Logistic Regression

Installing all the packages:-

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(readr)  
library(ggplot2)  
library(DescTools)  
library(moments)  
library(ISLR)  
library(aod)

Reading <Data:->

data<-read.csv("Placement\_Data\_Full\_Class.csv")  
head(data)

## sl\_no gender ssc\_p ssc\_b hsc\_p hsc\_b hsc\_s degree\_p degree\_t workex  
## 1 1 M 67.00 Others 91.00 Others Commerce 58.00 Sci&Tech No  
## 2 2 M 79.33 Central 78.33 Others Science 77.48 Sci&Tech Yes  
## 3 3 M 65.00 Central 68.00 Central Arts 64.00 Comm&Mgmt No  
## 4 4 M 56.00 Central 52.00 Central Science 52.00 Sci&Tech No  
## 5 5 M 85.80 Central 73.60 Central Commerce 73.30 Comm&Mgmt No  
## 6 6 M 55.00 Others 49.80 Others Science 67.25 Sci&Tech Yes  
## etest\_p specialisation mba\_p status salary  
## 1 55.0 Mkt&HR 58.80 Placed 270000  
## 2 86.5 Mkt&Fin 66.28 Placed 200000  
## 3 75.0 Mkt&Fin 57.80 Placed 250000  
## 4 66.0 Mkt&HR 59.43 Not Placed NA  
## 5 96.8 Mkt&Fin 55.50 Placed 425000  
## 6 55.0 Mkt&Fin 51.58 Not Placed NA

Gaining insights from the <data:->

summary(data)

## sl\_no gender ssc\_p ssc\_b   
## Min. : 1.0 Length:215 Min. :40.89 Length:215   
## 1st Qu.: 54.5 Class :character 1st Qu.:60.60 Class :character   
## Median :108.0 Mode :character Median :67.00 Mode :character   
## Mean :108.0 Mean :67.30   
## 3rd Qu.:161.5 3rd Qu.:75.70   
## Max. :215.0 Max. :89.40   
##   
## hsc\_p hsc\_b hsc\_s degree\_p   
## Min. :37.00 Length:215 Length:215 Min. :50.00   
## 1st Qu.:60.90 Class :character Class :character 1st Qu.:61.00   
## Median :65.00 Mode :character Mode :character Median :66.00   
## Mean :66.33 Mean :66.37   
## 3rd Qu.:73.00 3rd Qu.:72.00   
## Max. :97.70 Max. :91.00   
##   
## degree\_t workex etest\_p specialisation   
## Length:215 Length:215 Min. :50.0 Length:215   
## Class :character Class :character 1st Qu.:60.0 Class :character   
## Mode :character Mode :character Median :71.0 Mode :character   
## Mean :72.1   
## 3rd Qu.:83.5   
## Max. :98.0   
##   
## mba\_p status salary   
## Min. :51.21 Length:215 Min. :200000   
## 1st Qu.:57.95 Class :character 1st Qu.:240000   
## Median :62.00 Mode :character Median :265000   
## Mean :62.28 Mean :288655   
## 3rd Qu.:66.25 3rd Qu.:300000   
## Max. :77.89 Max. :940000   
## NA's :67

str(data)

## 'data.frame': 215 obs. of 15 variables:  
## $ sl\_no : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ gender : chr "M" "M" "M" "M" ...  
## $ ssc\_p : num 67 79.3 65 56 85.8 ...  
## $ ssc\_b : chr "Others" "Central" "Central" "Central" ...  
## $ hsc\_p : num 91 78.3 68 52 73.6 ...  
## $ hsc\_b : chr "Others" "Others" "Central" "Central" ...  
## $ hsc\_s : chr "Commerce" "Science" "Arts" "Science" ...  
## $ degree\_p : num 58 77.5 64 52 73.3 ...  
## $ degree\_t : chr "Sci&Tech" "Sci&Tech" "Comm&Mgmt" "Sci&Tech" ...  
## $ workex : chr "No" "Yes" "No" "No" ...  
## $ etest\_p : num 55 86.5 75 66 96.8 ...  
## $ specialisation: chr "Mkt&HR" "Mkt&Fin" "Mkt&Fin" "Mkt&HR" ...  
## $ mba\_p : num 58.8 66.3 57.8 59.4 55.5 ...  
## $ status : chr "Placed" "Placed" "Placed" "Not Placed" ...  
## $ salary : int 270000 200000 250000 NA 425000 NA NA 252000 231000 NA ...

Removing NA values from salary(if not placed) to 0:-

data$salary[is.na(data$salary)]<-0;  
head(data)

## sl\_no gender ssc\_p ssc\_b hsc\_p hsc\_b hsc\_s degree\_p degree\_t workex  
## 1 1 M 67.00 Others 91.00 Others Commerce 58.00 Sci&Tech No  
## 2 2 M 79.33 Central 78.33 Others Science 77.48 Sci&Tech Yes  
## 3 3 M 65.00 Central 68.00 Central Arts 64.00 Comm&Mgmt No  
## 4 4 M 56.00 Central 52.00 Central Science 52.00 Sci&Tech No  
## 5 5 M 85.80 Central 73.60 Central Commerce 73.30 Comm&Mgmt No  
## 6 6 M 55.00 Others 49.80 Others Science 67.25 Sci&Tech Yes  
## etest\_p specialisation mba\_p status salary  
## 1 55.0 Mkt&HR 58.80 Placed 270000  
## 2 86.5 Mkt&Fin 66.28 Placed 200000  
## 3 75.0 Mkt&Fin 57.80 Placed 250000  
## 4 66.0 Mkt&HR 59.43 Not Placed 0  
## 5 96.8 Mkt&Fin 55.50 Placed 425000  
## 6 55.0 Mkt&Fin 51.58 Not Placed 0

Feature selection:-

data<-subset(data,select = -c(sl\_no,ssc\_b,hsc\_b))  
head(data)

## gender ssc\_p hsc\_p hsc\_s degree\_p degree\_t workex etest\_p specialisation  
## 1 M 67.00 91.00 Commerce 58.00 Sci&Tech No 55.0 Mkt&HR  
## 2 M 79.33 78.33 Science 77.48 Sci&Tech Yes 86.5 Mkt&Fin  
## 3 M 65.00 68.00 Arts 64.00 Comm&Mgmt No 75.0 Mkt&Fin  
## 4 M 56.00 52.00 Science 52.00 Sci&Tech No 66.0 Mkt&HR  
## 5 M 85.80 73.60 Commerce 73.30 Comm&Mgmt No 96.8 Mkt&Fin  
## 6 M 55.00 49.80 Science 67.25 Sci&Tech Yes 55.0 Mkt&Fin  
## mba\_p status salary  
## 1 58.80 Placed 270000  
## 2 66.28 Placed 200000  
## 3 57.80 Placed 250000  
## 4 59.43 Not Placed 0  
## 5 55.50 Placed 425000  
## 6 51.58 Not Placed 0

Checking for outliers:-

boxplot(data$ssc\_p,xlab="10th Marks")

boxplot(data$degree\_p,xlab="degree")

boxplot(data$mba\_p,xlab="MBA")

boxplot(data$etest\_p,xlab="Employbality test")

boxplot(data$hsc\_p,xlab="12th Marks")

so there is outliers in column hsc\_p. No other columns have a significant amount of outliers.

Removing outliers from hsc\_p column:-

q1<-quantile(data$hsc\_p,0.25)  
q3<-quantile(data$hsc\_p,0.75)  
iqr<-1.5\*(q3-q1)  
data<- data %>% filter(data$hsc\_p>=q1-iqr,data$hsc\_p<=q3+iqr)  
boxplot(data$hsc\_p,xlab="after Cleaning 12th Marks")

Preparing the data for the model Logistic Regression

Identifying the unique data in the categorical columns:-

unique(data$hsc\_s,incomparables = FALSE)

## [1] "Commerce" "Science" "Arts"

unique(data$degree\_t,incomparables = FALSE)

## [1] "Sci&Tech" "Comm&Mgmt" "Others"

unique(data$specialisation,incomparables = FALSE)

## [1] "Mkt&HR" "Mkt&Fin"

Replacing the categorical values with corresponding numbers:-

data$workex<-as.numeric(as.factor(data$workex))-1  
data$gender<-as.numeric(as.factor(data$gender))-1  
data$specialisation<-as.numeric(as.factor(data$specialisation))-1  
data$status<-as.numeric(as.factor(data$status))-1  
data

## gender ssc\_p hsc\_p hsc\_s degree\_p degree\_t workex etest\_p  
## 1 1 67.00 91.00 Commerce 58.00 Sci&Tech 0 55.00  
## 2 1 79.33 78.33 Science 77.48 Sci&Tech 1 86.50  
## 3 1 65.00 68.00 Arts 64.00 Comm&Mgmt 0 75.00  
## 4 1 56.00 52.00 Science 52.00 Sci&Tech 0 66.00  
## 5 1 85.80 73.60 Commerce 73.30 Comm&Mgmt 0 96.80  
## 6 1 55.00 49.80 Science 67.25 Sci&Tech 1 55.00  
## 7 0 46.00 49.20 Commerce 79.00 Comm&Mgmt 0 74.28  
## 8 1 82.00 64.00 Science 66.00 Sci&Tech 1 67.00  
## 9 1 73.00 79.00 Commerce 72.00 Comm&Mgmt 0 91.34  
## 10 1 58.00 70.00 Commerce 61.00 Comm&Mgmt 0 54.00  
## 11 1 58.00 61.00 Commerce 60.00 Comm&Mgmt 1 62.00  
## 12 1 69.60 68.40 Commerce 78.30 Comm&Mgmt 1 60.00  
## 13 0 47.00 55.00 Science 65.00 Comm&Mgmt 0 62.00  
## 14 0 77.00 87.00 Commerce 59.00 Comm&Mgmt 0 68.00  
## 15 1 62.00 47.00 Commerce 50.00 Comm&Mgmt 0 76.00  
## 16 0 65.00 75.00 Commerce 69.00 Comm&Mgmt 1 72.00  
## 17 1 63.00 66.20 Commerce 65.60 Comm&Mgmt 1 60.00  
## 18 0 55.00 67.00 Commerce 64.00 Comm&Mgmt 0 60.00  
## 19 0 63.00 66.00 Commerce 64.00 Comm&Mgmt 0 68.00  
## 20 1 60.00 67.00 Arts 70.00 Comm&Mgmt 1 50.48  
## 21 1 62.00 65.00 Commerce 66.00 Comm&Mgmt 0 50.00  
## 22 0 79.00 76.00 Commerce 85.00 Comm&Mgmt 0 95.00  
## 23 0 69.80 60.80 Science 72.23 Sci&Tech 0 55.53  
## 24 0 77.40 60.00 Science 64.74 Sci&Tech 1 92.00  
## 25 0 52.58 54.60 Commerce 50.20 Comm&Mgmt 1 76.00  
## 26 1 71.00 79.00 Commerce 66.00 Comm&Mgmt 1 94.00  
## 27 1 63.00 67.00 Commerce 66.00 Comm&Mgmt 0 68.00  
## 28 1 76.76 76.50 Commerce 67.50 Comm&Mgmt 1 73.35  
## 29 1 62.00 67.00 Commerce 58.00 Comm&Mgmt 0 77.00  
## 30 0 64.00 73.50 Commerce 73.00 Comm&Mgmt 0 52.00  
## 31 0 67.00 53.00 Science 65.00 Sci&Tech 0 64.00  
## 32 0 61.00 81.00 Commerce 66.40 Comm&Mgmt 0 50.89  
## 33 0 87.00 65.00 Science 81.00 Comm&Mgmt 1 88.00  
## 34 1 62.00 51.00 Science 52.00 Others 0 68.44  
## 35 0 69.00 78.00 Commerce 72.00 Comm&Mgmt 0 71.00  
## 36 1 51.00 44.00 Commerce 57.00 Comm&Mgmt 0 64.00  
## 37 0 79.00 76.00 Science 65.60 Sci&Tech 0 58.00  
## 38 0 73.00 58.00 Science 66.00 Comm&Mgmt 0 53.70  
## 39 1 81.00 68.00 Science 64.00 Sci&Tech 0 93.00  
## 40 0 78.00 77.00 Commerce 80.00 Comm&Mgmt 0 60.00  
## 41 0 74.00 63.16 Commerce 65.00 Comm&Mgmt 1 65.00  
## 42 1 87.00 87.00 Commerce 68.00 Comm&Mgmt 0 95.00  
## 43 0 77.00 73.00 Commerce 81.00 Comm&Mgmt 1 89.00  
## 44 0 76.00 64.00 Science 72.00 Sci&Tech 0 58.00  
## 45 0 70.89 71.98 Science 65.60 Comm&Mgmt 0 68.00  
## 46 1 63.00 60.00 Commerce 57.00 Comm&Mgmt 1 78.00  
## 47 1 63.00 62.00 Commerce 68.00 Comm&Mgmt 0 64.00  
## 48 0 75.20 73.20 Science 68.40 Comm&Mgmt 0 65.00  
## 49 1 54.40 61.12 Commerce 56.20 Comm&Mgmt 0 67.00  
## 50 0 40.89 45.83 Commerce 53.00 Comm&Mgmt 0 71.20  
## 51 1 80.00 70.00 Science 72.00 Sci&Tech 0 87.00  
## 52 0 74.00 60.00 Science 69.00 Comm&Mgmt 0 78.00  
## 53 1 60.40 66.60 Science 65.00 Comm&Mgmt 0 71.00  
## 54 1 63.00 71.40 Commerce 61.40 Comm&Mgmt 0 68.00  
## 55 1 68.00 76.00 Commerce 74.00 Comm&Mgmt 0 80.00  
## 56 1 74.00 62.00 Science 68.00 Comm&Mgmt 0 74.00  
## 57 1 52.60 65.58 Science 72.11 Sci&Tech 0 57.60  
## 58 1 74.00 70.00 Science 72.00 Comm&Mgmt 1 60.00  
## 59 1 84.20 73.40 Commerce 66.89 Comm&Mgmt 0 61.60  
## 60 0 86.50 64.20 Science 67.40 Sci&Tech 0 59.00  
## 61 1 61.00 70.00 Commerce 64.00 Comm&Mgmt 0 68.50  
## 62 1 80.00 73.00 Commerce 75.00 Comm&Mgmt 0 61.00  
## 63 1 54.00 47.00 Science 57.00 Comm&Mgmt 0 89.69  
## 64 1 83.00 74.00 Science 66.00 Comm&Mgmt 0 68.92  
## 65 1 80.92 78.50 Commerce 67.00 Comm&Mgmt 0 68.71  
## 66 0 69.70 47.00 Commerce 72.70 Sci&Tech 0 79.00  
## 67 1 73.00 73.00 Science 66.00 Sci&Tech 1 70.00  
## 68 1 82.00 61.00 Science 62.00 Sci&Tech 0 89.00  
## 69 1 75.00 70.29 Commerce 71.00 Comm&Mgmt 0 95.00  
## 70 1 84.86 67.00 Science 78.00 Comm&Mgmt 0 95.50  
## 71 1 64.60 83.83 Commerce 71.72 Comm&Mgmt 0 86.00  
## 72 1 56.60 64.80 Commerce 70.20 Comm&Mgmt 0 84.27  
## 73 0 59.00 62.00 Commerce 77.50 Comm&Mgmt 0 74.00  
## 74 0 66.50 70.40 Arts 71.93 Comm&Mgmt 0 61.00  
## 75 1 64.00 80.00 Science 65.00 Sci&Tech 1 69.00  
## 76 1 84.00 90.90 Science 64.50 Sci&Tech 0 86.04  
## 77 0 69.00 62.00 Science 66.00 Sci&Tech 0 75.00  
## 78 0 69.00 62.00 Commerce 69.00 Comm&Mgmt 1 67.00  
## 79 1 81.70 63.00 Science 67.00 Comm&Mgmt 1 86.00  
## 80 1 63.00 67.00 Commerce 74.00 Comm&Mgmt 0 82.00  
## 81 1 84.00 79.00 Science 68.00 Sci&Tech 1 84.00  
## 82 1 70.00 63.00 Science 70.00 Sci&Tech 1 55.00  
## 83 0 83.84 89.83 Commerce 77.20 Comm&Mgmt 1 78.74  
## 84 1 62.00 63.00 Commerce 64.00 Comm&Mgmt 0 67.00  
## 85 1 59.60 51.00 Science 60.00 Others 0 75.00  
## 86 0 66.00 62.00 Commerce 73.00 Comm&Mgmt 0 58.00  
## 87 0 84.00 75.00 Science 69.00 Sci&Tech 1 62.00  
## 88 0 85.00 90.00 Commerce 82.00 Comm&Mgmt 0 92.00  
## 89 1 52.00 57.00 Commerce 50.80 Comm&Mgmt 0 67.00  
## 90 0 60.23 69.00 Science 66.00 Comm&Mgmt 0 72.00  
## 91 1 52.00 62.00 Commerce 54.00 Comm&Mgmt 0 72.00  
## 92 1 58.00 62.00 Commerce 64.00 Comm&Mgmt 0 53.88  
## 93 1 73.00 78.00 Commerce 65.00 Comm&Mgmt 1 95.46  
## 94 0 76.00 70.00 Science 76.00 Comm&Mgmt 1 66.00  
## 95 0 70.50 62.50 Commerce 61.00 Comm&Mgmt 0 93.91  
## 96 0 69.00 73.00 Commerce 65.00 Comm&Mgmt 0 70.00  
## 97 1 54.00 82.00 Commerce 63.00 Sci&Tech 0 50.00  
## 98 0 45.00 57.00 Commerce 58.00 Comm&Mgmt 1 56.39  
## 99 1 63.00 72.00 Commerce 68.00 Comm&Mgmt 0 78.00  
## 100 0 77.00 61.00 Commerce 68.00 Comm&Mgmt 1 57.50  
## 101 1 73.00 78.00 Science 73.00 Sci&Tech 1 85.00  
## 102 1 69.00 63.00 Science 65.00 Comm&Mgmt 1 55.00  
## 103 1 59.00 64.00 Science 58.00 Sci&Tech 0 85.00  
## 104 1 61.08 50.00 Science 54.00 Sci&Tech 0 71.00  
## 105 1 82.00 90.00 Commerce 83.00 Comm&Mgmt 0 80.00  
## 106 1 61.00 82.00 Commerce 69.00 Comm&Mgmt 0 84.00  
## 107 1 52.00 63.00 Science 65.00 Sci&Tech 1 86.00  
## 108 0 69.50 70.00 Science 72.00 Sci&Tech 0 57.20  
## 109 1 51.00 54.00 Science 61.00 Sci&Tech 0 60.00  
## 110 1 58.00 61.00 Commerce 61.00 Comm&Mgmt 0 58.00  
## 111 0 73.96 79.00 Commerce 67.00 Comm&Mgmt 0 72.15  
## 112 1 65.00 68.00 Science 69.00 Comm&Mgmt 0 53.70  
## 113 0 73.00 63.00 Science 66.00 Comm&Mgmt 0 89.00  
## 114 1 68.20 72.80 Commerce 66.60 Comm&Mgmt 1 96.00  
## 115 1 77.00 75.00 Science 73.00 Sci&Tech 0 80.00  
## 116 1 76.00 80.00 Science 78.00 Sci&Tech 1 97.00  
## 117 1 60.80 68.40 Commerce 64.60 Comm&Mgmt 1 82.66  
## 118 0 64.00 67.00 Science 69.60 Sci&Tech 1 55.67  
## 119 0 66.50 66.80 Arts 69.30 Comm&Mgmt 1 80.40  
## 120 1 74.00 59.00 Commerce 73.00 Comm&Mgmt 1 60.00  
## 121 1 67.00 71.00 Science 64.33 Others 1 64.00  
## 122 0 84.00 73.00 Commerce 73.00 Comm&Mgmt 0 75.00  
## 123 0 79.00 61.00 Science 75.50 Sci&Tech 1 70.00  
## 124 0 72.00 60.00 Science 69.00 Comm&Mgmt 0 55.50  
## 125 1 80.40 73.40 Science 77.72 Sci&Tech 1 81.20  
## 126 1 76.70 89.70 Commerce 66.00 Comm&Mgmt 1 90.00  
## 127 1 62.00 65.00 Commerce 60.00 Comm&Mgmt 0 84.00  
## 128 0 74.90 57.00 Science 62.00 Others 1 80.00  
## 129 1 67.00 68.00 Commerce 64.00 Comm&Mgmt 1 74.40  
## 130 1 73.00 64.00 Commerce 77.00 Comm&Mgmt 1 65.00  
## 131 0 72.00 56.00 Science 69.00 Comm&Mgmt 0 55.60  
## 132 0 47.00 59.00 Arts 64.00 Comm&Mgmt 0 78.00  
## 133 1 67.00 63.00 Commerce 72.00 Comm&Mgmt 0 56.00  
## 134 0 82.00 64.00 Science 73.00 Sci&Tech 1 96.00  
## 135 1 77.00 70.00 Commerce 59.00 Comm&Mgmt 1 58.00  
## 136 1 65.00 64.80 Commerce 69.50 Comm&Mgmt 1 56.00  
## 137 1 66.00 64.00 Science 60.00 Comm&Mgmt 0 60.00  
## 138 1 85.00 60.00 Science 73.43 Sci&Tech 1 60.00  
## 139 1 77.67 64.89 Commerce 70.67 Comm&Mgmt 0 89.00  
## 140 1 52.00 50.00 Arts 61.00 Comm&Mgmt 0 60.00  
## 141 1 89.40 65.66 Science 71.25 Sci&Tech 0 72.00  
## 142 1 62.00 63.00 Science 66.00 Comm&Mgmt 0 85.00  
## 143 1 70.00 74.00 Commerce 65.00 Comm&Mgmt 0 83.00  
## 144 0 77.00 86.00 Arts 56.00 Others 0 57.00  
## 145 1 44.00 58.00 Arts 55.00 Comm&Mgmt 1 64.25  
## 146 1 71.00 58.66 Science 58.00 Sci&Tech 1 56.00  
## 147 1 65.00 65.00 Commerce 75.00 Comm&Mgmt 0 83.00  
## 148 0 75.40 60.50 Science 84.00 Sci&Tech 0 98.00  
## 149 1 49.00 59.00 Science 65.00 Sci&Tech 1 86.00  
## 150 1 53.00 63.00 Science 60.00 Comm&Mgmt 1 70.00  
## 151 1 51.57 74.66 Commerce 59.90 Comm&Mgmt 1 56.15  
## 152 1 84.20 69.40 Science 65.00 Sci&Tech 1 80.00  
## 153 1 66.50 62.50 Commerce 60.90 Comm&Mgmt 0 93.40  
## 154 1 67.00 63.00 Science 64.00 Sci&Tech 0 60.00  
## 155 1 52.00 49.00 Commerce 58.00 Comm&Mgmt 0 62.00  
## 156 1 87.00 74.00 Science 65.00 Sci&Tech 1 75.00  
## 157 1 55.60 51.00 Commerce 57.50 Comm&Mgmt 0 57.63  
## 158 1 74.20 87.60 Commerce 77.25 Comm&Mgmt 1 75.20  
## 159 1 63.00 67.00 Science 64.00 Sci&Tech 0 75.00  
## 160 0 67.16 72.50 Commerce 63.35 Comm&Mgmt 0 53.04  
## 161 0 63.30 78.33 Commerce 74.00 Comm&Mgmt 0 80.00  
## 162 1 62.00 62.00 Commerce 60.00 Comm&Mgmt 1 63.00  
## 163 1 67.90 62.00 Science 67.00 Sci&Tech 1 58.10  
## 164 0 48.00 51.00 Commerce 58.00 Comm&Mgmt 1 60.00  
## 165 0 63.40 67.20 Commerce 60.00 Comm&Mgmt 0 58.06  
## 166 1 80.00 80.00 Commerce 72.00 Comm&Mgmt 1 63.79  
## 167 1 73.00 58.00 Commerce 56.00 Comm&Mgmt 0 84.00  
## 168 0 52.00 52.00 Science 55.00 Sci&Tech 0 67.00  
## 169 1 73.24 50.83 Science 64.27 Sci&Tech 1 64.00  
## 170 1 63.00 62.00 Science 65.00 Sci&Tech 0 87.50  
## 171 0 59.00 60.00 Commerce 56.00 Comm&Mgmt 0 55.00  
## 172 1 68.00 56.00 Science 68.00 Sci&Tech 0 73.00  
## 173 0 77.80 64.00 Science 64.20 Sci&Tech 0 75.50  
## 174 1 65.00 71.50 Commerce 62.80 Comm&Mgmt 1 57.00  
## 175 1 62.00 60.33 Science 64.21 Sci&Tech 0 63.00  
## 176 1 52.00 65.00 Arts 57.00 Others 1 75.00  
## 177 1 65.00 77.00 Commerce 69.00 Comm&Mgmt 0 60.00  
## 178 0 56.28 62.83 Commerce 59.79 Comm&Mgmt 0 60.00  
## 179 0 88.00 72.00 Science 78.00 Others 0 82.00  
## 180 0 52.00 64.00 Commerce 61.00 Comm&Mgmt 0 55.00  
## 181 1 78.50 65.50 Science 67.00 Sci&Tech 1 95.00  
## 182 1 61.80 47.00 Commerce 54.38 Comm&Mgmt 0 57.00  
## 183 0 54.00 77.60 Commerce 69.20 Comm&Mgmt 0 95.65  
## 184 0 64.00 70.20 Commerce 61.00 Comm&Mgmt 0 50.00  
## 185 1 67.00 61.00 Science 72.00 Comm&Mgmt 0 72.00  
## 186 1 65.20 61.40 Commerce 64.80 Comm&Mgmt 1 93.40  
## 187 0 60.00 63.00 Arts 56.00 Others 1 80.00  
## 188 1 52.00 55.00 Commerce 56.30 Comm&Mgmt 0 59.00  
## 189 1 66.00 76.00 Commerce 72.00 Comm&Mgmt 1 84.00  
## 190 1 72.00 63.00 Science 77.50 Sci&Tech 1 78.00  
## 191 0 83.96 53.00 Science 91.00 Sci&Tech 0 59.32  
## 192 0 67.00 70.00 Commerce 65.00 Others 0 88.00  
## 193 1 69.00 65.00 Commerce 57.00 Comm&Mgmt 0 73.00  
## 194 1 69.00 60.00 Commerce 65.00 Comm&Mgmt 0 87.55  
## 195 1 54.20 63.00 Science 58.00 Comm&Mgmt 0 79.00  
## 196 1 70.00 63.00 Science 66.00 Sci&Tech 0 61.28  
## 197 1 55.68 61.33 Commerce 56.87 Comm&Mgmt 0 66.00  
## 198 0 74.00 73.00 Commerce 73.00 Comm&Mgmt 1 80.00  
## 199 1 61.00 62.00 Commerce 65.00 Comm&Mgmt 0 62.00  
## 200 1 83.33 78.00 Commerce 61.00 Comm&Mgmt 1 88.56  
## 201 0 43.00 60.00 Science 65.00 Comm&Mgmt 0 92.66  
## 202 1 62.00 72.00 Commerce 65.00 Comm&Mgmt 0 67.00  
## 203 1 80.60 82.00 Commerce 77.60 Comm&Mgmt 0 91.00  
## 204 1 58.00 60.00 Science 72.00 Sci&Tech 0 74.00  
## 205 1 67.00 67.00 Commerce 73.00 Comm&Mgmt 1 59.00  
## 206 0 74.00 66.00 Commerce 58.00 Comm&Mgmt 0 70.00  
## 207 1 62.00 58.00 Science 53.00 Comm&Mgmt 0 89.00  
## specialisation mba\_p status salary  
## 1 1 58.80 1 270000  
## 2 0 66.28 1 200000  
## 3 0 57.80 1 250000  
## 4 1 59.43 0 0  
## 5 0 55.50 1 425000  
## 6 0 51.58 0 0  
## 7 0 53.29 0 0  
## 8 0 62.14 1 252000  
## 9 0 61.29 1 231000  
## 10 0 52.21 0 0  
## 11 1 60.85 1 260000  
## 12 0 63.70 1 250000  
## 13 1 65.04 0 0  
## 14 0 68.63 1 218000  
## 15 1 54.96 0 0  
## 16 0 64.66 1 200000  
## 17 0 62.54 1 300000  
## 18 0 67.28 0 0  
## 19 1 64.08 0 0  
## 20 0 77.89 1 236000  
## 21 1 56.70 1 265000  
## 22 0 69.06 1 393000  
## 23 1 68.81 1 360000  
## 24 0 63.62 1 300000  
## 25 0 65.33 0 0  
## 26 0 57.55 1 240000  
## 27 1 57.69 1 265000  
## 28 0 64.15 1 350000  
## 29 0 51.29 0 0  
## 30 1 56.70 1 250000  
## 31 1 58.32 0 0  
## 32 1 62.21 1 278000  
## 33 0 72.78 1 260000  
## 34 1 62.77 0 0  
## 35 1 62.74 1 300000  
## 36 0 51.45 0 0  
## 37 1 55.47 1 320000  
## 38 1 56.86 1 240000  
## 39 0 62.56 1 411000  
## 40 0 66.72 1 287000  
## 41 1 69.76 0 0  
## 42 1 62.90 1 300000  
## 43 0 69.70 1 200000  
## 44 1 66.53 0 0  
## 45 1 71.63 0 0  
## 46 0 54.55 1 204000  
## 47 0 62.46 1 250000  
## 48 1 62.98 1 200000  
## 49 1 62.65 0 0  
## 50 1 65.49 0 0  
## 51 1 71.04 1 450000  
## 52 1 65.56 1 216000  
## 53 1 52.71 1 220000  
## 54 0 66.88 1 240000  
## 55 0 63.59 1 360000  
## 56 0 57.99 1 268000  
## 57 0 56.66 1 265000  
## 58 0 57.24 1 260000  
## 59 0 62.48 1 300000  
## 60 0 59.69 1 240000  
## 61 1 59.50 0 0  
## 62 0 58.78 1 240000  
## 63 1 57.10 0 0  
## 64 1 58.46 1 275000  
## 65 0 60.99 1 275000  
## 66 1 59.24 0 0  
## 67 0 68.07 1 275000  
## 68 0 65.45 1 360000  
## 69 0 66.94 1 240000  
## 70 0 68.53 1 240000  
## 71 0 59.75 1 218000  
## 72 0 67.20 1 336000  
## 73 1 67.00 0 0  
## 74 0 64.27 1 230000  
## 75 0 57.65 1 500000  
## 76 0 59.42 1 270000  
## 77 1 67.99 0 0  
## 78 1 62.35 1 240000  
## 79 0 70.20 1 300000  
## 80 0 60.44 0 0  
## 81 0 66.69 1 300000  
## 82 0 62.00 1 300000  
## 83 0 76.18 1 400000  
## 84 0 57.03 1 220000  
## 85 1 59.08 0 0  
## 86 1 64.36 1 210000  
## 87 1 62.36 1 210000  
## 88 0 68.03 1 300000  
## 89 1 62.79 0 0  
## 90 0 59.47 1 230000  
## 91 1 55.41 0 0  
## 92 0 54.97 1 260000  
## 93 0 62.16 1 420000  
## 94 0 64.44 1 300000  
## 95 0 69.03 0 0  
## 96 0 57.31 1 220000  
## 97 0 59.47 0 0  
## 98 1 64.95 0 0  
## 99 1 60.44 1 380000  
## 100 0 61.31 1 300000  
## 101 1 65.83 1 240000  
## 102 1 58.23 1 360000  
## 103 1 55.30 0 0  
## 104 0 65.69 0 0  
## 105 1 73.52 1 200000  
## 106 0 58.31 1 300000  
## 107 1 56.09 0 0  
## 108 1 54.80 1 250000  
## 109 1 60.64 0 0  
## 110 1 53.94 1 250000  
## 111 0 63.08 1 280000  
## 112 1 55.01 1 250000  
## 113 0 60.50 1 216000  
## 114 0 70.85 1 300000  
## 115 0 67.05 1 240000  
## 116 1 70.48 1 276000  
## 117 0 64.34 1 940000  
## 118 1 71.49 1 250000  
## 119 0 71.00 1 236000  
## 120 1 56.70 1 240000  
## 121 1 61.26 1 250000  
## 122 0 73.33 1 350000  
## 123 0 68.20 1 210000  
## 124 1 58.40 1 250000  
## 125 1 76.26 1 400000  
## 126 0 68.55 1 250000  
## 127 0 64.15 0 0  
## 128 0 60.78 1 360000  
## 129 1 53.49 1 300000  
## 130 1 60.98 1 250000  
## 131 1 65.63 1 200000  
## 132 0 61.58 0 0  
## 133 1 60.41 1 225000  
## 134 0 71.77 1 250000  
## 135 0 54.43 1 220000  
## 136 0 56.94 1 265000  
## 137 1 61.90 0 0  
## 138 0 61.29 1 260000  
## 139 0 60.39 1 300000  
## 140 0 58.52 0 0  
## 141 1 63.23 1 400000  
## 142 1 55.14 1 233000  
## 143 0 62.28 1 300000  
## 144 0 64.08 1 240000  
## 145 1 58.54 0 0  
## 146 0 61.30 1 690000  
## 147 0 58.87 1 270000  
## 148 0 65.25 1 240000  
## 149 0 62.48 1 340000  
## 150 0 53.20 1 250000  
## 151 1 65.99 0 0  
## 152 1 52.72 1 255000  
## 153 0 55.03 1 300000  
## 154 0 61.87 0 0  
## 155 1 60.59 0 0  
## 156 1 72.29 1 300000  
## 157 1 62.72 0 0  
## 158 0 66.06 1 285000  
## 159 0 66.46 1 500000  
## 160 0 65.52 1 250000  
## 161 0 74.56 0 0  
## 162 1 52.38 1 240000  
## 163 0 75.71 0 0  
## 164 1 58.79 0 0  
## 165 1 69.28 0 0  
## 166 0 66.04 1 290000  
## 167 1 52.64 1 300000  
## 168 1 59.32 0 0  
## 169 0 66.23 1 500000  
## 170 1 60.69 0 0  
## 171 1 57.90 1 220000  
## 172 1 68.07 1 350000  
## 173 1 72.14 0 0  
## 174 0 56.60 1 265000  
## 175 1 60.02 0 0  
## 176 0 59.81 0 0  
## 177 1 61.82 1 276000  
## 178 1 57.29 0 0  
## 179 1 71.43 1 252000  
## 180 0 62.93 0 0  
## 181 0 64.86 1 280000  
## 182 0 56.13 0 0  
## 183 0 66.94 0 0  
## 184 0 62.50 0 0  
## 185 0 61.01 1 264000  
## 186 0 57.34 1 270000  
## 187 1 56.63 1 300000  
## 188 0 64.74 0 0  
## 189 1 58.95 1 275000  
## 190 0 54.48 1 250000  
## 191 1 69.71 1 260000  
## 192 1 71.96 0 0  
## 193 1 55.80 1 265000  
## 194 0 52.81 1 300000  
## 195 1 58.44 0 0  
## 196 1 60.11 1 240000  
## 197 1 58.30 1 260000  
## 198 0 67.69 1 210000  
## 199 0 56.81 1 250000  
## 200 0 71.55 1 300000  
## 201 1 62.92 0 0  
## 202 0 56.49 1 216000  
## 203 0 74.49 1 400000  
## 204 0 53.62 1 275000  
## 205 0 69.72 1 295000  
## 206 1 60.23 1 204000  
## 207 1 60.22 0 0

#library(caret)  
#scales <- list(x=list(relation="free"), y=list(relation="free"))  
#featurePlot(x=data[,2:3], y=data$status, plot="density")

Creating dummy variables for qualitative columns:-

data<- data %>% mutate(  
 Commerce=case\_when(hsc\_s=="Commerce"~1,TRUE~0),  
 Arts=case\_when(hsc\_s=="Arts"~1,TRUE~0),  
 Science=case\_when(hsc\_s=="Science"~1,TRUE~0),  
 SciTech=case\_when(degree\_t=="Sci&Tech"~1,TRUE~0),  
 CommMgmt=case\_when(degree\_t=="Comm&Mgmt"~1,TRUE~0),   
 Others=case\_when(degree\_t=="Others"~1,TRUE~0),  
 )  
data<-subset(data,select = -c(hsc\_s,degree\_t,salary))  
head(data)

## gender ssc\_p hsc\_p degree\_p workex etest\_p specialisation mba\_p status  
## 1 1 67.00 91.00 58.00 0 55.0 1 58.80 1  
## 2 1 79.33 78.33 77.48 1 86.5 0 66.28 1  
## 3 1 65.00 68.00 64.00 0 75.0 0 57.80 1  
## 4 1 56.00 52.00 52.00 0 66.0 1 59.43 0  
## 5 1 85.80 73.60 73.30 0 96.8 0 55.50 1  
## 6 1 55.00 49.80 67.25 1 55.0 0 51.58 0  
## Commerce Arts Science SciTech CommMgmt Others  
## 1 1 0 0 1 0 0  
## 2 0 0 1 1 0 0  
## 3 0 1 0 0 1 0  
## 4 0 0 1 1 0 0  
## 5 1 0 0 0 1 0  
## 6 0 0 1 1 0 0

Testing the model locally

Model Creation with all parameters:-

model1<-glm(status~gender+ssc\_p+hsc\_p+degree\_p+workex+etest\_p+specialisation+mba\_p+Commerce+Arts+Science+SciTech+CommMgmt+Others , data=data, family=binomial)  
summary(model1)

##   
## Call:  
## glm(formula = status ~ gender + ssc\_p + hsc\_p + degree\_p + workex +   
## etest\_p + specialisation + mba\_p + Commerce + Arts + Science +   
## SciTech + CommMgmt + Others, family = binomial, data = data)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.41915 -0.14397 0.08541 0.30871 2.36440   
##   
## Coefficients: (2 not defined because of singularities)  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -19.76027 5.47774 -3.607 0.000309 \*\*\*  
## gender 1.34242 0.66305 2.025 0.042906 \*   
## ssc\_p 0.22431 0.04291 5.228 1.72e-07 \*\*\*  
## hsc\_p 0.09889 0.03676 2.690 0.007143 \*\*   
## degree\_p 0.18482 0.05560 3.324 0.000887 \*\*\*  
## workex 2.01495 0.70164 2.872 0.004082 \*\*   
## etest\_p -0.01581 0.02189 -0.722 0.470211   
## specialisation -0.28472 0.54797 -0.520 0.603341   
## mba\_p -0.20210 0.05569 -3.629 0.000285 \*\*\*  
## Commerce -0.64701 0.75226 -0.860 0.389738   
## Arts 0.78504 1.42262 0.552 0.581068   
## Science NA NA NA NA   
## SciTech -0.71639 1.64054 -0.437 0.662345   
## CommMgmt 1.01313 1.50869 0.672 0.501884   
## Others NA NA NA NA   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 252.73 on 206 degrees of freedom  
## Residual deviance: 100.22 on 194 degrees of freedom  
## AIC: 126.22  
##   
## Number of Fisher Scoring iterations: 7

We see that gender,ssc\_p,hsc\_p,degree\_p,workex,mba\_p are statistically significant at 0.05 (ie.. 95% confidence interval)

Interpretation For every one unit change in gender, the log odds of admission (versus non-admission) increases by 0.042906. For every one unit change in hsc\_p, the log odds of admission (versus non-admission) increases by 0.007143. For every one unit change in degree\_p, the log odds of admission (versus non-admission) increases by 0.000887. For every one unit change in mba\_p, the log odds of admission (versus non-admission) decreases by 0.000285.

ANOVA test for the overall logistic model:-

anova(model1,test = 'Chisq')

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: status  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 206 252.73   
## gender 1 2.618 205 250.11 0.1056389   
## ssc\_p 1 92.103 204 158.01 < 2.2e-16 \*\*\*  
## hsc\_p 1 14.981 203 143.03 0.0001086 \*\*\*  
## degree\_p 1 8.722 202 134.31 0.0031445 \*\*   
## workex 1 8.112 201 126.19 0.0043976 \*\*   
## etest\_p 1 0.945 200 125.25 0.3309516   
## specialisation 1 0.768 199 124.48 0.3808594   
## mba\_p 1 17.012 198 107.47 3.714e-05 \*\*\*  
## Commerce 1 0.023 197 107.44 0.8790968   
## Arts 1 1.571 196 105.87 0.2100312   
## Science 0 0.000 196 105.87   
## SciTech 1 5.175 195 100.70 0.0229093 \*   
## CommMgmt 1 0.479 194 100.22 0.4886626   
## Others 0 0.000 194 100.22   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

We see that ssc\_p,hsc\_p,degree\_p,workex,mba\_p are statistically significant.

Developing a model with these parameters:-

model2<-glm(status~gender+ssc\_p+hsc\_p+degree\_p+workex+mba\_p,data=data,family=binomial)  
summary(model2)

##   
## Call:  
## glm(formula = status ~ gender + ssc\_p + hsc\_p + degree\_p + workex +   
## mba\_p, family = binomial, data = data)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.25412 -0.15955 0.09118 0.36926 2.34286   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -19.09656 4.72283 -4.043 5.27e-05 \*\*\*  
## gender 1.08697 0.57625 1.886 0.05925 .   
## ssc\_p 0.20384 0.04080 4.996 5.84e-07 \*\*\*  
## hsc\_p 0.11553 0.03571 3.235 0.00122 \*\*   
## degree\_p 0.16321 0.05028 3.246 0.00117 \*\*   
## workex 2.06250 0.64625 3.191 0.00142 \*\*   
## mba\_p -0.20172 0.05132 -3.931 8.47e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 252.73 on 206 degrees of freedom  
## Residual deviance: 108.62 on 200 degrees of freedom  
## AIC: 122.62  
##   
## Number of Fisher Scoring iterations: 7

ANOVA test:-

anova(model1,model2,test = 'Chisq')

## Analysis of Deviance Table  
##   
## Model 1: status ~ gender + ssc\_p + hsc\_p + degree\_p + workex + etest\_p +   
## specialisation + mba\_p + Commerce + Arts + Science + SciTech +   
## CommMgmt + Others  
## Model 2: status ~ gender + ssc\_p + hsc\_p + degree\_p + workex + mba\_p  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1 194 100.22   
## 2 200 108.62 -6 -8.4023 0.2101

Now we got a general idea of what parameters to consider for the model.

Splitting data set into training and testing <data:->

library(caret)

## Loading required package: lattice

##   
## Attaching package: 'caret'

## The following objects are masked from 'package:DescTools':  
##   
## MAE, RMSE

set.seed(3456)  
split <- createDataPartition(y = data$status,p = 0.67,list = FALSE)

new\_train <- data[split,]   
new\_train

## gender ssc\_p hsc\_p degree\_p workex etest\_p specialisation mba\_p status  
## 2 1 79.33 78.33 77.48 1 86.50 0 66.28 1  
## 6 1 55.00 49.80 67.25 1 55.00 0 51.58 0  
## 8 1 82.00 64.00 66.00 1 67.00 0 62.14 1  
## 10 1 58.00 70.00 61.00 0 54.00 0 52.21 0  
## 11 1 58.00 61.00 60.00 1 62.00 1 60.85 1  
## 12 1 69.60 68.40 78.30 1 60.00 0 63.70 1  
## 13 0 47.00 55.00 65.00 0 62.00 1 65.04 0  
## 14 0 77.00 87.00 59.00 0 68.00 0 68.63 1  
## 15 1 62.00 47.00 50.00 0 76.00 1 54.96 0  
## 16 0 65.00 75.00 69.00 1 72.00 0 64.66 1  
## 19 0 63.00 66.00 64.00 0 68.00 1 64.08 0  
## 21 1 62.00 65.00 66.00 0 50.00 1 56.70 1  
## 23 0 69.80 60.80 72.23 0 55.53 1 68.81 1  
## 26 1 71.00 79.00 66.00 1 94.00 0 57.55 1  
## 27 1 63.00 67.00 66.00 0 68.00 1 57.69 1  
## 30 0 64.00 73.50 73.00 0 52.00 1 56.70 1  
## 31 0 67.00 53.00 65.00 0 64.00 1 58.32 0  
## 33 0 87.00 65.00 81.00 1 88.00 0 72.78 1  
## 34 1 62.00 51.00 52.00 0 68.44 1 62.77 0  
## 35 0 69.00 78.00 72.00 0 71.00 1 62.74 1  
## 37 0 79.00 76.00 65.60 0 58.00 1 55.47 1  
## 38 0 73.00 58.00 66.00 0 53.70 1 56.86 1  
## 40 0 78.00 77.00 80.00 0 60.00 0 66.72 1  
## 41 0 74.00 63.16 65.00 1 65.00 1 69.76 0  
## 42 1 87.00 87.00 68.00 0 95.00 1 62.90 1  
## 43 0 77.00 73.00 81.00 1 89.00 0 69.70 1  
## 44 0 76.00 64.00 72.00 0 58.00 1 66.53 0  
## 46 1 63.00 60.00 57.00 1 78.00 0 54.55 1  
## 49 1 54.40 61.12 56.20 0 67.00 1 62.65 0  
## 50 0 40.89 45.83 53.00 0 71.20 1 65.49 0  
## 51 1 80.00 70.00 72.00 0 87.00 1 71.04 1  
## 52 0 74.00 60.00 69.00 0 78.00 1 65.56 1  
## 53 1 60.40 66.60 65.00 0 71.00 1 52.71 1  
## 55 1 68.00 76.00 74.00 0 80.00 0 63.59 1  
## 58 1 74.00 70.00 72.00 1 60.00 0 57.24 1  
## 61 1 61.00 70.00 64.00 0 68.50 1 59.50 0  
## 63 1 54.00 47.00 57.00 0 89.69 1 57.10 0  
## 64 1 83.00 74.00 66.00 0 68.92 1 58.46 1  
## 65 1 80.92 78.50 67.00 0 68.71 0 60.99 1  
## 68 1 82.00 61.00 62.00 0 89.00 0 65.45 1  
## 70 1 84.86 67.00 78.00 0 95.50 0 68.53 1  
## 71 1 64.60 83.83 71.72 0 86.00 0 59.75 1  
## 72 1 56.60 64.80 70.20 0 84.27 0 67.20 1  
## 73 0 59.00 62.00 77.50 0 74.00 1 67.00 0  
## 74 0 66.50 70.40 71.93 0 61.00 0 64.27 1  
## 75 1 64.00 80.00 65.00 1 69.00 0 57.65 1  
## 76 1 84.00 90.90 64.50 0 86.04 0 59.42 1  
## 78 0 69.00 62.00 69.00 1 67.00 1 62.35 1  
## 79 1 81.70 63.00 67.00 1 86.00 0 70.20 1  
## 81 1 84.00 79.00 68.00 1 84.00 0 66.69 1  
## 82 1 70.00 63.00 70.00 1 55.00 0 62.00 1  
## 83 0 83.84 89.83 77.20 1 78.74 0 76.18 1  
## 84 1 62.00 63.00 64.00 0 67.00 0 57.03 1  
## 85 1 59.60 51.00 60.00 0 75.00 1 59.08 0  
## 86 0 66.00 62.00 73.00 0 58.00 1 64.36 1  
## 88 0 85.00 90.00 82.00 0 92.00 0 68.03 1  
## 89 1 52.00 57.00 50.80 0 67.00 1 62.79 0  
## 91 1 52.00 62.00 54.00 0 72.00 1 55.41 0  
## 93 1 73.00 78.00 65.00 1 95.46 0 62.16 1  
## 94 0 76.00 70.00 76.00 1 66.00 0 64.44 1  
## 95 0 70.50 62.50 61.00 0 93.91 0 69.03 0  
## 98 0 45.00 57.00 58.00 1 56.39 1 64.95 0  
## 99 1 63.00 72.00 68.00 0 78.00 1 60.44 1  
## 100 0 77.00 61.00 68.00 1 57.50 0 61.31 1  
## 103 1 59.00 64.00 58.00 0 85.00 1 55.30 0  
## 104 1 61.08 50.00 54.00 0 71.00 0 65.69 0  
## 105 1 82.00 90.00 83.00 0 80.00 1 73.52 1  
## 106 1 61.00 82.00 69.00 0 84.00 0 58.31 1  
## 108 0 69.50 70.00 72.00 0 57.20 1 54.80 1  
## 110 1 58.00 61.00 61.00 0 58.00 1 53.94 1  
## 111 0 73.96 79.00 67.00 0 72.15 0 63.08 1  
## 113 0 73.00 63.00 66.00 0 89.00 0 60.50 1  
## 118 0 64.00 67.00 69.60 1 55.67 1 71.49 1  
## 119 0 66.50 66.80 69.30 1 80.40 0 71.00 1  
## 120 1 74.00 59.00 73.00 1 60.00 1 56.70 1  
## 122 0 84.00 73.00 73.00 0 75.00 0 73.33 1  
## 123 0 79.00 61.00 75.50 1 70.00 0 68.20 1  
## 124 0 72.00 60.00 69.00 0 55.50 1 58.40 1  
## 125 1 80.40 73.40 77.72 1 81.20 1 76.26 1  
## 127 1 62.00 65.00 60.00 0 84.00 0 64.15 0  
## 128 0 74.90 57.00 62.00 1 80.00 0 60.78 1  
## 129 1 67.00 68.00 64.00 1 74.40 1 53.49 1  
## 130 1 73.00 64.00 77.00 1 65.00 1 60.98 1  
## 132 0 47.00 59.00 64.00 0 78.00 0 61.58 0  
## 133 1 67.00 63.00 72.00 0 56.00 1 60.41 1  
## 134 0 82.00 64.00 73.00 1 96.00 0 71.77 1  
## 135 1 77.00 70.00 59.00 1 58.00 0 54.43 1  
## 136 1 65.00 64.80 69.50 1 56.00 0 56.94 1  
## 137 1 66.00 64.00 60.00 0 60.00 1 61.90 0  
## 138 1 85.00 60.00 73.43 1 60.00 0 61.29 1  
## 140 1 52.00 50.00 61.00 0 60.00 0 58.52 0  
## 141 1 89.40 65.66 71.25 0 72.00 1 63.23 1  
## 142 1 62.00 63.00 66.00 0 85.00 1 55.14 1  
## 144 0 77.00 86.00 56.00 0 57.00 0 64.08 1  
## 145 1 44.00 58.00 55.00 1 64.25 1 58.54 0  
## 147 1 65.00 65.00 75.00 0 83.00 0 58.87 1  
## 148 0 75.40 60.50 84.00 0 98.00 0 65.25 1  
## 149 1 49.00 59.00 65.00 1 86.00 0 62.48 1  
## 150 1 53.00 63.00 60.00 1 70.00 0 53.20 1  
## 152 1 84.20 69.40 65.00 1 80.00 1 52.72 1  
## 153 1 66.50 62.50 60.90 0 93.40 0 55.03 1  
## 154 1 67.00 63.00 64.00 0 60.00 0 61.87 0  
## 155 1 52.00 49.00 58.00 0 62.00 1 60.59 0  
## 156 1 87.00 74.00 65.00 1 75.00 1 72.29 1  
## 157 1 55.60 51.00 57.50 0 57.63 1 62.72 0  
## 158 1 74.20 87.60 77.25 1 75.20 0 66.06 1  
## 159 1 63.00 67.00 64.00 0 75.00 0 66.46 1  
## 160 0 67.16 72.50 63.35 0 53.04 0 65.52 1  
## 161 0 63.30 78.33 74.00 0 80.00 0 74.56 0  
## 163 1 67.90 62.00 67.00 1 58.10 0 75.71 0  
## 166 1 80.00 80.00 72.00 1 63.79 0 66.04 1  
## 167 1 73.00 58.00 56.00 0 84.00 1 52.64 1  
## 168 0 52.00 52.00 55.00 0 67.00 1 59.32 0  
## 170 1 63.00 62.00 65.00 0 87.50 1 60.69 0  
## 172 1 68.00 56.00 68.00 0 73.00 1 68.07 1  
## 175 1 62.00 60.33 64.21 0 63.00 1 60.02 0  
## 177 1 65.00 77.00 69.00 0 60.00 1 61.82 1  
## 178 0 56.28 62.83 59.79 0 60.00 1 57.29 0  
## 179 0 88.00 72.00 78.00 0 82.00 1 71.43 1  
## 181 1 78.50 65.50 67.00 1 95.00 0 64.86 1  
## 183 0 54.00 77.60 69.20 0 95.65 0 66.94 0  
## 184 0 64.00 70.20 61.00 0 50.00 0 62.50 0  
## 185 1 67.00 61.00 72.00 0 72.00 0 61.01 1  
## 186 1 65.20 61.40 64.80 1 93.40 0 57.34 1  
## 187 0 60.00 63.00 56.00 1 80.00 1 56.63 1  
## 188 1 52.00 55.00 56.30 0 59.00 0 64.74 0  
## 189 1 66.00 76.00 72.00 1 84.00 1 58.95 1  
## 192 0 67.00 70.00 65.00 0 88.00 1 71.96 0  
## 194 1 69.00 60.00 65.00 0 87.55 0 52.81 1  
## 196 1 70.00 63.00 66.00 0 61.28 1 60.11 1  
## 197 1 55.68 61.33 56.87 0 66.00 1 58.30 1  
## 198 0 74.00 73.00 73.00 1 80.00 0 67.69 1  
## 199 1 61.00 62.00 65.00 0 62.00 0 56.81 1  
## 200 1 83.33 78.00 61.00 1 88.56 0 71.55 1  
## 201 0 43.00 60.00 65.00 0 92.66 1 62.92 0  
## 202 1 62.00 72.00 65.00 0 67.00 0 56.49 1  
## 204 1 58.00 60.00 72.00 0 74.00 0 53.62 1  
## 205 1 67.00 67.00 73.00 1 59.00 0 69.72 1  
## 207 1 62.00 58.00 53.00 0 89.00 1 60.22 0  
## Commerce Arts Science SciTech CommMgmt Others  
## 2 0 0 1 1 0 0  
## 6 0 0 1 1 0 0  
## 8 0 0 1 1 0 0  
## 10 1 0 0 0 1 0  
## 11 1 0 0 0 1 0  
## 12 1 0 0 0 1 0  
## 13 0 0 1 0 1 0  
## 14 1 0 0 0 1 0  
## 15 1 0 0 0 1 0  
## 16 1 0 0 0 1 0  
## 19 1 0 0 0 1 0  
## 21 1 0 0 0 1 0  
## 23 0 0 1 1 0 0  
## 26 1 0 0 0 1 0  
## 27 1 0 0 0 1 0  
## 30 1 0 0 0 1 0  
## 31 0 0 1 1 0 0  
## 33 0 0 1 0 1 0  
## 34 0 0 1 0 0 1  
## 35 1 0 0 0 1 0  
## 37 0 0 1 1 0 0  
## 38 0 0 1 0 1 0  
## 40 1 0 0 0 1 0  
## 41 1 0 0 0 1 0  
## 42 1 0 0 0 1 0  
## 43 1 0 0 0 1 0  
## 44 0 0 1 1 0 0  
## 46 1 0 0 0 1 0  
## 49 1 0 0 0 1 0  
## 50 1 0 0 0 1 0  
## 51 0 0 1 1 0 0  
## 52 0 0 1 0 1 0  
## 53 0 0 1 0 1 0  
## 55 1 0 0 0 1 0  
## 58 0 0 1 0 1 0  
## 61 1 0 0 0 1 0  
## 63 0 0 1 0 1 0  
## 64 0 0 1 0 1 0  
## 65 1 0 0 0 1 0  
## 68 0 0 1 1 0 0  
## 70 0 0 1 0 1 0  
## 71 1 0 0 0 1 0  
## 72 1 0 0 0 1 0  
## 73 1 0 0 0 1 0  
## 74 0 1 0 0 1 0  
## 75 0 0 1 1 0 0  
## 76 0 0 1 1 0 0  
## 78 1 0 0 0 1 0  
## 79 0 0 1 0 1 0  
## 81 0 0 1 1 0 0  
## 82 0 0 1 1 0 0  
## 83 1 0 0 0 1 0  
## 84 1 0 0 0 1 0  
## 85 0 0 1 0 0 1  
## 86 1 0 0 0 1 0  
## 88 1 0 0 0 1 0  
## 89 1 0 0 0 1 0  
## 91 1 0 0 0 1 0  
## 93 1 0 0 0 1 0  
## 94 0 0 1 0 1 0  
## 95 1 0 0 0 1 0  
## 98 1 0 0 0 1 0  
## 99 1 0 0 0 1 0  
## 100 1 0 0 0 1 0  
## 103 0 0 1 1 0 0  
## 104 0 0 1 1 0 0  
## 105 1 0 0 0 1 0  
## 106 1 0 0 0 1 0  
## 108 0 0 1 1 0 0  
## 110 1 0 0 0 1 0  
## 111 1 0 0 0 1 0  
## 113 0 0 1 0 1 0  
## 118 0 0 1 1 0 0  
## 119 0 1 0 0 1 0  
## 120 1 0 0 0 1 0  
## 122 1 0 0 0 1 0  
## 123 0 0 1 1 0 0  
## 124 0 0 1 0 1 0  
## 125 0 0 1 1 0 0  
## 127 1 0 0 0 1 0  
## 128 0 0 1 0 0 1  
## 129 1 0 0 0 1 0  
## 130 1 0 0 0 1 0  
## 132 0 1 0 0 1 0  
## 133 1 0 0 0 1 0  
## 134 0 0 1 1 0 0  
## 135 1 0 0 0 1 0  
## 136 1 0 0 0 1 0  
## 137 0 0 1 0 1 0  
## 138 0 0 1 1 0 0  
## 140 0 1 0 0 1 0  
## 141 0 0 1 1 0 0  
## 142 0 0 1 0 1 0  
## 144 0 1 0 0 0 1  
## 145 0 1 0 0 1 0  
## 147 1 0 0 0 1 0  
## 148 0 0 1 1 0 0  
## 149 0 0 1 1 0 0  
## 150 0 0 1 0 1 0  
## 152 0 0 1 1 0 0  
## 153 1 0 0 0 1 0  
## 154 0 0 1 1 0 0  
## 155 1 0 0 0 1 0  
## 156 0 0 1 1 0 0  
## 157 1 0 0 0 1 0  
## 158 1 0 0 0 1 0  
## 159 0 0 1 1 0 0  
## 160 1 0 0 0 1 0  
## 161 1 0 0 0 1 0  
## 163 0 0 1 1 0 0  
## 166 1 0 0 0 1 0  
## 167 1 0 0 0 1 0  
## 168 0 0 1 1 0 0  
## 170 0 0 1 1 0 0  
## 172 0 0 1 1 0 0  
## 175 0 0 1 1 0 0  
## 177 1 0 0 0 1 0  
## 178 1 0 0 0 1 0  
## 179 0 0 1 0 0 1  
## 181 0 0 1 1 0 0  
## 183 1 0 0 0 1 0  
## 184 1 0 0 0 1 0  
## 185 0 0 1 0 1 0  
## 186 1 0 0 0 1 0  
## 187 0 1 0 0 0 1  
## 188 1 0 0 0 1 0  
## 189 1 0 0 0 1 0  
## 192 1 0 0 0 0 1  
## 194 1 0 0 0 1 0  
## 196 0 0 1 1 0 0  
## 197 1 0 0 0 1 0  
## 198 1 0 0 0 1 0  
## 199 1 0 0 0 1 0  
## 200 1 0 0 0 1 0  
## 201 0 0 1 0 1 0  
## 202 1 0 0 0 1 0  
## 204 0 0 1 1 0 0  
## 205 1 0 0 0 1 0  
## 207 0 0 1 0 1 0

new\_test <-data[-split,]  
new\_test

## gender ssc\_p hsc\_p degree\_p workex etest\_p specialisation mba\_p status  
## 1 1 67.00 91.00 58.00 0 55.00 1 58.80 1  
## 3 1 65.00 68.00 64.00 0 75.00 0 57.80 1  
## 4 1 56.00 52.00 52.00 0 66.00 1 59.43 0  
## 5 1 85.80 73.60 73.30 0 96.80 0 55.50 1  
## 7 0 46.00 49.20 79.00 0 74.28 0 53.29 0  
## 9 1 73.00 79.00 72.00 0 91.34 0 61.29 1  
## 17 1 63.00 66.20 65.60 1 60.00 0 62.54 1  
## 18 0 55.00 67.00 64.00 0 60.00 0 67.28 0  
## 20 1 60.00 67.00 70.00 1 50.48 0 77.89 1  
## 22 0 79.00 76.00 85.00 0 95.00 0 69.06 1  
## 24 0 77.40 60.00 64.74 1 92.00 0 63.62 1  
## 25 0 52.58 54.60 50.20 1 76.00 0 65.33 0  
## 28 1 76.76 76.50 67.50 1 73.35 0 64.15 1  
## 29 1 62.00 67.00 58.00 0 77.00 0 51.29 0  
## 32 0 61.00 81.00 66.40 0 50.89 1 62.21 1  
## 36 1 51.00 44.00 57.00 0 64.00 0 51.45 0  
## 39 1 81.00 68.00 64.00 0 93.00 0 62.56 1  
## 45 0 70.89 71.98 65.60 0 68.00 1 71.63 0  
## 47 1 63.00 62.00 68.00 0 64.00 0 62.46 1  
## 48 0 75.20 73.20 68.40 0 65.00 1 62.98 1  
## 54 1 63.00 71.40 61.40 0 68.00 0 66.88 1  
## 56 1 74.00 62.00 68.00 0 74.00 0 57.99 1  
## 57 1 52.60 65.58 72.11 0 57.60 0 56.66 1  
## 59 1 84.20 73.40 66.89 0 61.60 0 62.48 1  
## 60 0 86.50 64.20 67.40 0 59.00 0 59.69 1  
## 62 1 80.00 73.00 75.00 0 61.00 0 58.78 1  
## 66 0 69.70 47.00 72.70 0 79.00 1 59.24 0  
## 67 1 73.00 73.00 66.00 1 70.00 0 68.07 1  
## 69 1 75.00 70.29 71.00 0 95.00 0 66.94 1  
## 77 0 69.00 62.00 66.00 0 75.00 1 67.99 0  
## 80 1 63.00 67.00 74.00 0 82.00 0 60.44 0  
## 87 0 84.00 75.00 69.00 1 62.00 1 62.36 1  
## 90 0 60.23 69.00 66.00 0 72.00 0 59.47 1  
## 92 1 58.00 62.00 64.00 0 53.88 0 54.97 1  
## 96 0 69.00 73.00 65.00 0 70.00 0 57.31 1  
## 97 1 54.00 82.00 63.00 0 50.00 0 59.47 0  
## 101 1 73.00 78.00 73.00 1 85.00 1 65.83 1  
## 102 1 69.00 63.00 65.00 1 55.00 1 58.23 1  
## 107 1 52.00 63.00 65.00 1 86.00 1 56.09 0  
## 109 1 51.00 54.00 61.00 0 60.00 1 60.64 0  
## 112 1 65.00 68.00 69.00 0 53.70 1 55.01 1  
## 114 1 68.20 72.80 66.60 1 96.00 0 70.85 1  
## 115 1 77.00 75.00 73.00 0 80.00 0 67.05 1  
## 116 1 76.00 80.00 78.00 1 97.00 1 70.48 1  
## 117 1 60.80 68.40 64.60 1 82.66 0 64.34 1  
## 121 1 67.00 71.00 64.33 1 64.00 1 61.26 1  
## 126 1 76.70 89.70 66.00 1 90.00 0 68.55 1  
## 131 0 72.00 56.00 69.00 0 55.60 1 65.63 1  
## 139 1 77.67 64.89 70.67 0 89.00 0 60.39 1  
## 143 1 70.00 74.00 65.00 0 83.00 0 62.28 1  
## 146 1 71.00 58.66 58.00 1 56.00 0 61.30 1  
## 151 1 51.57 74.66 59.90 1 56.15 1 65.99 0  
## 162 1 62.00 62.00 60.00 1 63.00 1 52.38 1  
## 164 0 48.00 51.00 58.00 1 60.00 1 58.79 0  
## 165 0 63.40 67.20 60.00 0 58.06 1 69.28 0  
## 169 1 73.24 50.83 64.27 1 64.00 0 66.23 1  
## 171 0 59.00 60.00 56.00 0 55.00 1 57.90 1  
## 173 0 77.80 64.00 64.20 0 75.50 1 72.14 0  
## 174 1 65.00 71.50 62.80 1 57.00 0 56.60 1  
## 176 1 52.00 65.00 57.00 1 75.00 0 59.81 0  
## 180 0 52.00 64.00 61.00 0 55.00 0 62.93 0  
## 182 1 61.80 47.00 54.38 0 57.00 0 56.13 0  
## 190 1 72.00 63.00 77.50 1 78.00 0 54.48 1  
## 191 0 83.96 53.00 91.00 0 59.32 1 69.71 1  
## 193 1 69.00 65.00 57.00 0 73.00 1 55.80 1  
## 195 1 54.20 63.00 58.00 0 79.00 1 58.44 0  
## 203 1 80.60 82.00 77.60 0 91.00 0 74.49 1  
## 206 0 74.00 66.00 58.00 0 70.00 1 60.23 1  
## Commerce Arts Science SciTech CommMgmt Others  
## 1 1 0 0 1 0 0  
## 3 0 1 0 0 1 0  
## 4 0 0 1 1 0 0  
## 5 1 0 0 0 1 0  
## 7 1 0 0 0 1 0  
## 9 1 0 0 0 1 0  
## 17 1 0 0 0 1 0  
## 18 1 0 0 0 1 0  
## 20 0 1 0 0 1 0  
## 22 1 0 0 0 1 0  
## 24 0 0 1 1 0 0  
## 25 1 0 0 0 1 0  
## 28 1 0 0 0 1 0  
## 29 1 0 0 0 1 0  
## 32 1 0 0 0 1 0  
## 36 1 0 0 0 1 0  
## 39 0 0 1 1 0 0  
## 45 0 0 1 0 1 0  
## 47 1 0 0 0 1 0  
## 48 0 0 1 0 1 0  
## 54 1 0 0 0 1 0  
## 56 0 0 1 0 1 0  
## 57 0 0 1 1 0 0  
## 59 1 0 0 0 1 0  
## 60 0 0 1 1 0 0  
## 62 1 0 0 0 1 0  
## 66 1 0 0 1 0 0  
## 67 0 0 1 1 0 0  
## 69 1 0 0 0 1 0  
## 77 0 0 1 1 0 0  
## 80 1 0 0 0 1 0  
## 87 0 0 1 1 0 0  
## 90 0 0 1 0 1 0  
## 92 1 0 0 0 1 0  
## 96 1 0 0 0 1 0  
## 97 1 0 0 1 0 0  
## 101 0 0 1 1 0 0  
## 102 0 0 1 0 1 0  
## 107 0 0 1 1 0 0  
## 109 0 0 1 1 0 0  
## 112 0 0 1 0 1 0  
## 114 1 0 0 0 1 0  
## 115 0 0 1 1 0 0  
## 116 0 0 1 1 0 0  
## 117 1 0 0 0 1 0  
## 121 0 0 1 0 0 1  
## 126 1 0 0 0 1 0  
## 131 0 0 1 0 1 0  
## 139 1 0 0 0 1 0  
## 143 1 0 0 0 1 0  
## 146 0 0 1 1 0 0  
## 151 1 0 0 0 1 0  
## 162 1 0 0 0 1 0  
## 164 1 0 0 0 1 0  
## 165 1 0 0 0 1 0  
## 169 0 0 1 1 0 0  
## 171 1 0 0 0 1 0  
## 173 0 0 1 1 0 0  
## 174 1 0 0 0 1 0  
## 176 0 1 0 0 0 1  
## 180 1 0 0 0 1 0  
## 182 1 0 0 0 1 0  
## 190 0 0 1 1 0 0  
## 191 0 0 1 1 0 0  
## 193 1 0 0 0 1 0  
## 195 0 0 1 0 1 0  
## 203 1 0 0 0 1 0  
## 206 1 0 0 0 1 0

Applying the model for the training data set

log\_model<-glm(status~gender+ssc\_p+hsc\_p+degree\_p+workex+mba\_p,data=new\_train,family=binomial)  
summary(log\_model)

##   
## Call:  
## glm(formula = status ~ gender + ssc\_p + hsc\_p + degree\_p + workex +   
## mba\_p, family = binomial, data = new\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.24802 -0.10399 0.06784 0.39247 2.12222   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -19.83519 6.28157 -3.158 0.001590 \*\*   
## gender 1.27813 0.75112 1.702 0.088823 .   
## ssc\_p 0.18453 0.05174 3.566 0.000362 \*\*\*  
## hsc\_p 0.13554 0.05347 2.535 0.011245 \*   
## degree\_p 0.21308 0.07085 3.008 0.002634 \*\*   
## workex 2.41382 0.85214 2.833 0.004616 \*\*   
## mba\_p -0.24509 0.07178 -3.415 0.000639 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 168.617 on 138 degrees of freedom  
## Residual deviance: 66.688 on 132 degrees of freedom  
## AIC: 80.688  
##   
## Number of Fisher Scoring iterations: 7

Predicting the model for the testing dataset

log\_predict <- predict(log\_model,newdata = new\_test,type = "response")  
log\_predict <- ifelse(log\_predict > 0.5,1,0)

Setting the threshold to see the max accuracy.

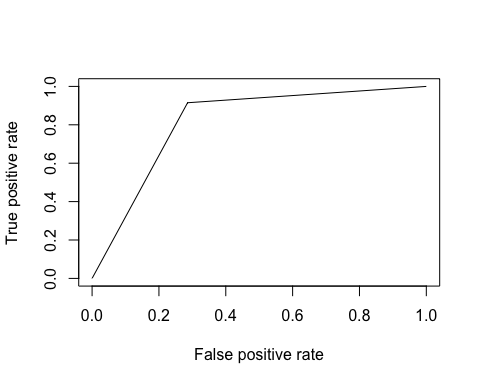
Plotting the ROC curve:-

library(ROCR)  
library(Metrics)

##   
## Attaching package: 'Metrics'

## The following objects are masked from 'package:caret':  
##   
## precision, recall

pr <- prediction(log\_predict,new\_test$status)  
perf <- performance(pr,measure = "tpr",x.measure = "fpr")   
plot(perf)



Accuracy:-

auc(new\_test$status,log\_predict)

## [1] 0.8145897

Confusion Matriv for threshold=0.5

library(caret)  
final<-factor(log\_predict)  
levels(final)

## [1] "0" "1"

attach(new\_test)  
test<-factor(status)  
levels(test)

## [1] "0" "1"

confusionMatrix(final, test)

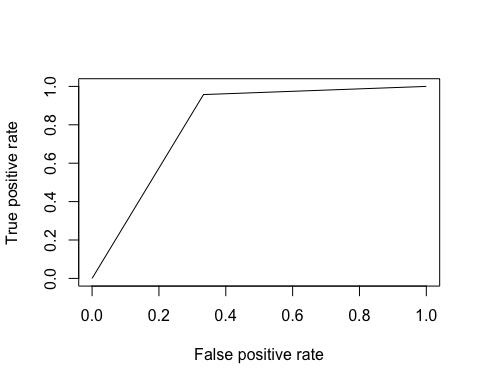
## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction 0 1  
## 0 15 4  
## 1 6 43  
##   
## Accuracy : 0.8529   
## 95% CI : (0.7461, 0.9272)  
## No Information Rate : 0.6912   
## P-Value [Acc > NIR] : 0.001773   
##   
## Kappa : 0.6462   
##   
## Mcnemar's Test P-Value : 0.751830   
##   
## Sensitivity : 0.7143   
## Specificity : 0.9149   
## Pos Pred Value : 0.7895   
## Neg Pred Value : 0.8776   
## Prevalence : 0.3088   
## Detection Rate : 0.2206   
## Detection Prevalence : 0.2794   
## Balanced Accuracy : 0.8146   
##   
## 'Positive' Class : 0   
##

Choosing a lower threshold (ie.. 0.4)

log\_predict <- predict(log\_model,newdata = new\_test,type = "response")  
log\_predict <- ifelse(log\_predict > 0.4,1,0)

Plotting ROC curve:-

library(ROCR)  
library(Metrics)  
pr <- prediction(log\_predict,new\_test$status)  
perf <- performance(pr,measure = "tpr",x.measure = "fpr")   
plot(perf)

 Accuracy:-

auc(new\_test$status,log\_predict)

## [1] 0.8120567

No significant change in Accuracy

Confusion matrix for thres=0.4

library(caret)  
final<-factor(log\_predict)  
levels(final)

## [1] "0" "1"

attach(new\_test)  
test<-factor(status)  
levels(test)

## [1] "0" "1"

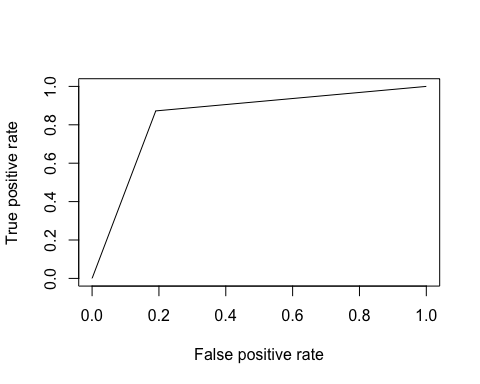
confusionMatrix(final, test)

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction 0 1  
## 0 14 2  
## 1 7 45  
##   
## Accuracy : 0.8676   
## 95% CI : (0.7636, 0.9377)  
## No Information Rate : 0.6912   
## P-Value [Acc > NIR] : 0.0006336   
##   
## Kappa : 0.6681   
##   
## Mcnemar's Test P-Value : 0.1824224   
##   
## Sensitivity : 0.6667   
## Specificity : 0.9574   
## Pos Pred Value : 0.8750   
## Neg Pred Value : 0.8654   
## Prevalence : 0.3088   
## Detection Rate : 0.2059   
## Detection Prevalence : 0.2353   
## Balanced Accuracy : 0.8121   
##   
## 'Positive' Class : 0   
##

Threshold =0.6

log\_predict <- predict(log\_model,newdata = new\_test,type = "response")  
log\_predict <- ifelse(log\_predict > 0.6,1,0)

library(ROCR)  
library(Metrics)  
pr <- prediction(log\_predict,new\_test$status)  
perf <- performance(pr,measure = "tpr",x.measure = "fpr")   
plot(perf)



auc(new\_test$status,log\_predict)

## [1] 0.8409321

library(caret)  
final<-factor(log\_predict)  
levels(final)

## [1] "0" "1"

attach(new\_test)  
test<-factor(status)  
levels(test)

## [1] "0" "1"

confusionMatrix(final, test)

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction 0 1  
## 0 17 6  
## 1 4 41  
##   
## Accuracy : 0.8529   
## 95% CI : (0.7461, 0.9272)  
## No Information Rate : 0.6912   
## P-Value [Acc > NIR] : 0.001773   
##   
## Kappa : 0.6644   
##   
## Mcnemar's Test P-Value : 0.751830   
##   
## Sensitivity : 0.8095   
## Specificity : 0.8723   
## Pos Pred Value : 0.7391   
## Neg Pred Value : 0.9111   
## Prevalence : 0.3088   
## Detection Rate : 0.2500   
## Detection Prevalence : 0.3382   
## Balanced Accuracy : 0.8409   
##   
## 'Positive' Class : 0   
##

Choosing threshold=0.06

Diagnostic test

McFadden’s R squared

null\_model<-glm(status~1,data=new\_train,family=binomial)  
1-logLik(log\_model)/logLik(null\_model)

## 'log Lik.' 0.6045018 (df=7)