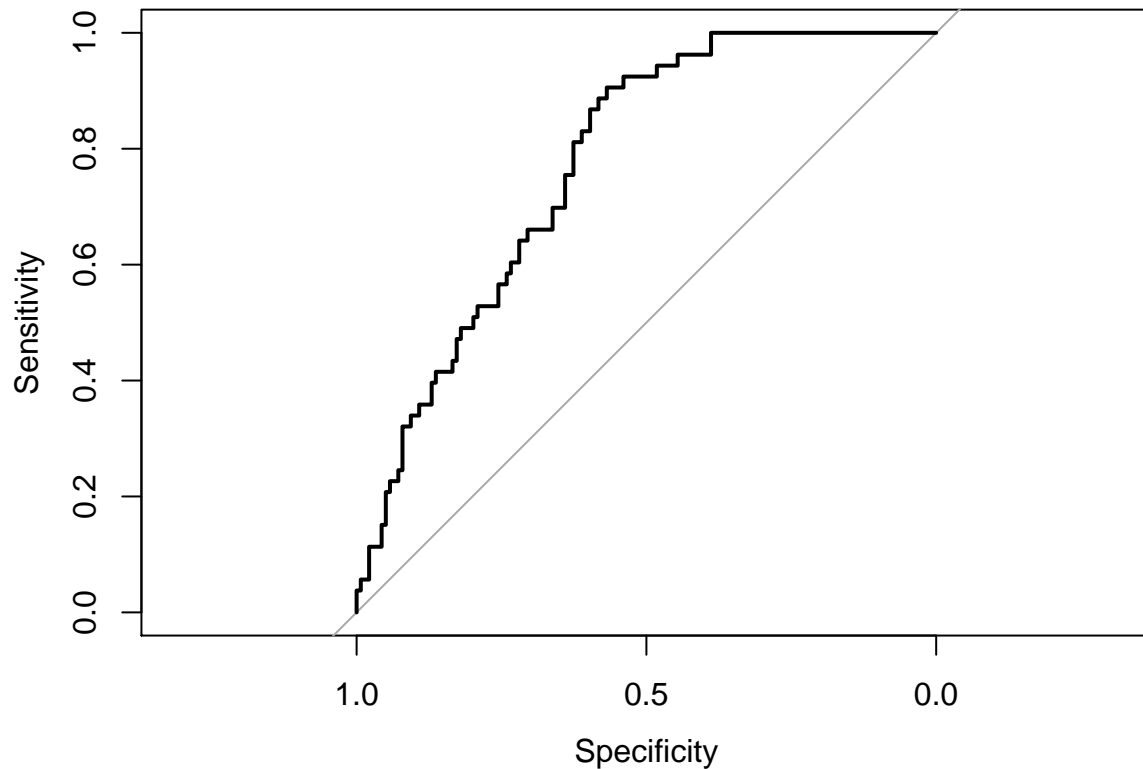


```
## [1] "AUC: 0.826449728827701"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 141 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.8264
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7963  -0.7243  -0.4229   0.6349   2.7511
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.446e+00  4.957e-01  -4.934 8.06e-07 ***
## INCOME        -9.131e-06  2.902e-06  -3.146 0.001655 **
## TRAVTIME       2.487e-02  6.988e-03   3.559 0.000372 ***
## BLUEBOOK       5.122e-05  1.563e-05   3.278 0.001046 **
## TIF          -3.927e-02  2.663e-02  -1.475 0.140265
## OLDCLAIM      1.459e-05  1.449e-05   1.007 0.313820
## PARENT1_Yes    1.262e+00  3.013e-01   4.189 2.80e-05 ***
## SEX_z_F       -9.800e-01  3.550e-01  -2.761 0.005762 **
## JOB_Manager   -1.036e+00  4.034e-01  -2.568 0.010226 *
```

```

## CAR_USE_Commercial          3.068e-01  2.360e-01  1.300 0.193582
## CAR_TYPE_Pickup             1.093e+00  3.222e-01  3.393 0.000690 ***
## CAR_TYPE_Sports.Car        1.993e+00  4.922e-01  4.049 5.15e-05 ***
## CAR_TYPE_z_SUV             1.538e+00  4.335e-01  3.548 0.000388 ***
## URBANICITY_z_Highly.Rural..Rural -2.308e+00  3.986e-01 -5.790 7.05e-09 ***
## HOME_VAL_NA                -2.375e-01  2.263e-01 -1.050 0.293831
## oldclaim_log               4.705e-02  3.241e-02  1.452 0.146575
## inter                      1.171e-02  4.196e-03  2.792 0.005239 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 717.37 on 624 degrees of freedom
## Residual deviance: 562.35 on 608 degrees of freedom
## AIC: 596.35
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0   1
##           0 128  38
##           1  11  15
##
##           Accuracy : 0.7448
##           95% CI : (0.677, 0.8048)
##           No Information Rate : 0.724
##           P-Value [Acc > NIR] : 0.2888995
##
##           Kappa : 0.242
##
## Mcnemar's Test P-Value : 0.0002038
##
##           Sensitivity : 0.9209
##           Specificity : 0.2830
##           Pos Pred Value : 0.7711
##           Neg Pred Value : 0.5769
##           Prevalence : 0.7240
##           Detection Rate : 0.6667
##           Detection Prevalence : 0.8646
##           Balanced Accuracy : 0.6019
##
##           'Positive' Class : 0
##

```

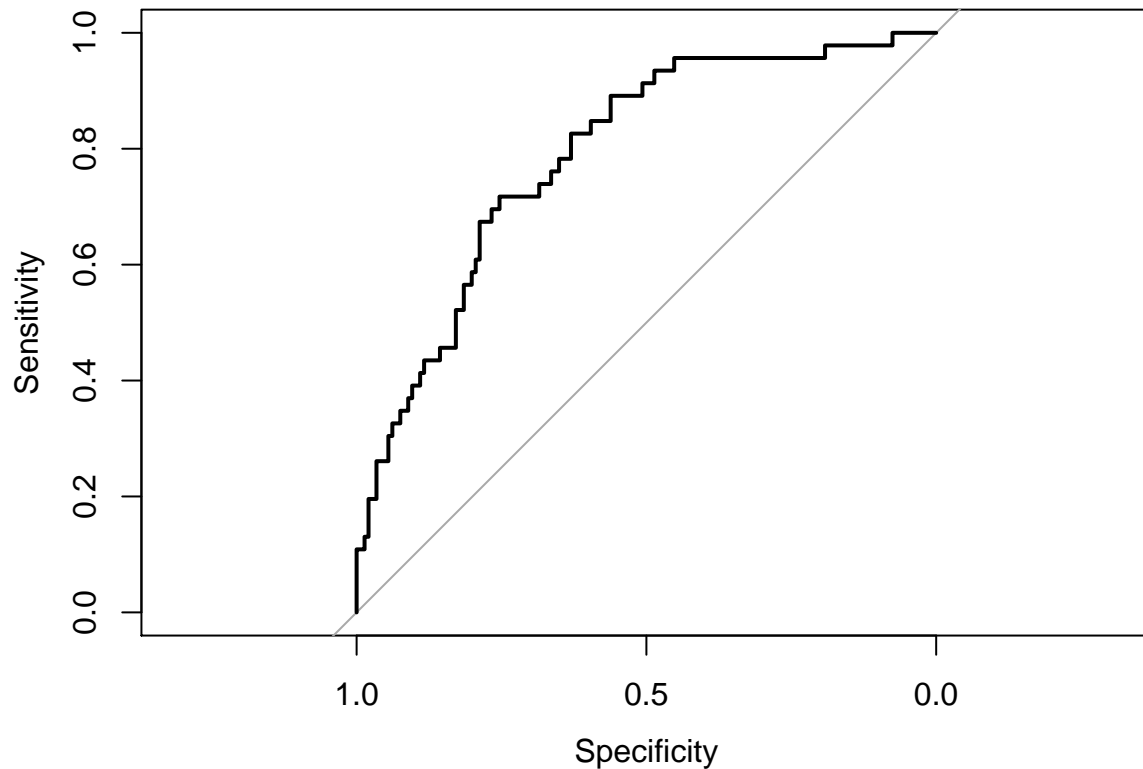


```
## [1] "AUC: 0.773584905660377"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 139 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.7736
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8360  -0.7240  -0.4095   0.7222   2.7537
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.919e+00  4.882e-01  -3.930 8.50e-05 ***
## INCOME        -1.002e-05  2.901e-06  -3.453 0.000553 ***
## TRAVTIME       2.303e-02  6.798e-03   3.388 0.000705 ***
## BLUEBOOK       2.278e-05  1.553e-05   1.466 0.142550
## TIF           -4.368e-02  2.713e-02  -1.610 0.107307
## OLDCLAIM       1.146e-05  1.424e-05   0.805 0.420821
## PARENT1_Yes    1.027e+00  2.977e-01   3.448 0.000565 ***
## SEX_z_F       -9.723e-01  3.585e-01  -2.712 0.006683 **
## JOB_Manager   -8.922e-01  3.847e-01  -2.319 0.020398 *
```

```

## CAR_USE_Commercial          4.727e-01  2.376e-01  1.990 0.046634 *
## CAR_TYPE_Pickup             9.773e-01  3.164e-01  3.088 0.002012 **
## CAR_TYPE_Sports.Car        2.017e+00  4.838e-01  4.169 3.06e-05 ***
## CAR_TYPE_z_SUV             1.706e+00  4.295e-01  3.972 7.14e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.211e+00  3.947e-01 -5.603 2.11e-08 ***
## HOME_VAL_NA                -2.182e-01  2.198e-01 -0.993 0.320859
## oldclaim_log               4.623e-02  3.128e-02  1.478 0.139412
## inter                      1.019e-02  4.088e-03  2.493 0.012673 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 731.55 on 624 degrees of freedom
## Residual deviance: 577.33 on 608 degrees of freedom
## AIC: 611.33
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 141  34
##           1   5  12
##
## Accuracy : 0.7969
## 95% CI : (0.733, 0.8514)
## No Information Rate : 0.7604
## P-Value [Acc > NIR] : 0.1351
##
## Kappa : 0.289
##
## McNemar's Test P-Value : 7.34e-06
##
## Sensitivity : 0.9658
## Specificity : 0.2609
## Pos Pred Value : 0.8057
## Neg Pred Value : 0.7059
## Prevalence : 0.7604
## Detection Rate : 0.7344
## Detection Prevalence : 0.9115
## Balanced Accuracy : 0.6133
##
## 'Positive' Class : 0
##

```



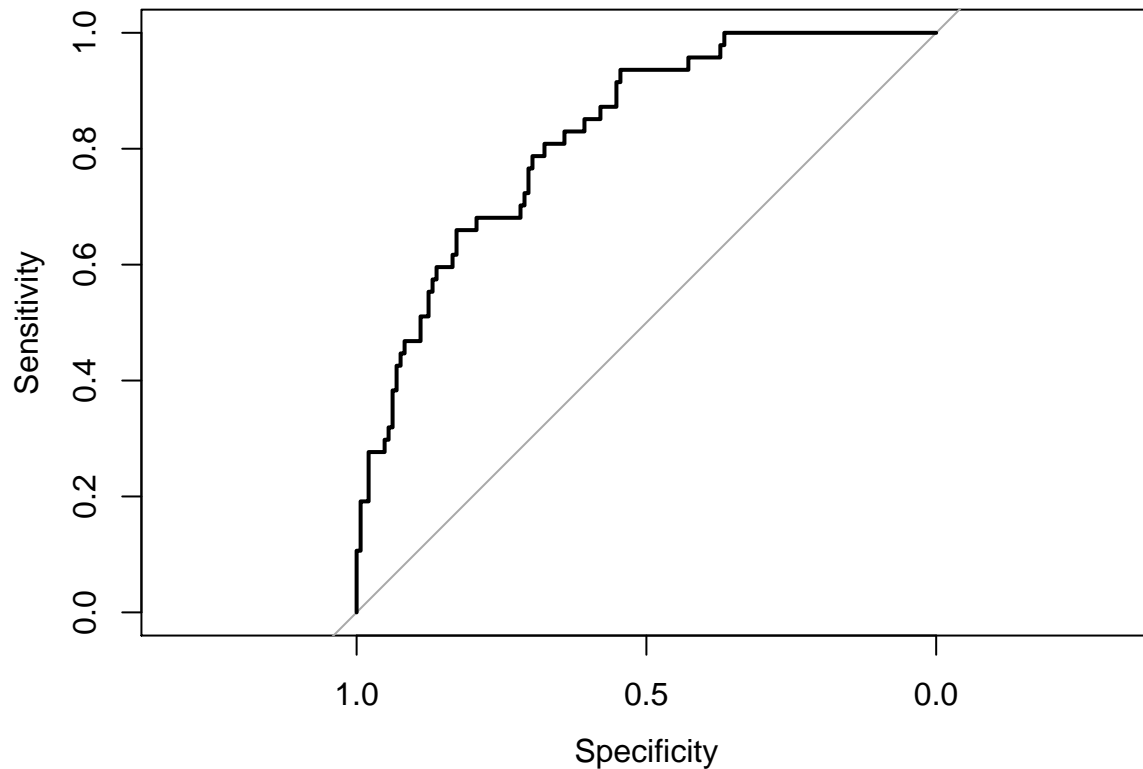
```
## [1] "AUC: 0.789309112567004"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7893
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9694  -0.7659  -0.4290   0.7507   2.7270
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.908e+00  4.805e-01  -3.970 7.18e-05 ***
## INCOME        -9.094e-06  2.847e-06  -3.195 0.001400 **
## TRAVTIME       2.189e-02  6.975e-03   3.139 0.001698 **
## BLUEBOOK       2.714e-05  1.580e-05   1.717 0.085899 .
## TIF           -4.552e-02  2.582e-02  -1.763 0.077829 .
## OLDCLAIM       1.148e-05  1.396e-05   0.822 0.411103
## PARENT1_Yes     8.431e-01  2.966e-01   2.843 0.004473 **
## SEX_z_F        -9.674e-01  3.475e-01  -2.784 0.005375 **
## JOB_Manager    -9.780e-01  4.016e-01  -2.435 0.014878 *
```



```

## CAR_USE_Commercial          4.027e-01  2.334e-01  1.725 0.084482 .
## CAR_TYPE_Pickup             1.062e+00  3.177e-01  3.343 0.000829 ***
## CAR_TYPE_Sports.Car         1.862e+00  4.883e-01  3.813 0.000137 ***
## CAR_TYPE_z_SUV              1.718e+00  4.213e-01  4.077 4.56e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.265e+00  3.932e-01 -5.759 8.47e-09 ***
## HOME_VAL_NA                 -1.962e-01  2.173e-01 -0.903 0.366506
## oldclaim_log                3.952e-02  3.114e-02  1.269 0.204440
## inter                       1.247e-02  3.990e-03  3.125 0.001776 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 729.57 on 624 degrees of freedom
## Residual deviance: 585.91 on 608 degrees of freedom
## AIC: 619.91
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 136  31
##           1   9  16
##
##           Accuracy : 0.7917
##           95% CI : (0.7273, 0.8468)
##           No Information Rate : 0.7552
##           P-Value [Acc > NIR] : 0.1369190
##
##           Kappa : 0.3307
##
## Mcnemar's Test P-Value : 0.0008989
##
##           Sensitivity : 0.9379
##           Specificity : 0.3404
##           Pos Pred Value : 0.8144
##           Neg Pred Value : 0.6400
##           Prevalence : 0.7552
##           Detection Rate : 0.7083
##           Detection Prevalence : 0.8698
##           Balanced Accuracy : 0.6392
##
##           'Positive' Class : 0
##

```

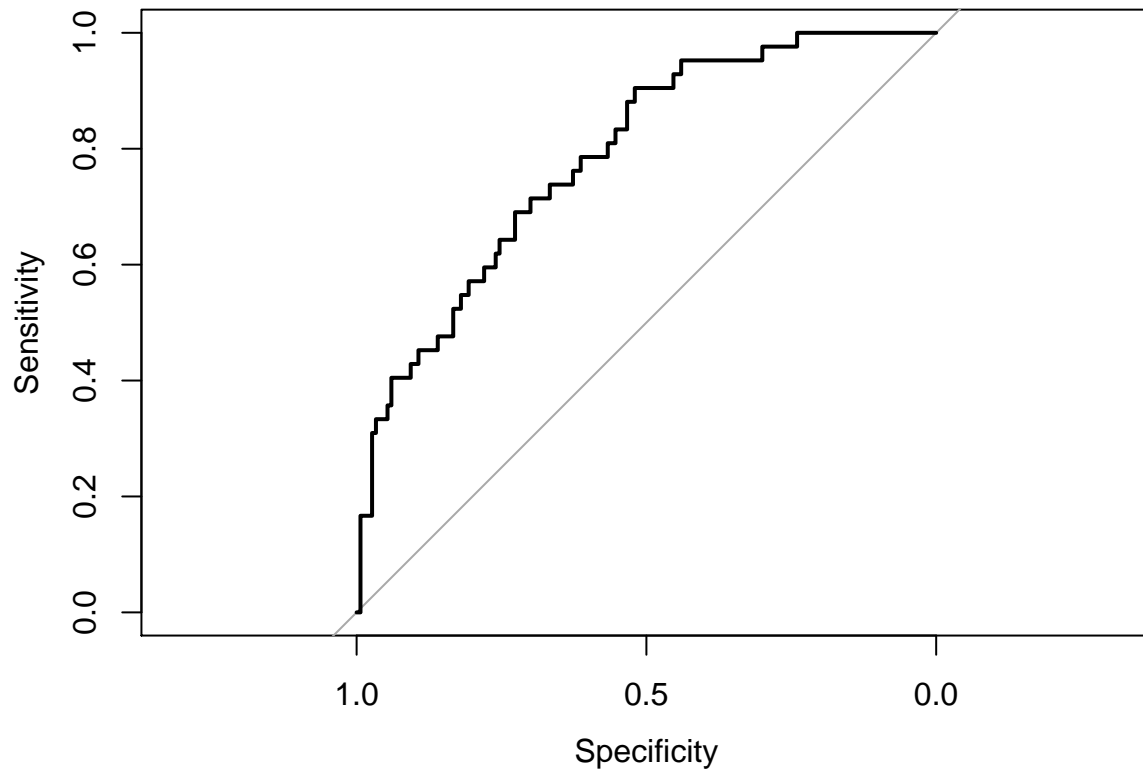


```
## [1] "AUC: 0.823330887747616"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 145 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.8233
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7934  -0.7504  -0.4167   0.7351   2.7612
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.103e+00  4.912e-01  -4.283 1.85e-05 ***
## INCOME        -9.529e-06  2.843e-06  -3.352 0.000803 ***
## TRAVTIME       2.466e-02  6.931e-03   3.557 0.000375 ***
## BLUEBOOK       3.550e-05  1.543e-05   2.301 0.021408 *
## TIF           -6.055e-02  2.763e-02  -2.192 0.028410 *
## OLDCLAIM       9.986e-06  1.409e-05   0.709 0.478470
## PARENT1_Yes    1.015e+00  3.003e-01   3.382 0.000721 ***
## SEX_z_F       -1.124e+00  3.798e-01  -2.959 0.003084 **
## JOB_Manager    -7.538e-01  3.587e-01  -2.101 0.035609 *
```

```

## CAR_USE_Commercial          5.779e-01  2.416e-01  2.392 0.016775 *
## CAR_TYPE_Pickup             9.485e-01  3.267e-01  2.903 0.003696 **
## CAR_TYPE_Sports.Car        2.174e+00  5.053e-01  4.303 1.69e-05 ***
## CAR_TYPE_z_SUV             1.975e+00  4.554e-01  4.338 1.44e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.282e+00  3.906e-01 -5.843 5.14e-09 ***
## HOME_VAL_NA                -2.898e-01  2.201e-01 -1.316 0.188019
## oldclaim_log               4.472e-02  3.129e-02  1.430 0.152849
## inter                      1.471e-02  4.081e-03  3.604 0.000313 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 739.29  on 624  degrees of freedom
## Residual deviance: 581.88  on 608  degrees of freedom
## AIC: 615.88
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 138  25
##           1  12  17
##
##           Accuracy : 0.8073
##           95% CI : (0.7443, 0.8605)
##      No Information Rate : 0.7812
##      P-Value [Acc > NIR] : 0.21792
##
##           Kappa : 0.3655
##
## Mcnemar's Test P-Value : 0.04852
##
##           Sensitivity : 0.9200
##           Specificity : 0.4048
##      Pos Pred Value : 0.8466
##      Neg Pred Value : 0.5862
##           Prevalence : 0.7812
##      Detection Rate : 0.7188
##      Detection Prevalence : 0.8490
##      Balanced Accuracy : 0.6624
##
##           'Positive' Class : 0
##

```

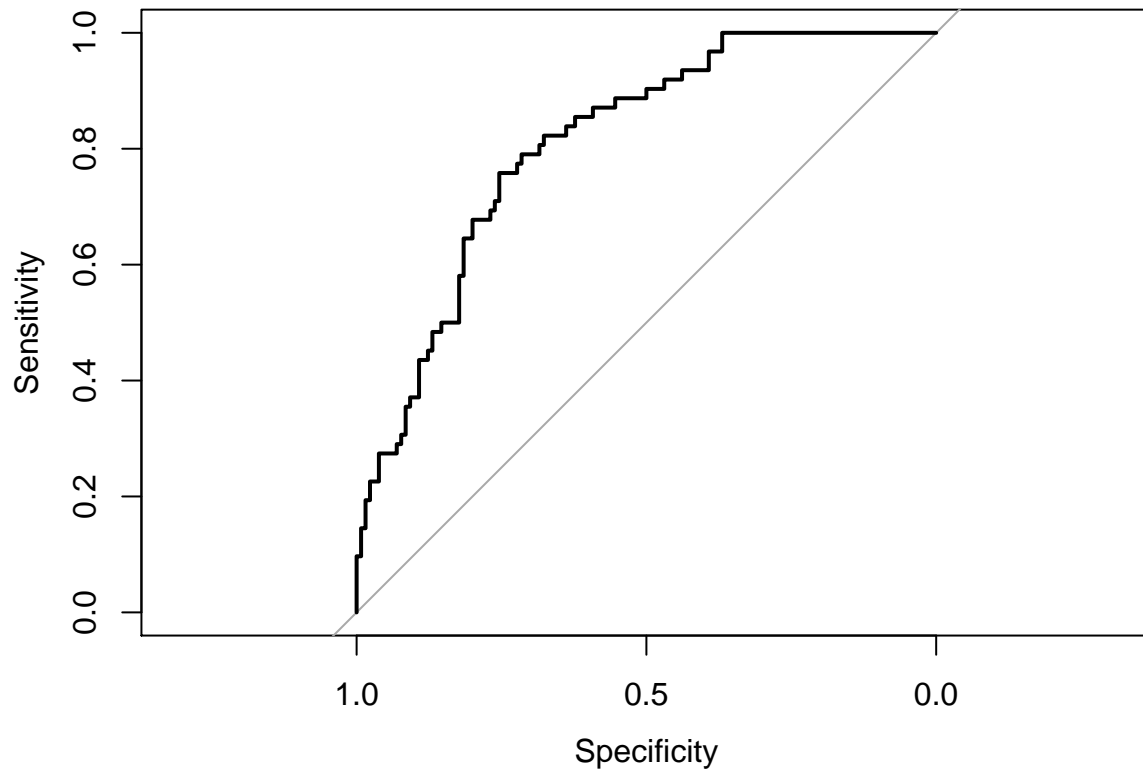


```
## [1] "AUC: 0.786507936507936"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 150 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.7865
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7807  -0.7152  -0.4387  -0.1162   2.7202
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.920e+00  5.055e-01  -3.798 0.000146 ***
## INCOME        -6.839e-06  2.912e-06  -2.349 0.018848 *
## TRAVTIME       1.920e-02  7.027e-03   2.732 0.006294 **
## BLUEBOOK       2.326e-05  1.593e-05   1.460 0.144346
## TIF           -6.100e-02  2.759e-02  -2.211 0.027019 *
## OLDCCLAIM      6.622e-06  1.477e-05   0.448 0.653806
## PARENT1_Yes    1.174e+00  3.258e-01   3.603 0.000315 ***
## SEX_z_F       -1.149e+00  3.769e-01  -3.049 0.002293 **
## JOB_Manager   -6.601e-01  3.901e-01  -1.692 0.090615 .
```

```

## CAR_USE_Commercial          4.223e-01  2.482e-01  1.701 0.088888 .
## CAR_TYPE_Pickup             1.136e+00  3.342e-01  3.398 0.000679 ***
## CAR_TYPE_Sports.Car        1.905e+00  4.984e-01  3.823 0.000132 ***
## CAR_TYPE_z_SUV             1.912e+00  4.565e-01  4.189 2.80e-05 ***
## URBANICITY_z_Highly.Rural..Rural -1.942e+00  3.826e-01 -5.075 3.88e-07 ***
## HOME_VAL_NA                -4.425e-01  2.237e-01 -1.978 0.047909 *
## oldclaim_log               6.370e-02  3.198e-02  1.992 0.046400 *
## inter                      1.256e-02  4.633e-03  2.711 0.006709 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 697.93  on 624  degrees of freedom
## Residual deviance: 561.72  on 608  degrees of freedom
## AIC: 595.72
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 119  41
##           1  11  21
##
##           Accuracy : 0.7292
##           95% CI : (0.6605, 0.7906)
##      No Information Rate : 0.6771
##      P-Value [Acc > NIR] : 0.06969
##
##           Kappa : 0.2909
##
## Mcnemar's Test P-Value : 5.781e-05
##
##           Sensitivity : 0.9154
##           Specificity : 0.3387
##      Pos Pred Value : 0.7438
##      Neg Pred Value : 0.6562
##           Prevalence : 0.6771
##      Detection Rate : 0.6198
## Detection Prevalence : 0.8333
##      Balanced Accuracy : 0.6270
##
##           'Positive' Class : 0
##

```

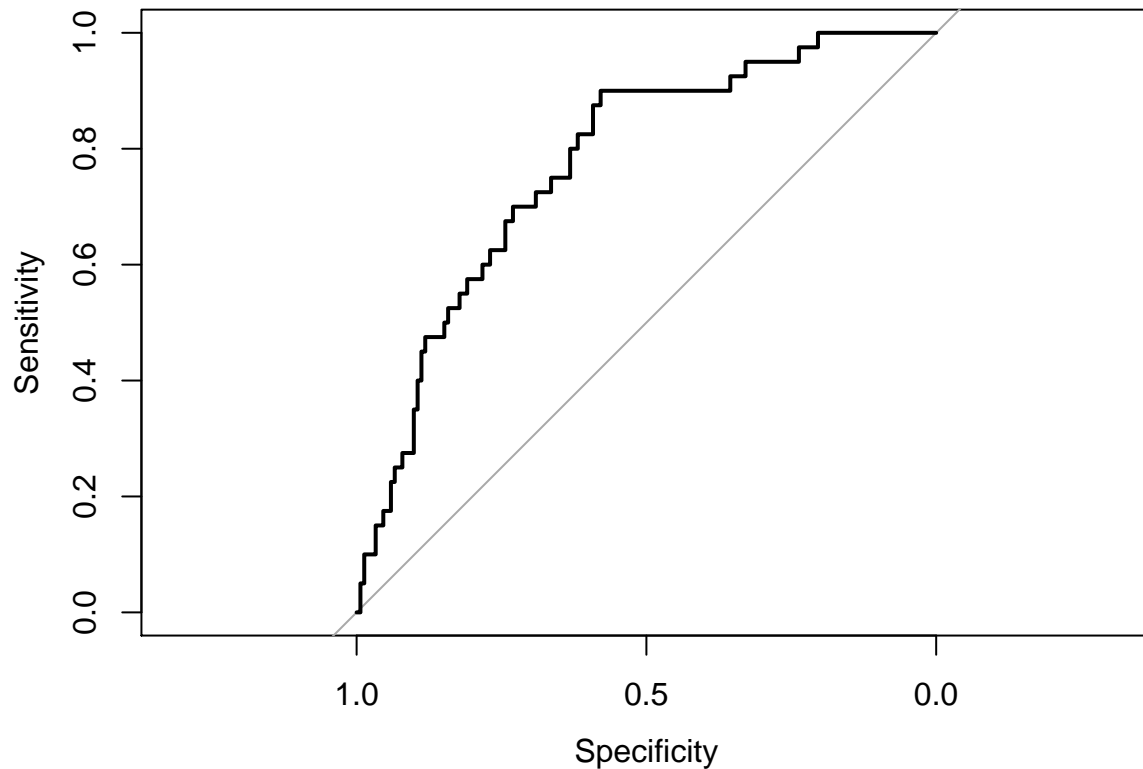


```
## [1] "AUC: 0.810049627791563"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 130 controls (dfPred_raw$class 0) < 62 cases (dfPred_raw$class 1).
## Area under the curve: 0.81
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9612  -0.7538  -0.4049   0.6972   2.8695
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.757e+00  4.804e-01  -3.658 0.000254 ***
## INCOME        -1.129e-05  3.020e-06  -3.739 0.000185 ***
## TRAVTIME       1.748e-02  6.946e-03   2.516 0.011873 *
## BLUEBOOK       3.070e-05  1.509e-05   2.035 0.041819 *
## TIF           -3.634e-02  2.658e-02  -1.367 0.171566
## OLDCLAIM       1.442e-05  1.409e-05   1.023 0.306205
## PARENT1_Yes    1.021e+00  3.124e-01   3.267 0.001089 **
## SEX_z_F       -9.815e-01  3.542e-01  -2.771 0.005588 **
## JOB_Manager   -5.161e-01  3.365e-01  -1.534 0.125115
```

```

## CAR_USE_Commercial          7.562e-01  2.432e-01  3.109 0.001878 **
## CAR_TYPE_Pickup             8.318e-01  3.150e-01  2.641 0.008277 **
## CAR_TYPE_Sports.Car         2.075e+00  4.789e-01  4.334 1.47e-05 ***
## CAR_TYPE_z_SUV              1.808e+00  4.311e-01  4.194 2.74e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.444e+00  3.883e-01 -6.294 3.09e-10 ***
## HOME_VAL_NA                 -3.620e-01  2.244e-01 -1.613 0.106707
## oldclaim_log                2.888e-02  3.157e-02  0.915 0.360264
## inter                       1.719e-02  4.042e-03  4.253 2.11e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 743.07 on 624 degrees of freedom
## Residual deviance: 576.48 on 608 degrees of freedom
## AIC: 610.48
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 135  22
##           1  17  18
##
## Accuracy : 0.7969
## 95% CI : (0.733, 0.8514)
## No Information Rate : 0.7917
## P-Value [Acc > NIR] : 0.4715
##
## Kappa : 0.3545
##
## McNemar's Test P-Value : 0.5218
##
## Sensitivity : 0.8882
## Specificity : 0.4500
## Pos Pred Value : 0.8599
## Neg Pred Value : 0.5143
## Prevalence : 0.7917
## Detection Rate : 0.7031
## Detection Prevalence : 0.8177
## Balanced Accuracy : 0.6691
##
## 'Positive' Class : 0
##

```



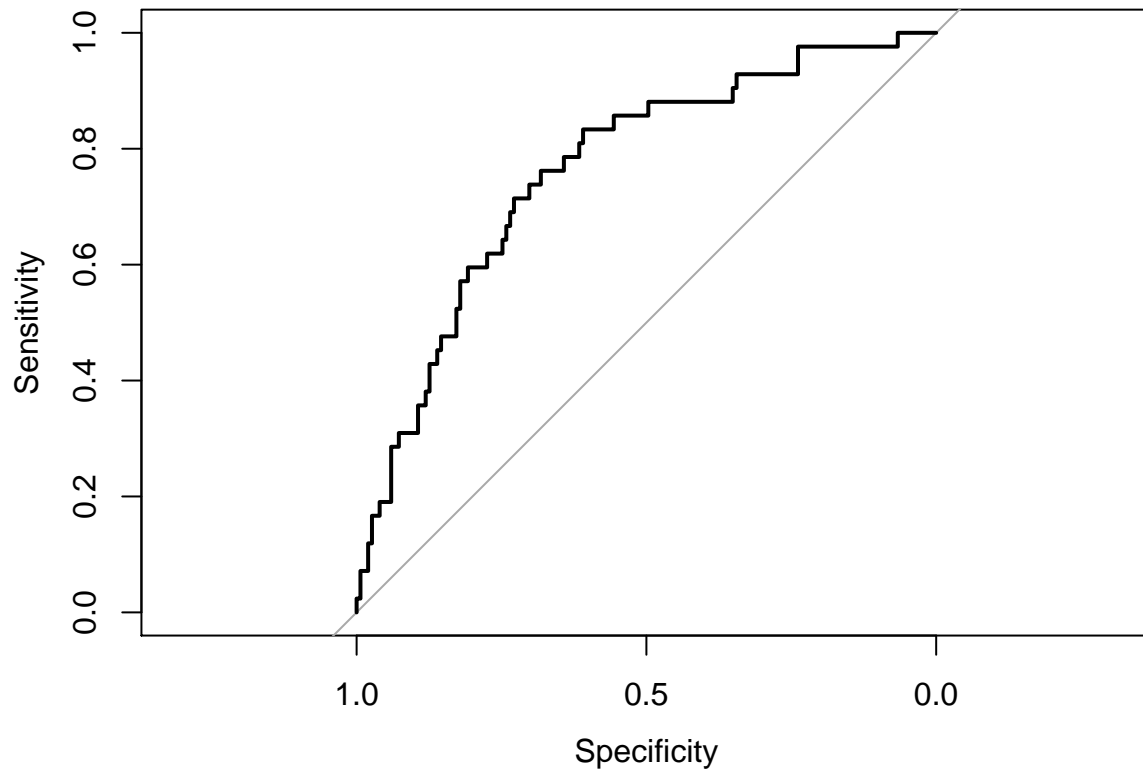
```
## [1] "AUC: 0.773848684210526"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 152 controls (dfPred_raw$class 0) < 40 cases (dfPred_raw$class 1).
## Area under the curve: 0.7738
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9706  -0.7294  -0.4062   0.6628   2.8382
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.835e+00  4.856e-01  -3.779 0.000158 ***
## INCOME        -1.115e-05  2.897e-06  -3.849 0.000119 ***
## TRAVTIME       2.260e-02  7.094e-03   3.186 0.001443 **
## BLUEBOOK       3.016e-05  1.518e-05   1.986 0.047005 *
## TIF           -4.036e-02  2.638e-02  -1.530 0.126033
## OLDCLAIM       6.443e-06  1.426e-05   0.452 0.651377
## PARENT1_Yes    1.151e+00  3.133e-01   3.673 0.000240 ***
## SEX_z_F       -8.940e-01  3.530e-01  -2.533 0.011323 *
## JOB_Manager   -5.802e-01  3.355e-01  -1.729 0.083769 .
```



```

## CAR_USE_Commercial          5.006e-01  2.410e-01  2.078 0.037742 *
## CAR_TYPE_Pickup             6.943e-01  3.215e-01  2.159 0.030822 *
## CAR_TYPE_Sports.Car        1.668e+00  4.760e-01  3.504 0.000459 ***
## CAR_TYPE_z_SUV             1.697e+00  4.238e-01  4.004 6.23e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.794e+00  4.300e-01 -6.497 8.17e-11 ***
## HOME_VAL_NA                -3.572e-01  2.257e-01 -1.583 0.113415
## oldclaim_log               6.795e-02  3.161e-02  2.150 0.031578 *
## inter                      1.776e-02  4.135e-03  4.295 1.74e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 738.64 on 623 degrees of freedom
## Residual deviance: 563.62 on 607 degrees of freedom
## AIC: 597.62
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 130  24
##           1  21  18
##
##           Accuracy : 0.7668
##           95% CI : (0.7007, 0.8246)
##           No Information Rate : 0.7824
##           P-Value [Acc > NIR] : 0.7326
##
##           Kappa : 0.2972
##
## Mcnemar's Test P-Value : 0.7656
##
##           Sensitivity : 0.8609
##           Specificity : 0.4286
##           Pos Pred Value : 0.8442
##           Neg Pred Value : 0.4615
##           Prevalence : 0.7824
##           Detection Rate : 0.6736
##           Detection Prevalence : 0.7979
##           Balanced Accuracy : 0.6447
##
##           'Positive' Class : 0
##

```

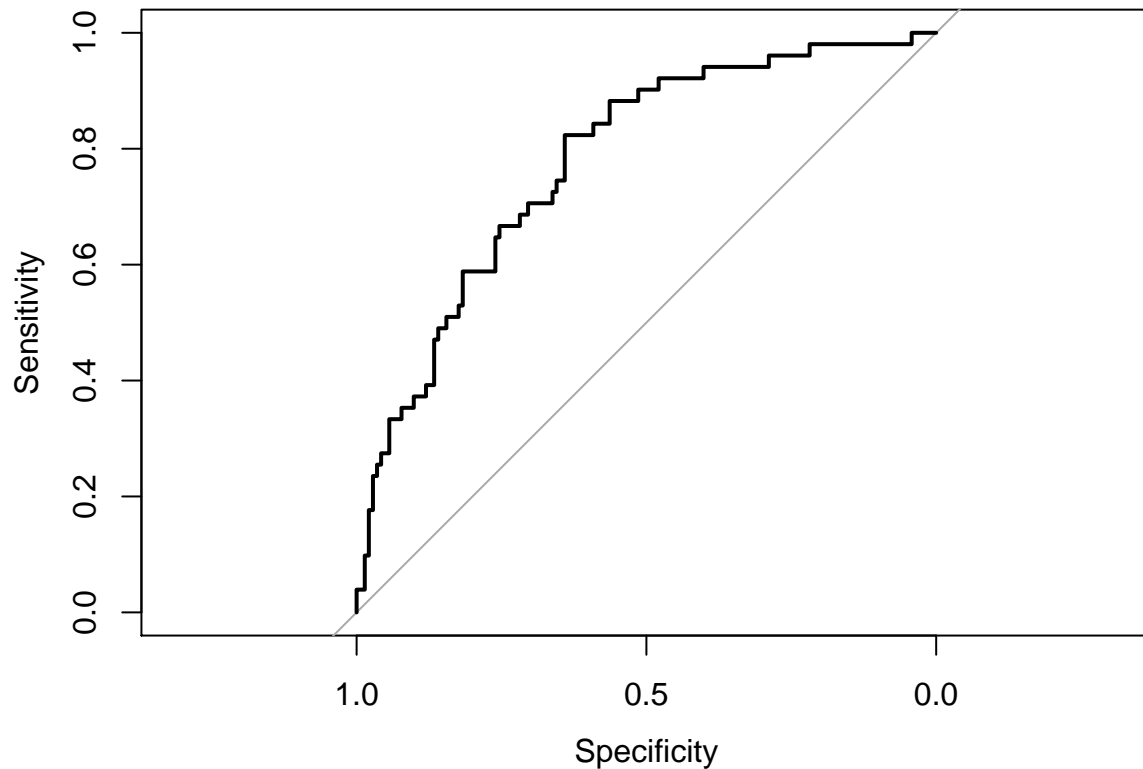


```
## [1] "AUC: 0.76316619362977"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 151 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.7632
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8322  -0.7186  -0.4018   0.6336   2.7047
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.744e+00  5.037e-01  -3.463 0.000534 ***
## INCOME         -1.212e-05  2.945e-06  -4.114 3.88e-05 ***
## TRAVTIME        1.768e-02  7.020e-03   2.518 0.011791 *
## BLUEBOOK        3.601e-05  1.549e-05   2.324 0.020101 *
## TIF            -5.171e-02  2.649e-02  -1.952 0.050907 .
## OLDCLAIM        1.472e-05  1.379e-05   1.068 0.285684
## PARENT1_Yes     9.409e-01  3.133e-01   3.004 0.002668 **
## SEX_z_F        -9.352e-01  3.497e-01  -2.674 0.007485 **
## JOB_Manager    -5.614e-01  3.484e-01  -1.611 0.107112
```

```

## CAR_USE_Commercial          5.511e-01  2.407e-01  2.290 0.022050 *
## CAR_TYPE_Pickup             8.726e-01  3.273e-01  2.667 0.007664 **
## CAR_TYPE_Sports.Car         1.927e+00  4.780e-01  4.032 5.54e-05 ***
## CAR_TYPE_z_SUV              1.731e+00  4.242e-01  4.080 4.51e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.453e+00  4.150e-01 -5.910 3.41e-09 ***
## HOME_VAL_NA                 -3.604e-01  2.241e-01 -1.608 0.107818
## oldclaim_log                5.015e-02  3.197e-02  1.569 0.116744
## inter                       1.236e-02  4.199e-03  2.944 0.003241 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 720.89  on 623  degrees of freedom
## Residual deviance: 560.39  on 607  degrees of freedom
## AIC: 594.39
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 131  33
##           1  11  18
##
##           Accuracy : 0.772
##           95% CI : (0.7063, 0.8292)
##      No Information Rate : 0.7358
##      P-Value [Acc > NIR] : 0.143868
##
##           Kappa : 0.3197
##
## Mcnemar's Test P-Value : 0.001546
##
##           Sensitivity : 0.9225
##           Specificity : 0.3529
##      Pos Pred Value : 0.7988
##      Neg Pred Value : 0.6207
##           Prevalence : 0.7358
##      Detection Rate : 0.6788
##      Detection Prevalence : 0.8497
##      Balanced Accuracy : 0.6377
##
##           'Positive' Class : 0
##

```

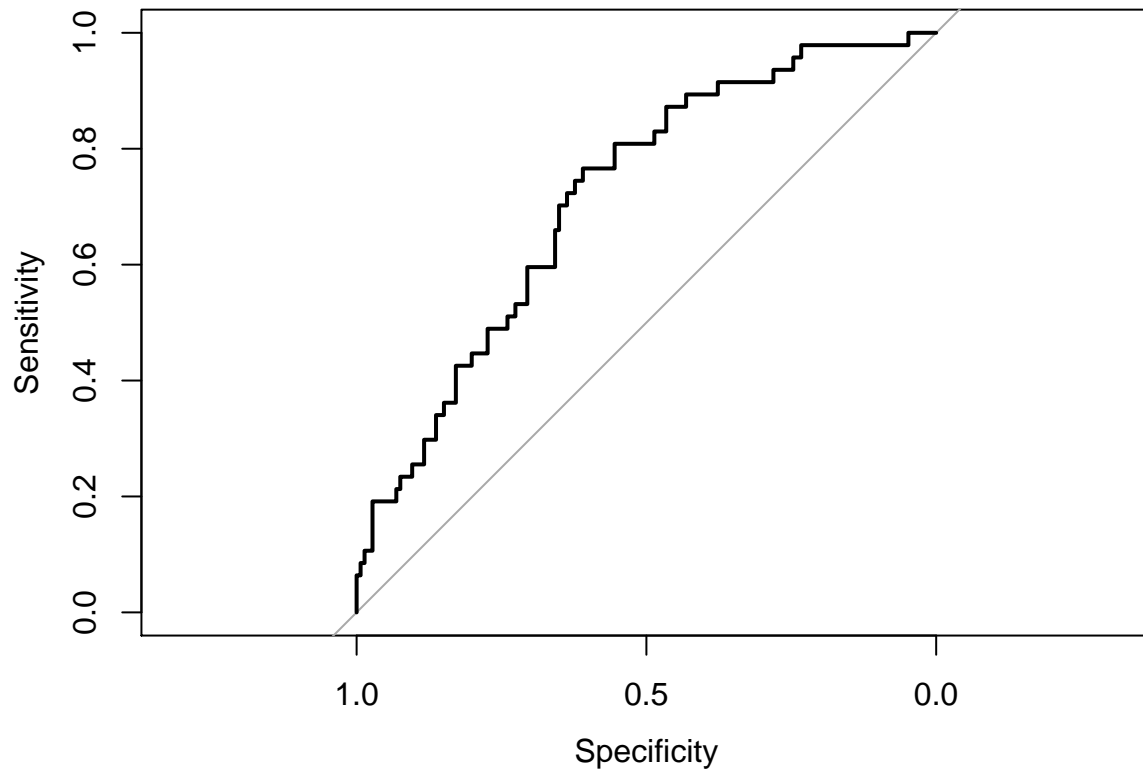


```
## [1] "AUC: 0.778238055785695"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 142 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.7782
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8874  -0.7054  -0.3838   0.6406   2.8889
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.928e+00  4.915e-01  -3.923  8.73e-05 ***
## INCOME        -1.333e-05  3.094e-06  -4.308  1.65e-05 ***
## TRAVTIME       1.846e-02  7.127e-03   2.591  0.009575 **
## BLUEBOOK       4.225e-05  1.607e-05   2.630  0.008545 **
## TIF           -1.364e-02  2.615e-02  -0.522  0.601891
## OLDCLAIM       5.092e-08  1.458e-05   0.003  0.997214
## PARENT1_Yes    1.075e+00  2.959e-01   3.635  0.000278 ***
## SEX_z_F       -8.365e-01  3.388e-01  -2.469  0.013541 *
## JOB_Manager   -5.071e-01  3.514e-01  -1.443  0.149047
```

```

## CAR_USE_Commercial          3.229e-01  2.421e-01  1.334 0.182256
## CAR_TYPE_Pickup             8.359e-01  3.228e-01  2.589 0.009612 **
## CAR_TYPE_Sports.Car        1.531e+00  4.654e-01  3.290 0.001003 **
## CAR_TYPE_z_SUV             1.524e+00  4.070e-01  3.745 0.000180 ***
## URBANICITY_z_Highly.Rural..Rural -2.744e+00  4.435e-01 -6.187 6.14e-10 ***
## HOME_VAL_NA                -4.466e-01  2.262e-01 -1.975 0.048324 *
## oldclaim_log               1.009e-01  3.149e-02  3.204 0.001353 **
## inter                      1.553e-02  4.171e-03  3.724 0.000196 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 728.94  on 623  degrees of freedom
## Residual deviance: 551.22  on 607  degrees of freedom
## AIC: 585.22
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 125  31
##           1  21  16
##
##              Accuracy : 0.7306
##              95% CI : (0.6621, 0.7918)
##      No Information Rate : 0.7565
##      P-Value [Acc > NIR] : 0.8224
##
##              Kappa : 0.2119
##
## Mcnemar's Test P-Value : 0.2120
##
##              Sensitivity : 0.8562
##              Specificity : 0.3404
##      Pos Pred Value : 0.8013
##      Neg Pred Value : 0.4324
##              Prevalence : 0.7565
##      Detection Rate : 0.6477
##      Detection Prevalence : 0.8083
##      Balanced Accuracy : 0.5983
##
##      'Positive' Class : 0
##

```

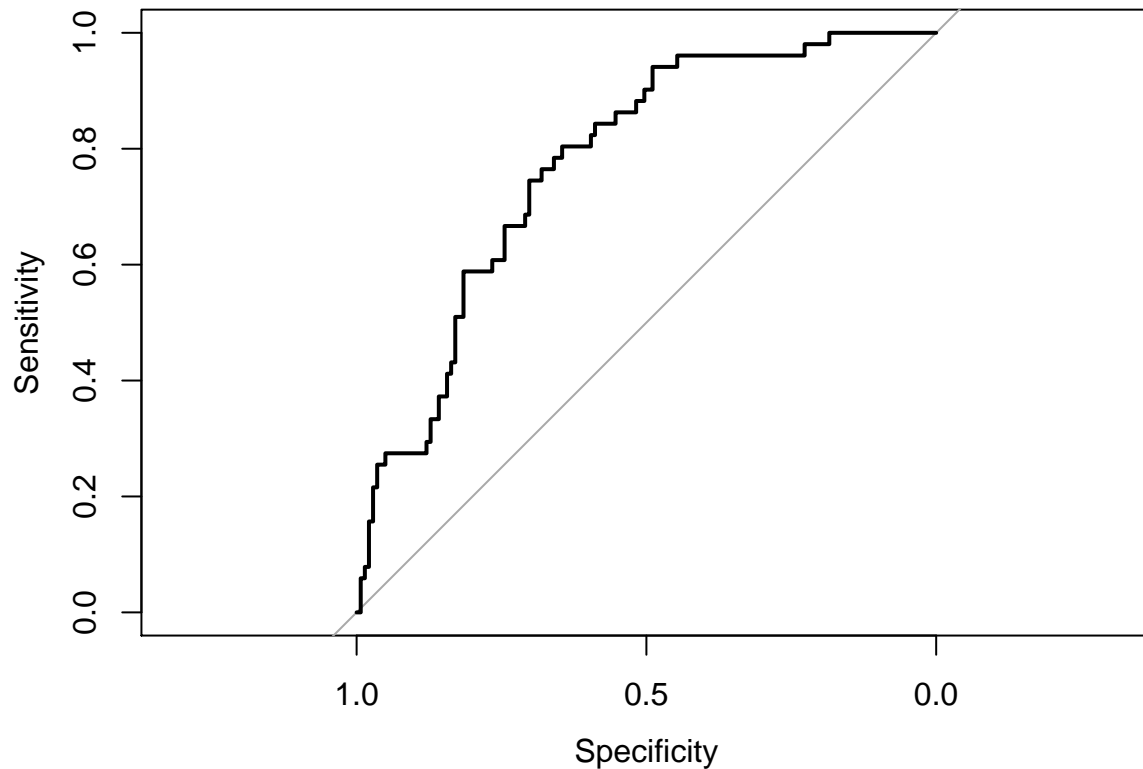


```
## [1] "AUC: 0.716554940250656"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.7166
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1087  -0.7270  -0.4066   0.5919   2.8857
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.221e+00  5.047e-01  -4.400 1.08e-05 ***
## INCOME        -9.844e-06  2.946e-06  -3.341 0.000834 ***
## TRAVTIME       2.568e-02  7.193e-03   3.571 0.000356 ***
## BLUEBOOK       3.919e-05  1.549e-05   2.531 0.011385 *
## TIF           -5.162e-02  2.746e-02  -1.880 0.060153 .
## OLDCLAIM       1.444e-05  1.382e-05   1.045 0.295897
## PARENT1_Yes     9.788e-01  3.003e-01   3.260 0.001116 **
## SEX_z_F       -7.258e-01  3.510e-01  -2.068 0.038626 *
## JOB_Manager    -7.540e-01  3.567e-01  -2.114 0.034524 *
```

```

## CAR_USE_Commercial          4.654e-01  2.466e-01  1.887 0.059130 .
## CAR_TYPE_Pickup             8.536e-01  3.258e-01  2.620 0.008791 **
## CAR_TYPE_Sports.Car        1.579e+00  4.825e-01  3.272 0.001067 **
## CAR_TYPE_z_SUV             1.474e+00  4.222e-01  3.490 0.000483 ***
## URBANICITY_z_Highly.Rural..Rural -2.498e+00  4.346e-01 -5.749 8.97e-09 ***
## HOME_VAL_NA                -4.087e-01  2.279e-01 -1.793 0.072965 .
## oldclaim_log               5.991e-02  3.141e-02  1.907 0.056508 .
## inter                      1.639e-02  4.310e-03  3.803 0.000143 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 721.50  on 624  degrees of freedom
## Residual deviance: 559.49  on 608  degrees of freedom
## AIC: 593.49
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 125  37
##           1  16  14
##
##           Accuracy : 0.724
##           95% CI : (0.655, 0.7859)
##      No Information Rate : 0.7344
##      P-Value [Acc > NIR] : 0.66239
##
##           Kappa : 0.1854
##
## Mcnemar's Test P-Value : 0.00601
##
##           Sensitivity : 0.8865
##           Specificity : 0.2745
##      Pos Pred Value : 0.7716
##      Neg Pred Value : 0.4667
##           Prevalence : 0.7344
##      Detection Rate : 0.6510
##      Detection Prevalence : 0.8438
##      Balanced Accuracy : 0.5805
##
##           'Positive' Class : 0
##

```



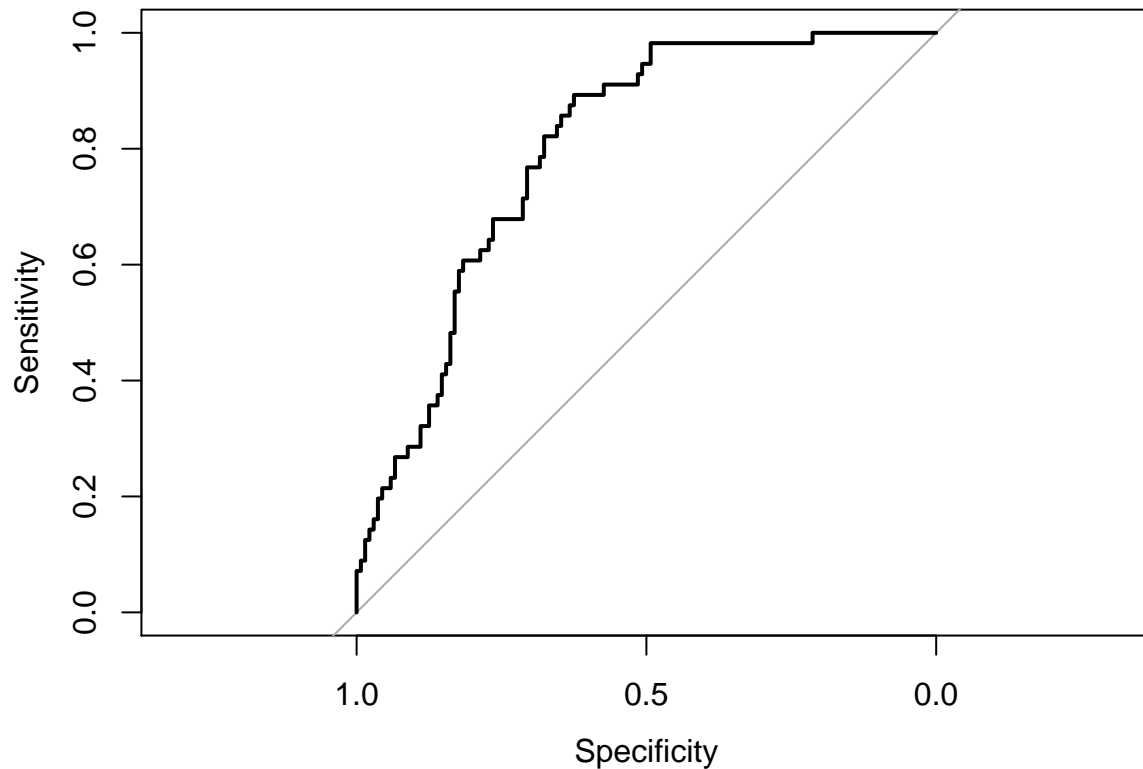
```
## [1] "AUC: 0.774857460714782"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 141 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.7749
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9114  -0.7300  -0.4137   0.5553   2.8509
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.971e+00  4.918e-01  -4.007 6.15e-05 ***
## INCOME        -1.007e-05  3.065e-06  -3.285 0.001020 **
## TRAVTIME       1.735e-02  6.793e-03   2.554 0.010656 *
## BLUEBOOK       2.657e-05  1.577e-05   1.685 0.092025 .
## TIF           -2.035e-02  2.653e-02  -0.767 0.443073
## OLDCLAIM      -2.699e-07  1.511e-05  -0.018 0.985749
## PARENT1_Yes    1.425e+00  3.239e-01   4.398 1.09e-05 ***
## SEX_z_F       -9.182e-01  3.509e-01  -2.617 0.008878 **
## JOB_Manager   -4.454e-01  3.592e-01  -1.240 0.215066
```



```

## CAR_USE_Commercial          4.141e-01  2.406e-01  1.721 0.085258 .
## CAR_TYPE_Pickup             1.059e+00  3.201e-01  3.308 0.000940 ***
## CAR_TYPE_Sports.Car        1.805e+00  4.703e-01  3.839 0.000124 ***
## CAR_TYPE_z_SUV             1.546e+00  4.263e-01  3.626 0.000288 ***
## URBANICITY_z_Highly.Rural..Rural -2.238e+00  3.887e-01 -5.757 8.54e-09 ***
## HOME_VAL_NA                -3.090e-01  2.262e-01 -1.366 0.171810
## oldclaim_log               8.283e-02  3.193e-02  2.595 0.009473 **
## inter                      1.130e-02  4.233e-03  2.668 0.007624 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 711.04  on 624  degrees of freedom
## Residual deviance: 563.20  on 608  degrees of freedom
## AIC: 597.2
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 119  37
##           1  17  19
##
##           Accuracy : 0.7188
##           95% CI : (0.6495, 0.7811)
##      No Information Rate : 0.7083
##      P-Value [Acc > NIR] : 0.409975
##
##           Kappa : 0.2394
##
## Mcnemar's Test P-Value : 0.009722
##
##           Sensitivity : 0.8750
##           Specificity : 0.3393
##      Pos Pred Value : 0.7628
##      Neg Pred Value : 0.5278
##           Prevalence : 0.7083
##      Detection Rate : 0.6198
## Detection Prevalence : 0.8125
##      Balanced Accuracy : 0.6071
##
##           'Positive' Class : 0
##

```

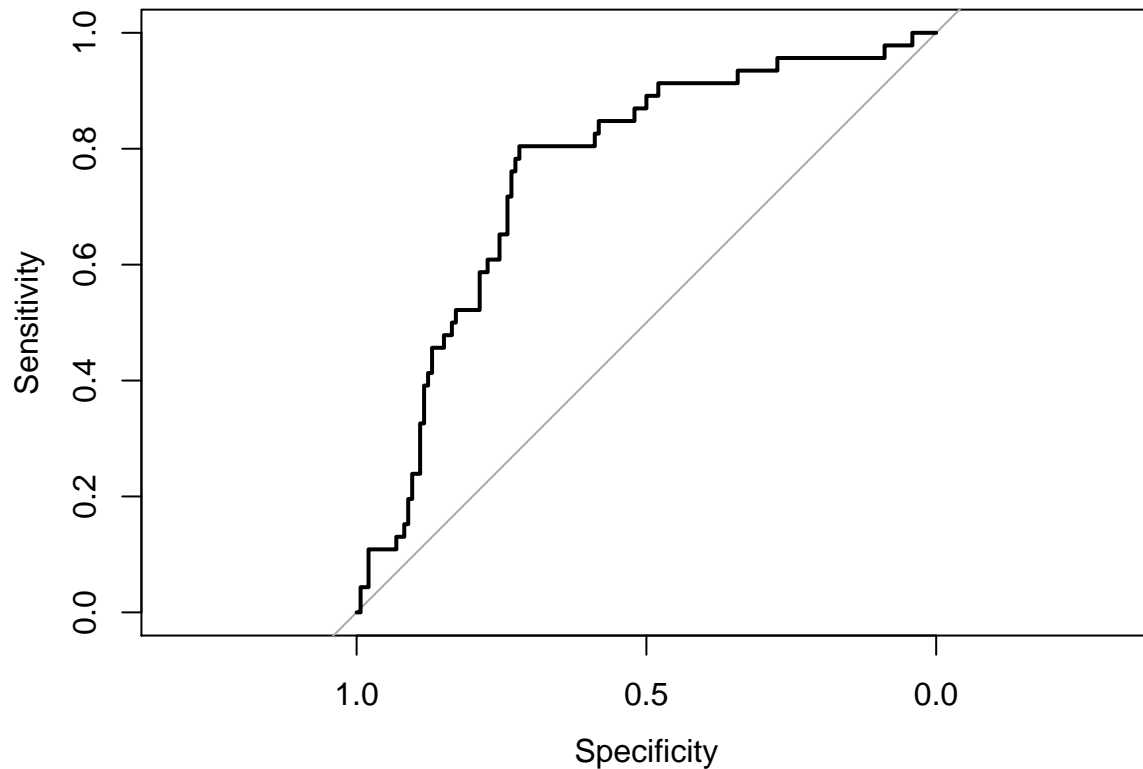


```
## [1] "AUC: 0.799632352941177"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 136 controls (dfPred_raw$class 0) < 56 cases (dfPred_raw$class 1).
## Area under the curve: 0.7996
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7941  -0.7112  -0.3851   0.6539   2.9299
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.698e+00  4.894e-01  -3.469 0.000522 ***
## INCOME        -1.077e-05  3.056e-06  -3.523 0.000426 ***
## TRAVTIME       2.140e-02  7.070e-03   3.026 0.002476 **
## BLUEBOOK       1.734e-05  1.546e-05   1.122 0.261961
## TIF           -3.874e-02  2.770e-02  -1.399 0.161951
## OLDCLAIM       1.421e-05  1.432e-05   0.992 0.321161
## PARENT1_Yes     8.233e-01  3.016e-01   2.730 0.006340 **
## SEX_z_F        -9.769e-01  3.617e-01  -2.701 0.006917 **
## JOB_Manager    -5.512e-01  3.550e-01  -1.552 0.120554
```

```

## CAR_USE_Commercial          5.354e-01  2.480e-01  2.159 0.030845 *
## CAR_TYPE_Pickup             1.015e+00  3.197e-01  3.176 0.001493 **
## CAR_TYPE_Sports.Car         1.866e+00  4.791e-01  3.896 9.80e-05 ***
## CAR_TYPE_z_SUV              1.780e+00  4.355e-01  4.086 4.39e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.436e+00  4.144e-01 -5.880 4.11e-09 ***
## HOME_VAL_NA                 -5.258e-01  2.268e-01 -2.318 0.020445 *
## oldclaim_log                5.344e-02  3.166e-02  1.688 0.091387 .
## inter                       1.999e-02  4.460e-03  4.483 7.37e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 731.55  on 624  degrees of freedom
## Residual deviance: 559.26  on 608  degrees of freedom
## AIC: 593.26
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 130  33
##           1  16  13
##
##           Accuracy : 0.7448
##           95% CI : (0.677, 0.8048)
##      No Information Rate : 0.7604
##      P-Value [Acc > NIR] : 0.72614
##
##           Kappa : 0.1981
##
## Mcnemar's Test P-Value : 0.02227
##
##           Sensitivity : 0.8904
##           Specificity : 0.2826
##      Pos Pred Value : 0.7975
##      Neg Pred Value : 0.4483
##           Prevalence : 0.7604
##      Detection Rate : 0.6771
##      Detection Prevalence : 0.8490
##      Balanced Accuracy : 0.5865
##
##           'Positive' Class : 0
##

```

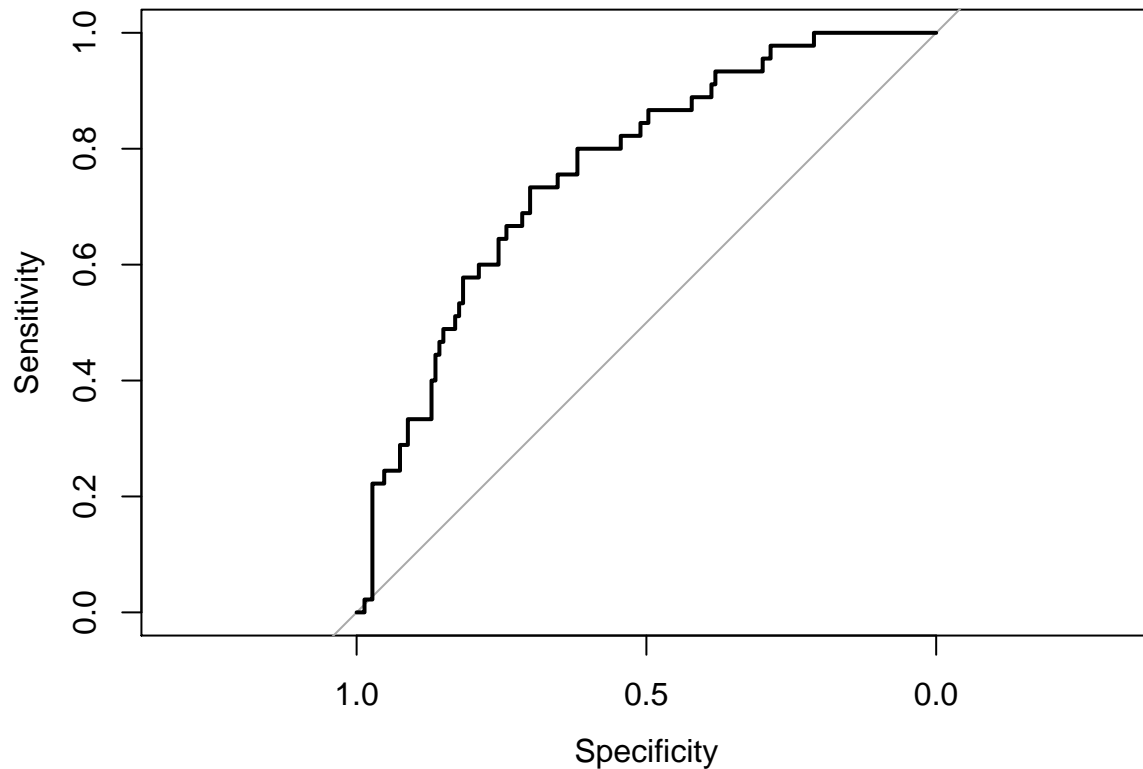


```
## [1] "AUC: 0.759529481834425"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 46 cases (dfPred_raw$class 1).
## Area under the curve: 0.7595
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3770  -0.7355  -0.4018   0.7419   2.8392
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.817e+00  4.827e-01  -3.764 0.000167 ***
## INCOME        -8.878e-06  2.842e-06  -3.124 0.001785 **
## TRAVTIME       2.415e-02  6.965e-03   3.467 0.000526 ***
## BLUEBOOK       2.432e-05  1.529e-05   1.591 0.111674
## TIF           -4.609e-02  2.598e-02  -1.774 0.076085 .
## OLDCLAIM       1.047e-05  1.403e-05   0.747 0.455298
## PARENT1_Yes     7.343e-01  3.008e-01   2.441 0.014631 *
## SEX_z_F        -8.382e-01  3.417e-01  -2.453 0.014172 *
## JOB_Manager    -1.193e+00  4.064e-01  -2.936 0.003324 **
```

```

## CAR_USE_Commercial          3.738e-01  2.378e-01  1.572 0.115951
## CAR_TYPE_Pickup             9.134e-01  3.192e-01  2.861 0.004220 **
## CAR_TYPE_Sports.Car        1.559e+00  4.831e-01  3.227 0.001249 **
## CAR_TYPE_z_SUV             1.531e+00  4.137e-01  3.700 0.000216 ***
## URBANICITY_z_Highly.Rural..Rural -2.598e+00  4.311e-01 -6.026 1.68e-09 ***
## HOME_VAL_NA                -3.111e-01  2.261e-01 -1.376 0.168915
## oldclaim_log               4.518e-02  3.137e-02  1.440 0.149834
## inter                      2.003e-02  4.372e-03  4.581 4.63e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 733.51  on 624  degrees of freedom
## Residual deviance: 571.09  on 608  degrees of freedom
## AIC: 605.09
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 133  30
##           1  14  15
##
##           Accuracy : 0.7708
##           95% CI : (0.7048, 0.8283)
##    No Information Rate : 0.7656
##    P-Value [Acc > NIR] : 0.47206
##
##           Kappa : 0.2716
##
## Mcnemar's Test P-Value : 0.02374
##
##           Sensitivity : 0.9048
##           Specificity : 0.3333
##           Pos Pred Value : 0.8160
##           Neg Pred Value : 0.5172
##           Prevalence : 0.7656
##           Detection Rate : 0.6927
##    Detection Prevalence : 0.8490
##           Balanced Accuracy : 0.6190
##
##           'Positive' Class : 0
##

```

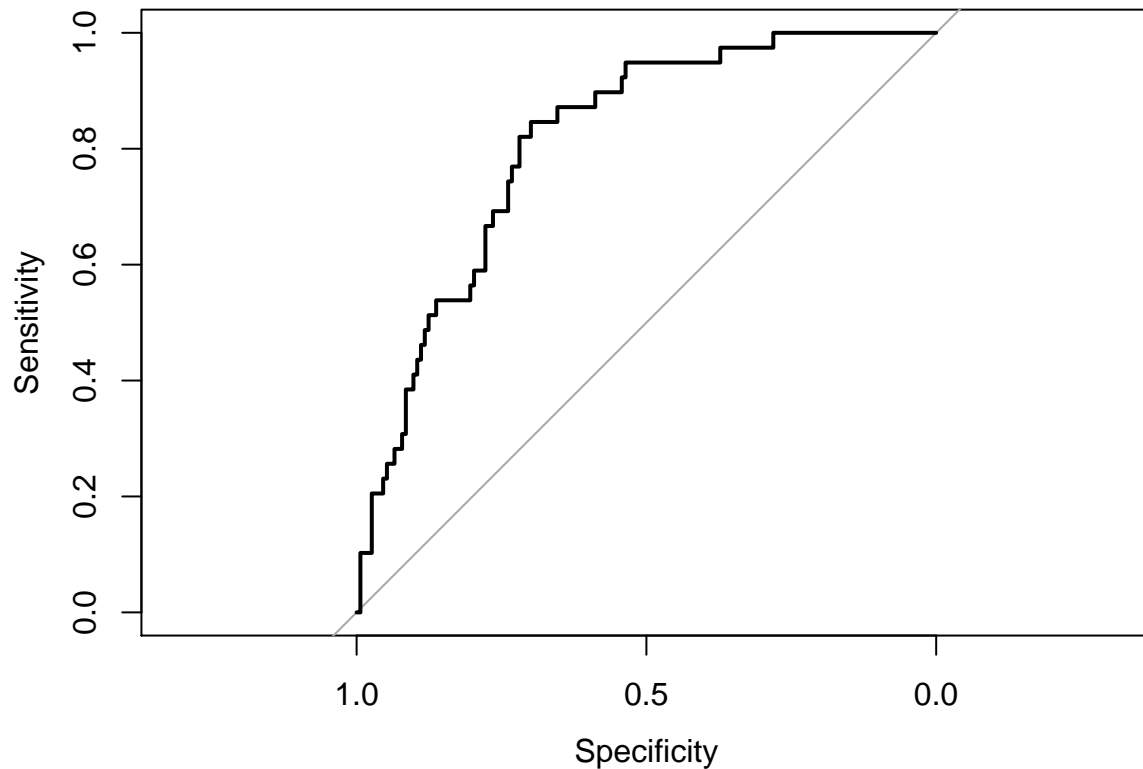


```
## [1] "AUC: 0.761904761904762"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 147 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7619
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8334  -0.7564  -0.4201   0.7303   2.8316
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.923e+00  4.740e-01  -4.058 4.95e-05 ***
## INCOME        -1.126e-05  2.989e-06  -3.766 0.000166 ***
## TRAVTIME       2.171e-02  7.006e-03   3.098 0.001945 **
## BLUEBOOK       3.074e-05  1.533e-05   2.005 0.045005 *
## TIF           -3.023e-02  2.629e-02  -1.150 0.250182
## OLDCLAIM       9.090e-06  1.379e-05   0.659 0.509807
## PARENT1_Yes    1.022e+00  3.017e-01   3.386 0.000709 ***
## SEX_z_F       -9.067e-01  3.482e-01  -2.604 0.009221 **
## JOB_Manager   -4.584e-01  3.366e-01  -1.362 0.173264
```

```

## CAR_USE_Commercial          4.853e-01  2.370e-01  2.048 0.040572 *
## CAR_TYPE_Pickup             8.751e-01  3.105e-01  2.818 0.004834 **
## CAR_TYPE_Sports.Car        1.832e+00  4.701e-01  3.896 9.78e-05 ***
## CAR_TYPE_z_SUV             1.672e+00  4.187e-01  3.994 6.49e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.382e+00  3.858e-01 -6.173 6.70e-10 ***
## HOME_VAL_NA                -3.053e-01  2.196e-01 -1.390 0.164427
## oldclaim_log                6.022e-02  3.109e-02  1.937 0.052730 .
## inter                       1.499e-02  4.006e-03  3.743 0.000182 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 744.94 on 624 degrees of freedom
## Residual deviance: 585.54 on 608 degrees of freedom
## AIC: 619.54
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 135  21
##           1  18  18
##
##           Accuracy : 0.7969
##           95% CI : (0.733, 0.8514)
##           No Information Rate : 0.7969
##           P-Value [Acc > NIR] : 0.5427
##
##           Kappa : 0.354
##
## Mcnemar's Test P-Value : 0.7488
##
##           Sensitivity : 0.8824
##           Specificity : 0.4615
##           Pos Pred Value : 0.8654
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7969
##           Detection Rate : 0.7031
##           Detection Prevalence : 0.8125
##           Balanced Accuracy : 0.6719
##
##           'Positive' Class : 0
##

```



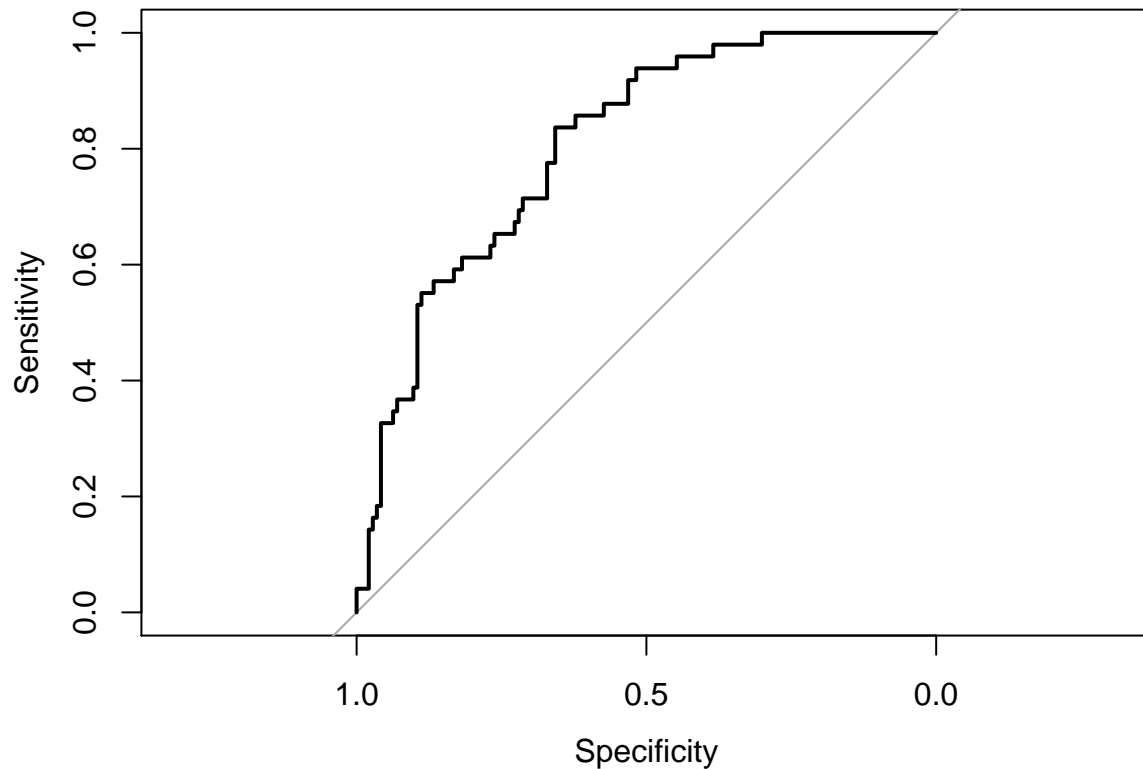
```
## [1] "AUC: 0.812803753980225"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 153 controls (dfPred_raw$class 0) < 39 cases (dfPred_raw$class 1).
## Area under the curve: 0.8128
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9881  -0.7467  -0.4160   0.6650   2.7817
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.847e+00  4.949e-01  -3.733 0.000190 ***
## INCOME        -1.048e-05  2.883e-06  -3.636 0.000276 ***
## TRAVTIME       1.914e-02  6.845e-03   2.796 0.005171 **
## BLUEBOOK       3.088e-05  1.553e-05   1.989 0.046694 *
## TIF           -5.339e-02  2.620e-02  -2.038 0.041598 *
## OLDCLAIM       1.619e-05  1.355e-05   1.195 0.232187
## PARENT1_Yes     7.986e-01  3.125e-01   2.555 0.010604 *
## SEX_z_F        -9.636e-01  3.475e-01  -2.773 0.005551 **
## JOB_Manager    -7.432e-01  3.666e-01  -2.027 0.042626 *
```



```

## CAR_USE_Commercial          5.644e-01  2.372e-01  2.380 0.017329 *
## CAR_TYPE_Pickup             1.103e+00  3.251e-01  3.393 0.000691 ***
## CAR_TYPE_Sports.Car        2.104e+00  4.865e-01  4.325 1.52e-05 ***
## CAR_TYPE_z_SUV             1.774e+00  4.265e-01  4.158 3.20e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.351e+00  3.966e-01 -5.926 3.10e-09 ***
## HOME_VAL_NA                 -2.063e-01  2.214e-01 -0.932 0.351500
## oldclaim_log                2.386e-02  3.177e-02  0.751 0.452528
## inter                       1.382e-02  4.160e-03  3.323 0.000892 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 725.57 on 624 degrees of freedom
## Residual deviance: 576.88 on 608 degrees of freedom
## AIC: 610.88
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 131  31
##           1  12  18
##
##           Accuracy : 0.776
##           95% CI : (0.7104, 0.8329)
##           No Information Rate : 0.7448
##           P-Value [Acc > NIR] : 0.181886
##
##           Kappa : 0.3248
##
## Mcnemar's Test P-Value : 0.006052
##
##           Sensitivity : 0.9161
##           Specificity : 0.3673
##           Pos Pred Value : 0.8086
##           Neg Pred Value : 0.6000
##           Prevalence : 0.7448
##           Detection Rate : 0.6823
##           Detection Prevalence : 0.8438
##           Balanced Accuracy : 0.6417
##
##           'Positive' Class : 0
##

```

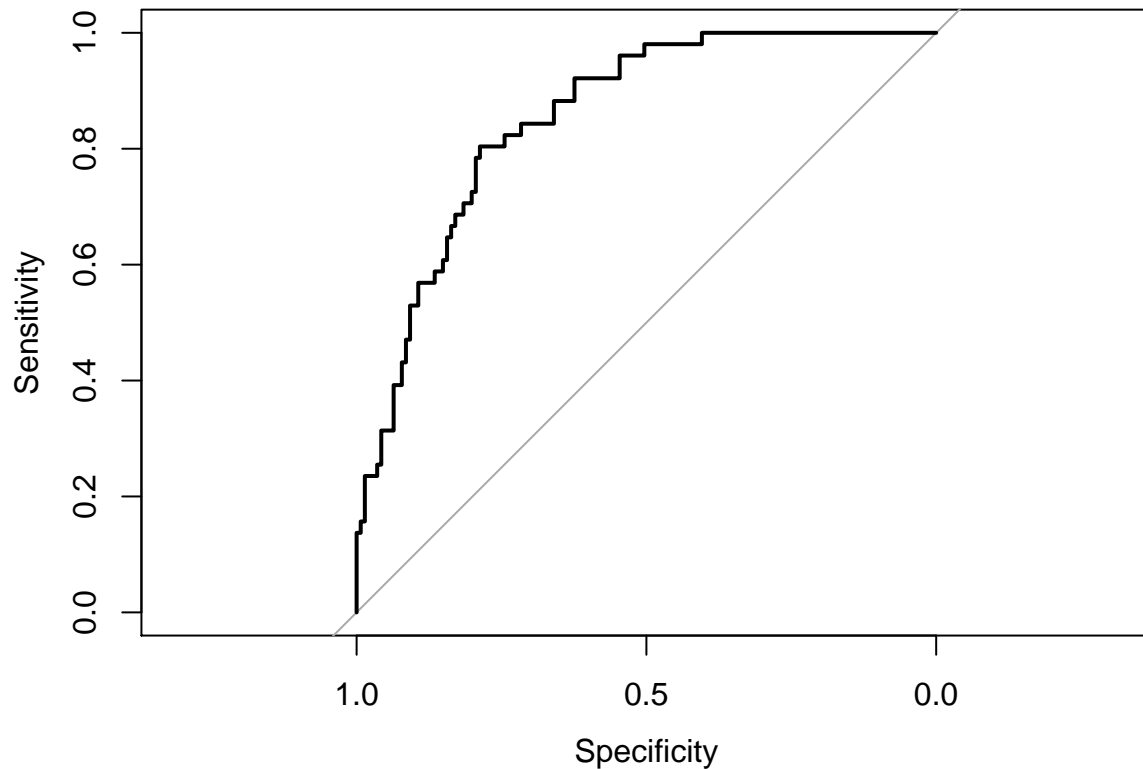


```
## [1] "AUC: 0.807478236049665"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 143 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.8075
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8026  -0.7476  -0.4501   0.6485   2.7165
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.913e+00  4.737e-01  -4.039 5.37e-05 ***
## INCOME        -8.604e-06  2.890e-06  -2.977 0.002908 **
## TRAVTIME       2.060e-02  7.012e-03   2.937 0.003312 **
## BLUEBOOK       2.648e-05  1.561e-05   1.696 0.089820 .
## TIF           -3.745e-02  2.560e-02  -1.463 0.143568
## OLDCLAIM       2.326e-06  1.498e-05   0.155 0.876603
## PARENT1_Yes    1.324e+00  3.088e-01   4.287 1.81e-05 ***
## SEX_z_F       -9.257e-01  3.479e-01  -2.661 0.007789 **
## JOB_Manager   -5.635e-01  3.668e-01  -1.536 0.124466
```

```

## CAR_USE_Commercial          4.116e-01  2.351e-01  1.751 0.079958 .
## CAR_TYPE_Pickup             9.038e-01  3.141e-01  2.877 0.004013 **
## CAR_TYPE_Sports.Car        1.622e+00  4.736e-01  3.425 0.000614 ***
## CAR_TYPE_z_SUV             1.643e+00  4.222e-01  3.890 0.000100 ***
## URBANICITY_z_Highly.Rural..Rural -2.091e+00  3.728e-01 -5.607 2.05e-08 ***
## HOME_VAL_NA                -3.154e-01  2.166e-01 -1.457 0.145209
## oldclaim_log               5.736e-02  3.135e-02  1.830 0.067311 .
## inter                      1.083e-02  4.115e-03  2.632 0.008482 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 721.50 on 624 degrees of freedom
## Residual deviance: 587.42 on 608 degrees of freedom
## AIC: 621.42
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 130  31
##           1  11  20
##
##           Accuracy : 0.7812
##           95% CI : (0.716, 0.8376)
##           No Information Rate : 0.7344
##           P-Value [Acc > NIR] : 0.08041
##
##           Kappa : 0.3591
##
## Mcnemar's Test P-Value : 0.00337
##
##           Sensitivity : 0.9220
##           Specificity : 0.3922
##           Pos Pred Value : 0.8075
##           Neg Pred Value : 0.6452
##           Prevalence : 0.7344
##           Detection Rate : 0.6771
##           Detection Prevalence : 0.8385
##           Balanced Accuracy : 0.6571
##
##           'Positive' Class : 0
##

```

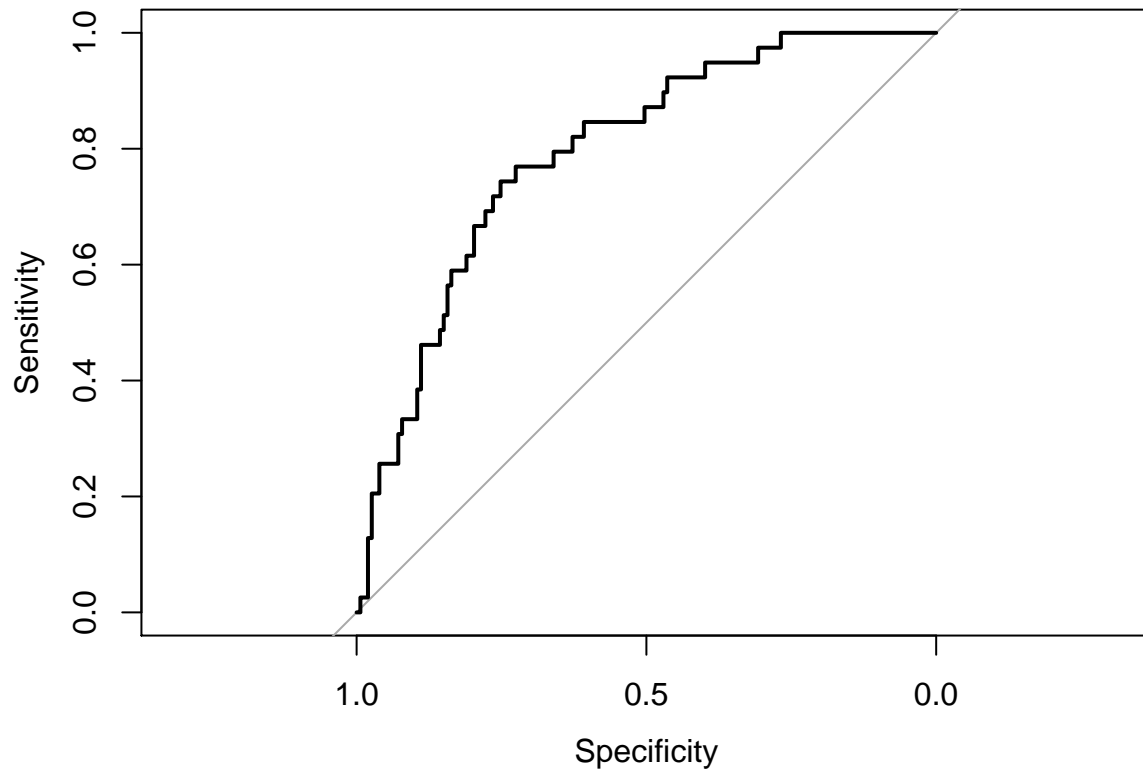


```
## [1] "AUC: 0.854679460436657"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 141 controls (dfPred_raw$class 0) < 51 cases (dfPred_raw$class 1).
## Area under the curve: 0.8547
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7724  -0.7407  -0.4217   0.7646   2.7838
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.892e+00  4.790e-01  -3.950 7.80e-05 ***
## INCOME        -1.002e-05  2.859e-06  -3.505 0.000456 ***
## TRAVTIME       2.733e-02  7.191e-03   3.800 0.000145 ***
## BLUEBOOK       3.040e-05  1.542e-05   1.972 0.048615 *
## TIF           -5.813e-02  2.636e-02  -2.206 0.027408 *
## OLDCLAIM       9.390e-06  1.337e-05   0.702 0.482499
## PARENT1_Yes    7.093e-01  2.953e-01   2.402 0.016300 *
## SEX_z_F       -8.738e-01  3.544e-01  -2.466 0.013678 *
## JOB_Manager   -6.226e-01  3.457e-01  -1.801 0.071731 .
```

```

## CAR_USE_Commercial          4.226e-01  2.390e-01  1.768 0.076983 .
## CAR_TYPE_Pickup             8.169e-01  3.209e-01  2.546 0.010895 *
## CAR_TYPE_Sports.Car        1.551e+00  4.794e-01  3.235 0.001215 **
## CAR_TYPE_z_SUV             1.726e+00  4.233e-01  4.076 4.58e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.374e+00  3.990e-01 -5.951 2.67e-09 ***
## HOME_VAL_NA                -4.443e-01  2.202e-01 -2.018 0.043567 *
## oldclaim_log               5.219e-02  3.122e-02  1.672 0.094552 .
## inter                      2.017e-02  4.449e-03  4.533 5.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 744.94  on 624  degrees of freedom
## Residual deviance: 585.46  on 608  degrees of freedom
## AIC: 619.46
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 136  23
##           1  17  16
##
##           Accuracy : 0.7917
##           95% CI : (0.7273, 0.8468)
##    No Information Rate : 0.7969
##    P-Value [Acc > NIR] : 0.6123
##
##           Kappa : 0.3173
##
## Mcnemar's Test P-Value : 0.4292
##
##           Sensitivity : 0.8889
##           Specificity : 0.4103
##    Pos Pred Value : 0.8553
##    Neg Pred Value : 0.4848
##           Prevalence : 0.7969
##    Detection Rate : 0.7083
##    Detection Prevalence : 0.8281
##    Balanced Accuracy : 0.6496
##
##           'Positive' Class : 0
##

```

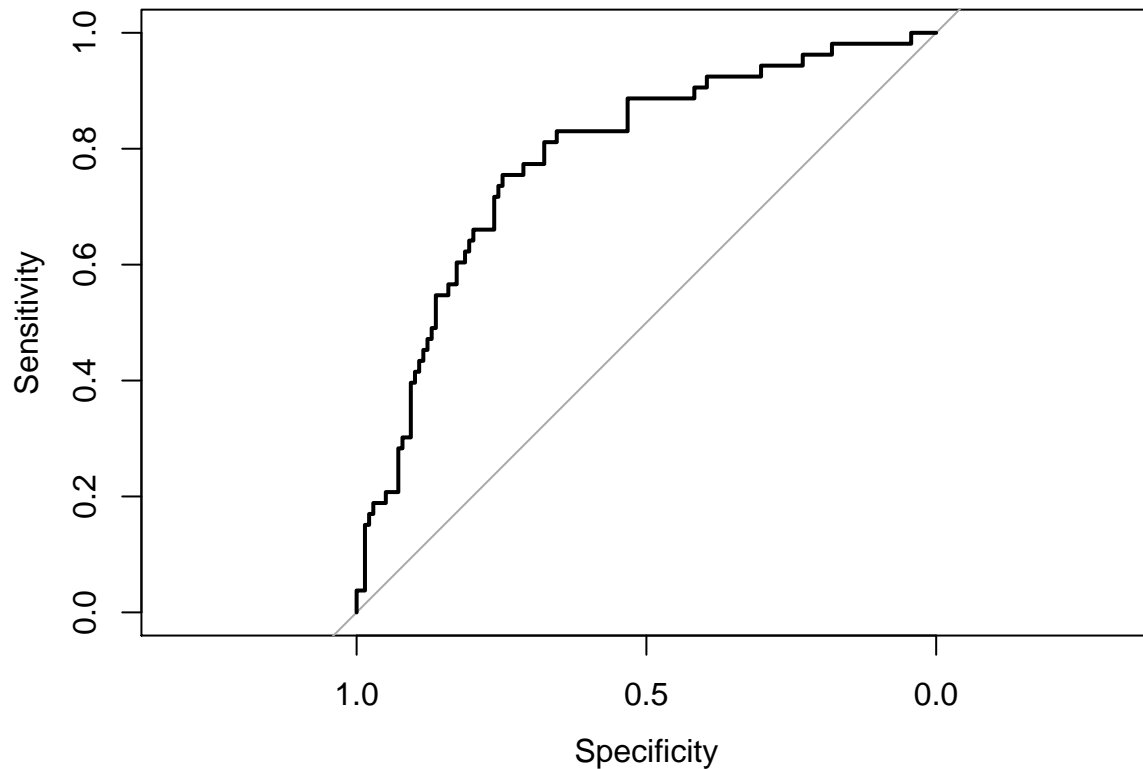


```
## [1] "AUC: 0.793698676051617"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 153 controls (dfPred_raw$class 0) < 39 cases (dfPred_raw$class 1).
## Area under the curve: 0.7937
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1486  -0.7266  -0.4031   0.5873   2.9770
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.763e+00  4.901e-01  -3.598 0.000320 ***
## INCOME        -9.336e-06  2.981e-06  -3.132 0.001738 **
## TRAVTIME       2.338e-02  7.109e-03   3.289 0.001005 **
## BLUEBOOK       1.135e-05  1.551e-05   0.732 0.464277
## TIF           -3.221e-02  2.624e-02  -1.227 0.219668
## OLDCLAIM       9.998e-06  1.403e-05   0.712 0.476205
## PARENT1_Yes    8.597e-01  2.982e-01   2.883 0.003944 **
## SEX_z_F       -5.769e-01  3.405e-01  -1.695 0.090157 .
## JOB_Manager   -5.771e-01  3.493e-01  -1.652 0.098469 .
```

```

## CAR_USE_Commercial          5.021e-01  2.430e-01  2.066 0.038802 *
## CAR_TYPE_Pickup             7.599e-01  3.193e-01  2.380 0.017318 *
## CAR_TYPE_Sports.Car        1.232e+00  4.643e-01  2.653 0.007983 **
## CAR_TYPE_z_SUV             1.308e+00  4.077e-01  3.209 0.001334 **
## URBANICITY_z_Highly.Rural..Rural -2.614e+00  4.607e-01 -5.675 1.39e-08 ***
## HOME_VAL_NA                -4.742e-01  2.261e-01 -2.097 0.036002 *
## oldclaim_log               5.731e-02  3.145e-02  1.822 0.068440 .
## inter                      1.749e-02  4.510e-03  3.878 0.000105 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 717.37 on 624 degrees of freedom
## Residual deviance: 559.97 on 608 degrees of freedom
## AIC: 593.97
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 129  39
##           1  10  14
##
##           Accuracy : 0.7448
##           95% CI : (0.677, 0.8048)
##           No Information Rate : 0.724
##           P-Value [Acc > NIR] : 0.2889
##
##           Kappa : 0.2314
##
## Mcnemar's Test P-Value : 6.334e-05
##
##           Sensitivity : 0.9281
##           Specificity : 0.2642
##           Pos Pred Value : 0.7679
##           Neg Pred Value : 0.5833
##           Prevalence : 0.7240
##           Detection Rate : 0.6719
##           Detection Prevalence : 0.8750
##           Balanced Accuracy : 0.5961
##
##           'Positive' Class : 0
##

```



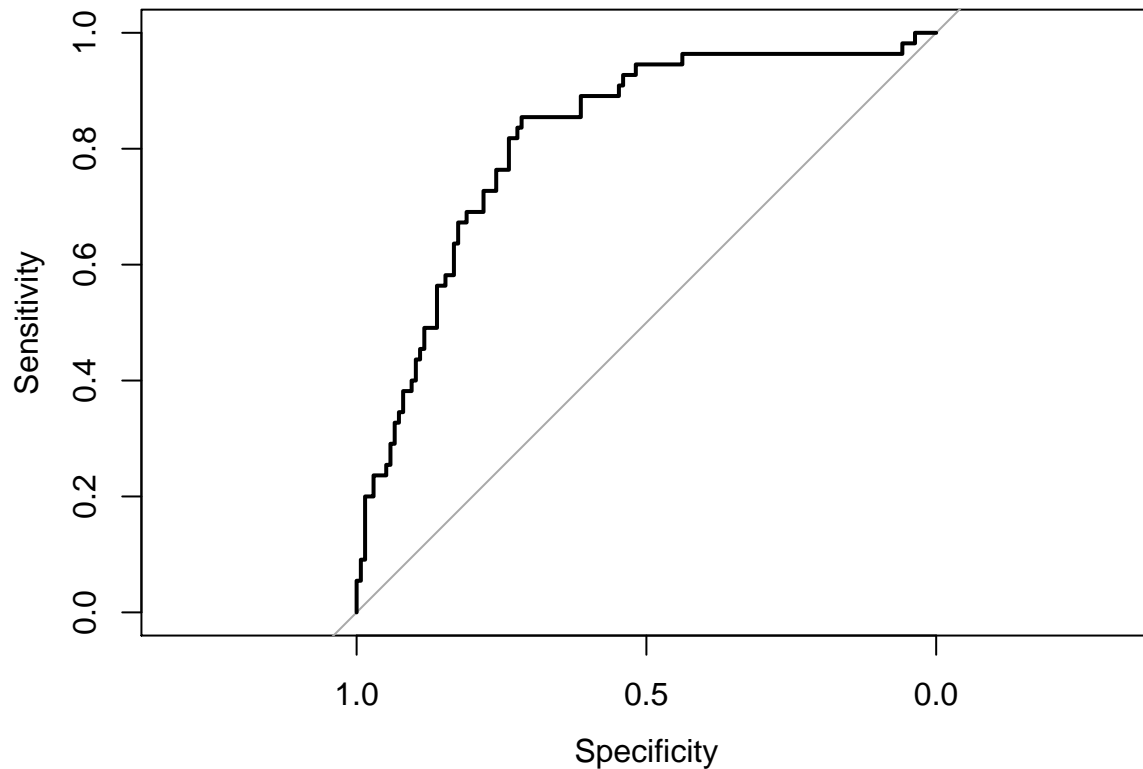
```
## [1] "AUC: 0.784579883263201"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 139 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.7846
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7357  -0.7280  -0.4306   0.6358   2.7874
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.868e+00  4.895e-01  -3.816 0.000136 ***
## INCOME        -1.020e-05  2.977e-06  -3.427 0.000610 ***
## TRAVTIME       2.180e-02  6.915e-03   3.153 0.001616 **
## BLUEBOOK       1.727e-05  1.615e-05   1.070 0.284779
## TIF           -3.781e-02  2.598e-02  -1.456 0.145523
## OLDCLAIM       4.286e-06  1.422e-05   0.301 0.763134
## PARENT1_Yes    1.146e+00  3.011e-01   3.806 0.000141 ***
## SEX_z_F       -7.986e-01  3.459e-01  -2.309 0.020938 *
## JOB_Manager   -1.573e-01  3.483e-01  -0.452 0.651546
```



```

## CAR_USE_Commercial          3.304e-01  2.360e-01  1.400 0.161610
## CAR_TYPE_Pickup             1.122e+00  3.203e-01  3.503 0.000459 ***
## CAR_TYPE_Sports.Car        1.596e+00  4.636e-01  3.442 0.000577 ***
## CAR_TYPE_z_SUV             1.605e+00  4.167e-01  3.851 0.000118 ***
## URBANICITY_z_Highly.Rural..Rural -2.068e+00  3.866e-01 -5.348 8.91e-08 ***
## HOME_VAL_NA                -3.577e-01  2.166e-01 -1.652 0.098632 .
## oldclaim_log               6.795e-02  3.188e-02  2.132 0.033040 *
## inter                      7.780e-03  4.596e-03  1.693 0.090469 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 713.16 on 624 degrees of freedom
## Residual deviance: 574.13 on 608 degrees of freedom
## AIC: 608.13
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 129  40
##           1   8  15
##
##           Accuracy : 0.75
##           95% CI : (0.6826, 0.8096)
##           No Information Rate : 0.7135
##           P-Value [Acc > NIR] : 0.1495
##
##           Kappa : 0.2595
##
## Mcnemar's Test P-Value : 7.66e-06
##
##           Sensitivity : 0.9416
##           Specificity : 0.2727
##           Pos Pred Value : 0.7633
##           Neg Pred Value : 0.6522
##           Prevalence : 0.7135
##           Detection Rate : 0.6719
##           Detection Prevalence : 0.8802
##           Balanced Accuracy : 0.6072
##
##           'Positive' Class : 0
##

```

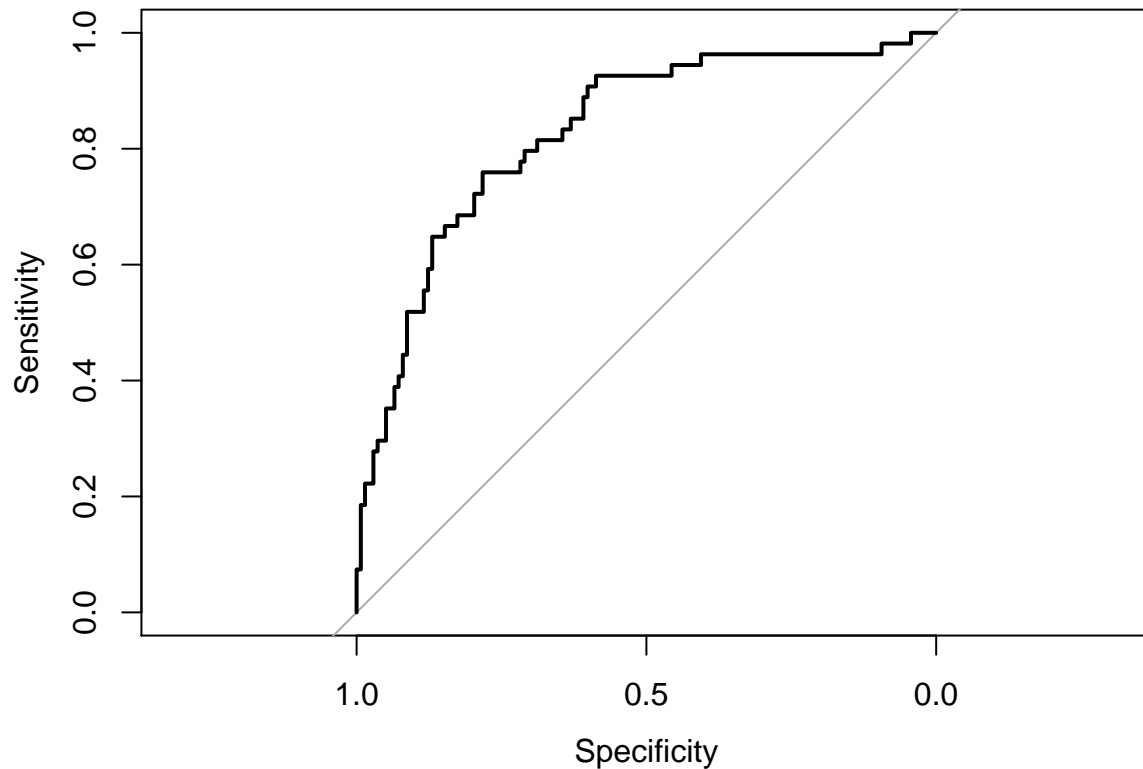


```
## [1] "AUC: 0.818314532183145"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 137 controls (dfPred_raw$class 0) < 55 cases (dfPred_raw$class 1).
## Area under the curve: 0.8183
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7807  -0.7358  -0.4273   0.6502   2.7823
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.737e+00  4.799e-01  -3.619 0.000296 ***
## INCOME        -8.284e-06  2.898e-06  -2.859 0.004252 **
## TRAVTIME       2.093e-02  6.847e-03   3.057 0.002234 **
## BLUEBOOK       1.217e-05  1.562e-05   0.779 0.436102
## TIF           -3.446e-02  2.591e-02  -1.330 0.183435
## OLDCLAIM       2.498e-06  1.557e-05   0.160 0.872593
## PARENT1_Yes    1.106e+00  2.998e-01   3.690 0.000224 ***
## SEX_z_F       -9.121e-01  3.530e-01  -2.584 0.009775 **
## JOB_Manager    -7.889e-01  4.028e-01  -1.958 0.050193 .
```

```

## CAR_USE_Commercial          4.924e-01  2.344e-01  2.100 0.035687 *
## CAR_TYPE_Pickup             9.821e-01  3.141e-01  3.126 0.001771 **
## CAR_TYPE_Sports.Car        1.745e+00  4.771e-01  3.658 0.000254 ***
## CAR_TYPE_z_SUV             1.627e+00  4.261e-01  3.818 0.000134 ***
## URBANICITY_z_Highly.Rural..Rural -2.183e+00  3.965e-01 -5.506 3.67e-08 ***
## HOME_VAL_NA                -3.529e-01  2.188e-01 -1.613 0.106840
## oldclaim_log               4.277e-02  3.189e-02  1.341 0.179875
## inter                      1.061e-02  4.219e-03  2.514 0.011940 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 715.27 on 624 degrees of freedom
## Residual deviance: 573.89 on 608 degrees of freedom
## AIC: 607.89
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 134  42
##           1   4  12
##
##           Accuracy : 0.7604
##           95% CI : (0.6937, 0.8189)
##           No Information Rate : 0.7188
##           P-Value [Acc > NIR] : 0.1132
##
##           Kappa : 0.2459
##
## Mcnemar's Test P-Value : 4.888e-08
##
##           Sensitivity : 0.9710
##           Specificity : 0.2222
##           Pos Pred Value : 0.7614
##           Neg Pred Value : 0.7500
##           Prevalence : 0.7188
##           Detection Rate : 0.6979
##           Detection Prevalence : 0.9167
##           Balanced Accuracy : 0.5966
##
##           'Positive' Class : 0
##

```

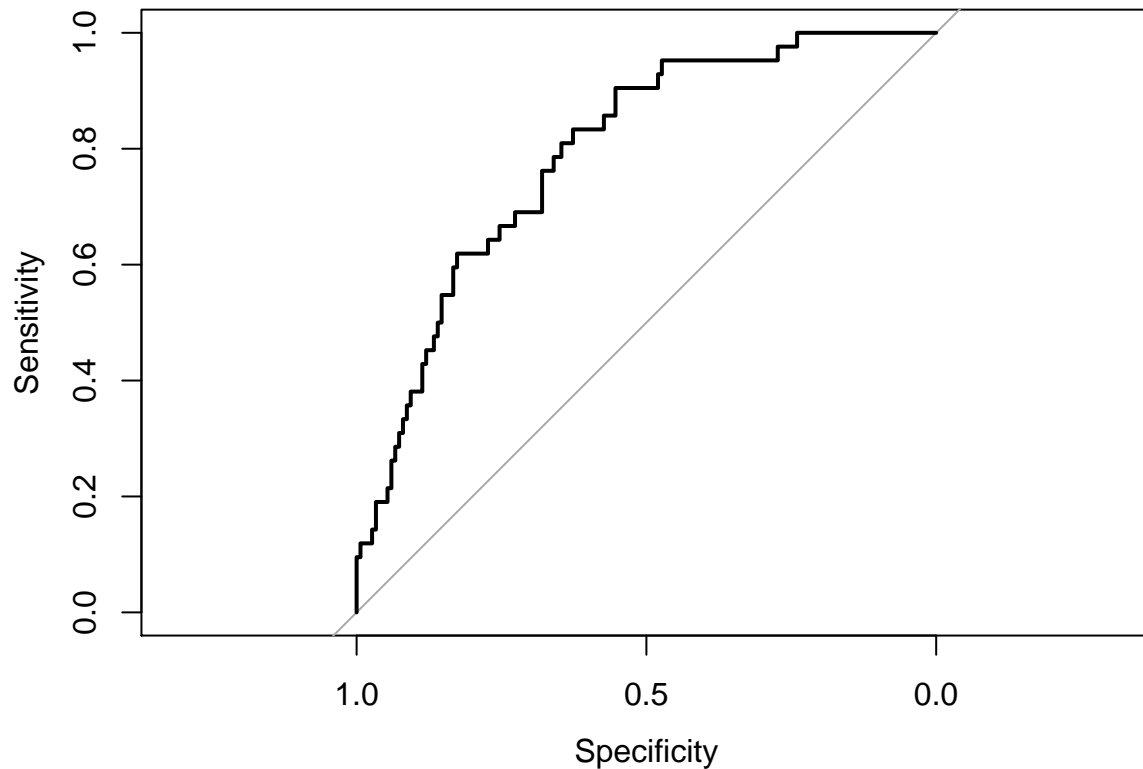


```
## [1] "AUC: 0.827831454643049"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 54 cases (dfPred_raw$class 1).
## Area under the curve: 0.8278
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0958  -0.7566  -0.4117   0.6614   2.8427
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.152e+00  4.874e-01  -4.414 1.01e-05 ***
## INCOME        -1.070e-05  2.885e-06  -3.710 0.000208 ***
## TRAVTIME       2.290e-02  6.844e-03   3.346 0.000819 ***
## BLUEBOOK       3.833e-05  1.537e-05   2.493 0.012663 *
## TIF           -3.745e-02  2.650e-02  -1.413 0.157683
## OLDCLAIM       7.830e-06  1.406e-05   0.557 0.577472
## PARENT1_Yes    1.230e+00  3.109e-01   3.958 7.55e-05 ***
## SEX_z_F       -8.921e-01  3.413e-01  -2.613 0.008965 **
## JOB_Manager   -8.559e-01  3.612e-01  -2.369 0.017817 *
```

```

## CAR_USE_Commercial          3.898e-01  2.385e-01  1.634 0.102215
## CAR_TYPE_Pickup             9.888e-01  3.180e-01  3.110 0.001871 **
## CAR_TYPE_Sports.Car        1.883e+00  4.798e-01  3.924 8.72e-05 ***
## CAR_TYPE_z_SUV             1.597e+00  4.163e-01  3.837 0.000125 ***
## URBANICITY_z_Highly.Rural..Rural -2.567e+00  4.046e-01 -6.345 2.22e-10 ***
## HOME_VAL_NA                -1.322e-01  2.245e-01 -0.589 0.555942
## oldclaim_log               6.070e-02  3.116e-02  1.948 0.051391 .
## inter                      1.559e-02  4.051e-03  3.848 0.000119 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 739.29 on 624 degrees of freedom
## Residual deviance: 576.91 on 608 degrees of freedom
## AIC: 610.91
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 131  23
##           1  19  19
##
##           Accuracy : 0.7812
##           95% CI : (0.716, 0.8376)
##           No Information Rate : 0.7812
##           P-Value [Acc > NIR] : 0.5412
##
##           Kappa : 0.3373
##
## Mcnemar's Test P-Value : 0.6434
##
##           Sensitivity : 0.8733
##           Specificity : 0.4524
##           Pos Pred Value : 0.8506
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7812
##           Detection Rate : 0.6823
##           Detection Prevalence : 0.8021
##           Balanced Accuracy : 0.6629
##
##           'Positive' Class : 0
##

```

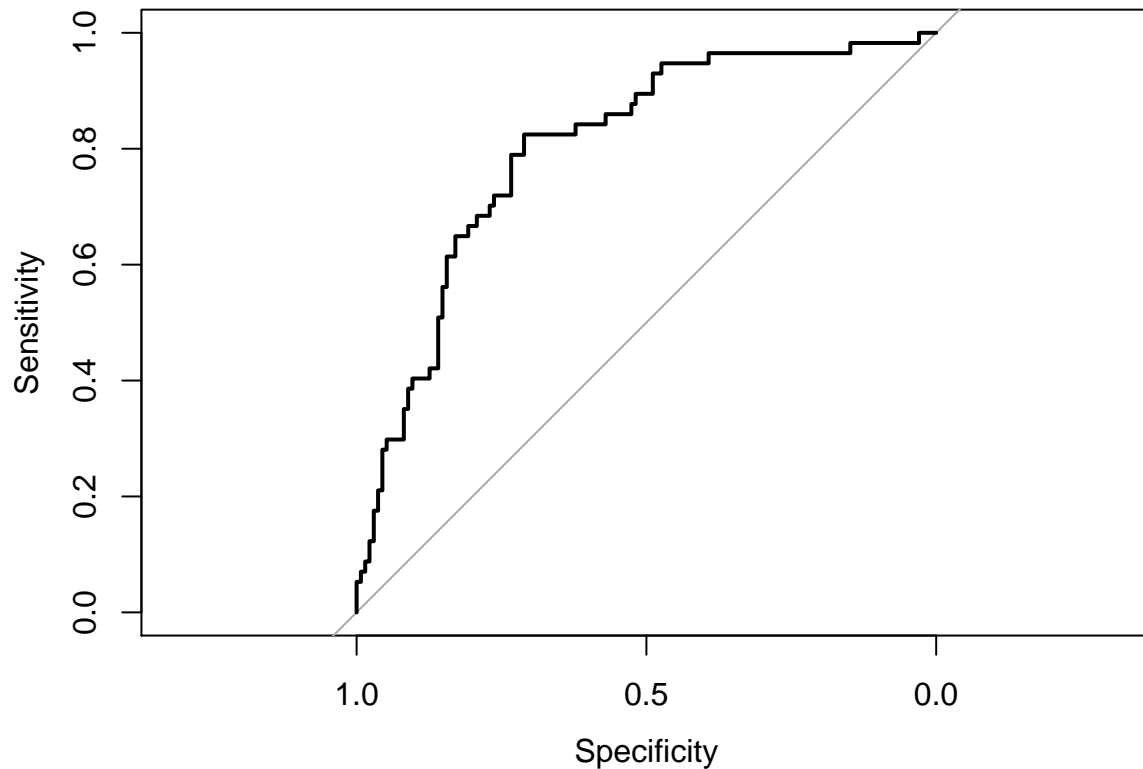


```
## [1] "AUC: 0.792380952380952"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 150 controls (dfPred_raw$class 0) < 42 cases (dfPred_raw$class 1).
## Area under the curve: 0.7924
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9804  -0.7108  -0.4146   0.4990   2.8306
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.832e+00  4.961e-01  -3.693 0.000222 ***
## INCOME        -8.948e-06  2.985e-06  -2.998 0.002720 **
## TRAVTIME       2.179e-02  6.926e-03   3.147 0.001650 **
## BLUEBOOK       1.584e-05  1.565e-05   1.012 0.311300
## TIF           -4.045e-02  2.680e-02  -1.509 0.131252
## OLDCLAIM       1.353e-05  1.345e-05   1.006 0.314378
## PARENT1_Yes     9.578e-01  3.157e-01   3.034 0.002412 **
## SEX_z_F        -9.359e-01  3.555e-01  -2.633 0.008465 **
## JOB_Manager    -6.437e-01  3.581e-01  -1.798 0.072227 .
```

```

## CAR_USE_Commercial          2.873e-01  2.429e-01  1.182 0.237014
## CAR_TYPE_Pickup             1.076e+00  3.253e-01  3.307 0.000944 ***
## CAR_TYPE_Sports.Car         1.696e+00  4.706e-01  3.605 0.000312 ***
## CAR_TYPE_z_SUV              1.512e+00  4.282e-01  3.531 0.000415 ***
## URBANICITY_z_Highly.Rural..Rural -2.261e+00  4.143e-01 -5.456 4.86e-08 ***
## HOME_VAL_NA                 -3.706e-01  2.260e-01 -1.640 0.101098
## oldclaim_log                7.145e-02  3.174e-02  2.251 0.024379 *
## inter                       1.585e-02  4.755e-03  3.333 0.000858 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 708.90  on 624  degrees of freedom
## Residual deviance: 559.98  on 608  degrees of freedom
## AIC: 593.98
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 123  35
##           1  12  22
##
##           Accuracy : 0.7552
##           95% CI : (0.6881, 0.8143)
##      No Information Rate : 0.7031
##      P-Value [Acc > NIR] : 0.064791
##
##           Kappa : 0.3363
##
## Mcnemar's Test P-Value : 0.001332
##
##           Sensitivity : 0.9111
##           Specificity : 0.3860
##      Pos Pred Value : 0.7785
##      Neg Pred Value : 0.6471
##           Prevalence : 0.7031
##      Detection Rate : 0.6406
##      Detection Prevalence : 0.8229
##      Balanced Accuracy : 0.6485
##
##           'Positive' Class : 0
##

```



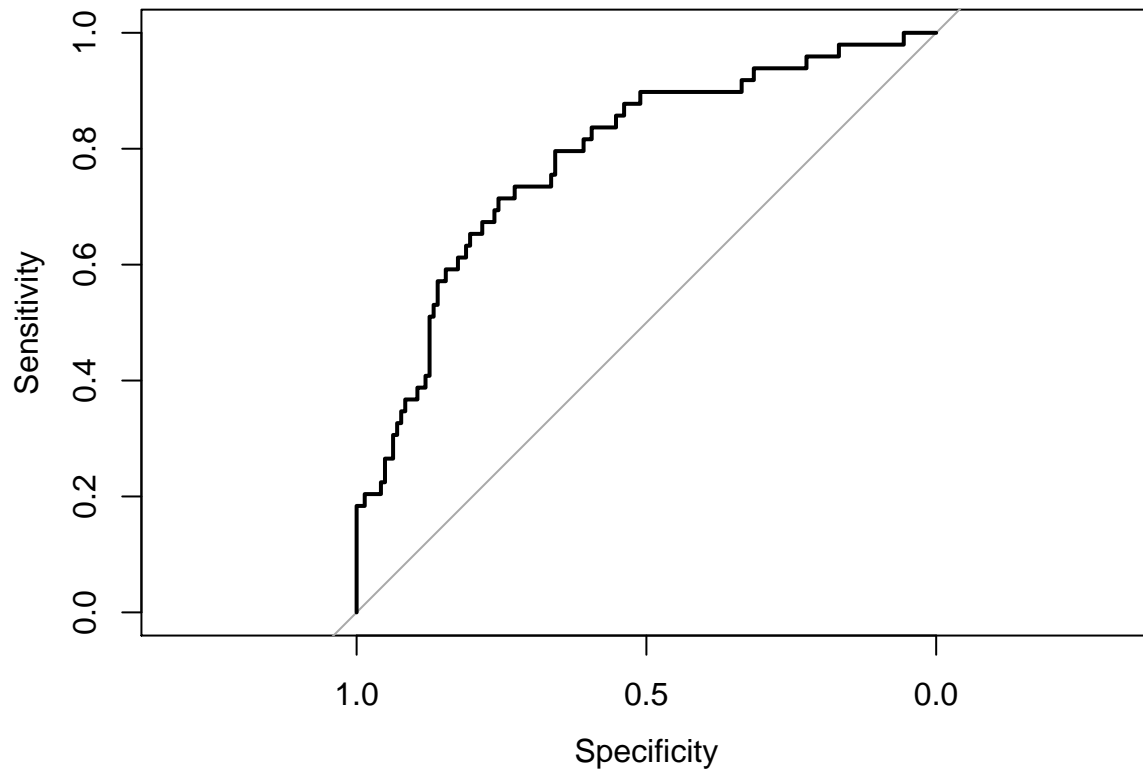
```
## [1] "AUC: 0.801559454191033"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 135 controls (dfPred_raw$class 0) < 57 cases (dfPred_raw$class 1).
## Area under the curve: 0.8016
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9903  -0.7556  -0.4162   0.7066   2.9744
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.794e+00  4.823e-01  -3.720 0.000199 ***
## INCOME        -1.019e-05  2.941e-06  -3.464 0.000533 ***
## TRAVTIME       2.433e-02  6.929e-03   3.512 0.000445 ***
## BLUEBOOK       1.113e-05  1.539e-05   0.723 0.469387
## TIF           -2.109e-02  2.542e-02  -0.830 0.406815
## OLDCLAIM       8.392e-06  1.402e-05   0.599 0.549389
## PARENT1_Yes     8.373e-01  2.901e-01   2.886 0.003897 **
## SEX_z_F        -5.201e-01  3.341e-01  -1.557 0.119565
## JOB_Manager    -6.269e-01  3.530e-01  -1.776 0.075747 .
```



```

## CAR_USE_Commercial          5.567e-01  2.336e-01  2.383 0.017175 *
## CAR_TYPE_Pickup             7.131e-01  3.128e-01  2.280 0.022635 *
## CAR_TYPE_Sports.Car        1.386e+00  4.629e-01  2.994 0.002755 **
## CAR_TYPE_z_SUV             1.221e+00  3.987e-01  3.062 0.002196 **
## URBANICITY_z_Highly.Rural..Rural -2.739e+00  4.653e-01 -5.887 3.94e-09 ***
## HOME_VAL_NA                -2.997e-01  2.224e-01 -1.348 0.177813
## oldclaim_log               4.241e-02  3.146e-02  1.348 0.177613
## inter                      1.298e-02  4.156e-03  3.123 0.001788 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 725.57 on 624 degrees of freedom
## Residual deviance: 571.16 on 608 degrees of freedom
## AIC: 605.16
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 140  39
##           1   3  10
##
##           Accuracy : 0.7812
##           95% CI : (0.716, 0.8376)
##           No Information Rate : 0.7448
##           P-Value [Acc > NIR] : 0.1404
##
##           Kappa : 0.2414
##
## Mcnemar's Test P-Value : 6.641e-08
##
##           Sensitivity : 0.9790
##           Specificity : 0.2041
##           Pos Pred Value : 0.7821
##           Neg Pred Value : 0.7692
##           Prevalence : 0.7448
##           Detection Rate : 0.7292
##           Detection Prevalence : 0.9323
##           Balanced Accuracy : 0.5916
##
##           'Positive' Class : 0
##

```

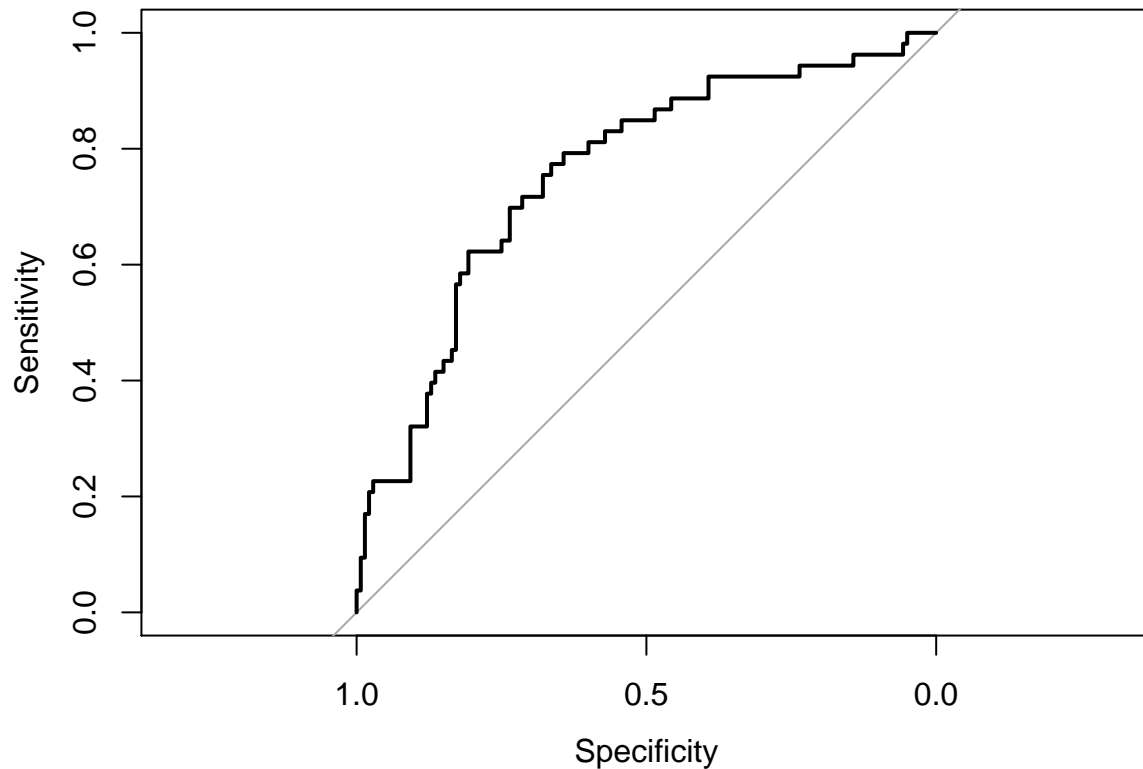


```
## [1] "AUC: 0.784073069787355"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 143 controls (dfPred_raw$class 0) < 49 cases (dfPred_raw$class 1).
## Area under the curve: 0.7841
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8342  -0.7083  -0.3882   0.5891   2.5739
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.639e+00  4.969e-01  -3.299 0.000969 ***
## INCOME        -1.185e-05  3.004e-06  -3.944 8.03e-05 ***
## TRAVTIME       1.842e-02  6.880e-03   2.677 0.007420 **
## BLUEBOOK       1.714e-05  1.580e-05   1.085 0.277988
## TIF           -3.009e-02  2.647e-02  -1.137 0.255666
## OLDCLAIM       2.310e-06  1.425e-05   0.162 0.871200
## PARENT1_Yes    1.079e+00  3.098e-01   3.482 0.000497 ***
## SEX_z_F       -7.933e-01  3.461e-01  -2.293 0.021873 *
## JOB_Manager   -3.478e-01  3.389e-01  -1.026 0.304810
```

```

## CAR_USE_Commercial          3.749e-01  2.428e-01  1.544 0.122525
## CAR_TYPE_Pickup             8.934e-01  3.270e-01  2.732 0.006297 **
## CAR_TYPE_Sports.Car        1.548e+00  4.620e-01  3.350 0.000809 ***
## CAR_TYPE_z_SUV             1.555e+00  4.138e-01  3.757 0.000172 ***
## URBANICITY_z_Highly.Rural..Rural -2.626e+00  4.461e-01 -5.886 3.95e-09 ***
## HOME_VAL_NA                -4.131e-01  2.257e-01 -1.830 0.067244 .
## oldclaim_log               9.087e-02  3.173e-02  2.864 0.004183 **
## inter                      1.303e-02  4.460e-03  2.922 0.003475 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 716.76  on 623  degrees of freedom
## Residual deviance: 551.60  on 607  degrees of freedom
## AIC: 585.6
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 123  36
##           1  17  17
##
##           Accuracy : 0.7254
##           95% CI : (0.6567, 0.787)
##      No Information Rate : 0.7254
##      P-Value [Acc > NIR] : 0.53692
##
##           Kappa : 0.2243
##
## Mcnemar's Test P-Value : 0.01342
##
##           Sensitivity : 0.8786
##           Specificity : 0.3208
##      Pos Pred Value : 0.7736
##      Neg Pred Value : 0.5000
##           Prevalence : 0.7254
##      Detection Rate : 0.6373
##      Detection Prevalence : 0.8238
##      Balanced Accuracy : 0.5997
##
##           'Positive' Class : 0
##

```

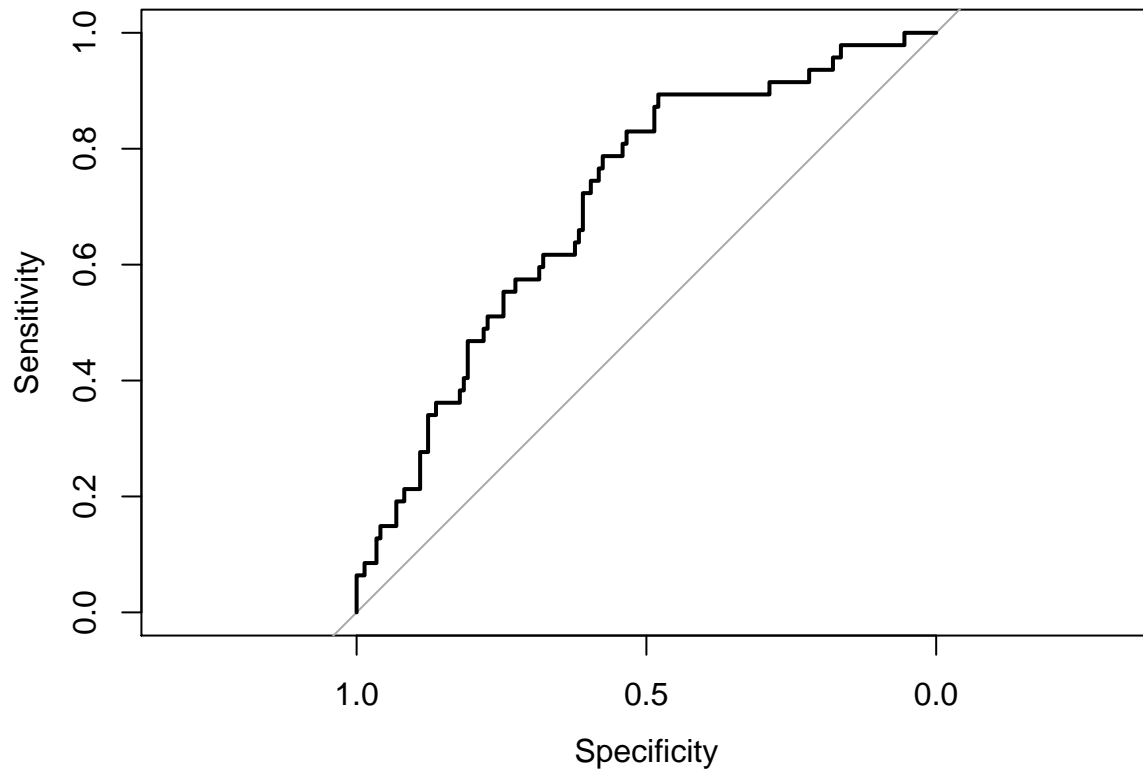


```
## [1] "AUC: 0.756873315363881"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 140 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.7569
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9610  -0.7185  -0.3700   0.5797   2.9317
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.817e+00  4.886e-01  -3.719 0.000200 ***
## INCOME         -1.261e-05  3.104e-06  -4.063 4.83e-05 ***
## TRAVTIME        1.737e-02  6.985e-03   2.487 0.012880 *
## BLUEBOOK        3.164e-05  1.556e-05   2.033 0.042041 *
## TIF            -6.651e-03  2.644e-02  -0.252 0.801369
## OLDCLAIM       -7.410e-07  1.518e-05  -0.049 0.961054
## PARENT1_Yes     1.321e+00  3.156e-01   4.187 2.83e-05 ***
## SEX_z_F        -8.423e-01  3.418e-01  -2.465 0.013715 *
## JOB_Manager    -6.349e-01  3.517e-01  -1.805 0.071064 .
```

```

## CAR_USE_Commercial          4.163e-01  2.447e-01  1.702 0.088845 .
## CAR_TYPE_Pickup             7.055e-01  3.203e-01  2.203 0.027623 *
## CAR_TYPE_Sports.Car        1.550e+00  4.665e-01  3.323 0.000892 ***
## CAR_TYPE_z_SUV             1.436e+00  4.105e-01  3.497 0.000470 ***
## URBANICITY_z_Highly.Rural..Rural -2.877e+00  4.487e-01 -6.411 1.44e-10 ***
## HOME_VAL_NA                -3.686e-01  2.310e-01 -1.596 0.110476
## oldclaim_log               9.913e-02  3.161e-02  3.136 0.001715 **
## inter                      1.645e-02  4.084e-03  4.028 5.63e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 728.94 on 623 degrees of freedom
## Residual deviance: 545.44 on 607 degrees of freedom
## AIC: 579.44
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 124  30
##           1  22  17
##
##           Accuracy : 0.7306
##           95% CI : (0.6621, 0.7918)
##           No Information Rate : 0.7565
##           P-Value [Acc > NIR] : 0.8224
##
##           Kappa : 0.2239
##
## Mcnemar's Test P-Value : 0.3317
##
##           Sensitivity : 0.8493
##           Specificity : 0.3617
##           Pos Pred Value : 0.8052
##           Neg Pred Value : 0.4359
##           Prevalence : 0.7565
##           Detection Rate : 0.6425
##           Detection Prevalence : 0.7979
##           Balanced Accuracy : 0.6055
##
##           'Positive' Class : 0
##

```

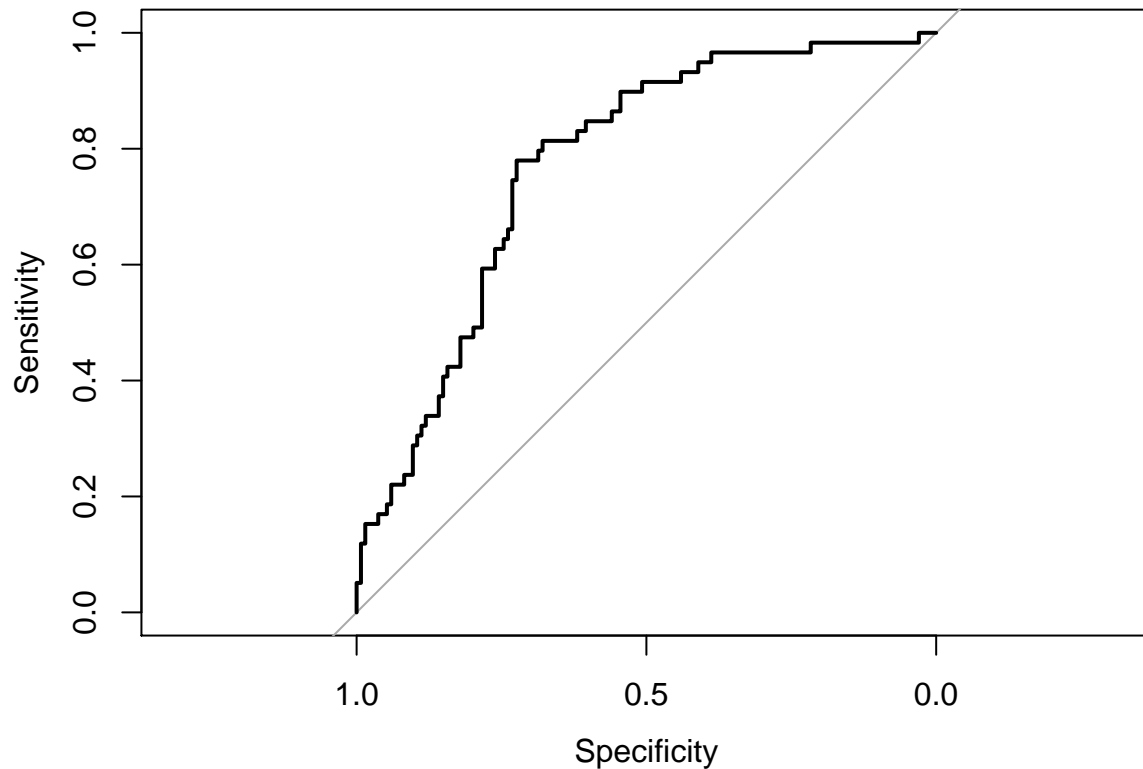


```
## [1] "AUC: 0.707956863888079"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 146 controls (dfPred_raw$class 0) < 47 cases (dfPred_raw$class 1).
## Area under the curve: 0.708
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9761  -0.7218  -0.3955   0.4248   2.7974
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.955e+00  5.061e-01  -3.864 0.000112 ***
## INCOME        -1.048e-05  2.979e-06  -3.517 0.000437 ***
## TRAVTIME       1.779e-02  7.044e-03   2.526 0.011550 *
## BLUEBOOK       3.369e-05  1.582e-05   2.130 0.033146 *
## TIF           -4.179e-02  2.666e-02  -1.567 0.117008
## OLDCLAIM       4.139e-06  1.502e-05   0.276 0.782917
## PARENT1_Yes    1.549e+00  3.331e-01   4.649 3.34e-06 ***
## SEX_z_F       -9.445e-01  3.533e-01  -2.673 0.007513 **
## JOB_Manager   -4.055e-01  3.519e-01  -1.152 0.249210
```

```

## CAR_USE_Commercial          3.726e-01  2.432e-01  1.532 0.125524
## CAR_TYPE_Pickup             1.012e+00  3.296e-01  3.070 0.002141 **
## CAR_TYPE_Sports.Car        1.705e+00  4.760e-01  3.583 0.000340 ***
## CAR_TYPE_z_SUV             1.664e+00  4.295e-01  3.875 0.000107 ***
## URBANICITY_z_Highly.Rural..Rural -2.432e+00  4.121e-01 -5.903 3.58e-09 ***
## HOME_VAL_NA                -3.720e-01  2.259e-01 -1.647 0.099564 .
## oldclaim_log               8.929e-02  3.245e-02  2.752 0.005931 **
## inter                      1.133e-02  4.346e-03  2.607 0.009133 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 703.98 on 623 degrees of freedom
## Residual deviance: 547.91 on 607 degrees of freedom
## AIC: 581.91
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 114  36
##           1  20  23
##
##           Accuracy : 0.7098
##           95% CI : (0.6403, 0.7728)
##           No Information Rate : 0.6943
##           P-Value [Acc > NIR] : 0.35129
##
##           Kappa : 0.2603
##
## Mcnemar's Test P-Value : 0.04502
##
##           Sensitivity : 0.8507
##           Specificity : 0.3898
##           Pos Pred Value : 0.7600
##           Neg Pred Value : 0.5349
##           Prevalence : 0.6943
##           Detection Rate : 0.5907
##           Detection Prevalence : 0.7772
##           Balanced Accuracy : 0.6203
##
##           'Positive' Class : 0
##

```



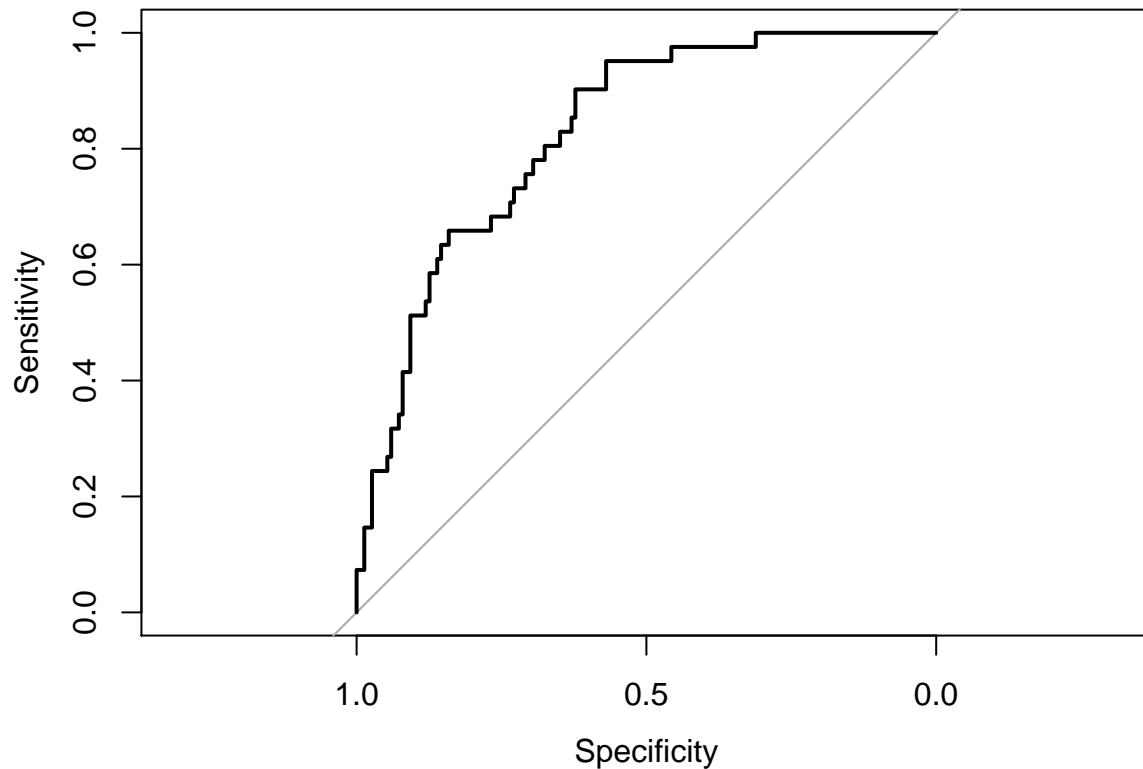
```
## [1] "AUC: 0.772704275234"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 134 controls (dfPred_raw$class 0) < 59 cases (dfPred_raw$class 1).
## Area under the curve: 0.7727
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0791  -0.7777  -0.4315   0.7588   2.7746
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.946e+00  4.731e-01  -4.113 3.90e-05 ***
## INCOME        -1.035e-05  2.853e-06  -3.627 0.000287 ***
## TRAVTIME       2.332e-02  6.935e-03   3.363 0.000770 ***
## BLUEBOOK       2.975e-05  1.539e-05   1.934 0.053138 .
## TIF           -3.691e-02  2.515e-02  -1.467 0.142332
## OLDCLAIM       1.056e-05  1.362e-05   0.775 0.438211
## PARENT1_Yes     9.612e-01  3.021e-01   3.181 0.001465 **
## SEX_z_F        -8.007e-01  3.311e-01  -2.418 0.015593 *
## JOB_Manager    -8.308e-01  3.672e-01  -2.263 0.023663 *
```



```

## CAR_USE_Commercial          3.586e-01  2.304e-01  1.557 0.119586
## CAR_TYPE_Pickup             9.246e-01  3.086e-01  2.996 0.002737 **
## CAR_TYPE_Sports.Car        1.716e+00  4.734e-01  3.626 0.000288 ***
## CAR_TYPE_z_SUV             1.521e+00  4.019e-01  3.785 0.000153 ***
## URBANICITY_z_Highly.Rural..Rural -2.456e+00  3.956e-01 -6.209 5.35e-10 ***
## HOME_VAL_NA                -1.323e-01  2.186e-01 -0.605 0.545107
## oldclaim_log               4.701e-02  3.108e-02  1.513 0.130350
## inter                      1.374e-02  3.974e-03  3.457 0.000547 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 741.19  on 624  degrees of freedom
## Residual deviance: 591.21  on 608  degrees of freedom
## AIC: 625.21
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0   1
##           0 137  20
##           1  14  21
##
##           Accuracy : 0.8229
##           95% CI : (0.7614, 0.8741)
##           No Information Rate : 0.7865
##           P-Value [Acc > NIR] : 0.1249
##
##           Kappa : 0.4431
##
## Mcnemar's Test P-Value : 0.3912
##
##           Sensitivity : 0.9073
##           Specificity : 0.5122
##           Pos Pred Value : 0.8726
##           Neg Pred Value : 0.6000
##           Prevalence : 0.7865
##           Detection Rate : 0.7135
##           Detection Prevalence : 0.8177
##           Balanced Accuracy : 0.7097
##
##           'Positive' Class : 0
##

```

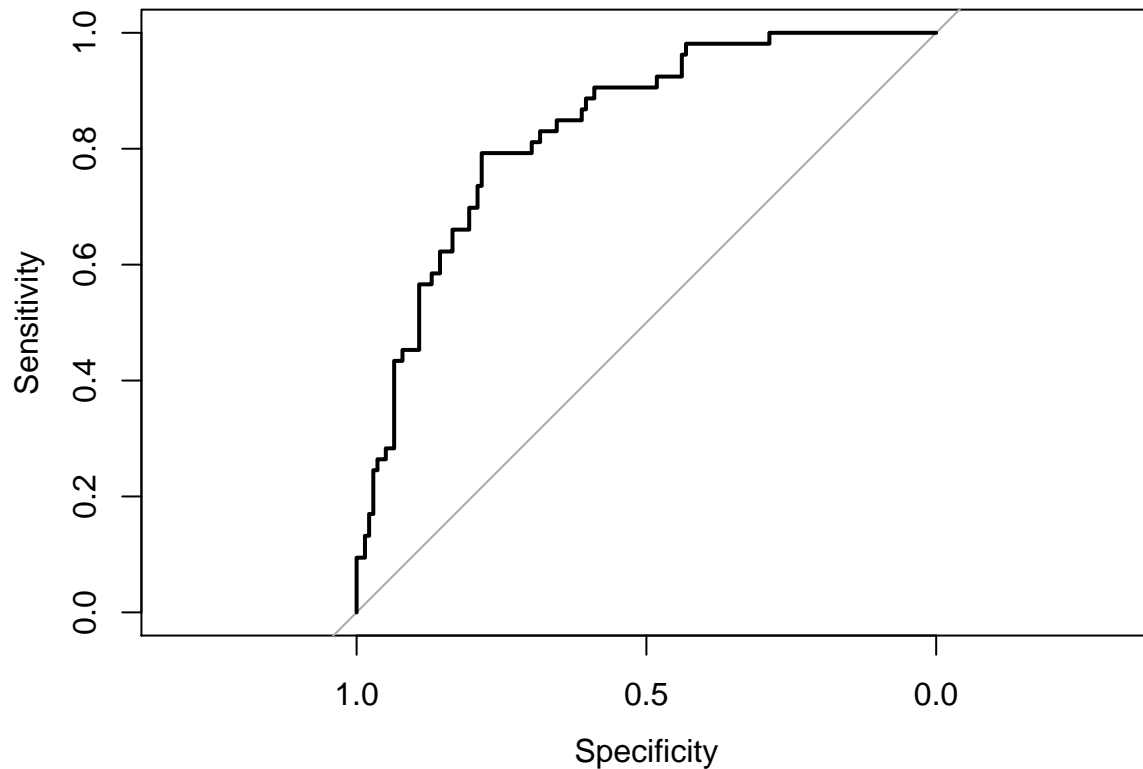


```
## [1] "AUC: 0.827491519948312"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 151 controls (dfPred_raw$class 0) < 41 cases (dfPred_raw$class 1).
## Area under the curve: 0.8275
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8643  -0.7419  -0.4436   0.6491   2.7611
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.128e+00  4.839e-01  -4.397 1.10e-05 ***
## INCOME        -9.777e-06  2.944e-06  -3.321 0.000898 ***
## TRAVTIME       1.977e-02  7.125e-03   2.775 0.005512 **
## BLUEBOOK       4.543e-05  1.568e-05   2.896 0.003774 **
## TIF          -3.501e-02  2.525e-02  -1.386 0.165611
## OLDCLAIM       1.315e-05  1.403e-05   0.937 0.348549
## PARENT1_Yes    1.153e+00  2.978e-01   3.873 0.000108 ***
## SEX_z_F       -8.990e-01  3.432e-01  -2.619 0.008808 **
## JOB_Manager   -5.310e-01  3.595e-01  -1.477 0.139617
```

```

## CAR_USE_Commercial          4.489e-01  2.357e-01  1.905 0.056810 .
## CAR_TYPE_Pickup             1.029e+00  3.199e-01  3.218 0.001292 **
## CAR_TYPE_Sports.Car         1.714e+00  4.740e-01  3.617 0.000298 ***
## CAR_TYPE_z_SUV              1.581e+00  4.224e-01  3.743 0.000182 ***
## URBANICITY_z_Highly.Rural..Rural -2.197e+00  3.866e-01 -5.683 1.32e-08 ***
## HOME_VAL_NA                 -3.735e-01  2.201e-01 -1.697 0.089735 .
## oldclaim_log                3.347e-02  3.204e-02  1.045 0.296213
## inter                       1.435e-02  4.343e-03  3.305 0.000951 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 717.37 on 624 degrees of freedom
## Residual deviance: 577.05 on 608 degrees of freedom
## AIC: 611.05
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 130  34
##           1   9  19
##
##           Accuracy : 0.776
##           95% CI : (0.7104, 0.8329)
##           No Information Rate : 0.724
##           P-Value [Acc > NIR] : 0.0603155
##
##           Kappa : 0.3439
##
## Mcnemar's Test P-Value : 0.0002522
##
##           Sensitivity : 0.9353
##           Specificity : 0.3585
##           Pos Pred Value : 0.7927
##           Neg Pred Value : 0.6786
##           Prevalence : 0.7240
##           Detection Rate : 0.6771
##           Detection Prevalence : 0.8542
##           Balanced Accuracy : 0.6469
##
##           'Positive' Class : 0
##

```

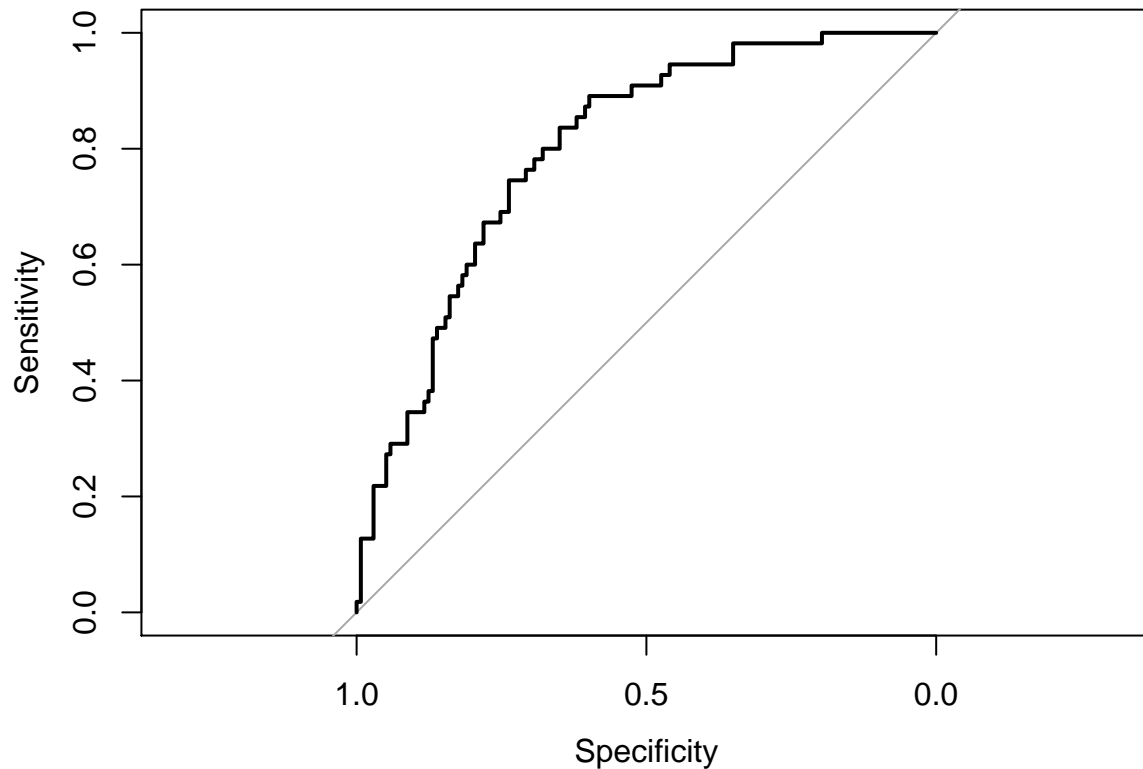


```
## [1] "AUC: 0.83398941224379"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 139 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.834
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7285  -0.7240  -0.4259   0.5578   2.7643
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.056e+00  5.078e-01  -4.049 5.14e-05 ***
## INCOME         -7.857e-06  2.876e-06  -2.732 0.006297 **
## TRAVTIME        2.373e-02  7.178e-03   3.307 0.000944 ***
## BLUEBOOK        3.038e-05  1.568e-05   1.937 0.052698 .
## TIF            -7.454e-02  2.774e-02  -2.687 0.007203 **
## OLDCLAIM        1.260e-05  1.416e-05   0.890 0.373343
## PARENT1_Yes     9.873e-01  3.189e-01   3.096 0.001962 **
## SEX_z_F        -1.113e+00  3.796e-01  -2.933 0.003362 **
## JOB_Manager    -4.622e-01  3.592e-01  -1.287 0.198231
```

```

## CAR_USE_Commercial          4.674e-01  2.451e-01  1.907 0.056567 .
## CAR_TYPE_Pickup             1.201e+00  3.345e-01  3.589 0.000332 ***
## CAR_TYPE_Sports.Car        1.991e+00  5.016e-01  3.968 7.24e-05 ***
## CAR_TYPE_z_SUV             1.993e+00  4.598e-01  4.334 1.46e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.110e+00  3.867e-01 -5.456 4.88e-08 ***
## HOME_VAL_NA                -4.777e-01  2.245e-01 -2.127 0.033396 *
## oldclaim_log               4.962e-02  3.239e-02  1.532 0.125579
## inter                      1.718e-02  4.679e-03  3.672 0.000241 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 713.16  on 624  degrees of freedom
## Residual deviance: 564.77  on 608  degrees of freedom
## AIC: 598.77
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 119  34
##           1  18  21
##
##               Accuracy : 0.7292
##               95% CI : (0.6605, 0.7906)
##       No Information Rate : 0.7135
##       P-Value [Acc > NIR] : 0.34853
##
##               Kappa : 0.2743
##
## Mcnemar's Test P-Value : 0.03751
##
##           Sensitivity : 0.8686
##           Specificity : 0.3818
##       Pos Pred Value : 0.7778
##       Neg Pred Value : 0.5385
##           Prevalence : 0.7135
##       Detection Rate : 0.6198
##   Detection Prevalence : 0.7969
##       Balanced Accuracy : 0.6252
##
##           'Positive' Class : 0
##

```

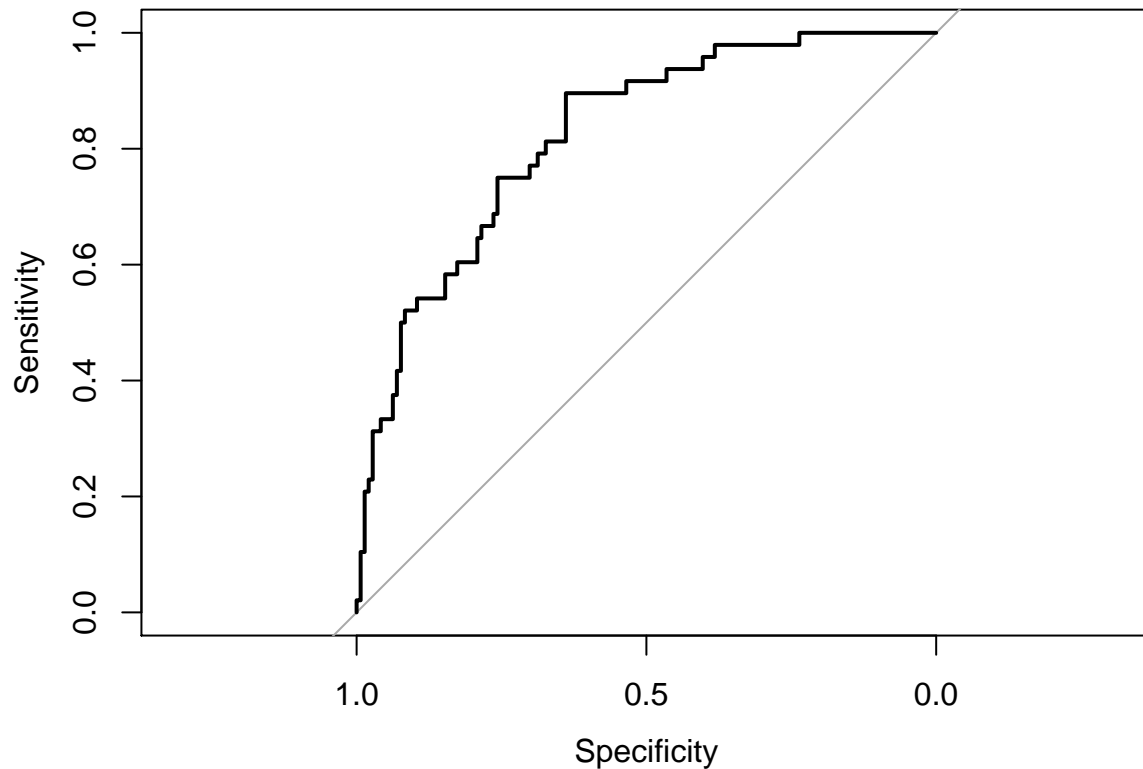


```
## [1] "AUC: 0.799203715992037"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 137 controls (dfPred_raw$class 0) < 55 cases (dfPred_raw$class 1).
## Area under the curve: 0.7992
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7101  -0.7537  -0.4344   0.7013   2.6902
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.877e+00  4.916e-01  -3.819 0.000134 ***
## INCOME        -9.307e-06  2.879e-06  -3.233 0.001225 **
## TRAVTIME       2.294e-02  7.044e-03   3.256 0.001130 **
## BLUEBOOK       2.344e-05  1.587e-05   1.477 0.139703
## TIF           -6.422e-02  2.727e-02  -2.355 0.018509 *
## OLDCLAIM       1.217e-05  1.335e-05   0.912 0.361828
## PARENT1_Yes     8.621e-01  3.075e-01   2.803 0.005059 **
## SEX_z_F       -1.153e+00  3.822e-01  -3.017 0.002552 **
## JOB_Manager    -4.098e-01  3.439e-01  -1.191 0.233475
```

```

## CAR_USE_Commercial          5.093e-01  2.393e-01  2.128 0.033316 *
## CAR_TYPE_Pickup             1.047e+00  3.296e-01  3.178 0.001483 **
## CAR_TYPE_Sports.Car        2.113e+00  4.987e-01  4.236 2.27e-05 ***
## CAR_TYPE_z_SUV             2.027e+00  4.535e-01  4.471 7.79e-06 ***
## URBANICITY_z_Highly.Rural..Rural -1.997e+00  3.732e-01 -5.351 8.75e-08 ***
## HOME_VAL_NA                -3.288e-01  2.151e-01 -1.529 0.126320
## oldclaim_log               5.072e-02  3.133e-02  1.619 0.105410
## inter                      1.084e-02  4.257e-03  2.547 0.010873 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 727.58 on 624 degrees of freedom
## Residual deviance: 586.56 on 608 degrees of freedom
## AIC: 620.56
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 132  24
##           1  12  24
##
##           Accuracy : 0.8125
##           95% CI : (0.75, 0.8651)
##           No Information Rate : 0.75
##           P-Value [Acc > NIR] : 0.02501
##
##           Kappa : 0.4545
##
## Mcnemar's Test P-Value : 0.06675
##
##           Sensitivity : 0.9167
##           Specificity : 0.5000
##           Pos Pred Value : 0.8462
##           Neg Pred Value : 0.6667
##           Prevalence : 0.7500
##           Detection Rate : 0.6875
##           Detection Prevalence : 0.8125
##           Balanced Accuracy : 0.7083
##
##           'Positive' Class : 0
##

```



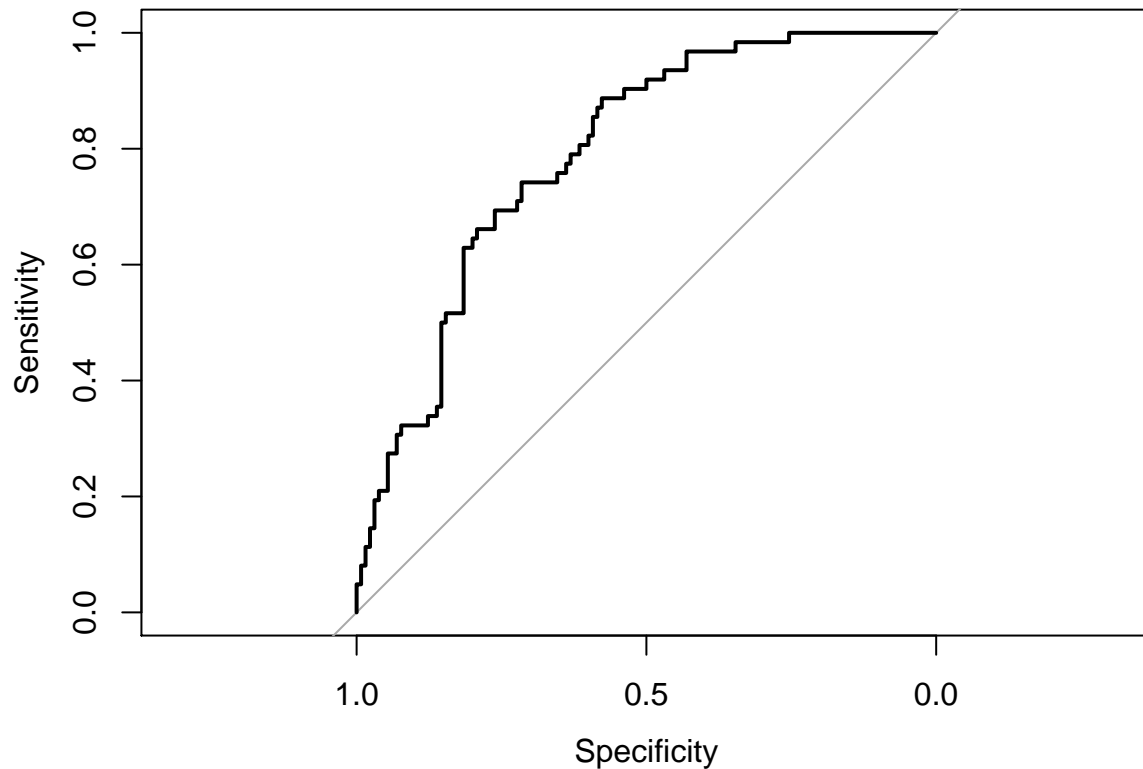
```
## [1] "AUC: 0.823640046296296"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 144 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.8236
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9888  -0.7271  -0.4037  -0.1084   2.8921
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.025e+00  5.196e-01  -3.896 9.78e-05 ***
## INCOME        -8.498e-06  2.955e-06  -2.876 0.004027 **
## TRAVTIME       2.106e-02  7.246e-03   2.906 0.003655 **
## BLUEBOOK       2.609e-05  1.595e-05   1.635 0.101970
## TIF           -6.067e-02  2.773e-02  -2.188 0.028684 *
## OLDCLAIM       2.035e-05  1.367e-05   1.489 0.136478
## PARENT1_Yes     8.802e-01  3.166e-01   2.780 0.005439 **
## SEX_z_F       -8.298e-01  3.552e-01  -2.336 0.019493 *
## JOB_Manager    -6.402e-01  3.629e-01  -1.764 0.077694 .
```



```

## CAR_USE_Commercial          4.318e-01  2.494e-01  1.731 0.083411 .
## CAR_TYPE_Pickup             1.213e+00  3.344e-01  3.626 0.000287 ***
## CAR_TYPE_Sports.Car         1.745e+00  4.907e-01  3.556 0.000377 ***
## CAR_TYPE_z_SUV              1.622e+00  4.333e-01  3.744 0.000181 ***
## URBANICITY_z_Highly.Rural..Rural -2.405e+00  4.327e-01 -5.558 2.74e-08 ***
## HOME_VAL_NA                 -3.846e-01  2.272e-01 -1.693 0.090423 .
## oldclaim_log                5.209e-02  3.188e-02  1.634 0.102282
## inter                       1.563e-02  4.462e-03  3.503 0.000460 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 697.93  on 624  degrees of freedom
## Residual deviance: 549.23  on 608  degrees of freedom
## AIC: 583.23
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 114  42
##           1  16  20
##
##           Accuracy : 0.6979
##           95% CI : (0.6277, 0.7619)
##      No Information Rate : 0.6771
##      P-Value [Acc > NIR] : 0.296831
##
##           Kappa : 0.2241
##
## Mcnemar's Test P-Value : 0.001028
##
##           Sensitivity : 0.8769
##           Specificity : 0.3226
##      Pos Pred Value : 0.7308
##      Neg Pred Value : 0.5556
##           Prevalence : 0.6771
##      Detection Rate : 0.5938
##      Detection Prevalence : 0.8125
##      Balanced Accuracy : 0.5998
##
##           'Positive' Class : 0
##

```

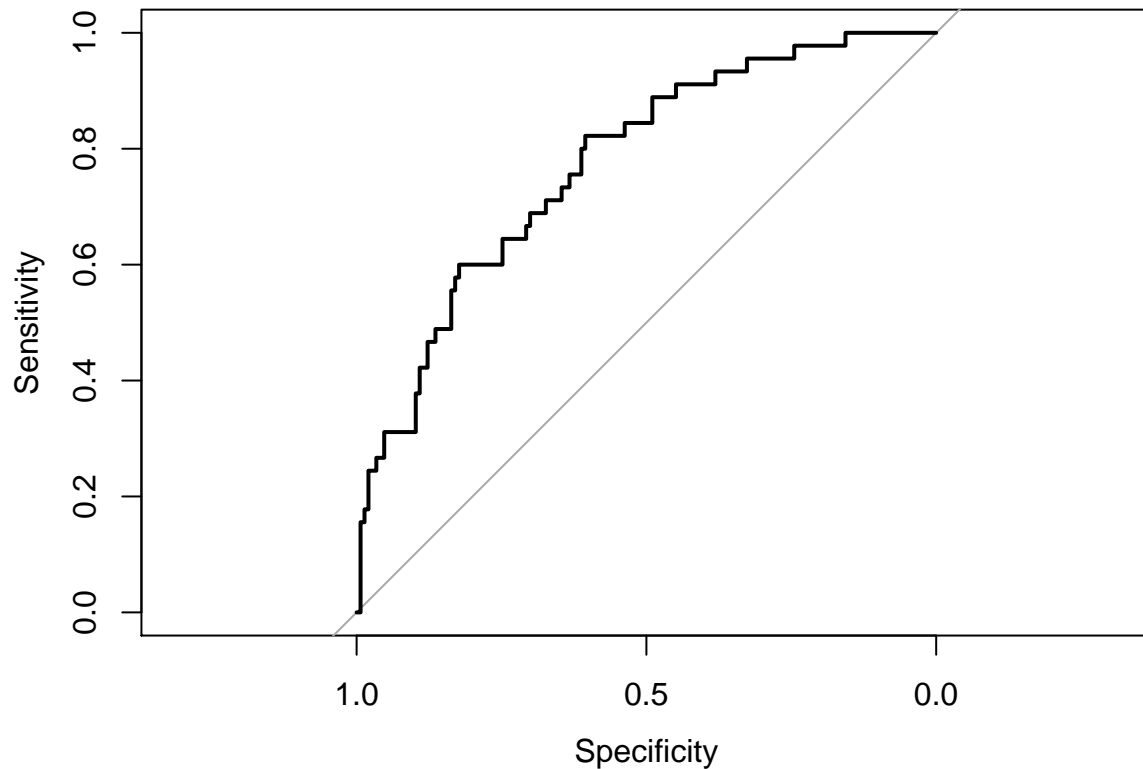


```
## [1] "AUC: 0.79106699751861"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 130 controls (dfPred_raw$class 0) < 62 cases (dfPred_raw$class 1).
## Area under the curve: 0.7911
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8008  -0.7632  -0.4137   0.7086   2.9141
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.025e+00  4.882e-01  -4.148 3.35e-05 ***
## INCOME        -9.453e-06  2.849e-06  -3.319 0.000905 ***
## TRAVTIME       2.809e-02  7.147e-03   3.931 8.47e-05 ***
## BLUEBOOK       2.396e-05  1.512e-05   1.585 0.113079
## TIF           -4.362e-02  2.614e-02  -1.669 0.095130 .
## OLDCLAIM       1.417e-05  1.370e-05   1.034 0.301082
## PARENT1_Yes     7.248e-01  2.931e-01   2.473 0.013416 *
## SEX_z_F        -6.613e-01  3.525e-01  -1.876 0.060646 .
## JOB_Manager    -7.061e-01  3.499e-01  -2.018 0.043612 *
```

```

## CAR_USE_Commercial          6.608e-01  2.351e-01  2.811 0.004940 **
## CAR_TYPE_Pickup             7.238e-01  3.194e-01  2.266 0.023430 *
## CAR_TYPE_Sports.Car        1.645e+00  4.844e-01  3.396 0.000684 ***
## CAR_TYPE_z_SUV             1.471e+00  4.211e-01  3.494 0.000476 ***
## URBANICITY_z_Highly.Rural..Rural -2.713e+00  4.452e-01 -6.094 1.10e-09 ***
## HOME_VAL_NA                -3.067e-01  2.223e-01 -1.380 0.167643
## oldclaim_log               2.487e-02  3.175e-02  0.783 0.433588
## inter                      1.543e-02  4.066e-03  3.794 0.000148 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 733.51 on 624 degrees of freedom
## Residual deviance: 575.09 on 608 degrees of freedom
## AIC: 609.09
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 137  31
##           1  10  14
##
##           Accuracy : 0.7865
##           95% CI : (0.7217, 0.8422)
##           No Information Rate : 0.7656
##           P-Value [Acc > NIR] : 0.278829
##
##           Kappa : 0.29
##
## Mcnemar's Test P-Value : 0.001787
##
##           Sensitivity : 0.9320
##           Specificity : 0.3111
##           Pos Pred Value : 0.8155
##           Neg Pred Value : 0.5833
##           Prevalence : 0.7656
##           Detection Rate : 0.7135
##           Detection Prevalence : 0.8750
##           Balanced Accuracy : 0.6215
##
##           'Positive' Class : 0
##

```

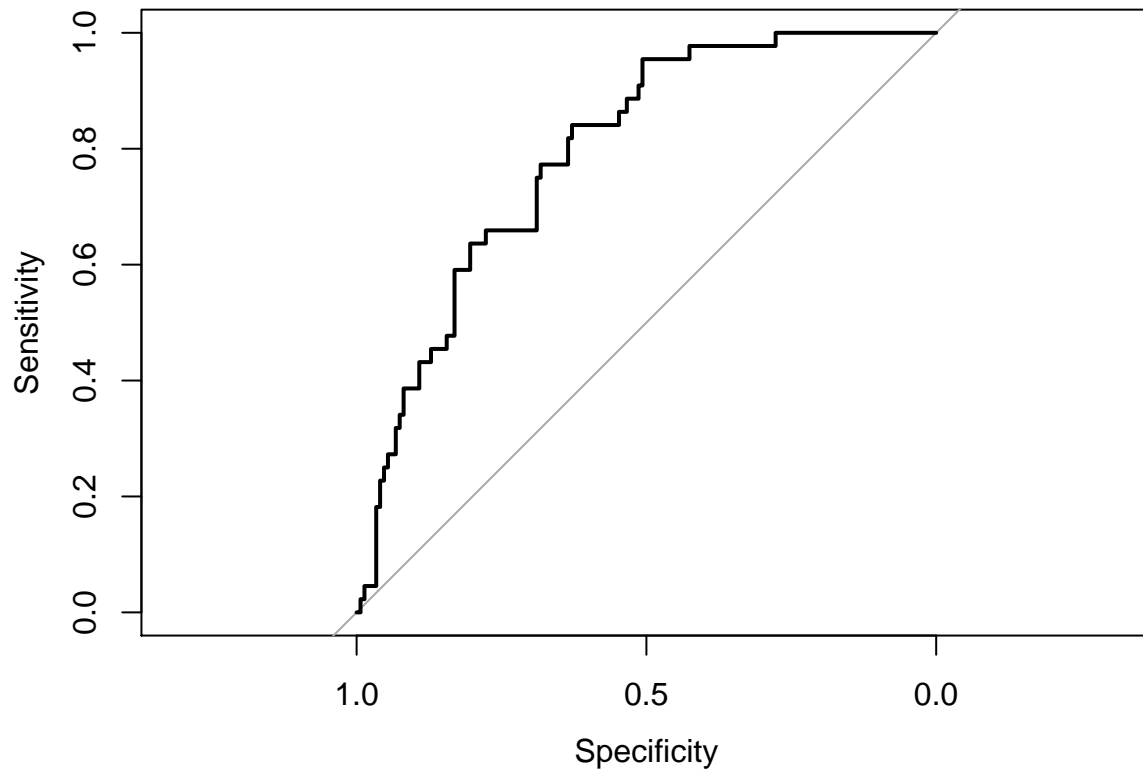


```
## [1] "AUC: 0.772637944066516"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 147 controls (dfPred_raw$class 0) < 45 cases (dfPred_raw$class 1).
## Area under the curve: 0.7726
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9857  -0.7382  -0.4292   0.6919   2.7672
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.392e+00  4.891e-01  -4.890 1.01e-06 ***
## INCOME         -1.088e-05  2.928e-06  -3.715 0.000203 ***
## TRAVTIME        2.609e-02  6.998e-03   3.728 0.000193 ***
## BLUEBOOK        5.057e-05  1.575e-05   3.211 0.001324 **
## TIF            -3.334e-02  2.607e-02  -1.279 0.200864
## OLDCLAIM        8.644e-06  1.336e-05   0.647 0.517681
## PARENT1_Yes     1.078e+00  2.937e-01   3.669 0.000243 ***
## SEX_z_F        -8.673e-01  3.441e-01  -2.521 0.011716 *
## JOB_Manager    -7.508e-01  3.549e-01  -2.116 0.034366 *
```

```

## CAR_USE_Commercial          2.304e-01  2.358e-01  0.977 0.328572
## CAR_TYPE_Pickup             9.411e-01  3.222e-01  2.920 0.003496 **
## CAR_TYPE_Sports.Car        1.754e+00  4.760e-01  3.684 0.000229 ***
## CAR_TYPE_z_SUV             1.489e+00  4.158e-01  3.582 0.000341 ***
## URBANICITY_z_Highly.Rural..Rural -2.368e+00  4.023e-01 -5.887 3.93e-09 ***
## HOME_VAL_NA                -2.077e-01  2.209e-01 -0.941 0.346957
## oldclaim_log               7.163e-02  3.121e-02  2.295 0.021732 *
## inter                      1.373e-02  4.249e-03  3.232 0.001228 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 735.45 on 624 degrees of freedom
## Residual deviance: 579.04 on 608 degrees of freedom
## AIC: 613.04
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 132  26
##           1  16  18
##
##           Accuracy : 0.7812
##           95% CI : (0.716, 0.8376)
##           No Information Rate : 0.7708
##           P-Value [Acc > NIR] : 0.4041
##
##           Kappa : 0.3271
##
## Mcnemar's Test P-Value : 0.1649
##
##           Sensitivity : 0.8919
##           Specificity : 0.4091
##           Pos Pred Value : 0.8354
##           Neg Pred Value : 0.5294
##           Prevalence : 0.7708
##           Detection Rate : 0.6875
##           Detection Prevalence : 0.8229
##           Balanced Accuracy : 0.6505
##
##           'Positive' Class : 0
##

```

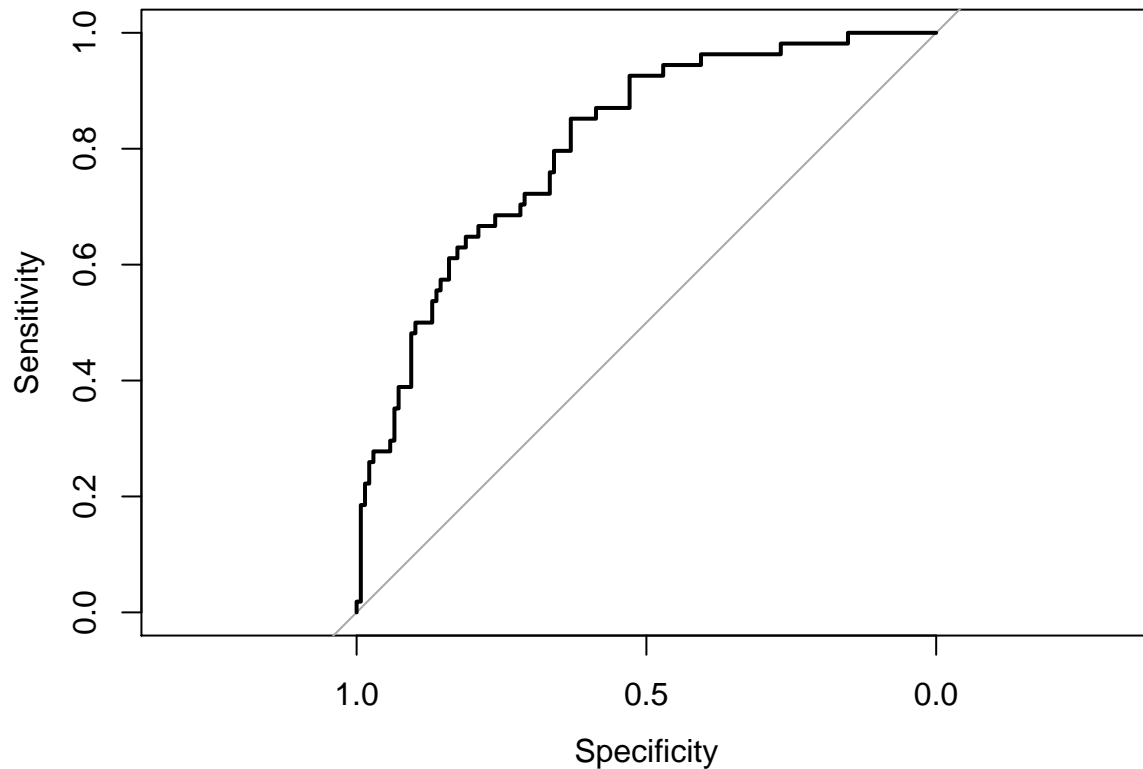


```
## [1] "AUC: 0.793458230958231"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 148 controls (dfPred_raw$class 0) < 44 cases (dfPred_raw$class 1).
## Area under the curve: 0.7935
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7897  -0.7343  -0.4160   0.6637   2.8366
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.852e+00  5.017e-01  -3.691 0.000223 ***
## INCOME        -1.062e-05  2.952e-06  -3.598 0.000320 ***
## TRAVTIME       2.032e-02  7.039e-03   2.887 0.003894 **
## BLUEBOOK       2.662e-05  1.546e-05   1.722 0.085109 .
## TIF           -4.978e-02  2.646e-02  -1.881 0.059919 .
## OLDCLAIM       1.878e-05  1.337e-05   1.404 0.160321
## PARENT1_Yes    7.849e-01  3.100e-01   2.532 0.011347 *
## SEX_z_F       -7.223e-01  3.478e-01  -2.077 0.037844 *
## JOB_Manager    -5.045e-01  3.451e-01  -1.462 0.143746
```

```

## CAR_USE_Commercial          7.078e-01  2.386e-01  2.966 0.003018 **
## CAR_TYPE_Pickup             8.367e-01  3.233e-01  2.588 0.009648 **
## CAR_TYPE_Sports.Car        1.836e+00  4.786e-01  3.836 0.000125 ***
## CAR_TYPE_z_SUV             1.517e+00  4.204e-01  3.609 0.000308 ***
## URBANICITY_z_Highly.Rural..Rural -2.369e+00  4.140e-01 -5.722 1.05e-08 ***
## HOME_VAL_NA                -2.977e-01  2.221e-01 -1.341 0.180077
## oldclaim_log               2.327e-02  3.180e-02  0.732 0.464411
## inter                      1.058e-02  4.162e-03  2.541 0.011057 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 715.27 on 624 degrees of freedom
## Residual deviance: 570.37 on 608 degrees of freedom
## AIC: 604.37
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 130  39
##           1   8  15
##
##           Accuracy : 0.7552
##           95% CI : (0.6881, 0.8143)
##           No Information Rate : 0.7188
##           P-Value [Acc > NIR] : 0.1481
##
##           Kappa : 0.2663
##
## Mcnemar's Test P-Value : 1.209e-05
##
##           Sensitivity : 0.9420
##           Specificity : 0.2778
##           Pos Pred Value : 0.7692
##           Neg Pred Value : 0.6522
##           Prevalence : 0.7188
##           Detection Rate : 0.6771
##           Detection Prevalence : 0.8802
##           Balanced Accuracy : 0.6099
##
##           'Positive' Class : 0
##

```



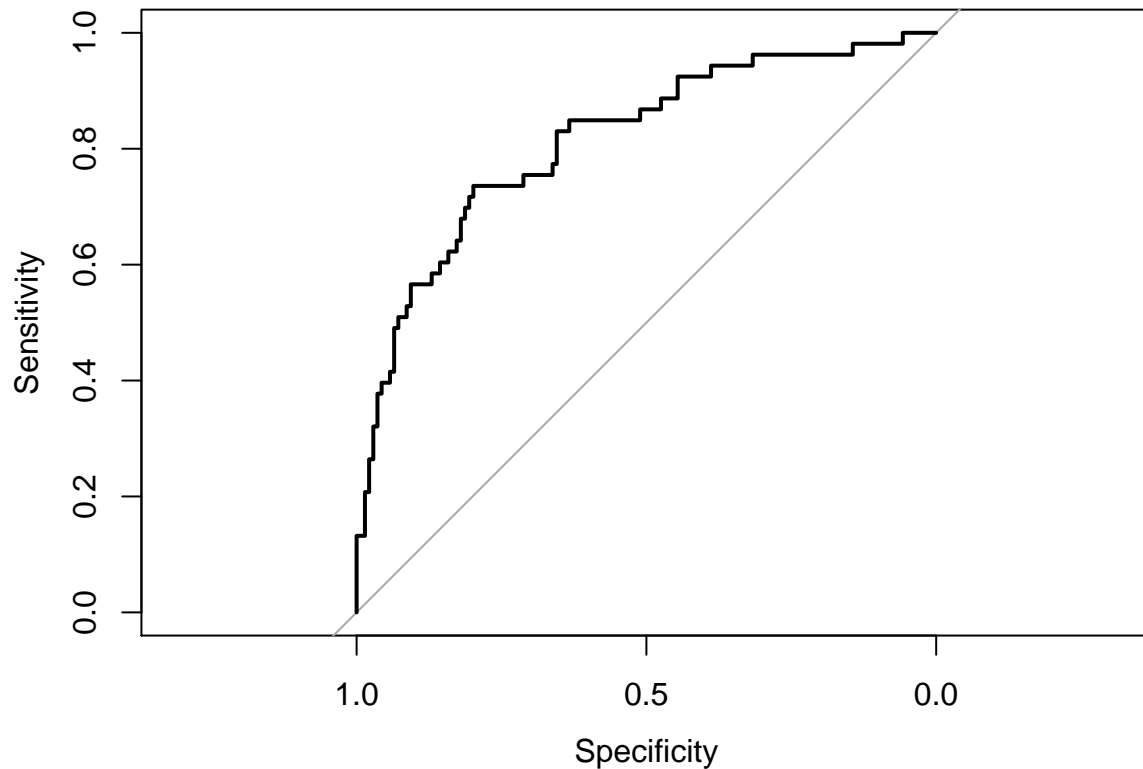
```
## [1] "AUC: 0.807971014492754"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 138 controls (dfPred_raw$class 0) < 54 cases (dfPred_raw$class 1).
## Area under the curve: 0.808
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8525  -0.7267  -0.4185   0.6233   2.8097
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.698e+00  4.855e-01  -3.497 0.000471 ***
## INCOME          -9.425e-06  2.999e-06  -3.142 0.001675 **
## TRAVTIME         1.676e-02  6.894e-03   2.431 0.015078 *
## BLUEBOOK         1.843e-05  1.546e-05   1.192 0.233398
## TIF             -2.868e-02  2.603e-02  -1.102 0.270538
## OLDCLAIM         1.611e-05  1.416e-05   1.138 0.255244
## PARENT1_Yes      1.059e+00  3.063e-01   3.458 0.000544 ***
## SEX_z_F         -9.887e-01  3.552e-01  -2.783 0.005379 **
## JOB_Manager     -6.653e-01  3.624e-01  -1.836 0.066343 .
```



```

## CAR_USE_Commercial          6.036e-01  2.403e-01  2.512 0.011994 *
## CAR_TYPE_Pickup             9.319e-01  3.202e-01  2.911 0.003607 **
## CAR_TYPE_Sports.Car        1.899e+00  4.766e-01  3.985 6.76e-05 ***
## CAR_TYPE_z_SUV             1.616e+00  4.324e-01  3.738 0.000186 ***
## URBANICITY_z_Highly.Rural..Rural -2.222e+00  3.992e-01 -5.567 2.60e-08 ***
## HOME_VAL_NA                -3.380e-01  2.209e-01 -1.530 0.126037
## oldclaim_log               2.815e-02  3.170e-02  0.888 0.374579
## inter                      1.469e-02  4.253e-03  3.454 0.000552 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 717.37 on 624 degrees of freedom
## Residual deviance: 572.29 on 608 degrees of freedom
## AIC: 606.29
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 134  34
##           1   5  19
##
##           Accuracy : 0.7969
##           95% CI : (0.733, 0.8514)
##           No Information Rate : 0.724
##           P-Value [Acc > NIR] : 0.01278
##
##           Kappa : 0.3882
##
## Mcnemar's Test P-Value : 7.34e-06
##
##           Sensitivity : 0.9640
##           Specificity : 0.3585
##           Pos Pred Value : 0.7976
##           Neg Pred Value : 0.7917
##           Prevalence : 0.7240
##           Detection Rate : 0.6979
##           Detection Prevalence : 0.8750
##           Balanced Accuracy : 0.6613
##
##           'Positive' Class : 0
##

```

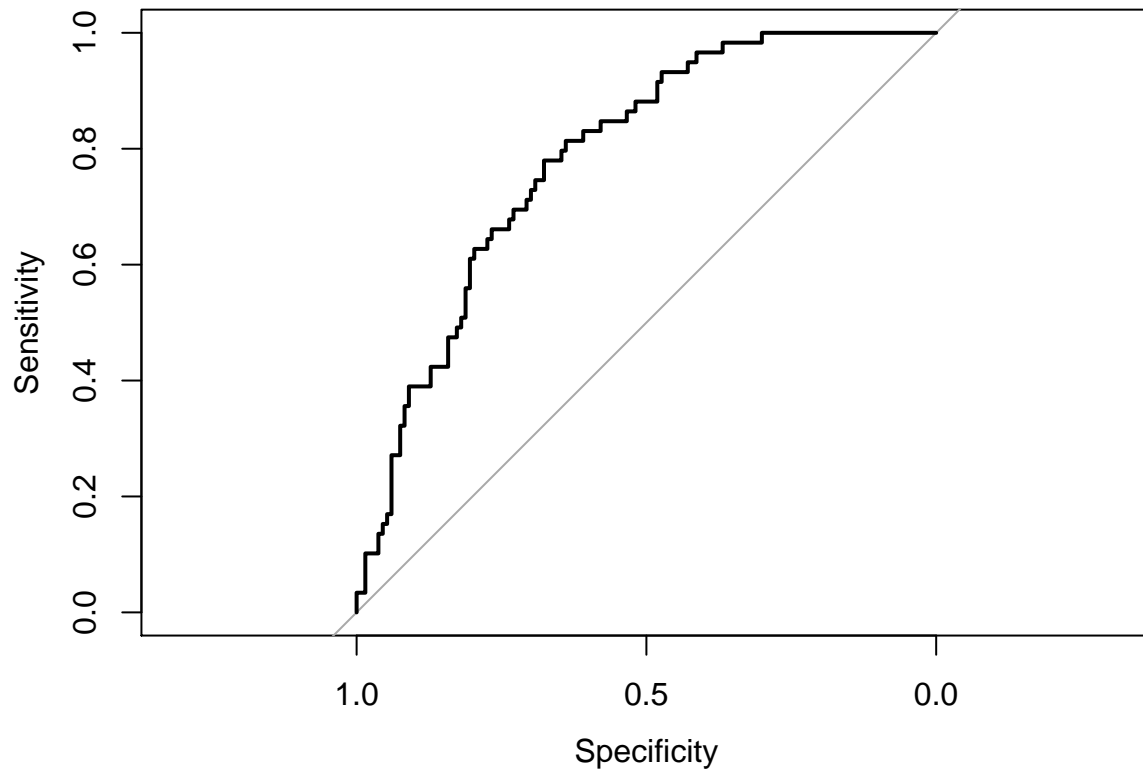


```
## [1] "AUC: 0.814850006787023"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 139 controls (dfPred_raw$class 0) < 53 cases (dfPred_raw$class 1).
## Area under the curve: 0.8149
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0998  -0.7157  -0.3994   0.2928   2.8205
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.125e+00  5.081e-01  -4.182 2.89e-05 ***
## INCOME        -8.046e-06  2.919e-06  -2.757 0.005839 **
## TRAVTIME       2.143e-02  7.037e-03   3.045 0.002323 **
## BLUEBOOK       3.216e-05  1.580e-05   2.035 0.041890 *
## TIF           -5.259e-02  2.708e-02  -1.942 0.052132 .
## OLDCLAIM       1.820e-05  1.350e-05   1.349 0.177367
## PARENT1_Yes     9.494e-01  3.209e-01   2.958 0.003096 **
## SEX_z_F       -1.063e+00  3.581e-01  -2.969 0.002991 **
## JOB_Manager    -9.124e-01  3.836e-01  -2.378 0.017397 *
```

```

## CAR_USE_Commercial          2.541e-01  2.456e-01  1.034 0.300916
## CAR_TYPE_Pickup             1.318e+00  3.356e-01  3.927 8.62e-05 ***
## CAR_TYPE_Sports.Car         1.936e+00  4.915e-01  3.939 8.17e-05 ***
## CAR_TYPE_z_SUV              1.666e+00  4.396e-01  3.790 0.000151 ***
## URBANICITY_z_Highly.Rural..Rural -2.344e+00  4.145e-01 -5.656 1.55e-08 ***
## HOME_VAL_NA                 -2.977e-01  2.278e-01 -1.307 0.191278
## oldclaim_log                6.251e-02  3.190e-02  1.959 0.050072 .
## inter                      1.857e-02  4.551e-03  4.081 4.49e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 704.56  on 624  degrees of freedom
## Residual deviance: 553.61  on 608  degrees of freedom
## AIC: 587.61
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 116  35
##           1  17  24
##
##           Accuracy : 0.7292
##           95% CI : (0.6605, 0.7906)
##    No Information Rate : 0.6927
##    P-Value [Acc > NIR] : 0.1545
##
##           Kappa : 0.3048
##
## Mcnemar's Test P-Value : 0.0184
##
##           Sensitivity : 0.8722
##           Specificity : 0.4068
##    Pos Pred Value : 0.7682
##    Neg Pred Value : 0.5854
##           Prevalence : 0.6927
##    Detection Rate : 0.6042
##    Detection Prevalence : 0.7865
##    Balanced Accuracy : 0.6395
##
##           'Positive' Class : 0
##

```

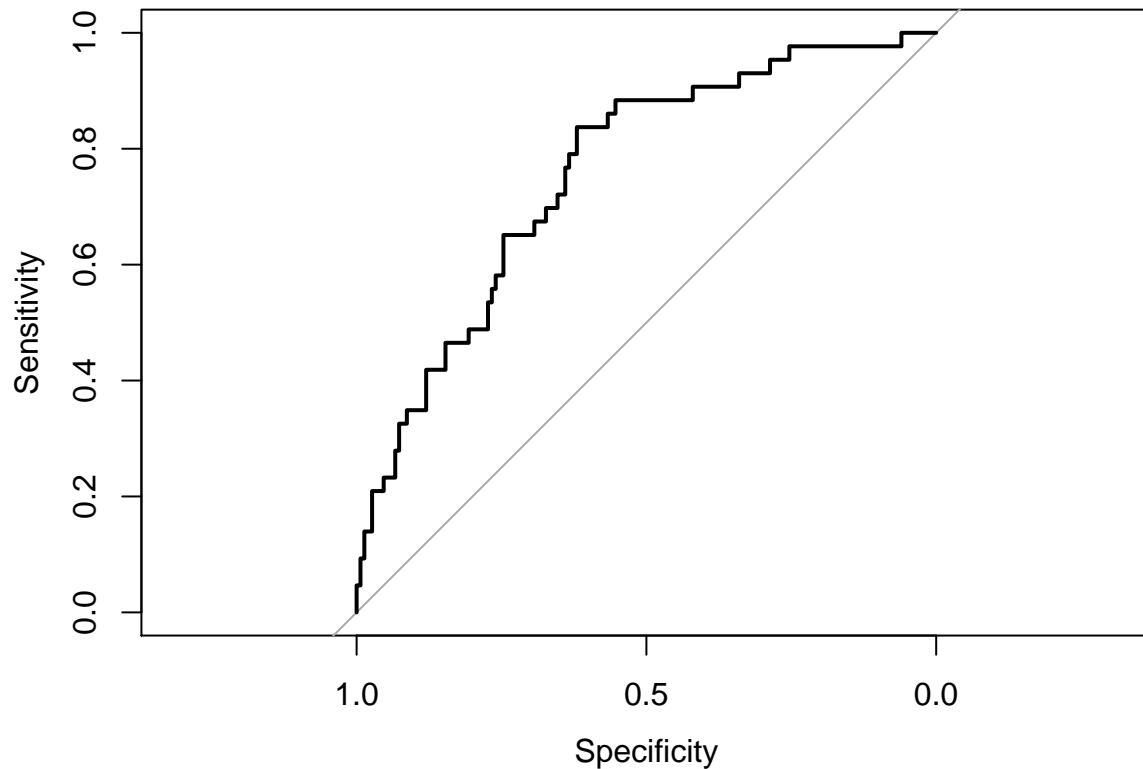


```
## [1] "AUC: 0.785268255384223"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 59 cases (dfPred_raw$class 1).
## Area under the curve: 0.7853
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0484  -0.7343  -0.4080   0.7572   2.7982
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.917e+00  4.835e-01  -3.965 7.34e-05 ***
## INCOME        -1.208e-05  2.899e-06  -4.165 3.12e-05 ***
## TRAVTIME       2.343e-02  7.119e-03   3.291 0.000999 ***
## BLUEBOOK       3.815e-05  1.564e-05   2.439 0.014739 *
## TIF           -3.291e-02  2.530e-02  -1.301 0.193307
## OLDCLAIM       4.017e-06  1.390e-05   0.289 0.772529
## PARENT1_Yes     9.209e-01  2.911e-01   3.164 0.001559 **
## SEX_z_F        -7.530e-01  3.348e-01  -2.249 0.024497 *
## JOB_Manager    -6.844e-01  3.556e-01  -1.925 0.054267 .
```

```

## CAR_USE_Commercial          3.310e-01  2.347e-01  1.410 0.158476
## CAR_TYPE_Pickup             6.926e-01  3.183e-01  2.176 0.029561 *
## CAR_TYPE_Sports.Car        1.467e+00  4.683e-01  3.132 0.001738 **
## CAR_TYPE_z_SUV             1.490e+00  4.007e-01  3.717 0.000201 ***
## URBANICITY_z_Highly.Rural..Rural -2.675e+00  4.373e-01 -6.118 9.49e-10 ***
## HOME_VAL_NA                -3.335e-01  2.219e-01 -1.503 0.132857
## oldclaim_log               7.581e-02  3.132e-02  2.421 0.015498 *
## inter                      1.412e-02  4.120e-03  3.426 0.000612 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 736.73  on 623  degrees of freedom
## Residual deviance: 570.57  on 607  degrees of freedom
## AIC: 604.57
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 135  28
##           1  15  15
##
##           Accuracy : 0.7772
##           95% CI : (0.7119, 0.8338)
##           No Information Rate : 0.7772
##           P-Value [Acc > NIR] : 0.54076
##
##           Kappa : 0.2789
##
## Mcnemar's Test P-Value : 0.06725
##
##           Sensitivity : 0.9000
##           Specificity : 0.3488
##           Pos Pred Value : 0.8282
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7772
##           Detection Rate : 0.6995
##           Detection Prevalence : 0.8446
##           Balanced Accuracy : 0.6244
##
##           'Positive' Class : 0
##

```

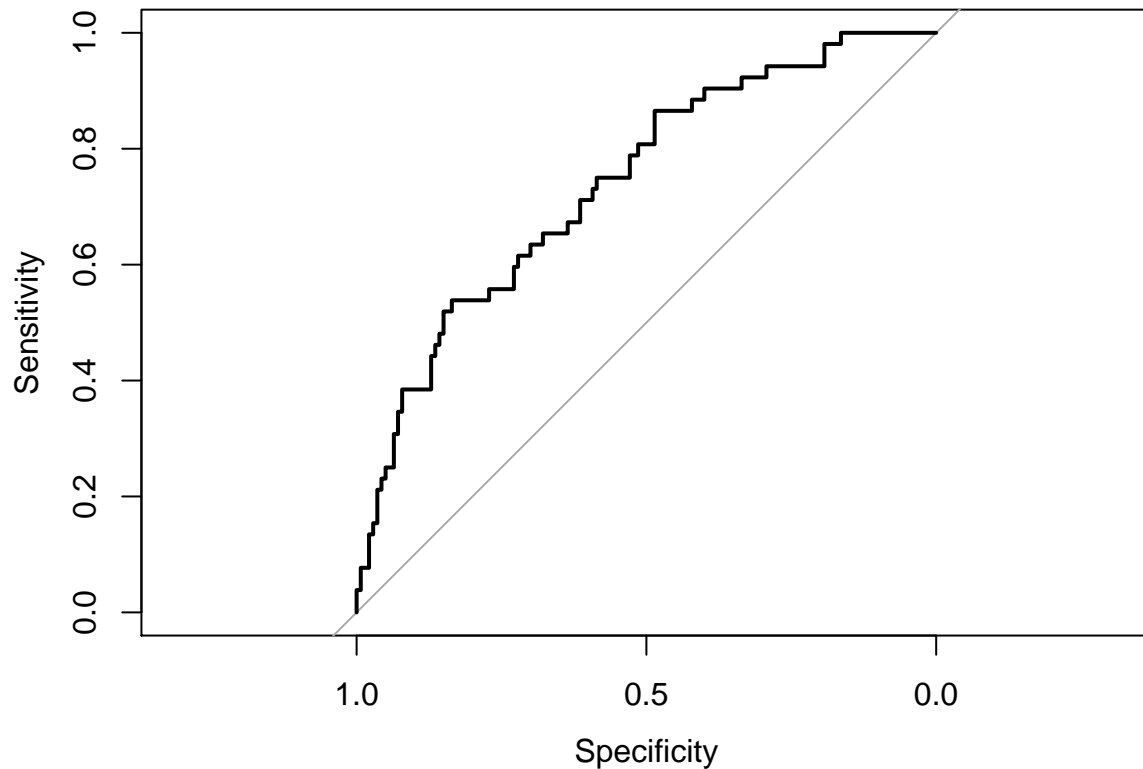


```
## [1] "AUC: 0.757519379844961"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 150 controls (dfPred_raw$class 0) < 43 cases (dfPred_raw$class 1).
## Area under the curve: 0.7575
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0179  -0.7213  -0.3723   0.5875   2.9822
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.136e+00  4.887e-01  -4.371 1.24e-05 ***
## INCOME        -9.572e-06  3.049e-06  -3.139 0.001693 **
## TRAVTIME       1.978e-02  6.897e-03   2.868 0.004129 **
## BLUEBOOK       3.192e-05  1.570e-05   2.033 0.042061 *
## TIF           -4.080e-03  2.634e-02  -0.155 0.876921
## OLDCLAIM       6.049e-06  1.440e-05   0.420 0.674545
## PARENT1_Yes    1.079e+00  3.039e-01   3.551 0.000384 ***
## SEX_z_F       -9.287e-01  3.427e-01  -2.710 0.006736 **
## JOB_Manager   -9.516e-01  3.863e-01  -2.463 0.013768 *
```

```

## CAR_USE_Commercial          3.411e-01  2.396e-01  1.424 0.154477
## CAR_TYPE_Pickup             1.114e+00  3.229e-01  3.449 0.000563 ***
## CAR_TYPE_Sports.Car         1.827e+00  4.786e-01  3.818 0.000134 ***
## CAR_TYPE_z_SUV              1.381e+00  4.213e-01  3.278 0.001045 **
## URBANICITY_z_Highly.Rural..Rural -2.751e+00  4.453e-01 -6.179 6.45e-10 ***
## HOME_VAL_NA                 -2.501e-01  2.296e-01 -1.089 0.276186
## oldclaim_log                7.092e-02  3.172e-02  2.236 0.025375 *
## inter                       1.900e-02  4.218e-03  4.506 6.62e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 719.44 on 624 degrees of freedom
## Residual deviance: 552.62 on 608 degrees of freedom
## AIC: 586.62
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 123  32
##           1  17  20
##
##           Accuracy : 0.7448
##           95% CI : (0.677, 0.8048)
##           No Information Rate : 0.7292
##           P-Value [Acc > NIR] : 0.3463
##
##           Kappa : 0.2894
##
## Mcnemar's Test P-Value : 0.0455
##
##           Sensitivity : 0.8786
##           Specificity : 0.3846
##           Pos Pred Value : 0.7935
##           Neg Pred Value : 0.5405
##           Prevalence : 0.7292
##           Detection Rate : 0.6406
##           Detection Prevalence : 0.8073
##           Balanced Accuracy : 0.6316
##
##           'Positive' Class : 0
##

```



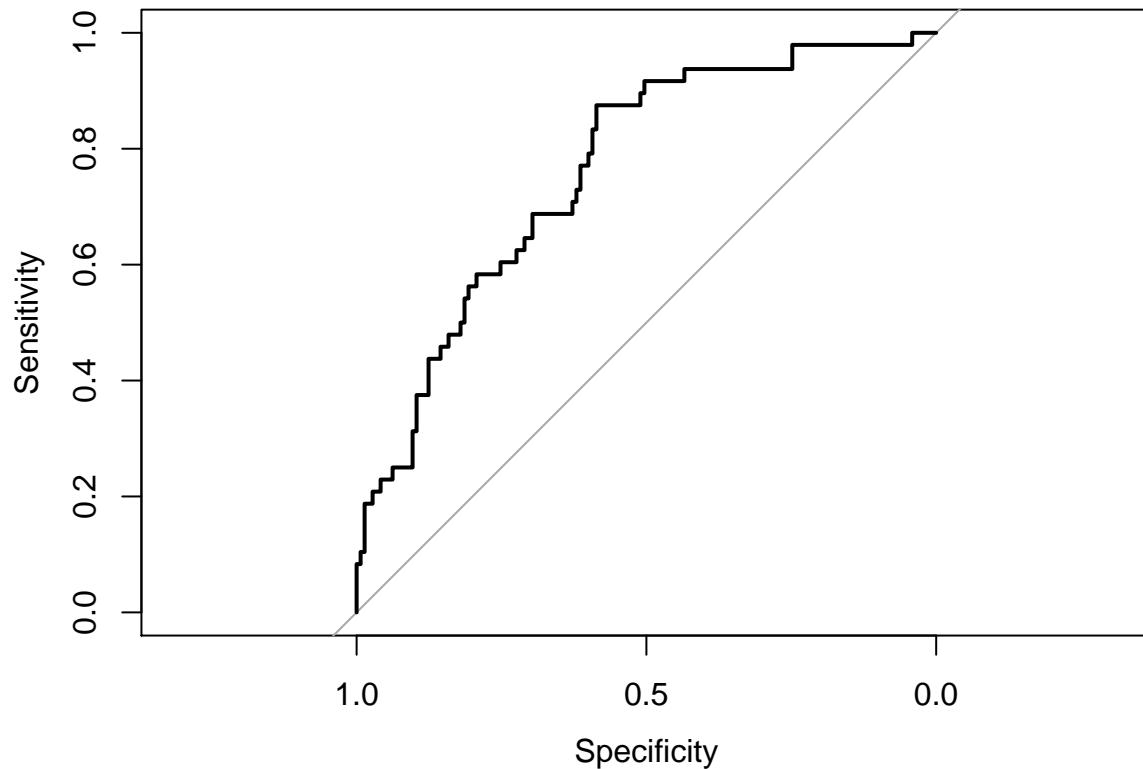
```
## [1] "AUC: 0.739835164835165"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 140 controls (dfPred_raw$class 0) < 52 cases (dfPred_raw$class 1).
## Area under the curve: 0.7398
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0152  -0.7202  -0.3946   0.7024   2.8032
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.033e+00  4.965e-01  -4.095 4.21e-05 ***
## INCOME        -1.185e-05  2.914e-06  -4.065 4.80e-05 ***
## TRAVTIME       2.186e-02  7.245e-03   3.018 0.002546 **
## BLUEBOOK       4.301e-05  1.591e-05   2.703 0.006876 **
## TIF           -4.299e-02  2.586e-02  -1.662 0.096440 .
## OLDCLAIM       1.430e-05  1.410e-05   1.014 0.310595
## PARENT1_Yes     9.659e-01  2.965e-01   3.258 0.001122 **
## SEX_z_F        -8.784e-01  3.377e-01  -2.601 0.009286 **
## JOB_Manager    -7.676e-01  3.734e-01  -2.056 0.039797 *
```



```

## CAR_USE_Commercial          3.058e-01  2.371e-01  1.290 0.197134
## CAR_TYPE_Pickup             1.051e+00  3.212e-01  3.273 0.001065 **
## CAR_TYPE_Sports.Car        1.765e+00  4.822e-01  3.660 0.000253 ***
## CAR_TYPE_z_SUV             1.655e+00  4.108e-01  4.028 5.63e-05 ***
## URBANICITY_z_Highly.Rural..Rural -2.684e+00  4.362e-01 -6.153 7.61e-10 ***
## HOME_VAL_NA                -2.828e-01  2.229e-01 -1.269 0.204474
## oldclaim_log               6.186e-02  3.177e-02  1.947 0.051475 .
## inter                      1.434e-02  4.085e-03  3.511 0.000447 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 726.95 on 623 degrees of freedom
## Residual deviance: 560.08 on 607 degrees of freedom
## AIC: 594.08
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 130  33
##           1  15  15
##
##           Accuracy : 0.7513
##           95% CI : (0.6841, 0.8106)
##           No Information Rate : 0.7513
##           P-Value [Acc > NIR] : 0.53868
##
##           Kappa : 0.239
##
## Mcnemar's Test P-Value : 0.01414
##
##           Sensitivity : 0.8966
##           Specificity : 0.3125
##           Pos Pred Value : 0.7975
##           Neg Pred Value : 0.5000
##           Prevalence : 0.7513
##           Detection Rate : 0.6736
##           Detection Prevalence : 0.8446
##           Balanced Accuracy : 0.6045
##
##           'Positive' Class : 0
##

```

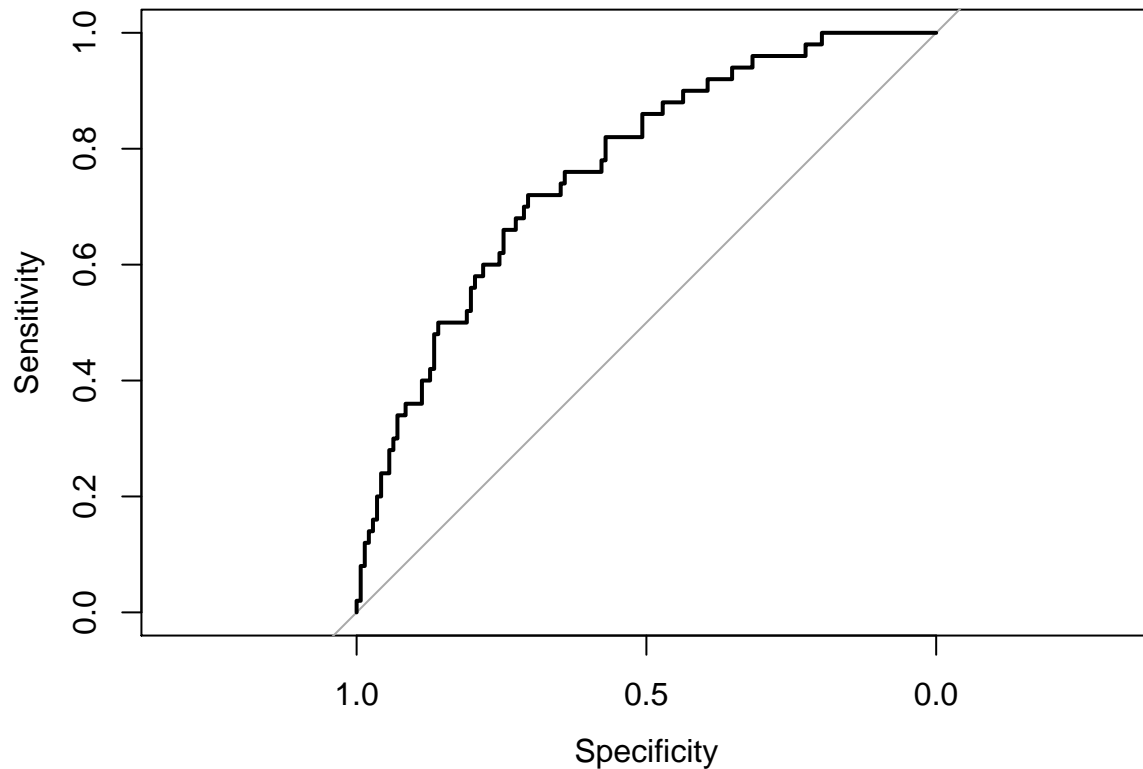


```
## [1] "AUC: 0.762068965517241"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 145 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.7621
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8959  -0.7594  -0.4191   0.5841   2.8492
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.067e+00  4.975e-01  -4.155 3.25e-05 ***
## INCOME         -9.172e-06  2.870e-06  -3.195 0.001396 **
## TRAVTIME        1.967e-02  7.175e-03   2.741 0.006128 **
## BLUEBOOK        4.081e-05  1.564e-05   2.609 0.009083 **
## TIF            -6.271e-02  2.676e-02  -2.344 0.019094 *
## OLDCLAIM        1.783e-05  1.475e-05   1.208 0.226950
## PARENT1_Yes     1.046e+00  3.073e-01   3.403 0.000668 ***
## SEX_z_F        -1.114e+00  3.639e-01  -3.062 0.002197 **
## JOB_Manager    -5.323e-01  3.641e-01  -1.462 0.143754
```

```

## CAR_USE_Commercial          6.856e-01  2.440e-01  2.810 0.004961 **
## CAR_TYPE_Pickup             1.295e+00  3.296e-01  3.929 8.52e-05 ***
## CAR_TYPE_Sports.Car         2.240e+00  5.045e-01  4.439 9.04e-06 ***
## CAR_TYPE_z_SUV              2.110e+00  4.538e-01  4.650 3.32e-06 ***
## URBANICITY_z_Highly.Rural..Rural -2.392e+00  3.954e-01 -6.050 1.45e-09 ***
## HOME_VAL_NA                 -4.106e-01  2.226e-01 -1.845 0.065106 .
## oldclaim_log                1.563e-03  3.279e-02  0.048 0.961985
## inter                       1.870e-02  4.310e-03  4.338 1.44e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 723.54  on 624  degrees of freedom
## Residual deviance: 565.72  on 608  degrees of freedom
## AIC: 599.72
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 127  32
##           1  15  18
##
##           Accuracy : 0.7552
##           95% CI : (0.6881, 0.8143)
##      No Information Rate : 0.7396
##      P-Value [Acc > NIR] : 0.3446
##
##           Kappa : 0.2858
##
## Mcnemar's Test P-Value : 0.0196
##
##           Sensitivity : 0.8944
##           Specificity : 0.3600
##      Pos Pred Value : 0.7987
##      Neg Pred Value : 0.5455
##           Prevalence : 0.7396
##      Detection Rate : 0.6615
##      Detection Prevalence : 0.8281
##      Balanced Accuracy : 0.6272
##
##           'Positive' Class : 0
##

```

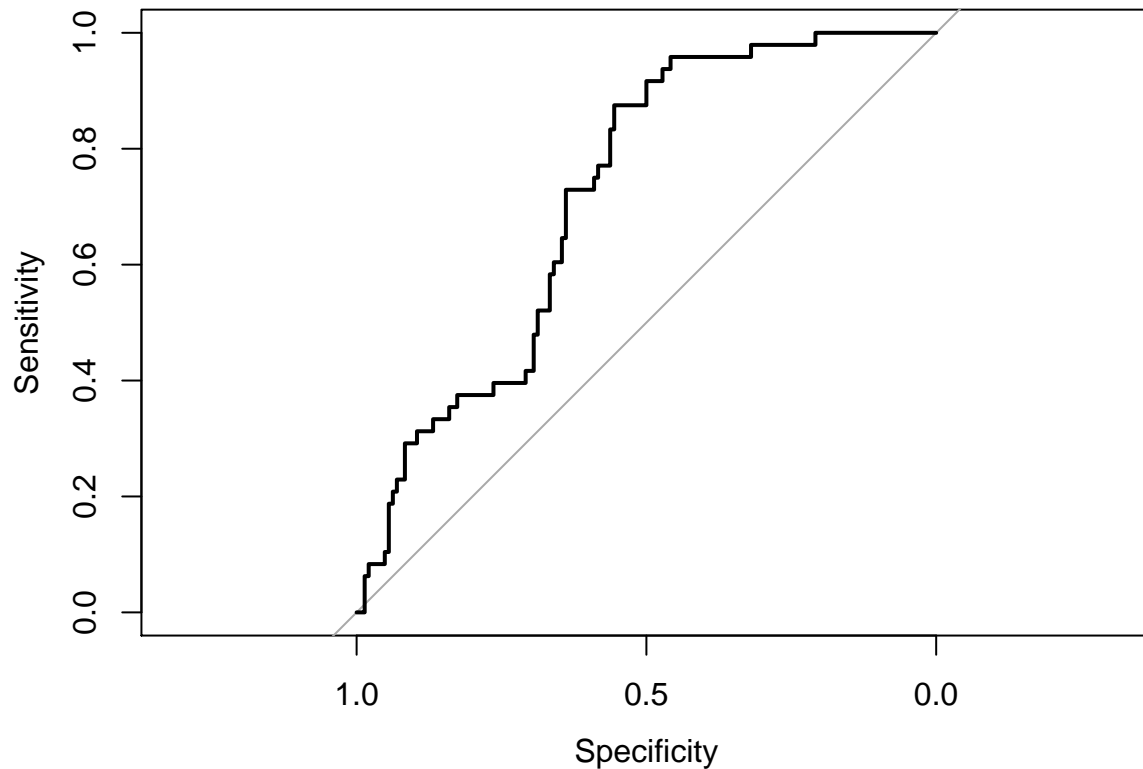


```
## [1] "AUC: 0.764929577464789"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg,      plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 142 controls (dfPred_raw$class 0) < 50 cases (dfPred_raw$class 1).
## Area under the curve: 0.7649
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8742  -0.7012  -0.3898   0.6547   2.8411
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.395e+00  4.987e-01  -4.802 1.57e-06 ***
## INCOME        -1.068e-05  3.044e-06  -3.510 0.000447 ***
## TRAVTIME       2.335e-02  6.918e-03   3.376 0.000736 ***
## BLUEBOOK       4.772e-05  1.563e-05   3.053 0.002269 **
## TIF           -1.767e-02  2.730e-02  -0.647 0.517565
## OLDCLAIM       1.445e-05  1.401e-05   1.032 0.302289
## PARENT1_Yes    1.135e+00  3.007e-01   3.774 0.000161 ***
## SEX_z_F       -9.992e-01  3.532e-01  -2.829 0.004671 **
## JOB_Manager   -1.062e+00  3.911e-01  -2.715 0.006630 **
```

```

## CAR_USE_Commercial          3.287e-01  2.396e-01  1.372 0.170205
## CAR_TYPE_Pickup             1.037e+00  3.197e-01  3.243 0.001181 **
## CAR_TYPE_Sports.Car         2.168e+00  4.900e-01  4.424 9.68e-06 ***
## CAR_TYPE_z_SUV              1.450e+00  4.286e-01  3.383 0.000718 ***
## URBANICITY_z_Highly.Rural..Rural -2.468e+00  4.141e-01 -5.961 2.51e-09 ***
## HOME_VAL_NA                 -1.947e-01  2.298e-01 -0.847 0.396838
## oldclaim_log                6.459e-02  3.177e-02  2.033 0.042071 *
## inter                       1.331e-02  4.063e-03  3.276 0.001055 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 727.58 on 624 degrees of freedom
## Residual deviance: 556.10 on 608 degrees of freedom
## AIC: 590.1
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 132  34
##           1  12  14
##
##           Accuracy : 0.7604
##           95% CI : (0.6937, 0.8189)
##           No Information Rate : 0.75
##           P-Value [Acc > NIR] : 0.40641
##
##           Kappa : 0.2459
##
## Mcnemar's Test P-Value : 0.00196
##
##           Sensitivity : 0.9167
##           Specificity : 0.2917
##           Pos Pred Value : 0.7952
##           Neg Pred Value : 0.5385
##           Prevalence : 0.7500
##           Detection Rate : 0.6875
##           Detection Prevalence : 0.8646
##           Balanced Accuracy : 0.6042
##
##           'Positive' Class : 0
##

```

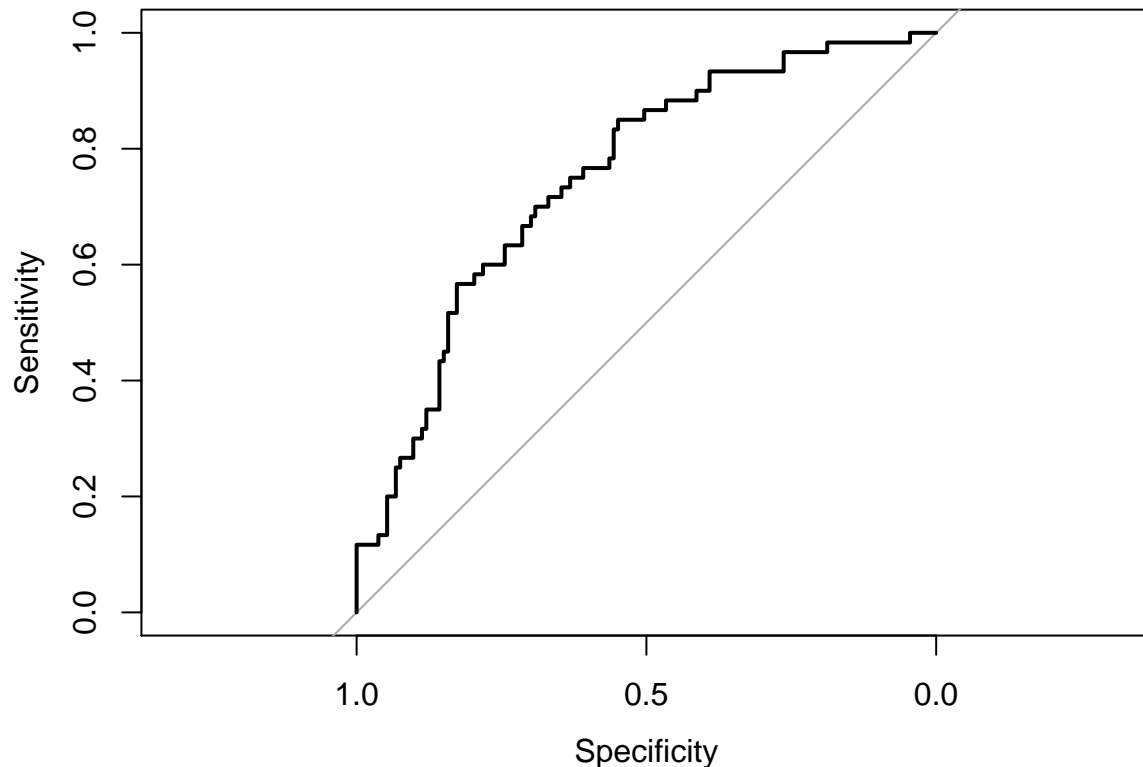


```
## [1] "AUC: 0.720486111111111"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 144 controls (dfPred_raw$class 0) < 48 cases (dfPred_raw$class 1).
## Area under the curve: 0.7205
##
## Call:
## glm(formula = fla, family = "binomial", data = train_reg)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.78814  -0.69731  -0.39259   0.05773   2.70262
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.124e+00  5.059e-01  -4.198 2.69e-05 ***
## INCOME        -1.036e-05  3.004e-06  -3.450 0.000560 ***
## TRAVTIME       1.659e-02  7.284e-03   2.278 0.022725 *
## BLUEBOOK       5.065e-05  1.581e-05   3.203 0.001358 **
## TIF           -3.512e-02  2.619e-02  -1.341 0.179898
## OLDCLAIM       1.624e-05  1.512e-05   1.074 0.282863
## PARENT1_Yes    1.375e+00  3.122e-01   4.405 1.06e-05 ***
## SEX_z_F       -1.005e+00  3.521e-01  -2.854 0.004323 **
## JOB_Manager   -5.935e-01  3.669e-01  -1.618 0.105717
```

```

## CAR_USE_Commercial          5.014e-01  2.434e-01  2.060 0.039415 *
## CAR_TYPE_Pickup             1.066e+00  3.294e-01  3.235 0.001215 **
## CAR_TYPE_Sports.Car         1.831e+00  4.849e-01  3.776 0.000159 ***
## CAR_TYPE_z_SUV              1.658e+00  4.359e-01  3.802 0.000143 ***
## URBANICITY_z_Highly.Rural..Rural -2.536e+00  4.261e-01 -5.952 2.64e-09 ***
## HOME_VAL_NA                 -4.965e-01  2.289e-01 -2.169 0.030061 *
## oldclaim_log                4.685e-02  3.305e-02  1.418 0.156307
## inter                       1.593e-02  4.314e-03  3.693 0.000221 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 701.79  on 623  degrees of freedom
## Residual deviance: 539.18  on 607  degrees of freedom
## AIC: 573.18
##
## Number of Fisher Scoring iterations: 5
##
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 120  43
##           1  13  17
##
##           Accuracy : 0.7098
##           95% CI : (0.6403, 0.7728)
##      No Information Rate : 0.6891
##      P-Value [Acc > NIR] : 0.2955652
##
##           Kappa : 0.2151
##
## Mcnemar's Test P-Value : 0.0001065
##
##           Sensitivity : 0.9023
##           Specificity : 0.2833
##      Pos Pred Value : 0.7362
##      Neg Pred Value : 0.5667
##           Prevalence : 0.6891
##      Detection Rate : 0.6218
##      Detection Prevalence : 0.8446
##      Balanced Accuracy : 0.5928
##
##           'Positive' Class : 0
##

```



```
## [1] "AUC: 0.751002506265664"
##
## Call:
## roc.default(response = dfPred_raw$class, predictor = dfPred_raw$predict_reg, plot = TRUE)
##
## Data: dfPred_raw$predict_reg in 133 controls (dfPred_raw$class 0) < 60 cases (dfPred_raw$class 1).
## Area under the curve: 0.751
## [1] "Accuracy: 0.755297387737478"
## [1] "AIC: 597.141230813576"
## [1] "AUC: 0.776814396431349"
```

This model has an accuracy of .755, an AIC of 597, and an AUC of .777. This is the best model so far.

#### 4. Select model

We select model 3

#### 5. Predict TARGET\_AMT

#### 10. Conclusion

We examined 466 records of town statistics to create a predictive model of whether crime rates were above the median or not. We used a logistic regression to do this, testing our models on an 80/20 split 100 times and taking the average accuracy and AIC.



Several enhancements to the model increased accuracy and lowered AIC. First, some predictors were transformed with the log or square to improve fit. Second, dummy variables were introduced to capture the fact that highly industrial areas appeared to operate by a different logic than mixed use areas. interaction terms to model this phenomenon did not improve the model. The final model 93% accurate.