Project\_2

Fabio\_Savorgnan

3/14/2020

library(dplyr)  
library(ggplot2)  
library(tidyr)  
library(SDMTools)  
library(readr)  
library(digest)  
library(ISLR)  
library(car)   
library(leaps)  
library( Matrix)  
library(foreach)  
library(glmnet)  
library(gridExtra)  
library(lsmeans)  
library(limma)  
library(Sleuth3)  
library(tseries)  
library(forecast)  
library(ggplot2)  
library(MASS)  
library(mvtnorm)  
library(epitools)  
library(samplesizeCMH)  
library(caret)  
library(GGally)  
library(glmnet)  
library(bestglm)  
library(data.table)  
library(broom)  
library(plyr)  
library(repr)  
library(ResourceSelection)  
library(ROCR)  
library(pROC)

# Loading the data

bank\_20 = read.csv("bank-additional-full.csv", sep=";")  
head(bank\_20)

## age job marital education default housing loan contact month  
## 1 56 housemaid married basic.4y no no no telephone may  
## 2 57 services married high.school unknown no no telephone may  
## 3 37 services married high.school no yes no telephone may  
## 4 40 admin. married basic.6y no no no telephone may  
## 5 56 services married high.school no no yes telephone may  
## 6 45 services married basic.9y unknown no no telephone may  
## day\_of\_week duration campaign pdays previous poutcome emp.var.rate  
## 1 mon 261 1 999 0 nonexistent 1.1  
## 2 mon 149 1 999 0 nonexistent 1.1  
## 3 mon 226 1 999 0 nonexistent 1.1  
## 4 mon 151 1 999 0 nonexistent 1.1  
## 5 mon 307 1 999 0 nonexistent 1.1  
## 6 mon 198 1 999 0 nonexistent 1.1  
## cons.price.idx cons.conf.idx euribor3m nr.employed y  
## 1 93.994 -36.4 4.857 5191 no  
## 2 93.994 -36.4 4.857 5191 no  
## 3 93.994 -36.4 4.857 5191 no  
## 4 93.994 -36.4 4.857 5191 no  
## 5 93.994 -36.4 4.857 5191 no  
## 6 93.994 -36.4 4.857 5191 no

summary(bank\_20)

## age job marital   
## Min. :17.00 admin. :10422 divorced: 4612   
## 1st Qu.:32.00 blue-collar: 9254 married :24928   
## Median :38.00 technician : 6743 single :11568   
## Mean :40.02 services : 3969 unknown : 80   
## 3rd Qu.:47.00 management : 2924   
## Max. :98.00 retired : 1720   
## (Other) : 6156   
## education default housing loan   
## university.degree :12168 no :32588 no :18622 no :33950   
## high.school : 9515 unknown: 8597 unknown: 990 unknown: 990   
## basic.9y : 6045 yes : 3 yes :21576 yes : 6248   
## professional.course: 5243   
## basic.4y : 4176   
## basic.6y : 2292   
## (Other) : 1749   
## contact month day\_of\_week duration   
## cellular :26144 may :13769 fri:7827 Min. : 0.0   
## telephone:15044 jul : 7174 mon:8514 1st Qu.: 102.0   
## aug : 6178 thu:8623 Median : 180.0   
## jun : 5318 tue:8090 Mean : 258.3   
## nov : 4101 wed:8134 3rd Qu.: 319.0   
## apr : 2632 Max. :4918.0   
## (Other): 2016   
## campaign pdays previous poutcome   
## Min. : 1.000 Min. : 0.0 Min. :0.000 failure : 4252   
## 1st Qu.: 1.000 1st Qu.:999.0 1st Qu.:0.000 nonexistent:35563   
## Median : 2.000 Median :999.0 Median :0.000 success : 1373   
## Mean : 2.568 Mean :962.5 Mean :0.173   
## 3rd Qu.: 3.000 3rd Qu.:999.0 3rd Qu.:0.000   
## Max. :56.000 Max. :999.0 Max. :7.000   
##   
## emp.var.rate cons.price.idx cons.conf.idx euribor3m   
## Min. :-3.40000 Min. :92.20 Min. :-50.8 Min. :0.634   
## 1st Qu.:-1.80000 1st Qu.:93.08 1st Qu.:-42.7 1st Qu.:1.344   
## Median : 1.10000 Median :93.75 Median :-41.8 Median :4.857   
## Mean : 0.08189 Mean :93.58 Mean :-40.5 Mean :3.621   
## 3rd Qu.: 1.40000 3rd Qu.:93.99 3rd Qu.:-36.4 3rd Qu.:4.961   
## Max. : 1.40000 Max. :94.77 Max. :-26.9 Max. :5.045   
##   
## nr.employed y   
## Min. :4964 no :36548   
## 1st Qu.:5099 yes: 4640   
## Median :5191   
## Mean :5167   
## 3rd Qu.:5228   
## Max. :5228   
##

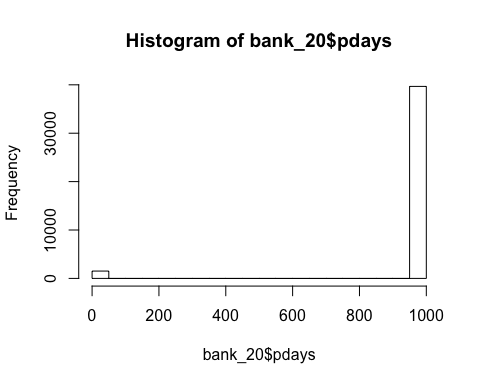
# Does not look like any NAs in either data set

#Does not look like any NAs in either data set  
sapply(bank\_20, function(x) sum(is.na(x)))

## age job marital education default   
## 0 0 0 0 0   
## housing loan contact month day\_of\_week   
## 0 0 0 0 0   
## duration campaign pdays previous poutcome   
## 0 0 0 0 0   
## emp.var.rate cons.price.idx cons.conf.idx euribor3m nr.employed   
## 0 0 0 0 0   
## y   
## 0

#pdays- about 40k of the 41k are at level 999, no previous contact #could bin this data

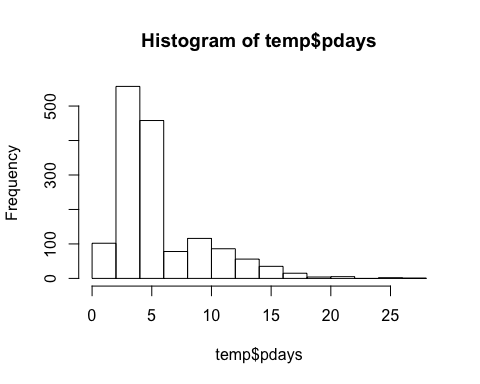
hist(bank\_20$pdays)



temp = bank\_20 %>% filter(pdays != 999)  
dim(temp)

## [1] 1515 21

hist(temp$pdays)



# Reshape for a balance data

yes\_answer = bank\_20 %>% filter(y == "yes")  
  
no\_answer\_all = bank\_20 %>% filter(y == "no")  
no\_indices = sample(dim(no\_answer\_all)[1], dim(yes\_answer)[1])  
no\_answer = no\_answer\_all[no\_indices,]  
  
balanced\_bank\_20 = rbind(yes\_answer, no\_answer)  
  
balanced\_indices = sample(dim(balanced\_bank\_20)[1], round(dim(balanced\_bank\_20)[1] \* .1 ))  
balanced\_test = balanced\_bank\_20[balanced\_indices,]  
head(balanced\_test)

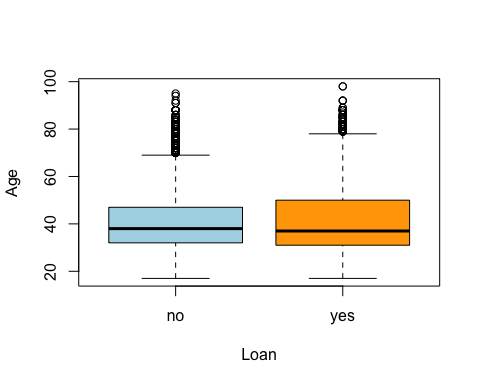
## age job marital education default housing loan  
## 34693 26 admin. single university.degree no unknown unknown  
## 1703 64 retired married university.degree no yes no  
## 14932 37 technician married professional.course no no no  
## 15914 32 services single basic.9y no no no  
## 5679 35 blue-collar married high.school no yes yes  
## 26127 36 admin. married university.degree no yes no  
## contact month day\_of\_week duration campaign pdays previous poutcome  
## 34693 cellular aug mon 115 1 999 0 nonexistent  
## 1703 cellular apr tue 146 4 999 0 nonexistent  
## 14932 cellular jul fri 127 3 999 0 nonexistent  
## 15914 cellular jul wed 711 1 999 0 nonexistent  
## 5679 telephone may mon 83 1 999 0 nonexistent  
## 26127 cellular nov fri 85 1 999 1 failure  
## emp.var.rate cons.price.idx cons.conf.idx euribor3m nr.employed y  
## 34693 -2.9 92.201 -31.4 0.861 5076.2 no  
## 1703 -1.8 93.075 -47.1 1.405 5099.1 yes  
## 14932 1.4 93.918 -42.7 4.957 5228.1 no  
## 15914 1.4 93.918 -42.7 4.963 5228.1 no  
## 5679 1.1 93.994 -36.4 4.857 5191.0 no  
## 26127 -0.1 93.200 -42.0 4.021 5195.8 no

balanced\_train = balanced\_bank\_20[-balanced\_indices,]  
head(balanced\_train)

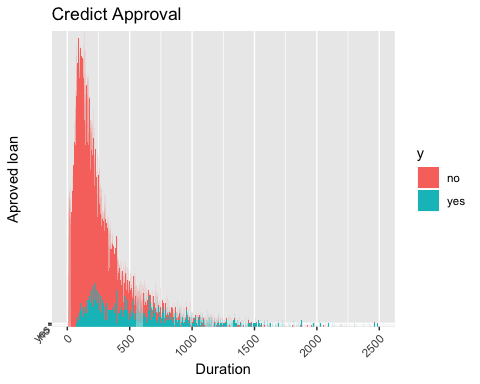
## age job marital education default housing loan contact  
## 1 41 blue-collar divorced basic.4y unknown yes no telephone  
## 2 49 entrepreneur married university.degree unknown yes no telephone  
## 3 49 technician married basic.9y no no no telephone  
## 4 41 technician married professional.course unknown yes no telephone  
## 5 45 blue-collar married basic.9y unknown yes no telephone  
## 6 42 blue-collar married basic.9y no yes yes telephone  
## month day\_of\_week duration campaign pdays previous poutcome emp.var.rate  
## 1 may mon 1575 1 999 0 nonexistent 1.1  
## 2 may mon 1042 1 999 0 nonexistent 1.1  
## 3 may mon 1467 1 999 0 nonexistent 1.1  
## 4 may mon 579 1 999 0 nonexistent 1.1  
## 5 may mon 461 1 999 0 nonexistent 1.1  
## 6 may mon 673 2 999 0 nonexistent 1.1  
## cons.price.idx cons.conf.idx euribor3m nr.employed y  
## 1 93.994 -36.4 4.857 5191 yes  
## 2 93.994 -36.4 4.857 5191 yes  
## 3 93.994 -36.4 4.857 5191 yes  
## 4 93.994 -36.4 4.857 5191 yes  
## 5 93.994 -36.4 4.857 5191 yes  
## 6 93.994 -36.4 4.857 5191 yes

# Plot EDA in general for continuous variables

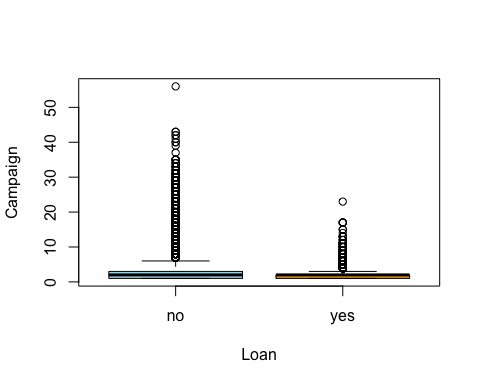
boxplot(bank\_20$age~bank\_20$y,xlab="Loan",ylab="Age",title="Credict Assessment",col=c("lightblue","orange"))



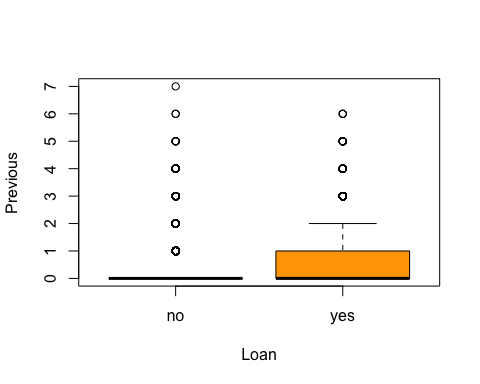
p1= bank\_20 %>% ggplot(aes(x = duration, y = y , fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Duration", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )  
p4= p3 + scale\_x\_continuous(limits = c(0, 2500))  
p4



boxplot(bank\_20$campaign~bank\_20$y,xlab="Loan",ylab="Campaign",title="Credict Assessment",col=c("lightblue","orange"))



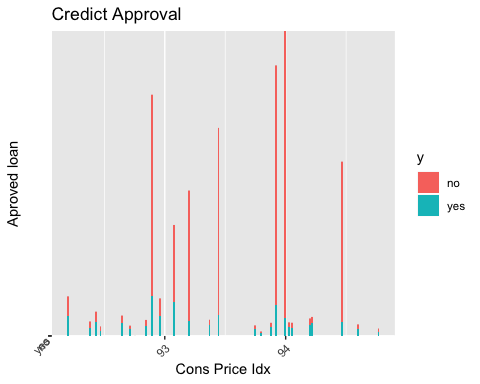
boxplot(bank\_20$previous~bank\_20$y,xlab="Loan",ylab="Previous",title="Credict Assessment",col=c("lightblue","orange"))



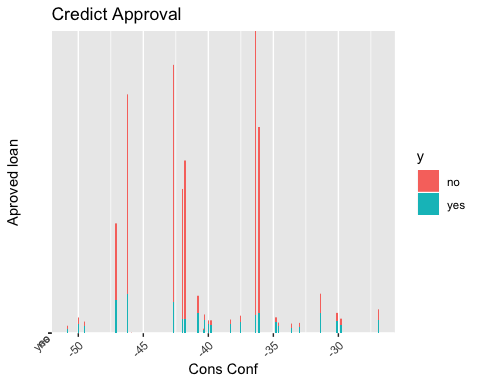
p1= bank\_20 %>% ggplot(aes(x = emp.var.rate, y = y , fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Emp Var Rate", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )  
p4= p3 + scale\_x\_continuous(limits = c(-3.5, 1.4))  
p4



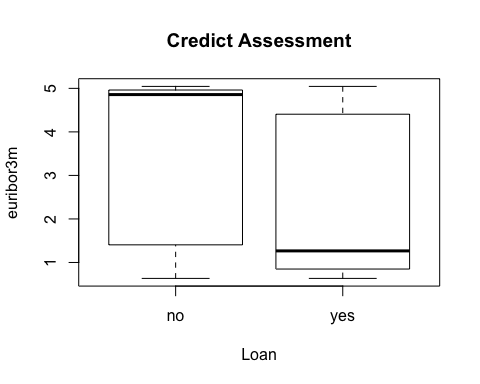
p1= bank\_20 %>% ggplot(aes(x = cons.price.idx, y = y , fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Cons Price Idx", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )  
p3



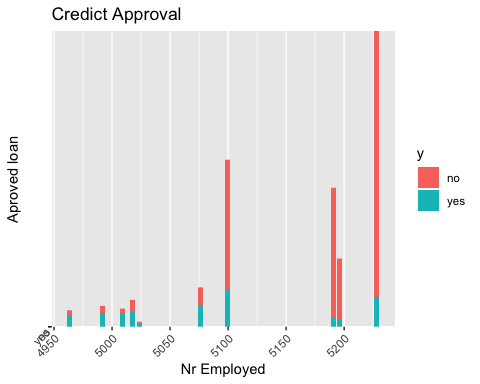
p1= bank\_20 %>% ggplot(aes(x = cons.conf.idx, y = y , fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Cons Conf", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )  
p3



plot(euribor3m~y,data=bank\_20, main="Credict Assessment",  
 xlab="Loan", ylab="euribor3m")



p1= bank\_20 %>% ggplot(aes(x = nr.employed, y = y , fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Nr Employed", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )  
p3

 # EDA for categorical values with proportion table

attach(bank\_20)  
ftable(addmargins(table(bank\_20$y,bank\_20$job)))

## admin. blue-collar entrepreneur housemaid management retired self-employed services student technician unemployed unknown Sum  
##   
## no 9070 8616 1332 954 2596 1286 1272 3646 600 6013 870 293 36548  
## yes 1352 638 124 106 328 434 149 323 275 730 144 37 4640  
## Sum 10422 9254 1456 1060 2924 1720 1421 3969 875 6743 1014 330 41188

job=factor(job)  
ftable(addmargins(table(y, marital)))

## marital divorced married single unknown Sum  
## y   
## no 4136 22396 9948 68 36548  
## yes 476 2532 1620 12 4640  
## Sum 4612 24928 11568 80 41188

marital=factor(marital)  
ftable(addmargins(table(y,education)))

## education basic.4y basic.6y basic.9y high.school illiterate professional.course university.degree unknown Sum  
## y   
## no 3748 2104 5572 8484 14 4648 10498 1480 36548  
## yes 428 188 473 1031 4 595 1670 251 4640  
## Sum 4176 2292 6045 9515 18 5243 12168 1731 41188

education=factor(education)  
ftable(addmargins(table(y,default)))

## default no unknown yes Sum  
## y   
## no 28391 8154 3 36548  
## yes 4197 443 0 4640  
## Sum 32588 8597 3 41188

education=factor(default)  
ftable(addmargins(table(y,housing)))

## housing no unknown yes Sum  
## y   
## no 16596 883 19069 36548  
## yes 2026 107 2507 4640  
## Sum 18622 990 21576 41188

job=factor(housing)  
ftable(addmargins(table(y, loan)))

## loan no unknown yes Sum  
## y   
## no 30100 883 5565 36548  
## yes 3850 107 683 4640  
## Sum 33950 990 6248 41188

marital=factor(loan)  
ftable(addmargins(table(y,contact)))

## contact cellular telephone Sum  
## y   
## no 22291 14257 36548  
## yes 3853 787 4640  
## Sum 26144 15044 41188

education=factor(contact)  
ftable(addmargins(table(y,month)))

## month apr aug dec jul jun mar may nov oct sep Sum  
## y   
## no 2093 5523 93 6525 4759 270 12883 3685 403 314 36548  
## yes 539 655 89 649 559 276 886 416 315 256 4640  
## Sum 2632 6178 182 7174 5318 546 13769 4101 718 570 41188

education=factor(month)  
ftable(addmargins(table(y,day\_of\_week)))

## day\_of\_week fri mon thu tue wed Sum  
## y   
## no 6981 7667 7578 7137 7185 36548  
## yes 846 847 1045 953 949 4640  
## Sum 7827 8514 8623 8090 8134 41188

education=factor(day\_of\_week)  
ftable(addmargins(table(y,poutcome)))

## poutcome failure nonexistent success Sum  
## y   
## no 3647 32422 479 36548  
## yes 605 3141 894 4640  
## Sum 4252 35563 1373 41188

education=factor(poutcome)  
  
#to get proportions that make sense  
prop.table(table(bank\_20$y,bank\_20$job),2)

##   
## admin. blue-collar entrepreneur housemaid management retired  
## no 0.87027442 0.93105684 0.91483516 0.90000000 0.88782490 0.74767442  
## yes 0.12972558 0.06894316 0.08516484 0.10000000 0.11217510 0.25232558  
##   
## self-employed services student technician unemployed unknown  
## no 0.89514426 0.91861930 0.68571429 0.89173958 0.85798817 0.88787879  
## yes 0.10485574 0.08138070 0.31428571 0.10826042 0.14201183 0.11212121

prop.table(table(y,education),2)

## education  
## y failure nonexistent success  
## no 0.85771402 0.91167787 0.34887109  
## yes 0.14228598 0.08832213 0.65112891

prop.table(table(y,default),2)

## default  
## y no unknown yes  
## no 0.8712103 0.9484704 1.0000000  
## yes 0.1287897 0.0515296 0.0000000

prop.table(table(y,housing),2)

## housing  
## y no unknown yes  
## no 0.8912040 0.8919192 0.8838061  
## yes 0.1087960 0.1080808 0.1161939

prop.table(table(y,loan),2)

## loan  
## y no unknown yes  
## no 0.8865979 0.8919192 0.8906850  
## yes 0.1134021 0.1080808 0.1093150

prop.table(table(y,contact),2)

## contact  
## y cellular telephone  
## no 0.85262393 0.94768679  
## yes 0.14737607 0.05231321

prop.table(table(y,month),2)

## month  
## y apr aug dec jul jun mar  
## no 0.79521277 0.89397863 0.51098901 0.90953443 0.89488530 0.49450549  
## yes 0.20478723 0.10602137 0.48901099 0.09046557 0.10511470 0.50549451  
## month  
## y may nov oct sep  
## no 0.93565255 0.89856133 0.56128134 0.55087719  
## yes 0.06434745 0.10143867 0.43871866 0.44912281

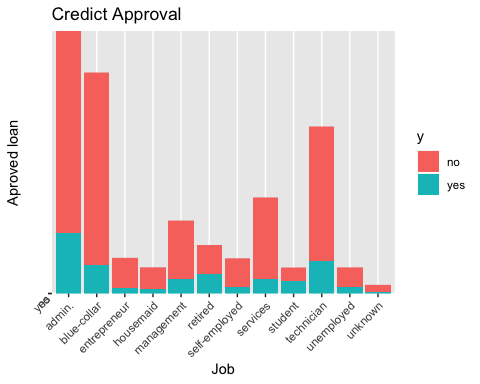
prop.table(table(y,day\_of\_week),2)

## day\_of\_week  
## y fri mon thu tue wed  
## no 0.8919126 0.9005168 0.8788125 0.8822002 0.8833292  
## yes 0.1080874 0.0994832 0.1211875 0.1177998 0.1166708

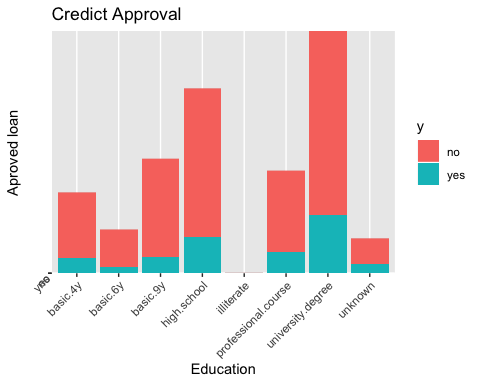
prop.table(table(y,poutcome),2)

## poutcome  
## y failure nonexistent success  
## no 0.85771402 0.91167787 0.34887109  
## yes 0.14228598 0.08832213 0.65112891

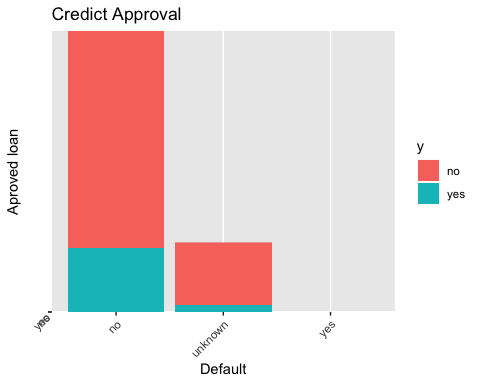
#Visualize  
p1= bank\_20 %>% ggplot(aes(x = job, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Job", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



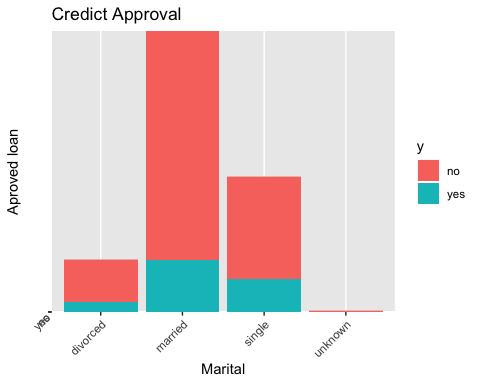
p1= bank\_20 %>% ggplot(aes(x = education, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Education", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



p1= bank\_20 %>% ggplot(aes(x = default, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Default", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



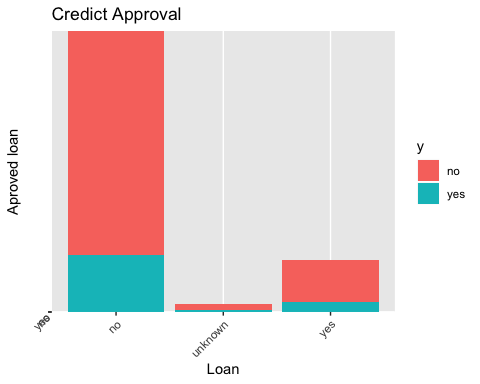
p1= bank\_20 %>% ggplot(aes(x = marital, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Marital", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



p1= bank\_20 %>% ggplot(aes(x = housing, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Housing", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



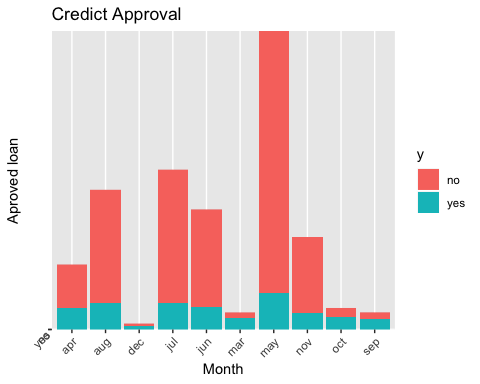
p1= bank\_20 %>% ggplot(aes(x = loan, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Loan", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



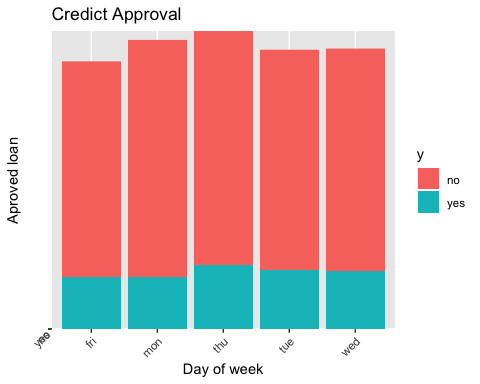
p1= bank\_20 %>% ggplot(aes(x = contact, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Contact", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



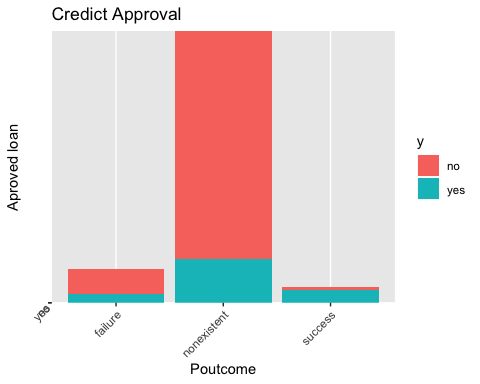
p1= bank\_20 %>% ggplot(aes(x = month, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Month", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



p1= bank\_20 %>% ggplot(aes(x = day\_of\_week, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Day of week", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3

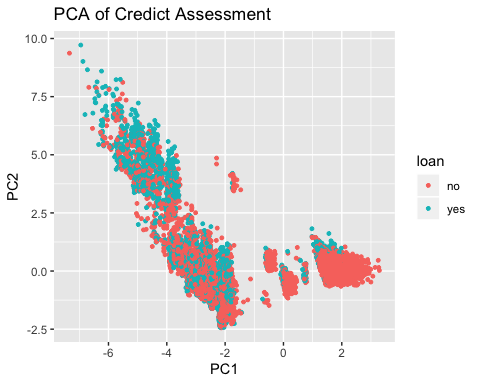


p1= bank\_20 %>% ggplot(aes(x = poutcome, y = y, fill = y))   
p2= p1+ geom\_col() + labs(title = "Credict Approval", x= "Poutcome", y= "Aproved loan")   
p3= p2+ theme( axis.text = element\_text(size = rel(0.8),angle =45, hjust = 1, vjust = 1) )   
p3



#PCA for continuous variables

pc.bc<-prcomp(bank\_20[,-c(2,3,4,5,6,7,8,9,10,15,21)],scale.=TRUE)  
pc.bc.scores<-pc.bc$x  
#Adding the response column to the PC's data frame  
pc.bc.scores<-data.frame(pc.bc.scores)  
pc.bc.scores$loan<-bank\_20$y  
  
#Use ggplot2 to plot the first few pc's  
library(ggplot2)  
ggplot(data = pc.bc.scores, aes(x = PC1, y = PC2)) +  
 geom\_point(aes(col= loan), size=1)+  
 ggtitle("PCA of Credict Assessment")

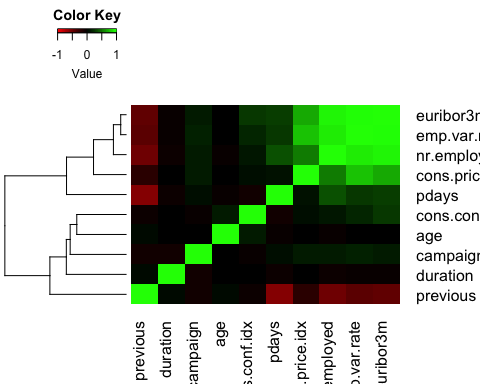


# Heat map for continuous variables

my.cor<-cor(bank\_20[,-c(2,3,4,5,6,7,8,9,10,15,21)])  
my.cor

## age duration campaign pdays previous  
## age 1.0000000000 -0.000865705 0.00459358 -0.03436895 0.02436474  
## duration -0.0008657050 1.000000000 -0.07169923 -0.04757702 0.02064035  
## campaign 0.0045935805 -0.071699226 1.00000000 0.05258357 -0.07914147  
## pdays -0.0343689512 -0.047577015 0.05258357 1.00000000 -0.58751386  
## previous 0.0243647409 0.020640351 -0.07914147 -0.58751386 1.00000000  
## emp.var.rate -0.0003706855 -0.027967884 0.15075381 0.27100417 -0.42048911  
## cons.price.idx 0.0008567150 0.005312268 0.12783591 0.07888911 -0.20312997  
## cons.conf.idx 0.1293716142 -0.008172873 -0.01373310 -0.09134235 -0.05093635  
## euribor3m 0.0107674295 -0.032896656 0.13513251 0.29689911 -0.45449365  
## nr.employed -0.0177251319 -0.044703223 0.14409489 0.37260474 -0.50133293  
## emp.var.rate cons.price.idx cons.conf.idx euribor3m  
## age -0.0003706855 0.000856715 0.129371614 0.01076743  
## duration -0.0279678845 0.005312268 -0.008172873 -0.03289666  
## campaign 0.1507538056 0.127835912 -0.013733099 0.13513251  
## pdays 0.2710041743 0.078889109 -0.091342354 0.29689911  
## previous -0.4204891094 -0.203129967 -0.050936351 -0.45449365  
## emp.var.rate 1.0000000000 0.775334171 0.196041268 0.97224467  
## cons.price.idx 0.7753341708 1.000000000 0.058986182 0.68823011  
## cons.conf.idx 0.1960412681 0.058986182 1.000000000 0.27768622  
## euribor3m 0.9722446712 0.688230107 0.277686220 1.00000000  
## nr.employed 0.9069701013 0.522033977 0.100513432 0.94515443  
## nr.employed  
## age -0.01772513  
## duration -0.04470322  
## campaign 0.14409489  
## pdays 0.37260474  
## previous -0.50133293  
## emp.var.rate 0.90697010  
## cons.price.idx 0.52203398  
## cons.conf.idx 0.10051343  
## euribor3m 0.94515443  
## nr.employed 1.00000000

library(gplots)  
library(ggplot2)  
heatmap.2(my.cor,col=redgreen(75),   
 density.info="none", trace="none", dendrogram=c("row"),   
 symm=F,symkey=T,symbreaks=T, scale="none")



# Pair correlation numerical variables

# Prepare some data  
df <- bank\_20[,-c(2,3,4,5,6,7,8,9,10,15,21)]  
# Correlation plot  
ggcorr(df, palette = "RdBu", label = TRUE)



# Running LDA

# Build X\_train, y\_train, X\_test, y\_test  
X\_train <- balanced\_train[,-c(2,3,4,5,6,7,8,9,10,15,21)]  
y\_train<-balanced\_train[,21]  
  
X\_test <- balanced\_test[,-c(2,3,4,5,6,7,8,9,10,15,21)]  
y\_test <- balanced\_test[,21]  
mylda<-lda(y ~ age + duration + campaign+pdays + previous + emp.var.rate + cons.price.idx + cons.conf.idx + euribor3m + nr.employed, data= balanced\_train)  
pred<-predict(mylda,newdata=X\_test )$class   
Truth<-y\_test  
x<-table(pred,Truth) # Creating a confusion matrix  
x

## Truth  
## pred no yes  
## no 412 71  
## yes 58 387

#Missclassification Error  
ME<-(x[2,1]+x[1,2])/114  
ME

## [1] 1.131579

#Calculating overall accuracy  
oneminusME<-1-ME  
oneminusME

## [1] -0.1315789

# Selection of the variable for LLR

# Forward

model.main<- glm(y~ ., data=balanced\_train,family = binomial(link="logit"))  
model.null<-glm(y ~ 1, data=balanced\_train,family = binomial(link="logit"))  
step(model.null,  
 scope = list(upper=model.main),  
 direction="forward",  
 test="Chisq",  
 data=balanced\_train)

## Start: AIC=11580.31  
## y ~ 1  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + duration 1 9102.6 9106.6 2475.73 < 2.2e-16 \*\*\*  
## + nr.employed 1 9647.9 9651.9 1930.44 < 2.2e-16 \*\*\*  
## + euribor3m 1 9838.1 9842.1 1740.26 < 2.2e-16 \*\*\*  
## + emp.var.rate 1 9966.0 9970.0 1612.27 < 2.2e-16 \*\*\*  
## + pdays 1 10718.7 10722.7 859.66 < 2.2e-16 \*\*\*  
## + month 9 10707.5 10727.5 870.85 < 2.2e-16 \*\*\*  
## + poutcome 2 10725.4 10731.4 852.91 < 2.2e-16 \*\*\*  
## + previous 1 10985.0 10989.0 593.33 < 2.2e-16 \*\*\*  
## + contact 1 11057.2 11061.2 521.16 < 2.2e-16 \*\*\*  
## + job 11 11228.2 11252.2 350.07 < 2.2e-16 \*\*\*  
## + cons.price.idx 1 11269.4 11273.4 308.90 < 2.2e-16 \*\*\*  
## + default 1 11301.7 11305.7 276.63 < 2.2e-16 \*\*\*  
## + campaign 1 11466.4 11470.4 111.93 < 2.2e-16 \*\*\*  
## + education 7 11471.8 11487.8 106.54 < 2.2e-16 \*\*\*  
## + marital 3 11517.8 11525.8 60.51 4.579e-13 \*\*\*  
## + cons.conf.idx 1 11548.7 11552.7 29.64 5.205e-08 \*\*\*  
## + age 1 11572.3 11576.3 6.06 0.01382 \*   
## + day\_of\_week 4 11569.7 11579.7 8.56 0.07297 .   
## <none> 11578.3 11580.3   
## + housing 2 11576.0 11582.0 2.33 0.31162   
## + loan 2 11578.0 11584.0 0.34 0.84322   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=9106.58  
## y ~ duration  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + nr.employed 1 6359.1 6365.1 2743.52 < 2.2e-16 \*\*\*  
## + euribor3m 1 6543.7 6549.7 2558.84 < 2.2e-16 \*\*\*  
## + emp.var.rate 1 6634.1 6640.1 2468.51 < 2.2e-16 \*\*\*  
## + month 9 7857.7 7879.7 1244.83 < 2.2e-16 \*\*\*  
## + poutcome 2 8049.5 8057.5 1053.06 < 2.2e-16 \*\*\*  
## + pdays 1 8088.7 8094.7 1013.93 < 2.2e-16 \*\*\*  
## + previous 1 8338.0 8344.0 764.54 < 2.2e-16 \*\*\*  
## + contact 1 8486.0 8492.0 616.62 < 2.2e-16 \*\*\*  
## + job 11 8557.4 8583.4 545.16 < 2.2e-16 \*\*\*  
## + cons.price.idx 1 8577.8 8583.8 524.76 < 2.2e-16 \*\*\*  
## + default 1 8679.6 8685.6 422.99 < 2.2e-16 \*\*\*  
## + education 7 8891.2 8909.2 211.41 < 2.2e-16 \*\*\*  
## + campaign 1 8980.9 8986.9 121.69 < 2.2e-16 \*\*\*  
## + cons.conf.idx 1 9012.9 9018.9 89.68 < 2.2e-16 \*\*\*  
## + marital 3 9015.4 9025.4 87.16 < 2.2e-16 \*\*\*  
## + age 1 9089.2 9095.2 13.38 0.0002547 \*\*\*  
## + housing 2 9094.3 9102.3 8.29 0.0158628 \*   
## <none> 9102.6 9106.6   
## + loan 2 9101.9 9109.9 0.69 0.7091192   
## + day\_of\_week 4 9099.6 9111.6 2.99 0.5600491   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=6365.06  
## y ~ duration + nr.employed  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + month 9 5855.7 5879.7 503.40 < 2.2e-16 \*\*\*  
## + poutcome 2 6199.0 6209.0 160.01 < 2.2e-16 \*\*\*  
## + pdays 1 6241.4 6249.4 117.68 < 2.2e-16 \*\*\*  
## + job 11 6230.9 6258.9 128.17 < 2.2e-16 \*\*\*  
## + education 7 6263.5 6283.5 95.58 < 2.2e-16 \*\*\*  
## + emp.var.rate 1 6275.7 6283.7 83.35 < 2.2e-16 \*\*\*  
## + contact 1 6282.4 6290.4 76.68 < 2.2e-16 \*\*\*  
## + default 1 6296.0 6304.0 63.10 1.966e-15 \*\*\*  
## + cons.price.idx 1 6311.2 6319.2 47.82 4.673e-12 \*\*\*  
## + marital 3 6334.4 6346.4 24.62 1.856e-05 \*\*\*  
## + euribor3m 1 6345.9 6353.9 13.19 0.0002816 \*\*\*  
## + cons.conf.idx 1 6346.4 6354.4 12.68 0.0003703 \*\*\*  
## + campaign 1 6352.8 6360.8 6.24 0.0124840 \*   
## <none> 6359.1 6365.1   
## + previous 1 6358.0 6366.0 1.11 0.2925308   
## + housing 2 6356.6 6366.6 2.44 0.2958667   
## + age 1 6358.9 6366.9 0.14 0.7128058   
## + loan 2 6358.6 6368.6 0.48 0.7850732   
## + day\_of\_week 4 6357.3 6371.3 1.73 0.7851344   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5879.66  
## y ~ duration + nr.employed + month  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + poutcome 2 5728.6 5756.6 127.083 < 2.2e-16 \*\*\*  
## + pdays 1 5762.9 5788.9 92.739 < 2.2e-16 \*\*\*  
## + emp.var.rate 1 5781.8 5807.8 73.820 < 2.2e-16 \*\*\*  
## + cons.price.idx 1 5808.4 5834.4 47.290 6.124e-12 \*\*\*  
## + default 1 5818.9 5844.9 36.791 1.315e-09 \*\*\*  
## + job 11 5805.8 5851.8 49.814 6.759e-07 \*\*\*  
## + euribor3m 1 5826.4 5852.4 29.291 6.228e-08 \*\*\*  
## + education 7 5818.8 5856.8 36.864 4.978e-06 \*\*\*  
## + marital 3 5839.0 5869.0 16.684 0.0008206 \*\*\*  
## + contact 1 5843.5 5869.5 12.120 0.0004987 \*\*\*  
## + campaign 1 5851.6 5877.6 4.096 0.0429811 \*   
## <none> 5855.7 5879.7   
## + age 1 5854.7 5880.7 0.918 0.3380923   
## + previous 1 5855.0 5881.0 0.623 0.4298194   
## + housing 2 5853.2 5881.2 2.492 0.2877041   
## + cons.conf.idx 1 5855.6 5881.6 0.055 0.8143075   
## + day\_of\_week 4 5849.9 5881.9 5.796 0.2148766   
## + loan 2 5855.6 5883.6 0.096 0.9532976   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5756.58  
## y ~ duration + nr.employed + month + poutcome  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + emp.var.rate 1 5629.3 5659.3 99.249 < 2.2e-16 \*\*\*  
## + cons.price.idx 1 5659.8 5689.8 68.778 < 2.2e-16 \*\*\*  
## + euribor3m 1 5681.4 5711.4 47.202 6.404e-12 \*\*\*  
## + default 1 5693.1 5723.1 35.455 2.610e-09 \*\*\*  
## + job 11 5680.6 5730.6 47.940 1.463e-06 \*\*\*  
## + education 7 5693.9 5735.9 34.635 1.310e-05 \*\*\*  
## + contact 1 5712.7 5742.7 15.853 6.845e-05 \*\*\*  
## + marital 3 5712.6 5746.6 16.020 0.001123 \*\*   
## + previous 1 5724.7 5754.7 3.852 0.049690 \*   
## + campaign 1 5725.2 5755.2 3.386 0.065758 .   
## <none> 5728.6 5756.6   
## + age 1 5727.3 5757.3 1.281 0.257799   
## + cons.conf.idx 1 5727.4 5757.4 1.166 0.280127   
## + pdays 1 5727.4 5757.4 1.147 0.284272   
## + housing 2 5726.4 5758.4 2.227 0.328471   
## + day\_of\_week 4 5722.7 5758.7 5.915 0.205560   
## + loan 2 5728.5 5760.5 0.105 0.949073   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5659.33  
## y ~ duration + nr.employed + month + poutcome + emp.var.rate  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + cons.price.idx 1 5582.1 5614.1 47.215 6.360e-12 \*\*\*  
## + job 11 5580.1 5632.1 49.257 8.509e-07 \*\*\*  
## + default 1 5600.9 5632.9 28.411 9.810e-08 \*\*\*  
## + education 7 5593.8 5637.8 35.544 8.837e-06 \*\*\*  
## + euribor3m 1 5615.0 5647.0 14.299 0.0001559 \*\*\*  
## + marital 3 5614.6 5650.6 14.753 0.0020403 \*\*   
## + pdays 1 5625.2 5657.2 4.119 0.0423945 \*   
## <none> 5629.3 5659.3   
## + campaign 1 5627.7 5659.7 1.608 0.2048216   
## + cons.conf.idx 1 5628.7 5660.7 0.649 0.4206393   
## + contact 1 5628.8 5660.8 0.516 0.4727304   
## + age 1 5628.9 5660.9 0.417 0.5185900   
## + previous 1 5629.2 5661.2 0.163 0.6864059   
## + housing 2 5627.3 5661.3 2.030 0.3623330   
## + day\_of\_week 4 5623.7 5661.7 5.603 0.2308107   
## + loan 2 5629.1 5663.1 0.238 0.8876953   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5614.11  
## y ~ duration + nr.employed + month + poutcome + emp.var.rate +   
## cons.price.idx  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + job 11 5534.9 5588.9 47.213 1.971e-06 \*\*\*  
## + default 1 5555.7 5589.7 26.404 2.769e-07 \*\*\*  
## + education 7 5545.8 5591.8 36.302 6.359e-06 \*\*\*  
## + marital 3 5566.9 5604.9 15.194 0.001658 \*\*   
## + contact 1 5573.0 5607.0 9.106 0.002547 \*\*   
## + euribor3m 1 5576.4 5610.4 5.687 0.017088 \*   
## + pdays 1 5579.3 5613.3 2.848 0.091465 .   
## <none> 5582.1 5614.1   
## + campaign 1 5580.9 5614.9 1.178 0.277839   
## + age 1 5581.7 5615.7 0.399 0.527759   
## + housing 2 5579.9 5615.9 2.178 0.336569   
## + cons.conf.idx 1 5582.0 5616.0 0.107 0.743911   
## + previous 1 5582.1 5616.1 0.031 0.861325   
## + day\_of\_week 4 5576.4 5616.4 5.700 0.222699   
## + loan 2 5582.0 5618.0 0.127 0.938403   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5588.9  
## y ~ duration + nr.employed + month + poutcome + emp.var.rate +   
## cons.price.idx + job  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + default 1 5513.8 5569.8 21.1159 4.323e-06 \*\*\*  
## + contact 1 5525.1 5581.1 9.8041 0.001741 \*\*   
## + education 7 5513.6 5581.6 21.3040 0.003345 \*\*   
## + marital 3 5525.6 5585.6 9.2579 0.026052 \*   
## + euribor3m 1 5531.3 5587.3 3.5731 0.058722 .   
## + age 1 5532.2 5588.2 2.6796 0.101644   
## + pdays 1 5532.4 5588.4 2.5009 0.113782   
## <none> 5534.9 5588.9   
## + campaign 1 5533.4 5589.4 1.4994 0.220760   
## + housing 2 5532.8 5590.8 2.1215 0.346203   
## + previous 1 5534.9 5590.9 0.0416 0.838372   
## + cons.conf.idx 1 5534.9 5590.9 0.0040 0.949439   
## + day\_of\_week 4 5529.7 5591.7 5.1742 0.269881   
## + loan 2 5534.7 5592.7 0.2207 0.895528   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5569.79  
## y ~ duration + nr.employed + month + poutcome + emp.var.rate +   
## cons.price.idx + job + default  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + contact 1 5504.8 5562.8 9.0352 0.002648 \*\*  
## + education 7 5495.5 5565.5 18.3298 0.010567 \*   
## + euribor3m 1 5510.0 5568.0 3.7932 0.051461 .   
## + marital 3 5506.6 5568.6 7.2275 0.064990 .   
## + pdays 1 5511.3 5569.3 2.4660 0.116331   
## <none> 5513.8 5569.8   
## + campaign 1 5512.3 5570.3 1.5168 0.218105   
## + age 1 5512.9 5570.9 0.9324 0.334252   
## + previous 1 5513.7 5571.7 0.0435 0.834802   
## + cons.conf.idx 1 5513.8 5571.8 0.0090 0.924401   
## + housing 2 5511.8 5571.8 2.0002 0.367848   
## + day\_of\_week 4 5509.1 5573.1 4.7313 0.315994   
## + loan 2 5513.5 5573.5 0.2636 0.876504   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5562.75  
## y ~ duration + nr.employed + month + poutcome + emp.var.rate +   
## cons.price.idx + job + default + contact  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + euribor3m 1 5493.2 5553.2 11.5947 0.0006614 \*\*\*  
## + education 7 5486.2 5558.2 18.5580 0.0096908 \*\*   
## + cons.conf.idx 1 5501.3 5561.3 3.4469 0.0633715 .   
## + marital 3 5498.0 5562.0 6.7871 0.0790032 .   
## + pdays 1 5502.3 5562.3 2.4392 0.1183334   
## <none> 5504.8 5562.8   
## + campaign 1 5503.7 5563.7 1.0246 0.3114414   
## + age 1 5503.9 5563.9 0.8287 0.3626460   
## + previous 1 5504.6 5564.6 0.1149 0.7345968   
## + housing 2 5503.0 5565.0 1.7718 0.4123370   
## + day\_of\_week 4 5499.9 5565.9 4.8185 0.3064340   
## + loan 2 5504.5 5566.5 0.2639 0.8763905   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5553.16  
## y ~ duration + nr.employed + month + poutcome + emp.var.rate +   
## cons.price.idx + job + default + contact + euribor3m  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + education 7 5474.4 5548.4 18.7148 0.009129 \*\*  
## + marital 3 5485.8 5551.8 7.3974 0.060255 .   
## + pdays 1 5491.0 5553.0 2.1495 0.142617   
## <none> 5493.2 5553.2   
## + age 1 5492.0 5554.0 1.1989 0.273533   
## + campaign 1 5492.4 5554.4 0.7949 0.372610   
## + cons.conf.idx 1 5492.4 5554.4 0.7736 0.379103   
## + previous 1 5493.0 5555.0 0.1850 0.667133   
## + housing 2 5491.4 5555.4 1.7556 0.415698   
## + day\_of\_week 4 5488.4 5556.4 4.7328 0.315831   
## + loan 2 5492.9 5556.9 0.2704 0.873555   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5548.44  
## y ~ duration + nr.employed + month + poutcome + emp.var.rate +   
## cons.price.idx + job + default + contact + euribor3m + education  
##   
## Df Deviance AIC LRT Pr(>Chi)  
## + pdays 1 5472.1 5548.1 2.3897 0.1221  
## <none> 5474.4 5548.4   
## + marital 3 5469.1 5549.1 5.3465 0.1481  
## + campaign 1 5473.5 5549.5 0.8919 0.3450  
## + cons.conf.idx 1 5473.9 5549.9 0.5824 0.4454  
## + age 1 5473.9 5549.9 0.4937 0.4823  
## + previous 1 5474.3 5550.3 0.1604 0.6888  
## + housing 2 5472.9 5550.9 1.5900 0.4516  
## + day\_of\_week 4 5469.5 5551.5 4.9329 0.2943  
## + loan 2 5474.2 5552.2 0.2804 0.8692  
##   
## Step: AIC=5548.05  
## y ~ duration + nr.employed + month + poutcome + emp.var.rate +   
## cons.price.idx + job + default + contact + euribor3m + education +   
## pdays  
##   
## Df Deviance AIC LRT Pr(>Chi)  
## <none> 5472.1 5548.1   
## + previous 1 5470.5 5548.5 1.5040 0.2201  
## + marital 3 5466.6 5548.6 5.4202 0.1435  
## + campaign 1 5471.1 5549.1 0.9309 0.3346  
## + cons.conf.idx 1 5471.4 5549.4 0.6132 0.4336  
## + age 1 5471.6 5549.6 0.4993 0.4798  
## + housing 2 5470.4 5550.4 1.6318 0.4422  
## + day\_of\_week 4 5467.1 5551.1 4.9752 0.2899  
## + loan 2 5471.7 5551.7 0.3161 0.8538

##   
## Call: glm(formula = y ~ duration + nr.employed + month + poutcome +   
## emp.var.rate + cons.price.idx + job + default + contact +   
## euribor3m + education + pdays, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Coefficients:  
## (Intercept) duration   
## -2.579e+02 7.013e-03   
## nr.employed monthaug   
## 5.514e-03 1.065e+00   
## monthdec monthjul   
## -2.495e-01 1.343e-01   
## monthjun monthmar   
## -8.776e-01 1.873e+00   
## monthmay monthnov   
## -8.085e-01 -6.803e-01   
## monthoct monthsep   
## 2.569e-01 3.117e-01   
## poutcomenonexistent poutcomesuccess   
## 5.584e-01 1.326e+00   
## emp.var.rate cons.price.idx   
## -2.227e+00 2.402e+00   
## jobblue-collar jobentrepreneur   
## -1.798e-01 -2.551e-01   
## jobhousemaid jobmanagement   
## 2.738e-01 3.607e-02   
## jobretired jobself-employed   
## 4.510e-01 -5.574e-03   
## jobservices jobstudent   
## -3.542e-02 6.010e-01   
## jobtechnician jobunemployed   
## -4.423e-02 1.735e-01   
## jobunknown defaultunknown   
## 5.944e-01 -4.471e-01   
## contacttelephone euribor3m   
## -4.967e-01 5.241e-01   
## educationbasic.6y educationbasic.9y   
## -9.108e-02 -1.700e-01   
## educationhigh.school educationilliterate   
## 4.365e-02 1.117e+00   
## educationprofessional.course educationuniversity.degree   
## 2.782e-01 3.272e-01   
## educationunknown pdays   
## 6.105e-03 -5.823e-04   
##   
## Degrees of Freedom: 8351 Total (i.e. Null); 8314 Residual  
## Null Deviance: 11580   
## Residual Deviance: 5472 AIC: 5548

# The forward selection select a lot of variables. Please see below.

forward<- glm(formula = y ~ duration + nr.employed + month + poutcome +   
 emp.var.rate + cons.price.idx + job + euribor3m + contact +   
 day\_of\_week + pdays + default + previous, family = binomial(link = "logit"),   
 data = balanced\_train)  
forward

##   
## Call: glm(formula = y ~ duration + nr.employed + month + poutcome +   
## emp.var.rate + cons.price.idx + job + euribor3m + contact +   
## day\_of\_week + pdays + default + previous, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Coefficients:  
## (Intercept) duration nr.employed   
## -2.567e+02 6.974e-03 5.124e-03   
## monthaug monthdec monthjul   
## 1.097e+00 -2.818e-01 1.159e-01   
## monthjun monthmar monthmay   
## -8.890e-01 1.878e+00 -8.531e-01   
## monthnov monthoct monthsep   
## -6.710e-01 2.329e-01 2.795e-01   
## poutcomenonexistent poutcomesuccess emp.var.rate   
## 4.069e-01 1.187e+00 -2.220e+00   
## cons.price.idx jobblue-collar jobentrepreneur   
## 2.417e+00 -4.241e-01 -2.617e-01   
## jobhousemaid jobmanagement jobretired   
## 1.394e-01 8.605e-02 3.135e-01   
## jobself-employed jobservices jobstudent   
## 2.160e-02 -1.794e-01 4.435e-01   
## jobtechnician jobunemployed jobunknown   
## -2.288e-02 5.994e-02 4.454e-01   
## euribor3m contacttelephone day\_of\_weekmon   
## 5.226e-01 -4.994e-01 -2.097e-02   
## day\_of\_weekthu day\_of\_weektue day\_of\_weekwed   
## -8.042e-02 -3.061e-02 1.416e-01   
## pdays defaultunknown previous   
## -7.758e-04 -4.711e-01 -1.532e-01   
##   
## Degrees of Freedom: 8351 Total (i.e. Null); 8316 Residual  
## Null Deviance: 11580   
## Residual Deviance: 5485 AIC: 5557

# Forward Odd ratio

summary(forward)

##   
## Call:  
## glm(formula = y ~ duration + nr.employed + month + poutcome +   
## emp.var.rate + cons.price.idx + job + euribor3m + contact +   
## day\_of\_week + pdays + default + previous, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -6.5179 -0.3786 0.0004 0.4923 2.9044   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.567e+02 4.944e+01 -5.193 2.07e-07 \*\*\*  
## duration 6.974e-03 1.776e-04 39.259 < 2e-16 \*\*\*  
## nr.employed 5.124e-03 3.586e-03 1.429 0.153034   
## monthaug 1.097e+00 2.254e-01 4.865 1.15e-06 \*\*\*  
## monthdec -2.818e-01 3.833e-01 -0.735 0.462186   
## monthjul 1.159e-01 1.616e-01 0.717 0.473194   
## monthjun -8.890e-01 1.970e-01 -4.513 6.39e-06 \*\*\*  
## monthmar 1.878e+00 2.433e-01 7.719 1.17e-14 \*\*\*  
## monthmay -8.531e-01 1.362e-01 -6.262 3.80e-10 \*\*\*  
## monthnov -6.710e-01 1.975e-01 -3.397 0.000682 \*\*\*  
## monthoct 2.329e-01 2.579e-01 0.903 0.366494   
## monthsep 2.795e-01 3.002e-01 0.931 0.351890   
## poutcomenonexistent 4.069e-01 1.751e-01 2.324 0.020137 \*   
## poutcomesuccess 1.187e+00 4.222e-01 2.811 0.004933 \*\*   
## emp.var.rate -2.220e+00 2.325e-01 -9.550 < 2e-16 \*\*\*  
## cons.price.idx 2.417e+00 3.527e-01 6.854 7.16e-12 \*\*\*  
## jobblue-collar -4.241e-01 1.084e-01 -3.913 9.10e-05 \*\*\*  
## jobentrepreneur -2.617e-01 1.991e-01 -1.314 0.188719   
## jobhousemaid 1.394e-01 2.437e-01 0.572 0.567167   
## jobmanagement 8.605e-02 1.435e-01 0.600 0.548838   
## jobretired 3.135e-01 1.505e-01 2.084 0.037198 \*   
## jobself-employed 2.160e-02 1.959e-01 0.110 0.912214   
## jobservices -1.794e-01 1.383e-01 -1.297 0.194477   
## jobstudent 4.435e-01 1.995e-01 2.223 0.026202 \*   
## jobtechnician -2.288e-02 1.084e-01 -0.211 0.832803   
## jobunemployed 5.994e-02 2.143e-01 0.280 0.779715   
## jobunknown 4.454e-01 3.924e-01 1.135 0.256284   
## euribor3m 5.226e-01 1.551e-01 3.368 0.000757 \*\*\*  
## contacttelephone -4.994e-01 1.202e-01 -4.153 3.28e-05 \*\*\*  
## day\_of\_weekmon -2.097e-02 1.113e-01 -0.188 0.850540   
## day\_of\_weekthu -8.042e-02 1.099e-01 -0.731 0.464503   
## day\_of\_weektue -3.061e-02 1.119e-01 -0.273 0.784537   
## day\_of\_weekwed 1.416e-01 1.115e-01 1.270 0.203916   
## pdays -7.758e-04 4.321e-04 -1.796 0.072563 .   
## defaultunknown -4.711e-01 1.066e-01 -4.418 9.96e-06 \*\*\*  
## previous -1.532e-01 1.235e-01 -1.241 0.214578   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 11578.3 on 8351 degrees of freedom  
## Residual deviance: 5484.8 on 8316 degrees of freedom  
## AIC: 5556.8  
##   
## Number of Fisher Scoring iterations: 6

exp(cbind("Odds ratio" = coef(forward), confint.default(forward, level = 0.95)))

## Odds ratio 2.5 % 97.5 %  
## (Intercept) 3.193196e-112 2.626168e-154 3.882653e-70  
## duration 1.006998e+00 1.006648e+00 1.007349e+00  
## nr.employed 1.005137e+00 9.980974e-01 1.012226e+00  
## monthaug 2.994208e+00 1.924793e+00 4.657790e+00  
## monthdec 7.543961e-01 3.558901e-01 1.599127e+00  
## monthjul 1.122890e+00 8.180754e-01 1.541278e+00  
## monthjun 4.110504e-01 2.793921e-01 6.047502e-01  
## monthmar 6.539026e+00 4.059190e+00 1.053384e+01  
## monthmay 4.260822e-01 3.262320e-01 5.564937e-01  
## monthnov 5.111951e-01 3.470921e-01 7.528848e-01  
## monthoct 1.262296e+00 7.613867e-01 2.092747e+00  
## monthsep 1.322406e+00 7.342517e-01 2.381687e+00  
## poutcomenonexistent 1.502184e+00 1.065782e+00 2.117277e+00  
## poutcomesuccess 3.277368e+00 1.432610e+00 7.497603e+00  
## emp.var.rate 1.085939e-01 6.885428e-02 1.712694e-01  
## cons.price.idx 1.121523e+01 5.618533e+00 2.238687e+01  
## jobblue-collar 6.543337e-01 5.291059e-01 8.092003e-01  
## jobentrepreneur 7.697325e-01 5.210180e-01 1.137174e+00  
## jobhousemaid 1.149616e+00 7.131004e-01 1.853340e+00  
## jobmanagement 1.089866e+00 8.226002e-01 1.443967e+00  
## jobretired 1.368258e+00 1.018777e+00 1.837624e+00  
## jobself-employed 1.021835e+00 6.959961e-01 1.500220e+00  
## jobservices 8.357612e-01 6.373492e-01 1.095941e+00  
## jobstudent 1.558186e+00 1.053919e+00 2.303730e+00  
## jobtechnician 9.773786e-01 7.903230e-01 1.208707e+00  
## jobunemployed 1.061770e+00 6.976202e-01 1.616002e+00  
## jobunknown 1.561169e+00 7.235300e-01 3.368554e+00  
## euribor3m 1.686342e+00 1.244185e+00 2.285633e+00  
## contacttelephone 6.068817e-01 4.794543e-01 7.681761e-01  
## day\_of\_weekmon 9.792519e-01 7.873760e-01 1.217886e+00  
## day\_of\_weekthu 9.227310e-01 7.438614e-01 1.144612e+00  
## day\_of\_weektue 9.698565e-01 7.787883e-01 1.207801e+00  
## day\_of\_weekwed 1.152109e+00 9.260345e-01 1.433376e+00  
## pdays 9.992245e-01 9.983786e-01 1.000071e+00  
## defaultunknown 6.242900e-01 5.065412e-01 7.694103e-01  
## previous 8.579462e-01 6.735611e-01 1.092806e+00

vif(forward)

## GVIF Df GVIF^(1/(2\*Df))  
## duration 1.414523 1 1.189337  
## nr.employed 73.952592 1 8.599569  
## month 41.683018 9 1.230262  
## poutcome 25.017444 2 2.236458  
## emp.var.rate 131.909513 1 11.485187  
## cons.price.idx 40.625286 1 6.373797  
## job 1.255600 11 1.010400  
## euribor3m 69.388682 1 8.329987  
## contact 2.176459 1 1.475283  
## day\_of\_week 1.065583 4 1.007972  
## pdays 9.663774 1 3.108661  
## default 1.117194 1 1.056974  
## previous 5.224567 1 2.285731

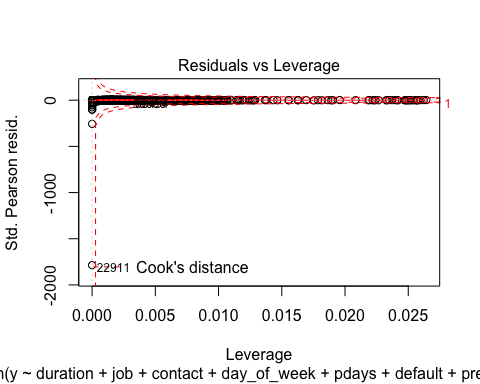
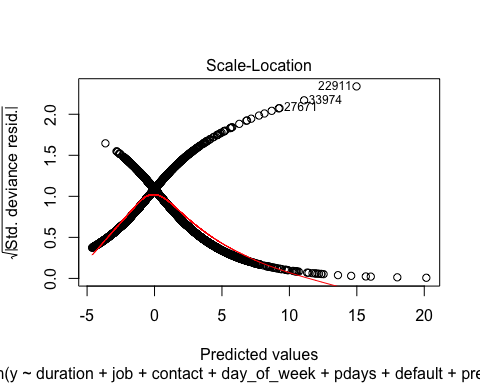
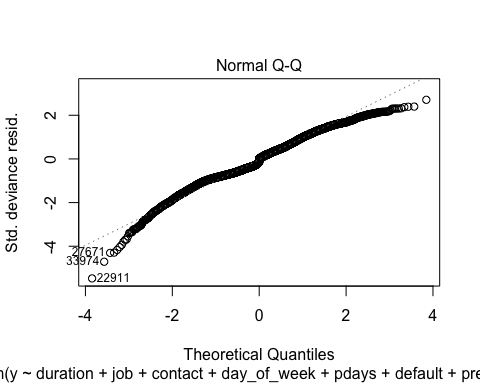
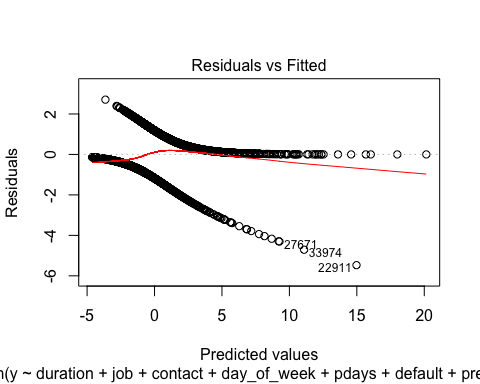
# Take off the high VIF variances

forward2<- glm(formula = y ~ duration + job + contact + day\_of\_week + pdays + default + previous, family = binomial(link = "logit"), data = balanced\_train)  
forward2

##   
## Call: glm(formula = y ~ duration + job + contact + day\_of\_week + pdays +   
## default + previous, family = binomial(link = "logit"), data = balanced\_train)  
##   
## Coefficients:  
## (Intercept) duration jobblue-collar jobentrepreneur   
## 0.7371724 0.0056303 -0.7027089 -0.5452956   
## jobhousemaid jobmanagement jobretired jobself-employed   
## -0.0009632 -0.0563133 0.9008874 -0.0617064   
## jobservices jobstudent jobtechnician jobunemployed   
## -0.4837288 1.2542740 -0.2292186 0.2978889   
## jobunknown contacttelephone day\_of\_weekmon day\_of\_weekthu   
## 0.7323284 -1.1485675 -0.0710180 -0.0466826   
## day\_of\_weektue day\_of\_weekwed pdays defaultunknown   
## -0.0386583 -0.0357929 -0.0023318 -1.1836297   
## previous   
## 0.2247782   
##   
## Degrees of Freedom: 8351 Total (i.e. Null); 8331 Residual  
## Null Deviance: 11580   
## Residual Deviance: 7172 AIC: 7214

# Plot forward

plot(forward2)



# Summary model forward

summary(forward2)

##   
## Call:  
## glm(formula = y ~ duration + job + contact + day\_of\_week + pdays +   
## default + previous, family = binomial(link = "logit"), data = balanced\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -5.4727 -0.6868 0.0021 0.6368 2.7081   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.7371724 0.1839245 4.008 6.12e-05 \*\*\*  
## duration 0.0056303 0.0001448 38.881 < 2e-16 \*\*\*  
## jobblue-collar -0.7027089 0.0935257 -7.514 5.76e-14 \*\*\*  
## jobentrepreneur -0.5452956 0.1742770 -3.129 0.001755 \*\*   
## jobhousemaid -0.0009632 0.1957079 -0.005 0.996073   
## jobmanagement -0.0563133 0.1229750 -0.458 0.647006   
## jobretired 0.9008874 0.1256851 7.168 7.62e-13 \*\*\*  
## jobself-employed -0.0617064 0.1672709 -0.369 0.712201   
## jobservices -0.4837288 0.1181435 -4.094 4.23e-05 \*\*\*  
## jobstudent 1.2542740 0.1784569 7.028 2.09e-12 \*\*\*  
## jobtechnician -0.2292186 0.0897349 -2.554 0.010637 \*   
## jobunemployed 0.2978889 0.1785784 1.668 0.095293 .   
## jobunknown 0.7323284 0.3189649 2.296 0.021679 \*   
## contacttelephone -1.1485675 0.0729853 -15.737 < 2e-16 \*\*\*  
## day\_of\_weekmon -0.0710180 0.0953979 -0.744 0.456610   
## day\_of\_weekthu -0.0466826 0.0932401 -0.501 0.616603   
## day\_of\_weektue -0.0386583 0.0942556 -0.410 0.681701   
## day\_of\_weekwed -0.0357929 0.0950760 -0.376 0.706570   
## pdays -0.0023318 0.0001596 -14.606 < 2e-16 \*\*\*  
## defaultunknown -1.1836297 0.0936555 -12.638 < 2e-16 \*\*\*  
## previous 0.2247782 0.0657258 3.420 0.000626 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 11578.3 on 8351 degrees of freedom  
## Residual deviance: 7172.4 on 8331 degrees of freedom  
## AIC: 7214.4  
##   
## Number of Fisher Scoring iterations: 5

# Prediction with the forward selection

logr<-glm(y ~ duration + job + contact + day\_of\_week + pdays + default + previous,data=balanced\_train, family=binomial)  
  
# 1 way  
logr.probs<-predict(logr, newdata=balanced\_test)  
logr.pred<-rep("No",928)  
logr.pred[logr.probs>.5]="Yes"  
Truth<-balanced\_test[,21]  
Pred<-logr.pred  
ftable(addmargins(table(Pred,Truth)))

## Truth no yes Sum  
## Pred   
## No 425 150 575  
## Yes 45 308 353  
## Sum 470 458 928

# 2 way  
pred = predict(logr, newdata=balanced\_test)  
accuracy <- table(pred, balanced\_test[,21])  
sum(diag(accuracy))/sum(accuracy)

## [1] 0.001077586

#confusionMatrix(factor(pred, levels = 1:928), factor(balanced\_test$y, levels = 1:928))

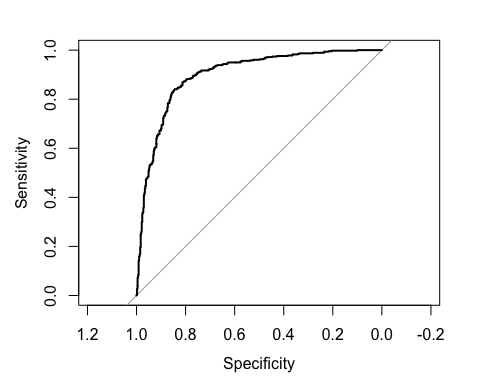
# Holmes test for forward selecttion

hoslem.test(forward$y, fitted(forward), g=10)

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: forward$y, fitted(forward)  
## X-squared = 641.79, df = 8, p-value < 2.2e-16

# ROC curve for the forward model

logr<-glm(y ~ duration + job + contact + day\_of\_week + pdays + default + previous,data=balanced\_train, family=binomial)  
logr.probs<-predict(logr, newdata=balanced\_test)  
roccurve <- roc(balanced\_test$y ~ logr.probs)  
plot(roccurve)



ME<-c()  
FP<-c()  
FN<-c()  
index<-0:928/928  
for (i in 1:928){  
logr.pred<-rep("no",928)  
logr.pred[logr.probs>index[i]]="yes"  
logr.pred<-factor(logr.pred,levels=c("no","yes"))  
Truth<-balanced\_test[,21]  
Pred<-logr.pred  
x<-table(Pred,Truth)  
ME[i]<-(x[1,2]+x[2,1])/sum(x)  
FP[i]<-x[2,1]/(x[2,1]+x[1,1])  
FN[i]<-x[1,2]/(x[1,2]+x[2,2])  
}  
TP= 1-FN  
ord.ind<-order(FP,decreasing=F)  
FP[ord.ind]

## [1] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [7] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [13] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [19] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [25] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [31] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [37] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [43] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [49] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [55] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [61] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [67] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [73] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [79] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [85] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277  
## [91] 0.07021277 0.07021277 0.07021277 0.07021277 0.07021277 0.07234043  
## [97] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [103] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [109] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [115] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [121] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [127] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [133] 0.07234043 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [139] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [145] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [151] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [157] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [163] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [169] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [175] 0.07446809 0.07446809 0.07446809 0.07659574 0.07659574 0.07659574  
## [181] 0.07659574 0.07659574 0.07659574 0.07659574 0.07872340 0.07872340  
## [187] 0.07872340 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [193] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [199] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [205] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [211] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [217] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [223] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [229] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [235] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [241] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [247] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [253] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [259] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [265] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [271] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [277] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [283] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [289] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [295] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [301] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [307] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [313] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [319] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [325] 0.08085106 0.08085106 0.08085106 0.08297872 0.08297872 0.08297872  
## [331] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [337] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [343] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [349] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [355] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [361] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08510638  
## [367] 0.08510638 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [373] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [379] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [385] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [391] 0.08936170 0.08936170 0.08936170 0.08936170 0.08936170 0.08936170  
## [397] 0.08936170 0.08936170 0.08936170 0.09148936 0.09148936 0.09148936  
## [403] 0.09148936 0.09148936 0.09148936 0.09148936 0.09148936 0.09148936  
## [409] 0.09148936 0.09148936 0.09148936 0.09148936 0.09148936 0.09148936  
## [415] 0.09148936 0.09148936 0.09148936 0.09148936 0.09361702 0.09361702  
## [421] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [427] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [433] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [439] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [445] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [451] 0.09361702 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [457] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [463] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [469] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09787234  
## [475] 0.09787234 0.09787234 0.09787234 0.09787234 0.09787234 0.09787234  
## [481] 0.09787234 0.09787234 0.09787234 0.09787234 0.10000000 0.10000000  
## [487] 0.10000000 0.10000000 0.10000000 0.10000000 0.10000000 0.10000000  
## [493] 0.10000000 0.10000000 0.10000000 0.10000000 0.10000000 0.10000000  
## [499] 0.10000000 0.10000000 0.10000000 0.10000000 0.10000000 0.10000000  
## [505] 0.10000000 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [511] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [517] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [523] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [529] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [535] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [541] 0.10425532 0.10425532 0.10425532 0.10425532 0.10638298 0.10638298  
## [547] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [553] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [559] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [565] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [571] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [577] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [583] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [589] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [595] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.11063830  
## [601] 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830  
## [607] 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830  
## [613] 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830  
## [619] 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830  
## [625] 0.11063830 0.11063830 0.11063830 0.11276596 0.11276596 0.11276596  
## [631] 0.11276596 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [637] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [643] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [649] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [655] 0.11489362 0.11489362 0.11489362 0.11489362 0.11702128 0.11702128  
## [661] 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128  
## [667] 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128  
## [673] 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128  
## [679] 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128  
## [685] 0.11702128 0.11702128 0.11702128 0.11702128 0.11914894 0.11914894  
## [691] 0.11914894 0.11914894 0.11914894 0.11914894 0.11914894 0.11914894  
## [697] 0.11914894 0.11914894 0.11914894 0.11914894 0.11914894 0.11914894  
## [703] 0.11914894 0.11914894 0.12127660 0.12127660 0.12127660 0.12127660  
## [709] 0.12127660 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426  
## [715] 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426  
## [721] 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426  
## [727] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [733] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [739] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [745] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [751] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [757] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [763] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [769] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [775] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [781] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [787] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [793] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [799] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [805] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [811] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [817] 0.12765957 0.12765957 0.12978723 0.12978723 0.12978723 0.12978723  
## [823] 0.12978723 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489  
## [829] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [835] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [841] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [847] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [853] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [859] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [865] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [871] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [877] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [883] 0.13617021 0.13617021 0.13617021 0.13617021 0.13617021 0.13617021  
## [889] 0.13617021 0.13617021 0.13617021 0.13617021 0.13617021 0.13829787  
## [895] 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787  
## [901] 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787  
## [907] 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787  
## [913] 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787  
## [919] 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787  
## [925] 0.14042553 0.14042553 0.14042553 0.14042553

TP[ord.ind]

## [1] 0.5873362 0.5873362 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528  
## [8] 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528  
## [15] 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528  
## [22] 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528 0.5851528  
## [29] 0.5851528 0.5829694 0.5829694 0.5829694 0.5829694 0.5829694 0.5829694  
## [36] 0.5829694 0.5829694 0.5829694 0.5829694 0.5829694 0.5829694 0.5829694  
## [43] 0.5807860 0.5807860 0.5807860 0.5807860 0.5786026 0.5786026 0.5786026  
## [50] 0.5786026 0.5786026 0.5786026 0.5786026 0.5786026 0.5786026 0.5786026  
## [57] 0.5786026 0.5786026 0.5786026 0.5786026 0.5786026 0.5786026 0.5786026  
## [64] 0.5786026 0.5786026 0.5786026 0.5786026 0.5786026 0.5764192 0.5764192  
## [71] 0.5764192 0.5764192 0.5742358 0.5742358 0.5742358 0.5742358 0.5742358  
## [78] 0.5742358 0.5742358 0.5742358 0.5742358 0.5742358 0.5742358 0.5742358  
## [85] 0.5720524 0.5720524 0.5720524 0.5720524 0.5720524 0.5720524 0.5720524  
## [92] 0.5720524 0.5720524 0.5720524 0.5720524 0.5982533 0.5982533 0.5982533  
## [99] 0.5982533 0.5982533 0.5982533 0.5982533 0.5960699 0.5960699 0.5960699  
## [106] 0.5938865 0.5938865 0.5917031 0.5917031 0.5917031 0.5917031 0.5917031  
## [113] 0.5917031 0.5917031 0.5917031 0.5917031 0.5917031 0.5917031 0.5895197  
## [120] 0.5895197 0.5895197 0.5895197 0.5895197 0.5895197 0.5895197 0.5895197  
## [127] 0.5895197 0.5895197 0.5895197 0.5895197 0.5895197 0.5895197 0.5895197  
## [134] 0.6026201 0.6026201 0.6026201 0.6026201 0.6026201 0.6026201 0.6004367  
## [141] 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367  
## [148] 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367  
## [155] 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367  
## [162] 0.6004367 0.6004367 0.6004367 0.6004367 0.5982533 0.5982533 0.5982533  
## [169] 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533  
## [176] 0.5982533 0.5982533 0.6026201 0.6026201 0.6026201 0.6026201 0.6026201  
## [183] 0.6026201 0.6026201 0.6048035 0.6048035 0.6026201 0.6397380 0.6397380  
## [190] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [197] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [204] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [211] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6375546  
## [218] 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546  
## [225] 0.6375546 0.6375546 0.6375546 0.6375546 0.6331878 0.6331878 0.6331878  
## [232] 0.6331878 0.6331878 0.6331878 0.6331878 0.6331878 0.6310044 0.6310044  
## [239] 0.6310044 0.6310044 0.6288210 0.6266376 0.6244541 0.6244541 0.6244541  
## [246] 0.6222707 0.6222707 0.6222707 0.6222707 0.6222707 0.6222707 0.6222707  
## [253] 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873  
## [260] 0.6200873 0.6200873 0.6200873 0.6179039 0.6179039 0.6157205 0.6157205  
## [267] 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205  
## [274] 0.6157205 0.6157205 0.6135371 0.6135371 0.6135371 0.6135371 0.6135371  
## [281] 0.6135371 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537  
## [288] 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537  
## [295] 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537  
## [302] 0.6113537 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869  
## [309] 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869  
## [316] 0.6069869 0.6069869 0.6069869 0.6048035 0.6048035 0.6048035 0.6048035  
## [323] 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6506550 0.6506550  
## [330] 0.6506550 0.6506550 0.6506550 0.6484716 0.6484716 0.6484716 0.6484716  
## [337] 0.6484716 0.6484716 0.6484716 0.6484716 0.6484716 0.6484716 0.6484716  
## [344] 0.6484716 0.6484716 0.6484716 0.6484716 0.6484716 0.6484716 0.6462882  
## [351] 0.6462882 0.6419214 0.6419214 0.6419214 0.6419214 0.6419214 0.6419214  
## [358] 0.6419214 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [365] 0.6397380 0.6506550 0.6506550 0.6572052 0.6572052 0.6572052 0.6572052  
## [372] 0.6572052 0.6572052 0.6572052 0.6550218 0.6550218 0.6550218 0.6528384  
## [379] 0.6528384 0.6528384 0.6506550 0.6506550 0.6506550 0.6506550 0.6506550  
## [386] 0.6506550 0.6506550 0.6506550 0.6506550 0.6506550 0.6572052 0.6572052  
## [393] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052  
## [400] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052  
## [407] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052  
## [414] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6724891 0.6724891  
## [421] 0.6724891 0.6724891 0.6681223 0.6681223 0.6681223 0.6681223 0.6681223  
## [428] 0.6681223 0.6681223 0.6681223 0.6681223 0.6659389 0.6637555 0.6637555  
## [435] 0.6615721 0.6615721 0.6615721 0.6615721 0.6615721 0.6615721 0.6593886  
## [442] 0.6593886 0.6593886 0.6593886 0.6593886 0.6593886 0.6593886 0.6572052  
## [449] 0.6572052 0.6572052 0.6572052 0.6724891 0.6724891 0.6724891 0.6724891  
## [456] 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891  
## [463] 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891  
## [470] 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891  
## [477] 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891  
## [484] 0.6724891 0.6812227 0.6812227 0.6790393 0.6790393 0.6790393 0.6768559  
## [491] 0.6768559 0.6768559 0.6768559 0.6746725 0.6746725 0.6746725 0.6724891  
## [498] 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891 0.6724891  
## [505] 0.6724891 0.6943231 0.6943231 0.6943231 0.6943231 0.6943231 0.6921397  
## [512] 0.6921397 0.6921397 0.6921397 0.6921397 0.6921397 0.6921397 0.6921397  
## [519] 0.6921397 0.6921397 0.6899563 0.6899563 0.6877729 0.6877729 0.6877729  
## [526] 0.6855895 0.6855895 0.6855895 0.6855895 0.6855895 0.6834061 0.6812227  
## [533] 0.6812227 0.6812227 0.6812227 0.6812227 0.6812227 0.6812227 0.6812227  
## [540] 0.6812227 0.6965066 0.6943231 0.6943231 0.6943231 0.6965066 0.6965066  
## [547] 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074  
## [554] 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074  
## [561] 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074  
## [568] 0.7227074 0.7227074 0.7227074 0.7205240 0.7205240 0.7205240 0.7205240  
## [575] 0.7205240 0.7205240 0.7205240 0.7205240 0.7205240 0.7183406 0.7183406  
## [582] 0.7161572 0.7139738 0.7117904 0.7117904 0.7117904 0.7117904 0.7117904  
## [589] 0.7117904 0.7117904 0.7117904 0.7117904 0.7096070 0.7074236 0.7074236  
## [596] 0.7030568 0.6986900 0.6965066 0.6965066 0.7314410 0.7314410 0.7314410  
## [603] 0.7314410 0.7270742 0.7270742 0.7248908 0.7248908 0.7248908 0.7248908  
## [610] 0.7248908 0.7248908 0.7248908 0.7248908 0.7248908 0.7248908 0.7248908  
## [617] 0.7248908 0.7248908 0.7248908 0.7248908 0.7227074 0.7227074 0.7227074  
## [624] 0.7227074 0.7227074 0.7227074 0.7227074 0.7314410 0.7314410 0.7314410  
## [631] 0.7314410 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079  
## [638] 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079  
## [645] 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079  
## [652] 0.7358079 0.7358079 0.7336245 0.7336245 0.7336245 0.7314410 0.7314410  
## [659] 0.7423581 0.7423581 0.7423581 0.7423581 0.7423581 0.7423581 0.7423581  
## [666] 0.7423581 0.7423581 0.7423581 0.7423581 0.7401747 0.7401747 0.7401747  
## [673] 0.7401747 0.7401747 0.7401747 0.7401747 0.7401747 0.7401747 0.7401747  
## [680] 0.7401747 0.7401747 0.7401747 0.7401747 0.7379913 0.7358079 0.7358079  
## [687] 0.7358079 0.7358079 0.7467249 0.7467249 0.7467249 0.7467249 0.7445415  
## [694] 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415  
## [701] 0.7445415 0.7445415 0.7445415 0.7423581 0.7489083 0.7489083 0.7489083  
## [708] 0.7489083 0.7489083 0.7510917 0.7510917 0.7510917 0.7510917 0.7510917  
## [715] 0.7510917 0.7510917 0.7510917 0.7510917 0.7510917 0.7510917 0.7510917  
## [722] 0.7510917 0.7510917 0.7510917 0.7510917 0.7489083 0.7663755 0.7663755  
## [729] 0.7663755 0.7663755 0.7663755 0.7663755 0.7663755 0.7663755 0.7663755  
## [736] 0.7663755 0.7663755 0.7663755 0.7641921 0.7641921 0.7641921 0.7641921  
## [743] 0.7641921 0.7620087 0.7620087 0.7598253 0.7598253 0.7576419 0.7576419  
## [750] 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585  
## [757] 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585  
## [764] 0.7532751 0.7532751 0.7532751 0.7532751 0.7532751 0.7532751 0.7532751  
## [771] 0.7532751 0.7532751 0.7532751 0.7532751 0.7532751 0.7532751 0.7510917  
## [778] 0.7510917 0.7510917 0.7510917 0.7838428 0.7838428 0.7838428 0.7838428  
## [785] 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428  
## [792] 0.7838428 0.7816594 0.7794760 0.7794760 0.7772926 0.7772926 0.7772926  
## [799] 0.7772926 0.7772926 0.7772926 0.7751092 0.7751092 0.7729258 0.7729258  
## [806] 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424  
## [813] 0.7707424 0.7685590 0.7685590 0.7663755 0.7663755 0.7663755 0.7838428  
## [820] 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428  
## [827] 0.7838428 0.7838428 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598  
## [834] 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598  
## [841] 0.7947598 0.7947598 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764  
## [848] 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764  
## [855] 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764  
## [862] 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764  
## [869] 0.7903930 0.7903930 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096  
## [876] 0.7882096 0.7860262 0.7860262 0.7838428 0.7838428 0.7838428 0.7838428  
## [883] 0.7991266 0.7991266 0.7991266 0.7991266 0.7991266 0.7991266 0.7991266  
## [890] 0.7969432 0.7947598 0.7947598 0.7947598 0.8078603 0.8078603 0.8078603  
## [897] 0.8078603 0.8078603 0.8056769 0.8034934 0.8034934 0.8013100 0.8013100  
## [904] 0.8013100 0.8013100 0.8013100 0.8013100 0.8013100 0.8013100 0.8013100  
## [911] 0.8013100 0.8013100 0.8013100 0.8013100 0.8013100 0.8013100 0.8013100  
## [918] 0.8013100 0.8013100 0.8013100 0.8013100 0.8013100 0.7991266 0.7991266  
## [925] 0.8122271 0.8100437 0.8078603 0.8078603

simple\_auc <- function(TPR, FPR){  
 # inputs already sorted, best scores first   
 dFPR <- c(diff(FPR), 0)  
 dTPR <- c(diff(TPR), 0)  
 sum(TPR \* dFPR) + sum(dTPR \* dFPR)/2  
}  
  
paste("AUC=", simple\_auc(TP[ord.ind],FP[ord.ind]),sep=" ")

## [1] "AUC= 0.0491939979559602"

# Stepwise selection

full.model <- glm(y ~., data = balanced\_train, family = binomial)  
coef(full.model)

## (Intercept) age   
## -2.325568e+02 2.876090e-04   
## jobblue-collar jobentrepreneur   
## -1.708717e-01 -2.164138e-01   
## jobhousemaid jobmanagement   
## 3.044249e-01 7.806889e-02   
## jobretired jobself-employed   
## 4.750497e-01 4.866471e-03   
## jobservices jobstudent   
## -2.526337e-02 4.998026e-01   
## jobtechnician jobunemployed   
## -4.699018e-02 1.785054e-01   
## jobunknown maritalmarried   
## 5.443983e-01 -1.149346e-01   
## maritalsingle maritalunknown   
## 7.220960e-02 2.790661e-01   
## educationbasic.6y educationbasic.9y   
## -6.835241e-02 -1.648517e-01   
## educationhigh.school educationilliterate   
## 2.135473e-02 1.161122e+00   
## educationprofessional.course educationuniversity.degree   
## 2.755826e-01 2.962529e-01   
## educationunknown defaultunknown   
## -8.423946e-03 -4.245768e-01   
## housingunknown housingyes   
## 1.271971e-01 8.209163e-02   
## loanunknown loanyes   
## NA -4.301952e-02   
## contacttelephone monthaug   
## -4.620588e-01 1.105625e+00   
## monthdec monthjul   
## -2.561386e-01 1.502030e-01   
## monthjun monthmar   
## -8.395005e-01 1.806863e+00   
## monthmay monthnov   
## -8.358067e-01 -7.186888e-01   
## monthoct monthsep   
## 2.007753e-01 2.568047e-01   
## day\_of\_weekmon day\_of\_weekthu   
## -3.257127e-02 -8.805456e-02   
## day\_of\_weektue day\_of\_weekwed   
## -4.906611e-02 1.355505e-01   
## duration campaign   
## 7.030507e-03 -1.691015e-02   
## pdays previous   
## -8.241879e-04 -1.476710e-01   
## poutcomenonexistent poutcomesuccess   
## 3.828285e-01 1.140048e+00   
## emp.var.rate cons.price.idx   
## -2.212706e+00 2.276281e+00   
## cons.conf.idx euribor3m   
## -9.503032e-03 6.446628e-01   
## nr.employed   
## 2.813439e-03

step.model <- full.model %>% stepAIC(trace = FALSE)  
coef(step.model)

## (Intercept) jobblue-collar   
## -1.970409e+02 -1.802049e-01   
## jobentrepreneur jobhousemaid   
## -2.602982e-01 2.727915e-01   
## jobmanagement jobretired   
## 3.306125e-02 4.513333e-01   
## jobself-employed jobservices   
## -6.708674e-03 -3.886706e-02   
## jobstudent jobtechnician   
## 5.967127e-01 -4.912401e-02   
## jobunemployed jobunknown   
## 1.697584e-01 5.901584e-01   
## educationbasic.6y educationbasic.9y   
## -9.243168e-02 -1.690657e-01   
## educationhigh.school educationilliterate   
## 4.575544e-02 1.121845e+00   
## educationprofessional.course educationuniversity.degree   
## 2.775349e-01 3.243953e-01   
## educationunknown defaultunknown   
## 1.153884e-02 -4.439485e-01   
## contacttelephone monthaug   
## -4.460769e-01 1.043680e+00   
## monthdec monthjul   
## -3.246304e-01 1.538821e-01   
## monthjun monthmar   
## -7.547137e-01 1.744570e+00   
## monthmay monthnov   
## -8.462151e-01 -7.640733e-01   
## monthoct monthsep   
## 1.581957e-01 1.605086e-01   
## duration pdays   
## 7.018519e-03 -5.857203e-04   
## poutcomenonexistent poutcomesuccess   
## 5.493895e-01 1.326284e+00   
## emp.var.rate cons.price.idx   
## -2.135243e+00 2.040748e+00   
## cons.conf.idx euribor3m   
## -1.629820e-02 7.391759e-01

# Make predictions  
probabilities <- full.model %>% predict(balanced\_test, type = "response")  
predicted.classes <- ifelse(probabilities > 0.5, "yes", "no")  
# Prediction accuracy  
observed.classes <- balanced\_test$y  
mean(predicted.classes == observed.classes)

## [1] 0.8803879

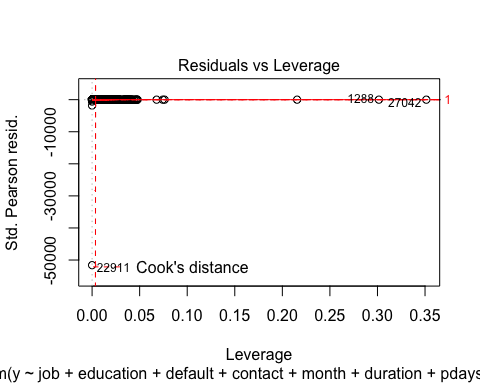
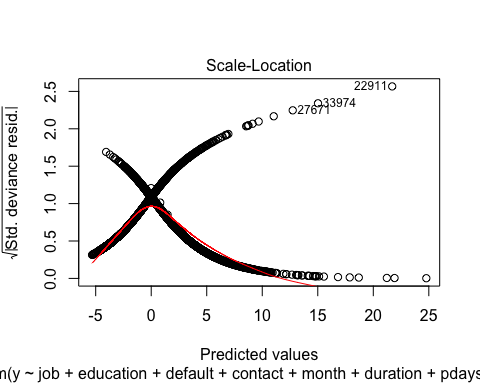
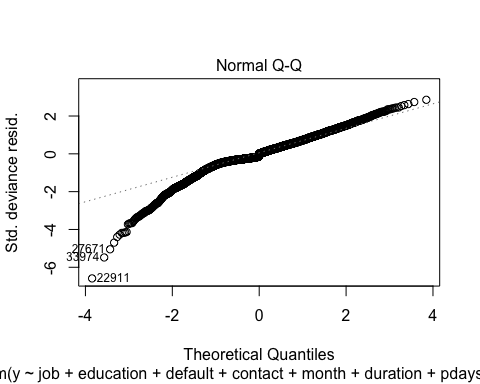
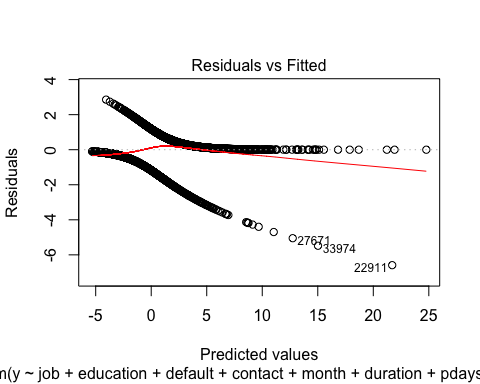
# Make predictions  
probabilities <- predict(step.model, balanced\_test, type = "response")  
predicted.classes <- ifelse(probabilities > 0.5, "yes", "no")  
# Prediction accuracy  
observed.classes <- balanced\_test$y  
mean(predicted.classes == observed.classes)

## [1] 0.8803879

summary(step.model)

##   
## Call:  
## glm(formula = y ~ job + education + default + contact + month +   
## duration + pdays + poutcome + emp.var.rate + cons.price.idx +   
## cons.conf.idx + euribor3m, family = binomial, data = balanced\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -6.5883 -0.3805 0.0005 0.4912 2.8568   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.970e+02 1.920e+01 -10.262 < 2e-16 \*\*\*  
## jobblue-collar -1.802e-01 1.318e-01 -1.367 0.17170   
## jobentrepreneur -2.603e-01 2.003e-01 -1.300 0.19377   
## jobhousemaid 2.728e-01 2.533e-01 1.077 0.28146   
## jobmanagement 3.306e-02 1.456e-01 0.227 0.82033   
## jobretired 4.513e-01 1.655e-01 2.726 0.00640 \*\*   
## jobself-employed -6.709e-03 1.989e-01 -0.034 0.97310   
## jobservices -3.887e-02 1.458e-01 -0.267 0.78984   
## jobstudent 5.967e-01 2.054e-01 2.905 0.00367 \*\*   
## jobtechnician -4.912e-02 1.230e-01 -0.399 0.68955   
## jobunemployed 1.698e-01 2.163e-01 0.785 0.43252   
## jobunknown 5.902e-01 3.971e-01 1.486 0.13722   
## educationbasic.6y -9.243e-02 2.031e-01 -0.455 0.64906   
## educationbasic.9y -1.691e-01 1.596e-01 -1.059 0.28942   
## educationhigh.school 4.576e-02 1.584e-01 0.289 0.77268   
## educationilliterate 1.122e+00 1.185e+00 0.947 0.34381   
## educationprofessional.course 2.775e-01 1.769e-01 1.569 0.11665   
## educationuniversity.degree 3.244e-01 1.595e-01 2.033 0.04201 \*   
## educationunknown 1.154e-02 2.090e-01 0.055 0.95597   
## defaultunknown -4.439e-01 1.080e-01 -4.111 3.94e-05 \*\*\*  
## contacttelephone -4.461e-01 1.237e-01 -3.607 0.00031 \*\*\*  
## monthaug 1.044e+00 2.082e-01 5.013 5.36e-07 \*\*\*  
## monthdec -3.246e-01 3.671e-01 -0.884 0.37653   
## monthjul 1.539e-01 1.636e-01 0.941 0.34695   
## monthjun -7.547e-01 1.786e-01 -4.227 2.37e-05 \*\*\*  
## monthmar 1.745e+00 2.149e-01 8.119 4.71e-16 \*\*\*  
## monthmay -8.462e-01 1.271e-01 -6.658 2.78e-11 \*\*\*  
## monthnov -7.641e-01 1.806e-01 -4.231 2.33e-05 \*\*\*  
## monthoct 1.582e-01 2.370e-01 0.668 0.50443   
## monthsep 1.605e-01 2.401e-01 0.669 0.50380   
## duration 7.019e-03 1.785e-04 39.312 < 2e-16 \*\*\*  
## pdays -5.857e-04 3.934e-04 -1.489 0.13653   
## poutcomenonexistent 5.494e-01 1.083e-01 5.074 3.90e-07 \*\*\*  
## poutcomesuccess 1.326e+00 4.036e-01 3.286 0.00102 \*\*   
## emp.var.rate -2.135e+00 1.977e-01 -10.802 < 2e-16 \*\*\*  
## cons.price.idx 2.041e+00 1.992e-01 10.244 < 2e-16 \*\*\*  
## cons.conf.idx -1.630e-02 9.868e-03 -1.652 0.09860 .   
## euribor3m 7.392e-01 1.441e-01 5.129 2.92e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 11578.3 on 8351 degrees of freedom  
## Residual deviance: 5471.7 on 8314 degrees of freedom  
## AIC: 5547.7  
##   
## Number of Fisher Scoring iterations: 6

plot(step.model)



# Hoslem test for stepwise selection

hoslem.test(step.model$y, fitted(step.model), g=10)

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: step.model$y, fitted(step.model)  
## X-squared = 653.7, df = 8, p-value < 2.2e-16

# VIF and odd ratio for stepwise selection

exp(cbind("Odds ratio" = coef(step.model), confint.default(step.model, level = 0.95)))

## Odds ratio 2.5 % 97.5 %  
## (Intercept) 2.668268e-86 1.210164e-102 5.883218e-70  
## jobblue-collar 8.350991e-01 6.449250e-01 1.081351e+00  
## jobentrepreneur 7.708217e-01 5.205403e-01 1.141441e+00  
## jobhousemaid 1.313626e+00 7.996145e-01 2.158057e+00  
## jobmanagement 1.033614e+00 7.770579e-01 1.374875e+00  
## jobretired 1.570405e+00 1.135282e+00 2.172299e+00  
## jobself-employed 9.933138e-01 6.725784e-01 1.467000e+00  
## jobservices 9.618786e-01 7.227451e-01 1.280134e+00  
## jobstudent 1.816139e+00 1.214276e+00 2.716319e+00  
## jobtechnician 9.520631e-01 7.481501e-01 1.211554e+00  
## jobunemployed 1.185019e+00 7.755722e-01 1.810623e+00  
## jobunknown 1.804274e+00 8.285169e-01 3.929196e+00  
## educationbasic.6y 9.117115e-01 6.123044e-01 1.357524e+00  
## educationbasic.9y 8.444535e-01 6.176413e-01 1.154556e+00  
## educationhigh.school 1.046818e+00 7.674469e-01 1.427889e+00  
## educationilliterate 3.070515e+00 3.009443e-01 3.132825e+01  
## educationprofessional.course 1.319872e+00 9.331779e-01 1.866807e+00  
## educationuniversity.degree 1.383194e+00 1.011787e+00 1.890936e+00  
## educationunknown 1.011606e+00 6.716029e-01 1.523737e+00  
## defaultunknown 6.414984e-01 5.191220e-01 7.927236e-01  
## contacttelephone 6.401346e-01 5.023376e-01 8.157309e-01  
## monthaug 2.839647e+00 1.888179e+00 4.270568e+00  
## monthdec 7.227945e-01 3.519958e-01 1.484199e+00  
## monthjul 1.166353e+00 8.463745e-01 1.607303e+00  
## monthjun 4.701452e-01 3.313202e-01 6.671387e-01  
## monthmar 5.723441e+00 3.756191e+00 8.721010e+00  
## monthmay 4.290357e-01 3.344325e-01 5.504000e-01  
## monthnov 4.657653e-01 3.269229e-01 6.635735e-01  
## monthoct 1.171395e+00 7.361780e-01 1.863907e+00  
## monthsep 1.174108e+00 7.333949e-01 1.879655e+00  
## duration 1.007043e+00 1.006691e+00 1.007396e+00  
## pdays 9.994145e-01 9.986441e-01 1.000185e+00  
## poutcomenonexistent 1.732195e+00 1.400974e+00 2.141725e+00  
## poutcomesuccess 3.767021e+00 1.707740e+00 8.309488e+00  
## emp.var.rate 1.182158e-01 8.024472e-02 1.741545e-01  
## cons.price.idx 7.696366e+00 5.208465e+00 1.137265e+01  
## cons.conf.idx 9.838339e-01 9.649889e-01 1.003047e+00  
## euribor3m 2.094209e+00 1.578850e+00 2.777788e+00

vif(step.model)

## GVIF Df GVIF^(1/(2\*Df))  
## job 3.580295 11 1.059688  
## education 3.306079 7 1.089165  
## default 1.141892 1 1.068594  
## contact 2.286740 1 1.512197  
## month 19.660207 9 1.179957  
## duration 1.424434 1 1.193496  
## pdays 7.929602 1 2.815955  
## poutcome 9.090223 2 1.736375  
## emp.var.rate 95.312291 1 9.762801  
## cons.price.idx 12.864382 1 3.586695  
## cons.conf.idx 2.601219 1 1.612830  
## euribor3m 59.830673 1 7.735029

# Prediction with the stepwise selection after taking care of the high vif

logr<- glm(y ~ job + education + default + contact +duration + poutcome + pdays + campaign, family = binomial(link = "logit"),data = balanced\_train)  
  
# 1 way  
logr.probs<-predict(logr, newdata=balanced\_test)  
logr.pred<-rep("No",928)  
logr.pred[logr.probs>.5]="Yes"  
Truth<-balanced\_test[,21]  
Pred<-logr.pred  
ftable(addmargins(table(Pred,Truth)))

## Truth no yes Sum  
## Pred   
## No 428 139 567  
## Yes 42 319 361  
## Sum 470 458 928

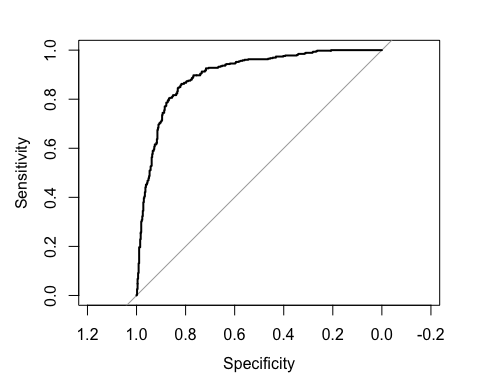
# 2 way  
pred = predict(logr, newdata=balanced\_test)  
accuracy <- table(pred, balanced\_test[,21])  
sum(diag(accuracy))/sum(accuracy)

## [1] 0.001077586

#confusionMatrix(factor(pred, levels = 1:928), factor(balanced\_test$y, levels = 1:928))

# ROC curve for the stepwise model

logr<-glm(y ~ job + education + default + contact +duration + poutcome + pdays + campaign, family = binomial(link = "logit"),data = balanced\_train)  
logr.probs<-predict(logr, newdata=balanced\_test)  
roccurve <- roc(balanced\_test$y ~ logr.probs)  
plot(roccurve)



ME<-c()  
FP<-c()  
FN<-c()  
index<-0:928/928  
for (i in 1:928){  
logr.pred<-rep("no",928)  
logr.pred[logr.probs>index[i]]="yes"  
logr.pred<-factor(logr.pred,levels=c("no","yes"))  
Truth<-balanced\_test[,21]  
Pred<-logr.pred  
x<-table(Pred,Truth)  
ME[i]<-(x[1,2]+x[2,1])/sum(x)  
FP[i]<-x[2,1]/(x[2,1]+x[1,1])  
FN[i]<-x[1,2]/(x[1,2]+x[2,2])  
}  
TP= 1-FN  
ord.ind<-order(FP,decreasing=F)  
FP[ord.ind]

## [1] 0.07021277 0.07021277 0.07021277 0.07234043 0.07234043 0.07234043  
## [7] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [13] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [19] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [25] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [31] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [37] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [43] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [49] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [55] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [61] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [67] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [73] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [79] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [85] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07446809  
## [91] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [97] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [103] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [109] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [115] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [121] 0.07446809 0.07446809 0.07446809 0.07659574 0.07659574 0.07659574  
## [127] 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574 0.07872340  
## [133] 0.07872340 0.07872340 0.07872340 0.07872340 0.07872340 0.07872340  
## [139] 0.07872340 0.07872340 0.07872340 0.07872340 0.07872340 0.07872340  
## [145] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [151] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [157] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [163] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [169] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [175] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [181] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [187] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [193] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [199] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [205] 0.08085106 0.08085106 0.08085106 0.08297872 0.08297872 0.08297872  
## [211] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [217] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [223] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [229] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [235] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [241] 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872 0.08297872  
## [247] 0.08297872 0.08297872 0.08297872 0.08510638 0.08510638 0.08510638  
## [253] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [259] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [265] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [271] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [277] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [283] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [289] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [295] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [301] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [307] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [313] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [319] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [325] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [331] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [337] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [343] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [349] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [355] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [361] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [367] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [373] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [379] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [385] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [391] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [397] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08723404  
## [403] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [409] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [415] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [421] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [427] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [433] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [439] 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [445] 0.08723404 0.08936170 0.08936170 0.08936170 0.08936170 0.08936170  
## [451] 0.08936170 0.08936170 0.08936170 0.08936170 0.08936170 0.08936170  
## [457] 0.08936170 0.08936170 0.08936170 0.08936170 0.08936170 0.08936170  
## [463] 0.08936170 0.08936170 0.08936170 0.08936170 0.09148936 0.09148936  
## [469] 0.09148936 0.09148936 0.09361702 0.09361702 0.09361702 0.09361702  
## [475] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [481] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [487] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [493] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [499] 0.09574468 0.09574468 0.09574468 0.09787234 0.09787234 0.09787234  
## [505] 0.09787234 0.09787234 0.09787234 0.09787234 0.09787234 0.09787234  
## [511] 0.09787234 0.09787234 0.09787234 0.09787234 0.09787234 0.09787234  
## [517] 0.09787234 0.09787234 0.09787234 0.09787234 0.09787234 0.09787234  
## [523] 0.09787234 0.09787234 0.10000000 0.10000000 0.10000000 0.10000000  
## [529] 0.10000000 0.10000000 0.10000000 0.10000000 0.10212766 0.10212766  
## [535] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [541] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [547] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10425532  
## [553] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [559] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [565] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [571] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [577] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [583] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [589] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [595] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [601] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [607] 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532 0.10425532  
## [613] 0.10425532 0.10425532 0.10425532 0.10425532 0.10638298 0.10638298  
## [619] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [625] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [631] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [637] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10851064  
## [643] 0.10851064 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830  
## [649] 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830  
## [655] 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830  
## [661] 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830 0.11063830  
## [667] 0.11063830 0.11063830 0.11276596 0.11276596 0.11276596 0.11276596  
## [673] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [679] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [685] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [691] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [697] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [703] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [709] 0.11276596 0.11276596 0.11276596 0.11489362 0.11489362 0.11489362  
## [715] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [721] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [727] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [733] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [739] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [745] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [751] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [757] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [763] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [769] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11702128  
## [775] 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128  
## [781] 0.11702128 0.11702128 0.11702128 0.11702128 0.11914894 0.11914894  
## [787] 0.11914894 0.11914894 0.11914894 0.11914894 0.11914894 0.11914894  
## [793] 0.11914894 0.11914894 0.11914894 0.11914894 0.11914894 0.11914894  
## [799] 0.11914894 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660  
## [805] 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660  
## [811] 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660  
## [817] 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426 0.12553191  
## [823] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [829] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [835] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [841] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [847] 0.12553191 0.12553191 0.12765957 0.12765957 0.12765957 0.12765957  
## [853] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [859] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [865] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [871] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [877] 0.12765957 0.12765957 0.12765957 0.12765957 0.12978723 0.13191489  
## [883] 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489  
## [889] 0.13191489 0.13191489 0.13191489 0.13191489 0.13404255 0.13404255  
## [895] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [901] 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255 0.13404255  
## [907] 0.13404255 0.13617021 0.13617021 0.13617021 0.13617021 0.13829787  
## [913] 0.13829787 0.14042553 0.14042553 0.14255319 0.14255319 0.14255319  
## [919] 0.14255319 0.14255319 0.14255319 0.14468085 0.14468085 0.14468085  
## [925] 0.14468085 0.14468085 0.14468085 0.14468085

TP[ord.ind]

## [1] 0.5960699 0.5960699 0.5960699 0.6069869 0.6069869 0.6069869 0.6069869  
## [8] 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869  
## [15] 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6048035  
## [22] 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035  
## [29] 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035  
## [36] 0.6048035 0.6048035 0.6026201 0.6026201 0.6004367 0.6004367 0.6004367  
## [43] 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367  
## [50] 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367  
## [57] 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367  
## [64] 0.6004367 0.6004367 0.6004367 0.5982533 0.5982533 0.5982533 0.5982533  
## [71] 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533  
## [78] 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533 0.5982533  
## [85] 0.5982533 0.5960699 0.5960699 0.5960699 0.5960699 0.6157205 0.6157205  
## [92] 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6135371 0.6135371  
## [99] 0.6135371 0.6135371 0.6135371 0.6135371 0.6113537 0.6113537 0.6113537  
## [106] 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6091703 0.6091703  
## [113] 0.6091703 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869  
## [120] 0.6069869 0.6069869 0.6069869 0.6069869 0.6157205 0.6157205 0.6157205  
## [127] 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205  
## [134] 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205  
## [141] 0.6157205 0.6157205 0.6157205 0.6157205 0.6222707 0.6222707 0.6222707  
## [148] 0.6222707 0.6222707 0.6222707 0.6222707 0.6222707 0.6222707 0.6222707  
## [155] 0.6222707 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873  
## [162] 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873  
## [169] 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873  
## [176] 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873  
## [183] 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873 0.6179039 0.6179039  
## [190] 0.6179039 0.6179039 0.6179039 0.6179039 0.6179039 0.6179039 0.6179039  
## [197] 0.6179039 0.6179039 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205  
## [204] 0.6157205 0.6157205 0.6157205 0.6157205 0.6375546 0.6375546 0.6375546  
## [211] 0.6375546 0.6375546 0.6375546 0.6353712 0.6353712 0.6353712 0.6331878  
## [218] 0.6331878 0.6331878 0.6331878 0.6331878 0.6331878 0.6310044 0.6310044  
## [225] 0.6310044 0.6288210 0.6288210 0.6288210 0.6288210 0.6288210 0.6288210  
## [232] 0.6288210 0.6288210 0.6288210 0.6266376 0.6266376 0.6266376 0.6266376  
## [239] 0.6266376 0.6244541 0.6244541 0.6244541 0.6244541 0.6244541 0.6244541  
## [246] 0.6244541 0.6222707 0.6222707 0.6222707 0.6746725 0.6724891 0.6703057  
## [253] 0.6703057 0.6703057 0.6703057 0.6703057 0.6703057 0.6703057 0.6681223  
## [260] 0.6681223 0.6681223 0.6681223 0.6681223 0.6681223 0.6659389 0.6659389  
## [267] 0.6659389 0.6659389 0.6659389 0.6659389 0.6659389 0.6659389 0.6637555  
## [274] 0.6637555 0.6615721 0.6615721 0.6615721 0.6615721 0.6615721 0.6615721  
## [281] 0.6615721 0.6615721 0.6615721 0.6615721 0.6615721 0.6615721 0.6593886  
## [288] 0.6593886 0.6593886 0.6593886 0.6593886 0.6593886 0.6572052 0.6572052  
## [295] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052  
## [302] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6550218 0.6550218  
## [309] 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218  
## [316] 0.6550218 0.6550218 0.6528384 0.6528384 0.6528384 0.6528384 0.6528384  
## [323] 0.6528384 0.6506550 0.6506550 0.6506550 0.6484716 0.6484716 0.6484716  
## [330] 0.6484716 0.6484716 0.6484716 0.6484716 0.6484716 0.6484716 0.6484716  
## [337] 0.6462882 0.6462882 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048  
## [344] 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048  
## [351] 0.6441048 0.6441048 0.6441048 0.6441048 0.6419214 0.6419214 0.6419214  
## [358] 0.6419214 0.6419214 0.6419214 0.6397380 0.6397380 0.6397380 0.6397380  
## [365] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [372] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [379] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [386] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [393] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6375546 0.6375546  
## [400] 0.6375546 0.6375546 0.6834061 0.6834061 0.6834061 0.6834061 0.6834061  
## [407] 0.6834061 0.6834061 0.6834061 0.6812227 0.6812227 0.6812227 0.6812227  
## [414] 0.6812227 0.6812227 0.6812227 0.6812227 0.6812227 0.6812227 0.6812227  
## [421] 0.6812227 0.6812227 0.6812227 0.6812227 0.6812227 0.6790393 0.6790393  
## [428] 0.6790393 0.6790393 0.6790393 0.6790393 0.6790393 0.6790393 0.6790393  
## [435] 0.6768559 0.6768559 0.6768559 0.6768559 0.6746725 0.6746725 0.6746725  
## [442] 0.6746725 0.6746725 0.6746725 0.6746725 0.6965066 0.6965066 0.6965066  
## [449] 0.6965066 0.6965066 0.6965066 0.6965066 0.6965066 0.6943231 0.6899563  
## [456] 0.6899563 0.6877729 0.6877729 0.6877729 0.6855895 0.6855895 0.6855895  
## [463] 0.6834061 0.6834061 0.6834061 0.6834061 0.6965066 0.6965066 0.6965066  
## [470] 0.6965066 0.7030568 0.7030568 0.7008734 0.7008734 0.7008734 0.7008734  
## [477] 0.7008734 0.7008734 0.7008734 0.6986900 0.7030568 0.7030568 0.7030568  
## [484] 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568  
## [491] 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568  
## [498] 0.7030568 0.7030568 0.7030568 0.7030568 0.7074236 0.7074236 0.7074236  
## [505] 0.7074236 0.7052402 0.7052402 0.7052402 0.7052402 0.7052402 0.7052402  
## [512] 0.7052402 0.7052402 0.7052402 0.7052402 0.7052402 0.7052402 0.7052402  
## [519] 0.7052402 0.7052402 0.7030568 0.7030568 0.7030568 0.7030568 0.7096070  
## [526] 0.7096070 0.7074236 0.7074236 0.7074236 0.7074236 0.7074236 0.7074236  
## [533] 0.7139738 0.7139738 0.7139738 0.7139738 0.7139738 0.7139738 0.7139738  
## [540] 0.7139738 0.7139738 0.7139738 0.7139738 0.7139738 0.7117904 0.7117904  
## [547] 0.7117904 0.7096070 0.7096070 0.7096070 0.7096070 0.7358079 0.7358079  
## [554] 0.7358079 0.7358079 0.7336245 0.7336245 0.7336245 0.7336245 0.7314410  
## [561] 0.7314410 0.7314410 0.7314410 0.7314410 0.7314410 0.7314410 0.7292576  
## [568] 0.7292576 0.7292576 0.7292576 0.7292576 0.7292576 0.7292576 0.7270742  
## [575] 0.7248908 0.7248908 0.7248908 0.7248908 0.7248908 0.7248908 0.7248908  
## [582] 0.7248908 0.7248908 0.7227074 0.7227074 0.7227074 0.7205240 0.7205240  
## [589] 0.7205240 0.7205240 0.7183406 0.7183406 0.7183406 0.7183406 0.7183406  
## [596] 0.7183406 0.7183406 0.7183406 0.7183406 0.7183406 0.7183406 0.7183406  
## [603] 0.7183406 0.7183406 0.7183406 0.7183406 0.7161572 0.7161572 0.7161572  
## [610] 0.7161572 0.7161572 0.7161572 0.7161572 0.7161572 0.7139738 0.7139738  
## [617] 0.7445415 0.7445415 0.7445415 0.7423581 0.7423581 0.7423581 0.7423581  
## [624] 0.7423581 0.7423581 0.7423581 0.7423581 0.7423581 0.7423581 0.7423581  
## [631] 0.7423581 0.7423581 0.7423581 0.7401747 0.7401747 0.7401747 0.7401747  
## [638] 0.7379913 0.7379913 0.7379913 0.7379913 0.7445415 0.7445415 0.7445415  
## [645] 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415  
## [652] 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415  
## [659] 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415  
## [666] 0.7445415 0.7445415 0.7445415 0.7554585 0.7554585 0.7554585 0.7532751  
## [673] 0.7532751 0.7532751 0.7510917 0.7510917 0.7510917 0.7510917 0.7510917  
## [680] 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083  
## [687] 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083 0.7467249  
## [694] 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249  
## [701] 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249  
## [708] 0.7445415 0.7445415 0.7445415 0.7445415 0.7707424 0.7707424 0.7707424  
## [715] 0.7707424 0.7707424 0.7707424 0.7707424 0.7685590 0.7685590 0.7685590  
## [722] 0.7663755 0.7663755 0.7663755 0.7663755 0.7663755 0.7641921 0.7641921  
## [729] 0.7641921 0.7641921 0.7641921 0.7641921 0.7620087 0.7620087 0.7620087  
## [736] 0.7620087 0.7620087 0.7620087 0.7620087 0.7620087 0.7620087 0.7620087  
## [743] 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253  
## [750] 0.7598253 0.7598253 0.7598253 0.7598253 0.7576419 0.7576419 0.7576419  
## [757] 0.7576419 0.7576419 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585  
## [764] 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585 0.7554585  
## [771] 0.7554585 0.7554585 0.7554585 0.7707424 0.7707424 0.7707424 0.7707424  
## [778] 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424  
## [785] 0.7729258 0.7729258 0.7729258 0.7729258 0.7707424 0.7707424 0.7707424  
## [792] 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424 0.7707424  
## [799] 0.7707424 0.7860262 0.7860262 0.7816594 0.7794760 0.7794760 0.7794760  
## [806] 0.7794760 0.7794760 0.7794760 0.7794760 0.7794760 0.7794760 0.7794760  
## [813] 0.7794760 0.7772926 0.7772926 0.7751092 0.7860262 0.7860262 0.7860262  
## [820] 0.7860262 0.7860262 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096  
## [827] 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096  
## [834] 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096  
## [841] 0.7882096 0.7882096 0.7860262 0.7860262 0.7860262 0.7860262 0.7860262  
## [848] 0.7860262 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598  
## [855] 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598  
## [862] 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598  
## [869] 0.7947598 0.7925764 0.7925764 0.7925764 0.7925764 0.7903930 0.7903930  
## [876] 0.7903930 0.7882096 0.7882096 0.7882096 0.7882096 0.7947598 0.8034934  
## [883] 0.7991266 0.7991266 0.7991266 0.7991266 0.7991266 0.7991266 0.7991266  
## [890] 0.7991266 0.7991266 0.7991266 0.8034934 0.8034934 0.8034934 0.8034934  
## [897] 0.8034934 0.8034934 0.8034934 0.8034934 0.8034934 0.8034934 0.8034934  
## [904] 0.8034934 0.8034934 0.8034934 0.8034934 0.8056769 0.8034934 0.8034934  
## [911] 0.8034934 0.8056769 0.8056769 0.8056769 0.8056769 0.8056769 0.8056769  
## [918] 0.8056769 0.8056769 0.8056769 0.8056769 0.8100437 0.8100437 0.8100437  
## [925] 0.8100437 0.8100437 0.8078603 0.8056769

simple\_auc <- function(TPR, FPR){  
 # inputs already sorted, best scores first   
 dFPR <- c(diff(FPR), 0)  
 dTPR <- c(diff(TPR), 0)  
 sum(TPR \* dFPR) + sum(dTPR \* dFPR)/2  
}  
  
paste("AUC=", simple\_auc(TP[ord.ind],FP[ord.ind]),sep=" ")

## [1] "AUC= 0.0540950478491127"

# Backward selection

model.null<-glm(y ~ ., data=balanced\_train,family = binomial(link="logit"))  
step(model.null,  
 scope = list(upper=model.main),  
 direction="backward",  
 test="Chisq",  
 data=balanced\_train)

## Start: AIC=5561.04  
## y ~ age + job + marital + education + default + housing + loan +   
## contact + month + day\_of\_week + duration + campaign + pdays +   
## previous + poutcome + emp.var.rate + cons.price.idx + cons.conf.idx +   
## euribor3m + nr.employed  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## - day\_of\_week 4 5462.0 5558.0 4.9 0.2933957   
## - age 1 5457.0 5559.0 0.0 0.9444433   
## - loan 1 5457.2 5559.2 0.2 0.6609909   
## - nr.employed 1 5457.3 5559.3 0.3 0.5891663   
## - cons.conf.idx 1 5457.5 5559.5 0.4 0.5079955   
## - campaign 1 5457.9 5559.9 0.8 0.3618930   
## - marital 3 5462.1 5560.1 5.1 0.1679739   
## - housing 1 5458.4 5560.4 1.3 0.2459109   
## - previous 1 5458.4 5560.4 1.4 0.2426784   
## <none> 5457.0 5561.0   
## - pdays 1 5461.0 5563.0 3.9 0.0479427 \*   
## - job 11 5481.4 5563.4 24.4 0.0112254 \*   
## - education 7 5474.0 5564.0 17.0 0.0176201 \*   
## - poutcome 2 5466.6 5566.6 9.6 0.0083646 \*\*   
## - euribor3m 1 5465.2 5567.2 8.1 0.0043693 \*\*   
## - contact 1 5469.9 5571.9 12.9 0.0003338 \*\*\*  
## - default 1 5472.4 5574.4 15.4 8.894e-05 \*\*\*  
## - cons.price.idx 1 5486.2 5588.2 29.2 6.592e-08 \*\*\*  
## - emp.var.rate 1 5543.9 5645.9 86.9 < 2.2e-16 \*\*\*  
## - month 9 5783.4 5869.4 326.3 < 2.2e-16 \*\*\*  
## - duration 1 8952.9 9054.9 3495.9 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5557.98  
## y ~ age + job + marital + education + default + housing + loan +   
## contact + month + duration + campaign + pdays + previous +   
## poutcome + emp.var.rate + cons.price.idx + cons.conf.idx +   
## euribor3m + nr.employed  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## - age 1 5462.0 5556.0 0.0 0.9639224   
## - loan 1 5462.2 5556.2 0.2 0.6582983   
## - nr.employed 1 5462.2 5556.2 0.3 0.6094426   
## - cons.conf.idx 1 5462.5 5556.5 0.5 0.4765937   
## - marital 3 5466.8 5556.8 4.8 0.1859371   
## - campaign 1 5462.9 5556.9 0.9 0.3420844   
## - previous 1 5463.4 5557.4 1.4 0.2389030   
## - housing 1 5463.4 5557.4 1.5 0.2267266   
## <none> 5462.0 5558.0   
## - pdays 1 5465.9 5559.9 3.9 0.0486025 \*   
## - education 7 5478.7 5560.7 16.7 0.0193397 \*   
## - job 11 5486.9 5560.9 24.9 0.0093865 \*\*   
## - poutcome 2 5471.6 5563.6 9.6 0.0082370 \*\*   
## - euribor3m 1 5470.3 5564.3 8.4 0.0038464 \*\*   
## - contact 1 5474.6 5568.6 12.6 0.0003859 \*\*\*  
## - default 1 5477.8 5571.8 15.8 7.140e-05 \*\*\*  
## - cons.price.idx 1 5490.9 5584.9 28.9 7.585e-08 \*\*\*  
## - emp.var.rate 1 5548.6 5642.6 86.6 < 2.2e-16 \*\*\*  
## - month 9 5785.6 5863.6 323.7 < 2.2e-16 \*\*\*  
## - duration 1 8964.9 9058.9 3502.9 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5555.98  
## y ~ job + marital + education + default + housing + loan + contact +   
## month + duration + campaign + pdays + previous + poutcome +   
## emp.var.rate + cons.price.idx + cons.conf.idx + euribor3m +   
## nr.employed  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## - loan 1 5462.2 5554.2 0.2 0.6580533   
## - nr.employed 1 5462.2 5554.2 0.3 0.6081041   
## - cons.conf.idx 1 5462.5 5554.5 0.5 0.4773046   
## - campaign 1 5462.9 5554.9 0.9 0.3424461   
## - marital 3 5467.2 5555.2 5.3 0.1538800   
## - previous 1 5463.4 5555.4 1.4 0.2388988   
## - housing 1 5463.5 5555.5 1.5 0.2255051   
## <none> 5462.0 5556.0   
## - pdays 1 5465.9 5557.9 3.9 0.0486137 \*   
## - education 7 5478.7 5558.7 16.8 0.0190733 \*   
## - job 11 5489.2 5561.2 27.2 0.0042730 \*\*   
## - poutcome 2 5471.6 5561.6 9.6 0.0082213 \*\*   
## - euribor3m 1 5470.3 5562.3 8.4 0.0038470 \*\*   
## - contact 1 5474.6 5566.6 12.6 0.0003779 \*\*\*  
## - default 1 5477.9 5569.9 15.9 6.553e-05 \*\*\*  
## - cons.price.idx 1 5491.0 5583.0 29.0 7.283e-08 \*\*\*  
## - emp.var.rate 1 5548.6 5640.6 86.7 < 2.2e-16 \*\*\*  
## - month 9 5785.9 5861.9 323.9 < 2.2e-16 \*\*\*  
## - duration 1 8965.4 9057.4 3503.4 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5554.18  
## y ~ job + marital + education + default + housing + contact +   
## month + duration + campaign + pdays + previous + poutcome +   
## emp.var.rate + cons.price.idx + cons.conf.idx + euribor3m +   
## nr.employed  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## - housing 2 5463.7 5551.7 1.5 0.4607213   
## - nr.employed 1 5462.4 5552.4 0.3 0.6068850   
## - cons.conf.idx 1 5462.7 5552.7 0.5 0.4736141   
## - campaign 1 5463.1 5553.1 0.9 0.3408507   
## - marital 3 5467.5 5553.5 5.3 0.1519424   
## - previous 1 5463.6 5553.6 1.4 0.2386093   
## <none> 5462.2 5554.2   
## - pdays 1 5466.0 5556.0 3.9 0.0497082 \*   
## - education 7 5478.9 5556.9 16.7 0.0194614 \*   
## - job 11 5489.3 5559.3 27.1 0.0043770 \*\*   
## - poutcome 2 5471.8 5559.8 9.6 0.0080967 \*\*   
## - euribor3m 1 5470.5 5560.5 8.4 0.0038294 \*\*   
## - contact 1 5474.8 5564.8 12.6 0.0003827 \*\*\*  
## - default 1 5478.1 5568.1 15.9 6.540e-05 \*\*\*  
## - cons.price.idx 1 5491.2 5581.2 29.0 7.112e-08 \*\*\*  
## - emp.var.rate 1 5549.0 5639.0 86.8 < 2.2e-16 \*\*\*  
## - month 9 5787.1 5861.1 324.9 < 2.2e-16 \*\*\*  
## - duration 1 8965.4 9055.4 3503.2 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5551.73  
## y ~ job + marital + education + default + contact + month + duration +   
## campaign + pdays + previous + poutcome + emp.var.rate + cons.price.idx +   
## cons.conf.idx + euribor3m + nr.employed  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## - nr.employed 1 5463.9 5549.9 0.2 0.6391773   
## - cons.conf.idx 1 5464.3 5550.3 0.6 0.4359112   
## - campaign 1 5464.7 5550.7 0.9 0.3306631   
## - marital 3 5469.0 5551.0 5.3 0.1529445   
## - previous 1 5465.1 5551.1 1.4 0.2442928   
## <none> 5463.7 5551.7   
## - pdays 1 5467.5 5553.5 3.8 0.0516414 .   
## - education 7 5480.6 5554.6 16.9 0.0183551 \*   
## - job 11 5490.9 5556.9 27.2 0.0043393 \*\*   
## - poutcome 2 5473.5 5557.5 9.8 0.0074670 \*\*   
## - euribor3m 1 5472.4 5558.4 8.6 0.0032790 \*\*   
## - contact 1 5476.4 5562.4 12.7 0.0003706 \*\*\*  
## - default 1 5479.8 5565.8 16.0 6.248e-05 \*\*\*  
## - cons.price.idx 1 5492.5 5578.5 28.7 8.310e-08 \*\*\*  
## - emp.var.rate 1 5550.3 5636.3 86.6 < 2.2e-16 \*\*\*  
## - month 9 5788.8 5858.8 325.1 < 2.2e-16 \*\*\*  
## - duration 1 8965.5 9051.5 3501.8 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5549.95  
## y ~ job + marital + education + default + contact + month + duration +   
## campaign + pdays + previous + poutcome + emp.var.rate + cons.price.idx +   
## cons.conf.idx + euribor3m  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## - campaign 1 5464.9 5548.9 1.0 0.3257905   
## - marital 3 5469.2 5549.2 5.3 0.1526387   
## - previous 1 5465.3 5549.3 1.4 0.2431010   
## <none> 5463.9 5549.9   
## - cons.conf.idx 1 5466.5 5550.5 2.6 0.1080447   
## - pdays 1 5467.7 5551.7 3.8 0.0518058 .   
## - education 7 5480.7 5552.7 16.8 0.0190678 \*   
## - job 11 5491.0 5555.0 27.1 0.0044736 \*\*   
## - poutcome 2 5473.7 5555.7 9.7 0.0077116 \*\*   
## - contact 1 5476.6 5560.6 12.6 0.0003791 \*\*\*  
## - default 1 5479.9 5563.9 16.0 6.347e-05 \*\*\*  
## - euribor3m 1 5490.5 5574.5 26.6 2.524e-07 \*\*\*  
## - cons.price.idx 1 5580.0 5664.0 116.0 < 2.2e-16 \*\*\*  
## - emp.var.rate 1 5584.8 5668.8 120.8 < 2.2e-16 \*\*\*  
## - month 9 5794.6 5862.6 330.7 < 2.2e-16 \*\*\*  
## - duration 1 8966.9 9050.9 3503.0 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5548.91  
## y ~ job + marital + education + default + contact + month + duration +   
## pdays + previous + poutcome + emp.var.rate + cons.price.idx +   
## cons.conf.idx + euribor3m  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## - marital 3 5470.2 5548.2 5.3 0.1524055   
## - previous 1 5466.3 5548.3 1.4 0.2441823   
## <none> 5464.9 5548.9   
## - cons.conf.idx 1 5467.5 5549.5 2.6 0.1058780   
## - pdays 1 5468.6 5550.6 3.7 0.0533225 .   
## - education 7 5481.6 5551.6 16.6 0.0198246 \*   
## - job 11 5491.8 5553.8 26.9 0.0047708 \*\*   
## - poutcome 2 5474.7 5554.7 9.8 0.0075337 \*\*   
## - contact 1 5478.2 5560.2 13.2 0.0002731 \*\*\*  
## - default 1 5480.9 5562.9 15.9 6.506e-05 \*\*\*  
## - euribor3m 1 5492.1 5574.1 27.2 1.838e-07 \*\*\*  
## - cons.price.idx 1 5582.5 5664.5 117.6 < 2.2e-16 \*\*\*  
## - emp.var.rate 1 5587.6 5669.6 122.7 < 2.2e-16 \*\*\*  
## - month 9 5796.0 5862.0 331.0 < 2.2e-16 \*\*\*  
## - duration 1 8974.1 9056.1 3509.2 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5548.19  
## y ~ job + education + default + contact + month + duration +   
## pdays + previous + poutcome + emp.var.rate + cons.price.idx +   
## cons.conf.idx + euribor3m  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## - previous 1 5471.7 5547.7 1.5 0.2211327   
## <none> 5470.2 5548.2   
## - cons.conf.idx 1 5472.8 5548.8 2.6 0.1038173   
## - pdays 1 5473.9 5549.9 3.8 0.0527801 .   
## - education 7 5488.9 5552.9 18.7 0.0092919 \*\*   
## - poutcome 2 5479.9 5553.9 9.8 0.0076179 \*\*   
## - job 11 5499.8 5555.8 29.6 0.0018527 \*\*   
## - contact 1 5483.7 5559.7 13.5 0.0002436 \*\*\*  
## - default 1 5487.4 5563.4 17.2 3.406e-05 \*\*\*  
## - euribor3m 1 5496.7 5572.7 26.5 2.611e-07 \*\*\*  
## - cons.price.idx 1 5586.8 5662.8 116.7 < 2.2e-16 \*\*\*  
## - emp.var.rate 1 5591.9 5667.9 121.7 < 2.2e-16 \*\*\*  
## - month 9 5801.8 5861.8 331.6 < 2.2e-16 \*\*\*  
## - duration 1 8977.7 9053.7 3507.5 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=5547.69  
## y ~ job + education + default + contact + month + duration +   
## pdays + poutcome + emp.var.rate + cons.price.idx + cons.conf.idx +   
## euribor3m  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 5471.7 5547.7   
## - pdays 1 5474.1 5548.1 2.4 0.1205407   
## - cons.conf.idx 1 5474.4 5548.4 2.7 0.0992358 .   
## - education 7 5490.3 5552.3 18.6 0.0094157 \*\*   
## - job 11 5501.2 5555.2 29.6 0.0018573 \*\*   
## - contact 1 5484.7 5558.7 13.0 0.0003098 \*\*\*  
## - default 1 5488.9 5562.9 17.2 3.372e-05 \*\*\*  
## - euribor3m 1 5498.2 5572.2 26.5 2.650e-07 \*\*\*  
## - poutcome 2 5506.2 5578.2 34.6 3.140e-08 \*\*\*  
## - cons.price.idx 1 5587.0 5661.0 115.3 < 2.2e-16 \*\*\*  
## - emp.var.rate 1 5592.4 5666.4 120.7 < 2.2e-16 \*\*\*  
## - month 9 5802.1 5860.1 330.4 < 2.2e-16 \*\*\*  
## - duration 1 8979.2 9053.2 3507.5 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Call: glm(formula = y ~ job + education + default + contact + month +   
## duration + pdays + poutcome + emp.var.rate + cons.price.idx +   
## cons.conf.idx + euribor3m, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Coefficients:  
## (Intercept) jobblue-collar   
## -1.970e+02 -1.802e-01   
## jobentrepreneur jobhousemaid   
## -2.603e-01 2.728e-01   
## jobmanagement jobretired   
## 3.306e-02 4.513e-01   
## jobself-employed jobservices   
## -6.709e-03 -3.887e-02   
## jobstudent jobtechnician   
## 5.967e-01 -4.912e-02   
## jobunemployed jobunknown   
## 1.698e-01 5.902e-01   
## educationbasic.6y educationbasic.9y   
## -9.243e-02 -1.691e-01   
## educationhigh.school educationilliterate   
## 4.576e-02 1.122e+00   
## educationprofessional.course educationuniversity.degree   
## 2.775e-01 3.244e-01   
## educationunknown defaultunknown   
## 1.154e-02 -4.439e-01   
## contacttelephone monthaug   
## -4.461e-01 1.044e+00   
## monthdec monthjul   
## -3.246e-01 1.539e-01   
## monthjun monthmar   
## -7.547e-01 1.745e+00   
## monthmay monthnov   
## -8.462e-01 -7.641e-01   
## monthoct monthsep   
## 1.582e-01 1.605e-01   
## duration pdays   
## 7.019e-03 -5.857e-04   
## poutcomenonexistent poutcomesuccess   
## 5.494e-01 1.326e+00   
## emp.var.rate cons.price.idx   
## -2.135e+00 2.041e+00   
## cons.conf.idx euribor3m   
## -1.630e-02 7.392e-01   
##   
## Degrees of Freedom: 8351 Total (i.e. Null); 8314 Residual  
## Null Deviance: 11580   
## Residual Deviance: 5472 AIC: 5548

# Backward final model

backward<- glm(formula = y ~ job + education + default + contact + month +   
 duration + campaign + pdays + poutcome + emp.var.rate + cons.price.idx +   
 euribor3m + nr.employed, family = binomial(link = "logit"),   
 data = balanced\_train)  
backward

##   
## Call: glm(formula = y ~ job + education + default + contact + month +   
## duration + campaign + pdays + poutcome + emp.var.rate + cons.price.idx +   
## euribor3m + nr.employed, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Coefficients:  
## (Intercept) jobblue-collar   
## -2.559e+02 -1.803e-01   
## jobentrepreneur jobhousemaid   
## -2.537e-01 2.690e-01   
## jobmanagement jobretired   
## 3.760e-02 4.552e-01   
## jobself-employed jobservices   
## -6.534e-03 -3.393e-02   
## jobstudent jobtechnician   
## 5.997e-01 -4.465e-02   
## jobunemployed jobunknown   
## 1.726e-01 6.034e-01   
## educationbasic.6y educationbasic.9y   
## -9.176e-02 -1.690e-01   
## educationhigh.school educationilliterate   
## 4.356e-02 1.112e+00   
## educationprofessional.course educationuniversity.degree   
## 2.804e-01 3.289e-01   
## educationunknown defaultunknown   
## 8.599e-03 -4.477e-01   
## contacttelephone monthaug   
## -4.875e-01 1.065e+00   
## monthdec monthjul   
## -2.393e-01 1.453e-01   
## monthjun monthmar   
## -8.682e-01 1.872e+00   
## monthmay monthnov   
## -8.061e-01 -6.813e-01   
## monthoct monthsep   
## 2.523e-01 3.063e-01   
## duration campaign   
## 7.011e-03 -1.780e-02   
## pdays poutcomenonexistent   
## -5.867e-04 5.591e-01   
## poutcomesuccess emp.var.rate   
## 1.319e+00 -2.212e+00   
## cons.price.idx euribor3m   
## 2.385e+00 5.190e-01   
## nr.employed   
## 5.439e-03   
##   
## Degrees of Freedom: 8351 Total (i.e. Null); 8313 Residual  
## Null Deviance: 11580   
## Residual Deviance: 5471 AIC: 5549

# VIF and odd ratio for backward selection

summary(backward)

##   
## Call:  
## glm(formula = y ~ job + education + default + contact + month +   
## duration + campaign + pdays + poutcome + emp.var.rate + cons.price.idx +   
## euribor3m + nr.employed, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -6.5914 -0.3807 0.0005 0.4878 2.8595   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.559e+02 4.944e+01 -5.176 2.27e-07 \*\*\*  
## jobblue-collar -1.803e-01 1.319e-01 -1.368 0.171424   
## jobentrepreneur -2.537e-01 2.004e-01 -1.266 0.205482   
## jobhousemaid 2.690e-01 2.537e-01 1.060 0.289049   
## jobmanagement 3.760e-02 1.456e-01 0.258 0.796248   
## jobretired 4.552e-01 1.658e-01 2.745 0.006048 \*\*   
## jobself-employed -6.534e-03 1.989e-01 -0.033 0.973789   
## jobservices -3.393e-02 1.457e-01 -0.233 0.815883   
## jobstudent 5.997e-01 2.052e-01 2.922 0.003477 \*\*   
## jobtechnician -4.465e-02 1.230e-01 -0.363 0.716614   
## jobunemployed 1.726e-01 2.164e-01 0.798 0.424937   
## jobunknown 6.034e-01 3.977e-01 1.517 0.129256   
## educationbasic.6y -9.176e-02 2.031e-01 -0.452 0.651408   
## educationbasic.9y -1.690e-01 1.596e-01 -1.059 0.289537   
## educationhigh.school 4.356e-02 1.583e-01 0.275 0.783240   
## educationilliterate 1.112e+00 1.185e+00 0.938 0.348157   
## educationprofessional.course 2.804e-01 1.768e-01 1.586 0.112766   
## educationuniversity.degree 3.289e-01 1.595e-01 2.062 0.039178 \*   
## educationunknown 8.599e-03 2.091e-01 0.041 0.967197   
## defaultunknown -4.477e-01 1.080e-01 -4.146 3.39e-05 \*\*\*  
## contacttelephone -4.875e-01 1.206e-01 -4.041 5.31e-05 \*\*\*  
## monthaug 1.065e+00 2.243e-01 4.747 2.06e-06 \*\*\*  
## monthdec -2.393e-01 3.837e-01 -0.624 0.532929   
## monthjul 1.453e-01 1.620e-01 0.897 0.369589   
## monthjun -8.682e-01 1.970e-01 -4.406 1.05e-05 \*\*\*  
## monthmar 1.872e+00 2.429e-01 7.706 1.29e-14 \*\*\*  
## monthmay -8.061e-01 1.361e-01 -5.923 3.17e-09 \*\*\*  
## monthnov -6.813e-01 1.977e-01 -3.446 0.000570 \*\*\*  
## monthoct 2.523e-01 2.585e-01 0.976 0.329001   
## monthsep 3.063e-01 2.999e-01 1.021 0.307081   
## duration 7.011e-03 1.784e-04 39.300 < 2e-16 \*\*\*  
## campaign -1.780e-02 1.855e-02 -0.959 0.337415   
## pdays -5.867e-04 3.927e-04 -1.494 0.135133   
## poutcomenonexistent 5.591e-01 1.082e-01 5.166 2.39e-07 \*\*\*  
## poutcomesuccess 1.319e+00 4.030e-01 3.273 0.001066 \*\*   
## emp.var.rate -2.212e+00 2.329e-01 -9.498 < 2e-16 \*\*\*  
## cons.price.idx 2.385e+00 3.521e-01 6.775 1.25e-11 \*\*\*  
## euribor3m 5.190e-01 1.550e-01 3.348 0.000814 \*\*\*  
## nr.employed 5.439e-03 3.584e-03 1.518 0.129081   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 11578.3 on 8351 degrees of freedom  
## Residual deviance: 5471.1 on 8313 degrees of freedom  
## AIC: 5549.1  
##   
## Number of Fisher Scoring iterations: 6

exp(cbind("Odds ratio" = coef(forward), confint.default(backward, level = 0.95)))

## Odds ratio 2.5 % 97.5 %  
## (Intercept) 3.193196e-112 6.210582e-154 9.043079e-70  
## jobblue-collar 1.006998e+00 6.448087e-01 1.081239e+00  
## jobentrepreneur 1.005137e+00 5.238887e-01 1.149179e+00  
## jobhousemaid 2.994208e+00 7.959044e-01 2.151678e+00  
## jobmanagement 7.543961e-01 7.805201e-01 1.381252e+00  
## jobretired 1.122890e+00 1.139056e+00 2.181920e+00  
## jobself-employed 4.110504e-01 6.728073e-01 1.467013e+00  
## jobservices 6.539026e+00 7.264776e-01 1.286191e+00  
## jobstudent 4.260822e-01 1.218304e+00 2.723658e+00  
## jobtechnician 5.111951e-01 7.514430e-01 1.217079e+00  
## jobunemployed 1.262296e+00 7.776961e-01 1.816061e+00  
## jobunknown 1.322406e+00 8.384771e-01 3.986650e+00  
## educationbasic.6y 1.502184e+00 6.127269e-01 1.358403e+00  
## educationbasic.9y 3.277368e+00 6.176547e-01 1.154613e+00  
## educationhigh.school 1.085939e-01 7.658288e-01 1.424641e+00  
## educationilliterate 1.121523e+01 2.978632e-01 3.103544e+01  
## educationprofessional.course 6.543337e-01 9.359924e-01 1.871986e+00  
## educationuniversity.degree 7.697325e-01 1.016458e+00 1.899333e+00  
## educationunknown 1.149616e+00 6.694945e-01 1.519574e+00  
## defaultunknown 1.089866e+00 5.171899e-01 7.897559e-01  
## contacttelephone 1.368258e+00 4.848321e-01 7.779508e-01  
## monthaug 1.021835e+00 1.868653e+00 4.502104e+00  
## monthdec 8.357612e-01 3.710524e-01 1.670040e+00  
## monthjul 1.558186e+00 8.418670e-01 1.588505e+00  
## monthjun 9.773786e-01 2.852511e-01 6.175665e-01  
## monthmar 1.061770e+00 4.037578e+00 1.046132e+01  
## monthmay 1.561169e+00 3.420462e-01 5.831505e-01  
## monthnov 1.686342e+00 3.433845e-01 7.454561e-01  
## monthoct 6.068817e-01 7.754385e-01 2.136130e+00  
## monthsep 9.792519e-01 7.546721e-01 2.445029e+00  
## duration 9.227310e-01 1.006684e+00 1.007388e+00  
## campaign 9.698565e-01 9.472839e-01 1.018738e+00  
## pdays 1.152109e+00 9.986445e-01 1.000183e+00  
## poutcomenonexistent 9.992245e-01 1.414782e+00 2.162253e+00  
## poutcomesuccess 6.242900e-01 1.697201e+00 8.237617e+00  
## emp.var.rate 8.579462e-01 6.932119e-02 1.727520e-01  
## cons.price.idx 3.193196e-112 5.447411e+00 2.165396e+01  
## euribor3m 1.006998e+00 1.240034e+00 2.276783e+00  
## nr.employed 1.005137e+00 9.984164e-01 1.012541e+00

vif(backward)

## GVIF Df GVIF^(1/(2\*Df))  
## job 3.591965 11 1.059845  
## education 3.315465 7 1.089386  
## default 1.141303 1 1.068318  
## contact 2.178743 1 1.476057  
## month 41.415360 9 1.229822  
## duration 1.421339 1 1.192199  
## campaign 1.054735 1 1.027003  
## pdays 7.898795 1 2.810480  
## poutcome 9.038089 2 1.733880  
## emp.var.rate 132.220385 1 11.498712  
## cons.price.idx 40.356353 1 6.352665  
## euribor3m 69.117795 1 8.313711  
## nr.employed 73.734304 1 8.586868

# After take of the high vif for backward selection

backwardfinal<- glm(formula = y ~ job + education + default + contact +duration + previous + pdays + campaign, family = binomial(link = "logit"),data = balanced\_train)  
backwardfinal

##   
## Call: glm(formula = y ~ job + education + default + contact + duration +   
## previous + pdays + campaign, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Coefficients:  
## (Intercept) jobblue-collar   
## 0.758488 -0.399162   
## jobentrepreneur jobhousemaid   
## -0.511366 0.091276   
## jobmanagement jobretired   
## -0.106068 1.005992   
## jobself-employed jobservices   
## -0.070597 -0.273144   
## jobstudent jobtechnician   
## 1.387810 -0.239660   
## jobunemployed jobunknown   
## 0.428734 0.794979   
## educationbasic.6y educationbasic.9y   
## -0.365705 -0.451945   
## educationhigh.school educationilliterate   
## -0.133472 1.380249   
## educationprofessional.course educationuniversity.degree   
## 0.164268 0.238962   
## educationunknown defaultunknown   
## 0.131748 -1.154987   
## contacttelephone duration   
## -1.100837 0.005714   
## previous pdays   
## 0.215515 -0.002309   
## campaign   
## -0.097103   
##   
## Degrees of Freedom: 8351 Total (i.e. Null); 8327 Residual  
## Null Deviance: 11580   
## Residual Deviance: 7085 AIC: 7135

summary(backwardfinal)

##   
## Call:  
## glm(formula = y ~ job + education + default + contact + duration +   
## previous + pdays + campaign, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -5.6182 -0.6726 0.0018 0.6228 2.6046   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.7584877 0.2169187 3.497 0.000471 \*\*\*  
## jobblue-collar -0.3991621 0.1147748 -3.478 0.000506 \*\*\*  
## jobentrepreneur -0.5113664 0.1767617 -2.893 0.003816 \*\*   
## jobhousemaid 0.0912759 0.2070616 0.441 0.659347   
## jobmanagement -0.1060683 0.1252734 -0.847 0.397166   
## jobretired 1.0059923 0.1404834 7.161 8.01e-13 \*\*\*  
## jobself-employed -0.0705975 0.1698743 -0.416 0.677713   
## jobservices -0.2731439 0.1254690 -2.177 0.029482 \*   
## jobstudent 1.3878104 0.1835617 7.560 4.02e-14 \*\*\*  
## jobtechnician -0.2396598 0.1032433 -2.321 0.020270 \*   
## jobunemployed 0.4287341 0.1822792 2.352 0.018669 \*   
## jobunknown 0.7949786 0.3251431 2.445 0.014485 \*   
## educationbasic.6y -0.3657051 0.1776217 -2.059 0.039504 \*   
## educationbasic.9y -0.4519454 0.1392102 -3.246 0.001168 \*\*   
## educationhigh.school -0.1334720 0.1365406 -0.978 0.328309   
## educationilliterate 1.3802491 1.1239962 1.228 0.219453   
## educationprofessional.course 0.1642676 0.1507088 1.090 0.275728   
## educationuniversity.degree 0.2389617 0.1363962 1.752 0.079779 .   
## educationunknown 0.1317485 0.1799180 0.732 0.464004   
## defaultunknown -1.1549868 0.0951570 -12.138 < 2e-16 \*\*\*  
## contacttelephone -1.1008370 0.0733572 -15.007 < 2e-16 \*\*\*  
## duration 0.0057135 0.0001467 38.934 < 2e-16 \*\*\*  
## previous 0.2155153 0.0662610 3.253 0.001144 \*\*   
## pdays -0.0023093 0.0001606 -14.379 < 2e-16 \*\*\*  
## campaign -0.0971026 0.0165710 -5.860 4.63e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 11578.3 on 8351 degrees of freedom  
## Residual deviance: 7085.2 on 8327 degrees of freedom  
## AIC: 7135.2  
##   
## Number of Fisher Scoring iterations: 6

exp(cbind("Odds ratio" = coef(forward), confint.default(backward, level = 0.95)))

## Odds ratio 2.5 % 97.5 %  
## (Intercept) 3.193196e-112 6.210582e-154 9.043079e-70  
## jobblue-collar 1.006998e+00 6.448087e-01 1.081239e+00  
## jobentrepreneur 1.005137e+00 5.238887e-01 1.149179e+00  
## jobhousemaid 2.994208e+00 7.959044e-01 2.151678e+00  
## jobmanagement 7.543961e-01 7.805201e-01 1.381252e+00  
## jobretired 1.122890e+00 1.139056e+00 2.181920e+00  
## jobself-employed 4.110504e-01 6.728073e-01 1.467013e+00  
## jobservices 6.539026e+00 7.264776e-01 1.286191e+00  
## jobstudent 4.260822e-01 1.218304e+00 2.723658e+00  
## jobtechnician 5.111951e-01 7.514430e-01 1.217079e+00  
## jobunemployed 1.262296e+00 7.776961e-01 1.816061e+00  
## jobunknown 1.322406e+00 8.384771e-01 3.986650e+00  
## educationbasic.6y 1.502184e+00 6.127269e-01 1.358403e+00  
## educationbasic.9y 3.277368e+00 6.176547e-01 1.154613e+00  
## educationhigh.school 1.085939e-01 7.658288e-01 1.424641e+00  
## educationilliterate 1.121523e+01 2.978632e-01 3.103544e+01  
## educationprofessional.course 6.543337e-01 9.359924e-01 1.871986e+00  
## educationuniversity.degree 7.697325e-01 1.016458e+00 1.899333e+00  
## educationunknown 1.149616e+00 6.694945e-01 1.519574e+00  
## defaultunknown 1.089866e+00 5.171899e-01 7.897559e-01  
## contacttelephone 1.368258e+00 4.848321e-01 7.779508e-01  
## monthaug 1.021835e+00 1.868653e+00 4.502104e+00  
## monthdec 8.357612e-01 3.710524e-01 1.670040e+00  
## monthjul 1.558186e+00 8.418670e-01 1.588505e+00  
## monthjun 9.773786e-01 2.852511e-01 6.175665e-01  
## monthmar 1.061770e+00 4.037578e+00 1.046132e+01  
## monthmay 1.561169e+00 3.420462e-01 5.831505e-01  
## monthnov 1.686342e+00 3.433845e-01 7.454561e-01  
## monthoct 6.068817e-01 7.754385e-01 2.136130e+00  
## monthsep 9.792519e-01 7.546721e-01 2.445029e+00  
## duration 9.227310e-01 1.006684e+00 1.007388e+00  
## campaign 9.698565e-01 9.472839e-01 1.018738e+00  
## pdays 1.152109e+00 9.986445e-01 1.000183e+00  
## poutcomenonexistent 9.992245e-01 1.414782e+00 2.162253e+00  
## poutcomesuccess 6.242900e-01 1.697201e+00 8.237617e+00  
## emp.var.rate 8.579462e-01 6.932119e-02 1.727520e-01  
## cons.price.idx 3.193196e-112 5.447411e+00 2.165396e+01  
## euribor3m 1.006998e+00 1.240034e+00 2.276783e+00  
## nr.employed 1.005137e+00 9.984164e-01 1.012541e+00

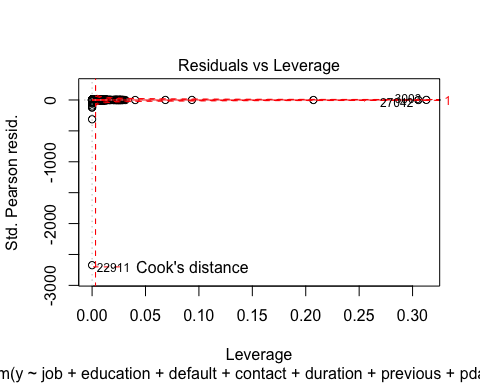
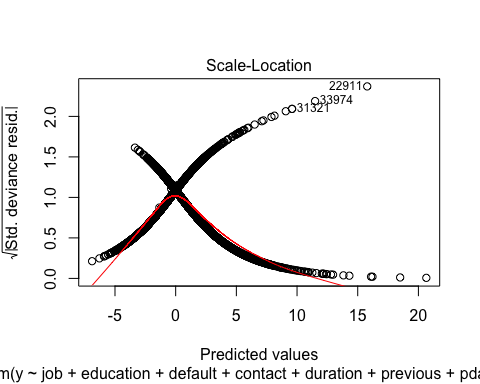
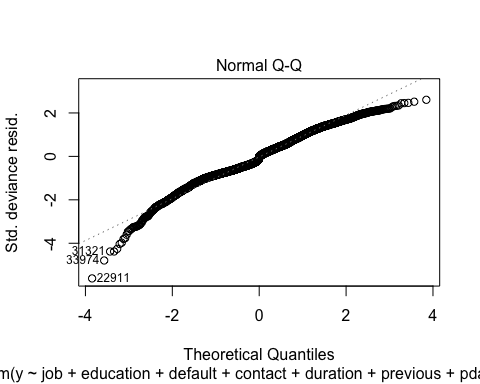
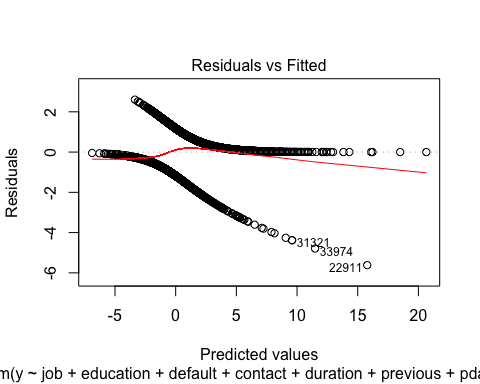
vif(backwardfinal)

## GVIF Df GVIF^(1/(2\*Df))  
## job 3.172482 11 1.053879  
## education 3.163795 7 1.085748  
## default 1.112657 1 1.054825  
## contact 1.079168 1 1.038830  
## duration 1.169913 1 1.081625  
## previous 1.514534 1 1.230664  
## pdays 1.477832 1 1.215661  
## campaign 1.016959 1 1.008444

summary(backwardfinal)

##   
## Call:  
## glm(formula = y ~ job + education + default + contact + duration +   
## previous + pdays + campaign, family = binomial(link = "logit"),   
## data = balanced\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -5.6182 -0.6726 0.0018 0.6228 2.6046   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.7584877 0.2169187 3.497 0.000471 \*\*\*  
## jobblue-collar -0.3991621 0.1147748 -3.478 0.000506 \*\*\*  
## jobentrepreneur -0.5113664 0.1767617 -2.893 0.003816 \*\*   
## jobhousemaid 0.0912759 0.2070616 0.441 0.659347   
## jobmanagement -0.1060683 0.1252734 -0.847 0.397166   
## jobretired 1.0059923 0.1404834 7.161 8.01e-13 \*\*\*  
## jobself-employed -0.0705975 0.1698743 -0.416 0.677713   
## jobservices -0.2731439 0.1254690 -2.177 0.029482 \*   
## jobstudent 1.3878104 0.1835617 7.560 4.02e-14 \*\*\*  
## jobtechnician -0.2396598 0.1032433 -2.321 0.020270 \*   
## jobunemployed 0.4287341 0.1822792 2.352 0.018669 \*   
## jobunknown 0.7949786 0.3251431 2.445 0.014485 \*   
## educationbasic.6y -0.3657051 0.1776217 -2.059 0.039504 \*   
## educationbasic.9y -0.4519454 0.1392102 -3.246 0.001168 \*\*   
## educationhigh.school -0.1334720 0.1365406 -0.978 0.328309   
## educationilliterate 1.3802491 1.1239962 1.228 0.219453   
## educationprofessional.course 0.1642676 0.1507088 1.090 0.275728   
## educationuniversity.degree 0.2389617 0.1363962 1.752 0.079779 .   
## educationunknown 0.1317485 0.1799180 0.732 0.464004   
## defaultunknown -1.1549868 0.0951570 -12.138 < 2e-16 \*\*\*  
## contacttelephone -1.1008370 0.0733572 -15.007 < 2e-16 \*\*\*  
## duration 0.0057135 0.0001467 38.934 < 2e-16 \*\*\*  
## previous 0.2155153 0.0662610 3.253 0.001144 \*\*   
## pdays -0.0023093 0.0001606 -14.379 < 2e-16 \*\*\*  
## campaign -0.0971026 0.0165710 -5.860 4.63e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 11578.3 on 8351 degrees of freedom  
## Residual deviance: 7085.2 on 8327 degrees of freedom  
## AIC: 7135.2  
##   
## Number of Fisher Scoring iterations: 6

plot(backwardfinal)



# Prediction with the backward selection

logr<- glm(y ~ job + education + default + contact +duration + previous + pdays + campaign, family = binomial(link = "logit"),data = balanced\_train)  
  
# 1 way  
logr.probs<-predict(logr, newdata=balanced\_test)  
logr.pred<-rep("No",928)  
logr.pred[logr.probs>.5]="Yes"  
Truth<-balanced\_test[,21]  
Pred<-logr.pred  
ftable(addmargins(table(Pred,Truth)))

## Truth no yes Sum  
## Pred   
## No 429 138 567  
## Yes 41 320 361  
## Sum 470 458 928

# 2 way  
pred = predict(logr, newdata=balanced\_test)  
accuracy <- table(pred, balanced\_test[,21])  
sum(diag(accuracy))/sum(accuracy)

## [1] 0.001077586

#confusionMatrix(factor(pred, levels = 1:928), factor(balanced\_test$y, levels = 1:928))

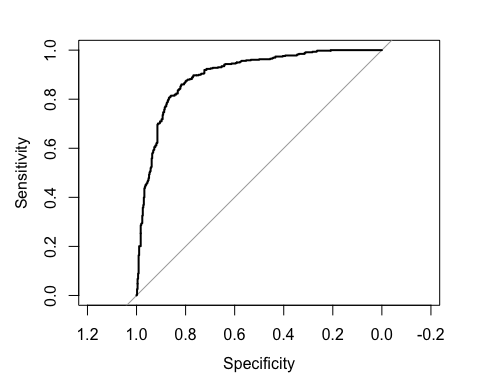
# Backward Hoslem test

hoslem.test(backwardfinal$y, fitted(backwardfinal), g=10)

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: backwardfinal$y, fitted(backwardfinal)  
## X-squared = 271.49, df = 8, p-value < 2.2e-16

# ROC curve for the backward model

logr<-glm(y ~ job + education + default + contact +duration + previous + pdays + campaign, family = binomial(link = "logit"),data = balanced\_train)  
logr.probs<-predict(logr, newdata=balanced\_test)  
roccurve <- roc(balanced\_test$y ~ logr.probs)  
plot(roccurve)



ME<-c()  
FP<-c()  
FN<-c()  
index<-0:928/928  
for (i in 1:928){  
logr.pred<-rep("no",928)  
logr.pred[logr.probs>index[i]]="yes"  
logr.pred<-factor(logr.pred,levels=c("no","yes"))  
Truth<-balanced\_test[,21]  
Pred<-logr.pred  
x<-table(Pred,Truth)  
ME[i]<-(x[1,2]+x[2,1])/sum(x)  
FP[i]<-x[2,1]/(x[2,1]+x[1,1])  
FN[i]<-x[1,2]/(x[1,2]+x[2,2])  
}  
TP= 1-FN  
ord.ind<-order(FP,decreasing=F)  
FP[ord.ind]

## [1] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [7] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [13] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [19] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [25] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [31] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [37] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [43] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [49] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [55] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [61] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [67] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [73] 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043 0.07234043  
## [79] 0.07234043 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [85] 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809 0.07446809  
## [91] 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574  
## [97] 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574  
## [103] 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574  
## [109] 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574  
## [115] 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574 0.07659574  
## [121] 0.07659574 0.07872340 0.07872340 0.07872340 0.07872340 0.07872340  
## [127] 0.07872340 0.07872340 0.08085106 0.08085106 0.08085106 0.08085106  
## [133] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [139] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [145] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [151] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [157] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [163] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [169] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [175] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [181] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [187] 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106 0.08085106  
## [193] 0.08085106 0.08085106 0.08085106 0.08085106 0.08297872 0.08510638  
## [199] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [205] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [211] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [217] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [223] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [229] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [235] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [241] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [247] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [253] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [259] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [265] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [271] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [277] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [283] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [289] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [295] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [301] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [307] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [313] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [319] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [325] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [331] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [337] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [343] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [349] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [355] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [361] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [367] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [373] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [379] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [385] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [391] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [397] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [403] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [409] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [415] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [421] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [427] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [433] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [439] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [445] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [451] 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638 0.08510638  
## [457] 0.08510638 0.08723404 0.08723404 0.08723404 0.08723404 0.08723404  
## [463] 0.08723404 0.08723404 0.08723404 0.08936170 0.08936170 0.08936170  
## [469] 0.08936170 0.08936170 0.08936170 0.08936170 0.09148936 0.09148936  
## [475] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [481] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [487] 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702 0.09361702  
## [493] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [499] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [505] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [511] 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468 0.09574468  
## [517] 0.09787234 0.10000000 0.10000000 0.10212766 0.10212766 0.10212766  
## [523] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [529] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [535] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [541] 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766 0.10212766  
## [547] 0.10425532 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [553] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [559] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [565] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [571] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [577] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [583] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [589] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [595] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298  
## [601] 0.10638298 0.10638298 0.10638298 0.10638298 0.10638298 0.10851064  
## [607] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [613] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [619] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [625] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [631] 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064 0.10851064  
## [637] 0.10851064 0.10851064 0.10851064 0.11063830 0.11063830 0.11063830  
## [643] 0.11063830 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [649] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [655] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [661] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [667] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [673] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [679] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [685] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [691] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [697] 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596 0.11276596  
## [703] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [709] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [715] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [721] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [727] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [733] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [739] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [745] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [751] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [757] 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362 0.11489362  
## [763] 0.11489362 0.11489362 0.11702128 0.11702128 0.11702128 0.11702128  
## [769] 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128  
## [775] 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128 0.11702128  
## [781] 0.11702128 0.11702128 0.11702128 0.11702128 0.11914894 0.11914894  
## [787] 0.11914894 0.11914894 0.12127660 0.12127660 0.12127660 0.12127660  
## [793] 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660  
## [799] 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660 0.12127660  
## [805] 0.12127660 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426  
## [811] 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426 0.12340426  
## [817] 0.12340426 0.12340426 0.12340426 0.12340426 0.12553191 0.12553191  
## [823] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [829] 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191 0.12553191  
## [835] 0.12553191 0.12553191 0.12553191 0.12553191 0.12765957 0.12765957  
## [841] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [847] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [853] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [859] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [865] 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957 0.12765957  
## [871] 0.12978723 0.12978723 0.12978723 0.13191489 0.13191489 0.13191489  
## [877] 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489  
## [883] 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489  
## [889] 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489 0.13191489  
## [895] 0.13191489 0.13191489 0.13191489 0.13404255 0.13404255 0.13404255  
## [901] 0.13404255 0.13404255 0.13404255 0.13404255 0.13617021 0.13617021  
## [907] 0.13617021 0.13617021 0.13617021 0.13829787 0.13829787 0.13829787  
## [913] 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787 0.13829787  
## [919] 0.14042553 0.14042553 0.14042553 0.14042553 0.14042553 0.14042553  
## [925] 0.14042553 0.14042553 0.14042553 0.14042553

TP[ord.ind]

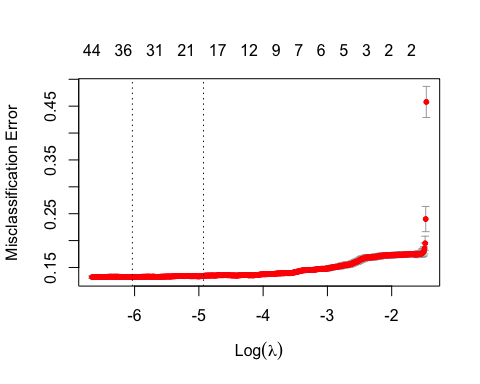
## [1] 0.6048035 0.6026201 0.6026201 0.6026201 0.6026201 0.6026201 0.6026201  
## [8] 0.6026201 0.6026201 0.6026201 0.6026201 0.6026201 0.6026201 0.6026201  
## [15] 0.6026201 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367  
## [22] 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367 0.6004367  
## [29] 0.6004367 0.6004367 0.6004367 0.6004367 0.5982533 0.5982533 0.5982533  
## [36] 0.5982533 0.5982533 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699  
## [43] 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699  
## [50] 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699  
## [57] 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699  
## [64] 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699  
## [71] 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699 0.5960699  
## [78] 0.5960699 0.5938865 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035  
## [85] 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6113537  
## [92] 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537  
## [99] 0.6113537 0.6113537 0.6113537 0.6069869 0.6069869 0.6069869 0.6069869  
## [106] 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6069869 0.6048035  
## [113] 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035 0.6048035  
## [120] 0.6048035 0.6048035 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537  
## [127] 0.6113537 0.6113537 0.6200873 0.6200873 0.6200873 0.6200873 0.6200873  
## [134] 0.6200873 0.6200873 0.6200873 0.6179039 0.6179039 0.6179039 0.6179039  
## [141] 0.6179039 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205  
## [148] 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205  
## [155] 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205 0.6157205  
## [162] 0.6157205 0.6157205 0.6157205 0.6157205 0.6135371 0.6135371 0.6135371  
## [169] 0.6135371 0.6135371 0.6135371 0.6135371 0.6135371 0.6135371 0.6135371  
## [176] 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537  
## [183] 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537  
## [190] 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537 0.6113537  
## [197] 0.6200873 0.6965066 0.6965066 0.6965066 0.6965066 0.6965066 0.6965066  
## [204] 0.6943231 0.6921397 0.6921397 0.6921397 0.6899563 0.6877729 0.6877729  
## [211] 0.6877729 0.6877729 0.6877729 0.6855895 0.6834061 0.6834061 0.6834061  
## [218] 0.6834061 0.6834061 0.6834061 0.6812227 0.6812227 0.6812227 0.6812227  
## [225] 0.6790393 0.6790393 0.6790393 0.6768559 0.6768559 0.6768559 0.6768559  
## [232] 0.6768559 0.6768559 0.6768559 0.6768559 0.6768559 0.6768559 0.6768559  
## [239] 0.6768559 0.6768559 0.6768559 0.6768559 0.6768559 0.6768559 0.6768559  
## [246] 0.6768559 0.6768559 0.6768559 0.6746725 0.6724891 0.6703057 0.6703057  
## [253] 0.6681223 0.6681223 0.6681223 0.6681223 0.6681223 0.6681223 0.6681223  
## [260] 0.6681223 0.6681223 0.6681223 0.6681223 0.6681223 0.6681223 0.6681223  
## [267] 0.6681223 0.6681223 0.6659389 0.6659389 0.6659389 0.6659389 0.6637555  
## [274] 0.6615721 0.6615721 0.6615721 0.6615721 0.6615721 0.6593886 0.6572052  
## [281] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052  
## [288] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052  
## [295] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052  
## [302] 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052 0.6572052  
## [309] 0.6572052 0.6572052 0.6572052 0.6550218 0.6550218 0.6550218 0.6550218  
## [316] 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218  
## [323] 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218 0.6550218  
## [330] 0.6550218 0.6528384 0.6506550 0.6506550 0.6506550 0.6506550 0.6462882  
## [337] 0.6462882 0.6462882 0.6462882 0.6462882 0.6462882 0.6462882 0.6462882  
## [344] 0.6462882 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048  
## [351] 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048  
## [358] 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048 0.6441048  
## [365] 0.6441048 0.6441048 0.6441048 0.6419214 0.6419214 0.6419214 0.6419214  
## [372] 0.6419214 0.6419214 0.6419214 0.6397380 0.6397380 0.6397380 0.6397380  
## [379] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [386] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [393] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [400] 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380 0.6397380  
## [407] 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546  
## [414] 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546 0.6375546  
## [421] 0.6353712 0.6353712 0.6353712 0.6353712 0.6331878 0.6331878 0.6331878  
## [428] 0.6331878 0.6331878 0.6331878 0.6331878 0.6310044 0.6288210 0.6288210  
## [435] 0.6288210 0.6288210 0.6266376 0.6266376 0.6266376 0.6266376 0.6266376  
## [442] 0.6266376 0.6266376 0.6266376 0.6266376 0.6244541 0.6244541 0.6244541  
## [449] 0.6244541 0.6244541 0.6244541 0.6244541 0.6244541 0.6222707 0.6222707  
## [456] 0.6222707 0.6222707 0.6986900 0.6986900 0.6986900 0.6986900 0.6986900  
## [463] 0.6986900 0.6965066 0.6965066 0.7008734 0.7008734 0.7008734 0.7008734  
## [470] 0.7008734 0.7008734 0.7008734 0.7008734 0.7008734 0.7030568 0.7030568  
## [477] 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568  
## [484] 0.7030568 0.7030568 0.7030568 0.7030568 0.7030568 0.7008734 0.7008734  
## [491] 0.7008734 0.7008734 0.7096070 0.7096070 0.7096070 0.7074236 0.7074236  
## [498] 0.7074236 0.7074236 0.7074236 0.7074236 0.7074236 0.7074236 0.7074236  
## [505] 0.7074236 0.7074236 0.7074236 0.7074236 0.7074236 0.7052402 0.7052402  
## [512] 0.7052402 0.7052402 0.7052402 0.7030568 0.7030568 0.7096070 0.7096070  
## [519] 0.7096070 0.7183406 0.7183406 0.7161572 0.7161572 0.7161572 0.7161572  
## [526] 0.7161572 0.7161572 0.7161572 0.7161572 0.7161572 0.7161572 0.7139738  
## [533] 0.7139738 0.7139738 0.7139738 0.7139738 0.7139738 0.7117904 0.7117904  
## [540] 0.7117904 0.7117904 0.7117904 0.7096070 0.7096070 0.7096070 0.7096070  
## [547] 0.7183406 0.7358079 0.7358079 0.7358079 0.7358079 0.7358079 0.7336245  
## [554] 0.7336245 0.7336245 0.7336245 0.7314410 0.7314410 0.7314410 0.7314410  
## [561] 0.7314410 0.7314410 0.7314410 0.7314410 0.7314410 0.7314410 0.7314410  
## [568] 0.7314410 0.7314410 0.7292576 0.7270742 0.7270742 0.7270742 0.7227074  
## [575] 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074 0.7227074  
## [582] 0.7227074 0.7227074 0.7227074 0.7227074 0.7205240 0.7205240 0.7205240  
## [589] 0.7205240 0.7205240 0.7205240 0.7183406 0.7183406 0.7183406 0.7183406  
## [596] 0.7183406 0.7183406 0.7183406 0.7183406 0.7183406 0.7183406 0.7183406  
## [603] 0.7183406 0.7183406 0.7183406 0.7467249 0.7467249 0.7467249 0.7467249  
## [610] 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249  
## [617] 0.7467249 0.7467249 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415  
## [624] 0.7445415 0.7445415 0.7445415 0.7445415 0.7445415 0.7423581 0.7423581  
## [631] 0.7423581 0.7423581 0.7401747 0.7401747 0.7379913 0.7379913 0.7379913  
## [638] 0.7379913 0.7379913 0.7467249 0.7467249 0.7467249 0.7467249 0.7576419  
## [645] 0.7576419 0.7576419 0.7576419 0.7576419 0.7576419 0.7576419 0.7576419  
## [652] 0.7532751 0.7532751 0.7532751 0.7532751 0.7532751 0.7510917 0.7510917  
## [659] 0.7510917 0.7510917 0.7510917 0.7489083 0.7489083 0.7489083 0.7489083  
## [666] 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083  
## [673] 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083 0.7489083  
## [680] 0.7489083 0.7489083 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249  
## [687] 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249  
## [694] 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249 0.7467249  
## [701] 0.7467249 0.7467249 0.7663755 0.7663755 0.7663755 0.7641921 0.7641921  
## [708] 0.7641921 0.7641921 0.7641921 0.7641921 0.7641921 0.7641921 0.7641921  
## [715] 0.7641921 0.7620087 0.7620087 0.7620087 0.7620087 0.7620087 0.7620087  
## [722] 0.7620087 0.7620087 0.7620087 0.7620087 0.7620087 0.7620087 0.7620087  
## [729] 0.7620087 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253  
## [736] 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253  
## [743] 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253 0.7598253  
## [750] 0.7598253 0.7576419 0.7576419 0.7576419 0.7576419 0.7576419 0.7576419  
## [757] 0.7576419 0.7576419 0.7576419 0.7576419 0.7576419 0.7576419 0.7576419  
## [764] 0.7576419 0.7685590 0.7685590 0.7685590 0.7685590 0.7685590 0.7685590  
## [771] 0.7685590 0.7685590 0.7685590 0.7685590 0.7685590 0.7685590 0.7685590  
## [778] 0.7685590 0.7685590 0.7685590 0.7685590 0.7685590 0.7685590 0.7663755  
## [785] 0.7707424 0.7707424 0.7707424 0.7707424 0.7794760 0.7794760 0.7794760  
## [792] 0.7772926 0.7751092 0.7751092 0.7751092 0.7751092 0.7751092 0.7729258  
## [799] 0.7729258 0.7729258 0.7729258 0.7729258 0.7707424 0.7707424 0.7707424  
## [806] 0.7860262 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428  
## [813] 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428 0.7838428  
## [820] 0.7816594 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096  
## [827] 0.7882096 0.7882096 0.7860262 0.7860262 0.7860262 0.7860262 0.7860262  
## [834] 0.7860262 0.7860262 0.7860262 0.7860262 0.7860262 0.7947598 0.7947598  
## [841] 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598 0.7947598  
## [848] 0.7947598 0.7947598 0.7947598 0.7947598 0.7925764 0.7925764 0.7925764  
## [855] 0.7925764 0.7925764 0.7925764 0.7925764 0.7925764 0.7903930 0.7882096  
## [862] 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096 0.7882096  
## [869] 0.7882096 0.7882096 0.7991266 0.7969432 0.7947598 0.8056769 0.8056769  
## [876] 0.8056769 0.8056769 0.8056769 0.8056769 0.8056769 0.8056769 0.8056769  
## [883] 0.8056769 0.8056769 0.8056769 0.8056769 0.8056769 0.8056769 0.8056769  
## [890] 0.8013100 0.8013100 0.8013100 0.8013100 0.7991266 0.7991266 0.7991266  
## [897] 0.7991266 0.8078603 0.8078603 0.8078603 0.8078603 0.8056769 0.8056769  
## [904] 0.8056769 0.8078603 0.8078603 0.8078603 0.8078603 0.8078603 0.8122271  
## [911] 0.8122271 0.8122271 0.8122271 0.8122271 0.8100437 0.8100437 0.8100437  
## [918] 0.8100437 0.8122271 0.8122271 0.8122271 0.8122271 0.8122271 0.8122271  
## [925] 0.8122271 0.8122271 0.8122271 0.8122271

simple\_auc <- function(TPR, FPR){  
 # inputs already sorted, best scores first   
 dFPR <- c(diff(FPR), 0)  
 dTPR <- c(diff(TPR), 0)  
 sum(TPR \* dFPR) + sum(dTPR \* dFPR)/2  
}  
  
paste("AUC=", simple\_auc(TP[ord.ind],FP[ord.ind]),sep=" ")

## [1] "AUC= 0.0493891108427019"

# Lasso Feature selection

dat.train.x <- model.matrix(y~ .,balanced\_train)  
dat.train.y<-balanced\_train[,21]  
cvfit <- cv.glmnet(dat.train.x, dat.train.y, family = "binomial", type.measure = "class", nlambda = 1000)  
plot(cvfit)



coef(cvfit, s = "lambda.min")

## 55 x 1 sparse Matrix of class "dgCMatrix"  
## 1  
## (Intercept) 27.2542586508  
## (Intercept) .   
## age .   
## jobblue-collar -0.2126127896  
## jobentrepreneur -0.1466695623  
## jobhousemaid 0.0236433517  
## jobmanagement .   
## jobretired 0.3460234786  
## jobself-employed .   
## jobservices -0.0053435272  
## jobstudent 0.4226282857  
## jobtechnician .   
## jobunemployed .   
## jobunknown 0.3529810361  
## maritalmarried -0.0922492182  
## maritalsingle 0.0416938408  
## maritalunknown .   
## educationbasic.6y -0.0175712219  
## educationbasic.9y -0.1513999249  
## educationhigh.school .   
## educationilliterate 0.3153258607  
## educationprofessional.course 0.1170906571  
## educationuniversity.degree 0.1944756079  
## educationunknown .   
## defaultunknown -0.4018353915  
## defaultyes .   
## housingunknown .   
## housingyes 0.0298379159  
## loanunknown .   
## loanyes .   
## contacttelephone -0.0546878212  
## monthaug 0.1286577329  
## monthdec -0.1228978712  
## monthjul 0.1144595522  
## monthjun .   
## monthmar 1.0788325873  
## monthmay -1.1502581289  
## monthnov -0.4579889571  
## monthoct 0.3791351537  
## monthsep -0.1376508803  
## day\_of\_weekmon .   
## day\_of\_weekthu -0.0343470298  
## day\_of\_weektue .   
## day\_of\_weekwed 0.0963429914  
## duration 0.0066208200  
## campaign -0.0116185776  
## pdays -0.0006339317  
## previous .   
## poutcomenonexistent 0.3945550761  
## poutcomesuccess 1.0800656704  
## emp.var.rate -0.4834166141  
## cons.price.idx 0.1157951717  
## cons.conf.idx .   
## euribor3m .   
## nr.employed -0.0078345716

#CV misclassification error rate   
cvfit$cvm[which(cvfit$lambda==cvfit$lambda.min)]

## [1] 0.1320642

#Optimal penalty  
cvfit$lambda.min

## [1] 0.002395375

# Prediction

# Final  
dat.test.y<- balanced\_test[,21]  
dat.test.y<-ifelse(dat.test.y == "yes", 1,0)  
dat.test.x<- model.matrix(y~ .,balanced\_test)  
finalmodel<-glmnet(dat.train.x, dat.train.y, family = "binomial",lambda=cvfit$lambda.min)  
  
pred <- predict(finalmodel, s = cvfit$lambda.min, newx = dat.test.x)  
Mypred<-ifelse(pred>.5, 1,0)  
  
final <- cbind(dat.test.y, Mypred)  
testMSE\_LASSO<-mean((dat.test.y-Mypred)^2)  
testMSE\_LASSO

## [1] 0.137931

# Checking the first six obs  
head(final)

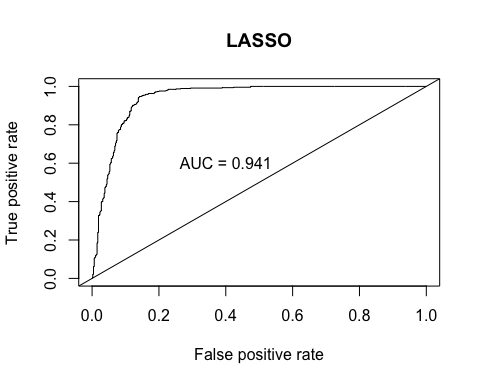
## dat.test.y 1  
## 34693 0 0  
## 1703 1 0  
## 14932 0 0  
## 15914 0 1  
## 5679 0 0  
## 26127 0 0

# RSQ  
actual <- dat.test.y  
preds <- pred  
rss <- sum((preds - actual) ^ 2)  
tss <- sum((actual - mean(actual)) ^ 2)  
rsq <- 1 - rss/tss  
rsq

## [1] -29.77997

# ROC for the lasso selection model

# Predict from model  
preds <- predict(finalmodel, newx = dat.test.x, type = 'response')  
  
# Calculate true positive rate and false positive rate on the prediction object  
perf <- performance(prediction(preds, dat.test.y), 'tpr', 'fpr')  
  
auc.train <- performance(prediction(preds, dat.test.y), measure = "auc")  
auc.train <- auc.train@y.values  
  
#Plot ROC  
plot(perf,main="LASSO")  
abline(a=0, b= 1) #Ref line indicating poor performance  
text(x = .40, y = .6,paste("AUC = ", round(auc.train[[1]],3), sep = ""))



# Extra code to use if we need later

# prediction with step method  
#eval\_results(y\_train, predictions\_train, balanced\_train)  
#predictions\_test <- predict(finalmodel, s = cvfit$lambda.min, newx = dat.test.x)  
#eval\_results(dat.test.y, predictions\_test, balanced\_test)  
#fit.pred <- predict(finalmodel, newy = dat.test, type = "response")  
  
# ROC curve with lasso  
#pred <- predict(finalmodel, s = cvfit$lambda.min, newx = dat.test.x)  
#roc.perf = performance(pred, measure = "tpr", x.measure = "fpr")  
#auc.train <- performance(pred, measure = "auc")  
#auc.train <- auc.train@y.values  
#Plot ROC  
#plot(roc.perf,main="LASSO")  
#abline(a=0, b= 1) #Ref line indicating poor performance  
#text(x = .40, y = .6,paste("AUC = ", round(auc.train[[1]],3), sep = ""))

```