Churn\_Prediction\_v2\_Final.R

Pranoy

Thu Aug 15 22:21:38 2019

## CHURN PREDICTION CLASSIFICATTION PROJECT  
  
## Classification project, we will do linear regression on the data to precdict wheather the customer   
## will churn 1 or not churn 0  
  
#importing Churn data from Churn.xls file  
setwd("D:/Data Analytics/R AcGg/Projects/Projects/Project 2 - Churn Prediction")  
library(readxl)  
Churn <- read\_excel("Churn.xls")  
class(Churn)

## [1] "tbl\_df" "tbl" "data.frame"

View(Churn)  
nrow(Churn)

## [1] 3333

str(Churn) # 3333 obs , 21 variables

## Classes 'tbl\_df', 'tbl' and 'data.frame': 3333 obs. of 21 variables:  
## $ Account Length: num 128 107 137 84 75 118 121 147 117 141 ...  
## $ VMail Message : num 25 26 0 0 0 0 24 0 0 37 ...  
## $ Day Mins : num 265 162 243 299 167 ...  
## $ Eve Mins : num 197.4 195.5 121.2 61.9 148.3 ...  
## $ Night Mins : num 245 254 163 197 187 ...  
## $ Intl Mins : num 10 13.7 12.2 6.6 10.1 6.3 7.5 7.1 8.7 11.2 ...  
## $ CustServ Calls: num 1 1 0 2 3 0 3 0 1 0 ...  
## $ Churn : num 0 0 0 0 0 0 0 0 0 0 ...  
## $ Int'l Plan : num 0 0 0 1 1 1 0 1 0 1 ...  
## $ VMail Plan : num 1 1 0 0 0 0 1 0 0 1 ...  
## $ Day Calls : num 110 123 114 71 113 98 88 79 97 84 ...  
## $ Day Charge : num 45.1 27.5 41.4 50.9 28.3 ...  
## $ Eve Calls : num 99 103 110 88 122 101 108 94 80 111 ...  
## $ Eve Charge : num 16.78 16.62 10.3 5.26 12.61 ...  
## $ Night Calls : num 91 103 104 89 121 118 118 96 90 97 ...  
## $ Night Charge : num 11.01 11.45 7.32 8.86 8.41 ...  
## $ Intl Calls : num 3 3 5 7 3 6 7 6 4 5 ...  
## $ Intl Charge : num 2.7 3.7 3.29 1.78 2.73 1.7 2.03 1.92 2.35 3.02 ...  
## $ State : chr "KS" "OH" "NJ" "OH" ...  
## $ Area Code : num 415 415 415 408 415 510 510 415 408 415 ...  
## $ Phone : chr "382-4657" "371-7191" "358-1921" "375-9999" ...

table(Churn$Churn)

##   
## 0 1   
## 2850 483

summary(Churn)

## Account Length VMail Message Day Mins Eve Mins   
## Min. : 1.0 Min. : 0.000 Min. : 0.0 Min. : 0.0   
## 1st Qu.: 74.0 1st Qu.: 0.000 1st Qu.:143.7 1st Qu.:166.6   
## Median :101.0 Median : 0.000 Median :179.4 Median :201.4   
## Mean :101.1 Mean : 8.099 Mean :179.8 Mean :201.0   
## 3rd Qu.:127.0 3rd Qu.:20.000 3rd Qu.:216.4 3rd Qu.:235.3   
## Max. :243.0 Max. :51.000 Max. :350.8 Max. :363.7   
## Night Mins Intl Mins CustServ Calls Churn   
## Min. : 23.2 Min. : 0.00 Min. :0.000 Min. :0.0000   
## 1st Qu.:167.0 1st Qu.: 8.50 1st Qu.:1.000 1st Qu.:0.0000   
## Median :201.2 Median :10.30 Median :1.000 Median :0.0000   
## Mean :200.9 Mean :10.24 Mean :1.563 Mean :0.1449   
## 3rd Qu.:235.3 3rd Qu.:12.10 3rd Qu.:2.000 3rd Qu.:0.0000   
## Max. :395.0 Max. :20.00 Max. :9.000 Max. :1.0000   
## Int'l Plan VMail Plan Day Calls Day Charge   
## Min. :0.00000 Min. :0.0000 Min. : 0.0 Min. : 0.00   
## 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.: 87.0 1st Qu.:24.43   
## Median :0.00000 Median :0.0000 Median :101.0 Median :30.50   
## Mean :0.09691 Mean :0.2766 Mean :100.4 Mean :30.56   
## 3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:114.0 3rd Qu.:36.79   
## Max. :1.00000 Max. :1.0000 Max. :165.0 Max. :59.64   
## Eve Calls Eve Charge Night Calls Night Charge   
## Min. : 0.0 Min. : 0.00 Min. : 33.0 Min. : 1.040   
## 1st Qu.: 87.0 1st Qu.:14.16 1st Qu.: 87.0 1st Qu.: 7.520   
## Median :100.0 Median :17.12 Median :100.0 Median : 9.050   
## Mean :100.1 Mean :17.08 Mean :100.1 Mean : 9.039   
## 3rd Qu.:114.0 3rd Qu.:20.00 3rd Qu.:113.0 3rd Qu.:10.590   
## Max. :170.0 Max. :30.91 Max. :175.0 Max. :17.770   
## Intl Calls Intl Charge State Area Code   
## Min. : 0.000 Min. :0.000 Length:3333 Min. :408.0   
## 1st Qu.: 3.000 1st Qu.:2.300 Class :character 1st Qu.:408.0   
## Median : 4.000 Median :2.780 Mode :character Median :415.0   
## Mean : 4.479 Mean :2.765 Mean :437.2   
## 3rd Qu.: 6.000 3rd Qu.:3.270 3rd Qu.:510.0   
## Max. :20.000 Max. :5.400 Max. :510.0   
## Phone   
## Length:3333   
## Class :character   
## Mode :character   
##   
##   
##

##Changing the names of the column for ROSE package:  
names(Churn) <- c("Account.Length", "VMail.Message", "Day.Mins", "Eve.Mins", "Night.Mins", "Intl.Mins",   
 "CustServ.Calls", "Churn", "Int.l.Plan", "VMail.Plan", "Day.Calls", "Day.Charge",   
 "Eve.Calls", "Eve.Charge", "Night.Calls", "Night.Charge", "Intl.Calls", "Intl.Charge" ,   
 "State" , "Area.Code" ,"Phone" )  
names(Churn)

## [1] "Account.Length" "VMail.Message" "Day.Mins" "Eve.Mins"   
## [5] "Night.Mins" "Intl.Mins" "CustServ.Calls" "Churn"   
## [9] "Int.l.Plan" "VMail.Plan" "Day.Calls" "Day.Charge"   
## [13] "Eve.Calls" "Eve.Charge" "Night.Calls" "Night.Charge"   
## [17] "Intl.Calls" "Intl.Charge" "State" "Area.Code"   
## [21] "Phone"

##obseravtions  
 #variable 8 is churn(binomial variable with 0 and 1)  
 #Variable 21 can be omitted as it is just a phone number and cannot be a predictor  
 #variable 19 (state) is string which needs to be converted to factor  
  
#### MISSING VALUES TREATMENT #####  
# visualize the missing values using the missing map from the Amelia package  
library(Amelia)

## Loading required package: Rcpp

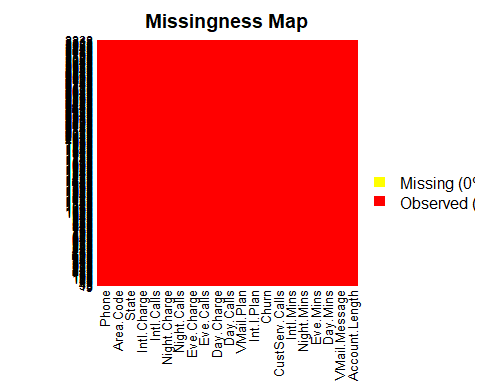
## ##   
## ## Amelia II: Multiple Imputation  
## ## (Version 1.7.5, built: 2018-05-07)  
## ## Copyright (C) 2005-2019 James Honaker, Gary King and Matthew Blackwell  
## ## Refer to http://gking.harvard.edu/amelia/ for more information  
## ##

missmap(Churn,col=c("yellow","red"))

## Warning in if (class(obj) == "amelia") {: the condition has length > 1 and  
## only the first element will be used

## Warning: Unknown or uninitialised column: 'arguments'.  
  
## Warning: Unknown or uninitialised column: 'arguments'.

## Warning: Unknown or uninitialised column: 'imputations'.



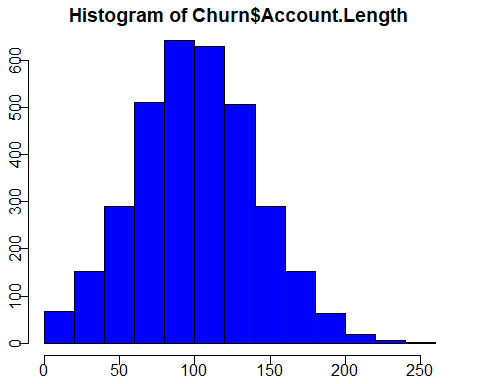
any(is.na(Churn))

## [1] FALSE

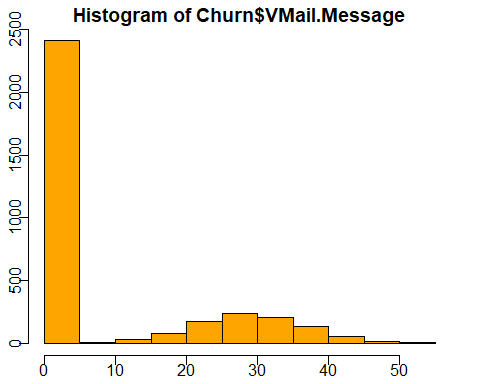
## no missing values were detected.  
sapply(Churn, function(x) sum(is.na(x)))

## Account.Length VMail.Message Day.Mins Eve.Mins Night.Mins   
## 0 0 0 0 0   
## Intl.Mins CustServ.Calls Churn Int.l.Plan VMail.Plan   
## 0 0 0 0 0   
## Day.Calls Day.Charge Eve.Calls Eve.Charge Night.Calls   
## 0 0 0 0 0   
## Night.Charge Intl.Calls Intl.Charge State Area.Code   
## 0 0 0 0 0   
## Phone   
## 0

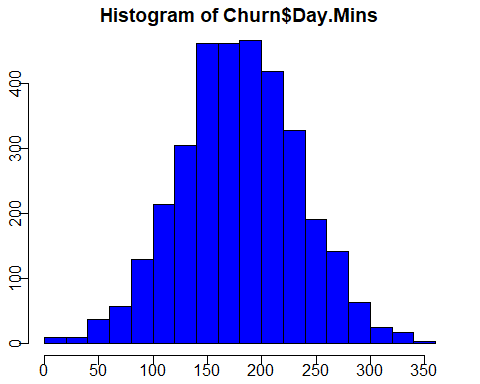
## NO missing values were found  
  
#for adjusting plots  
par(mfcol=c(1,1)) #1 graph per plot  
par(cex=1, mai=c(0.3,0.3,0.3,0.3)) #adjusting margins and borders  
  
###variable exploration  
library(ggplot2)  
#histogram for all variables  
hist(Churn$`Account.Length`, col = "blue")



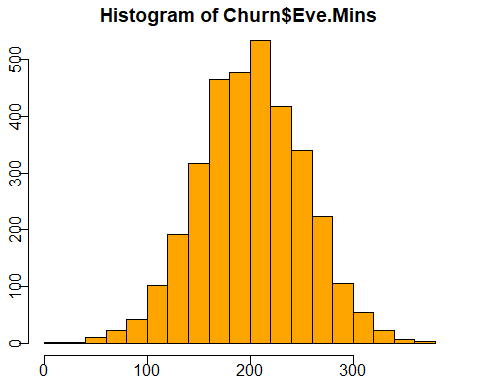
hist(Churn$`VMail.Message`, col = "orange")



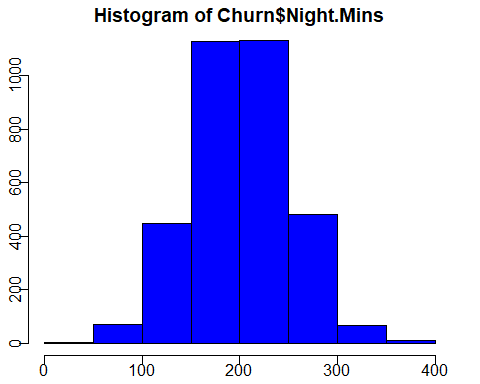
hist(Churn$`Day.Mins`, col = "blue")



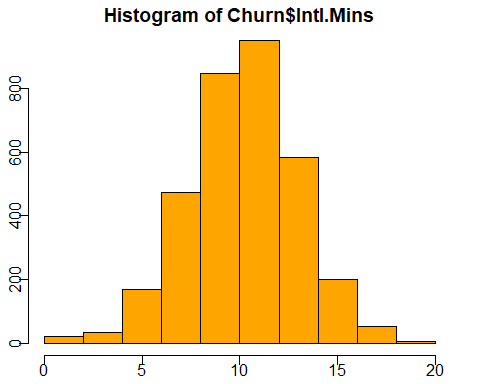
hist(Churn$`Eve.Mins`, col = "orange")



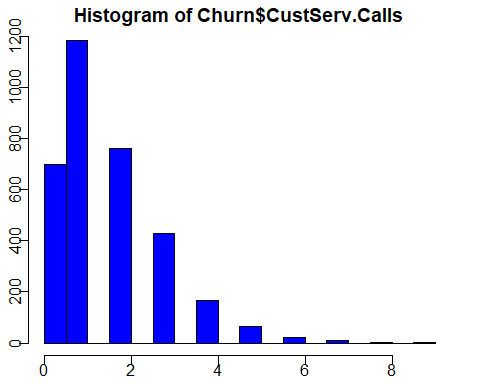
hist(Churn$`Night.Mins`, col = "blue")



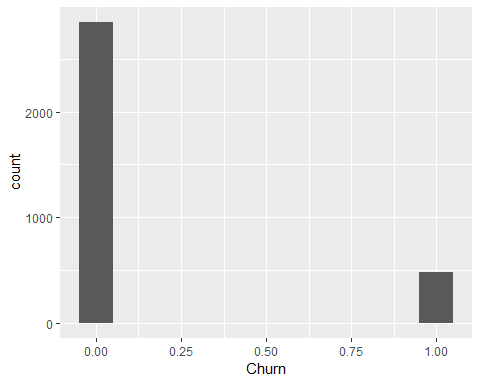
hist(Churn$`Intl.Mins`, col = "orange")



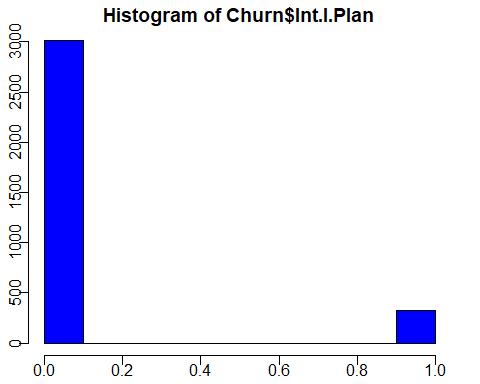
hist(Churn$`CustServ.Calls`, col = "blue")



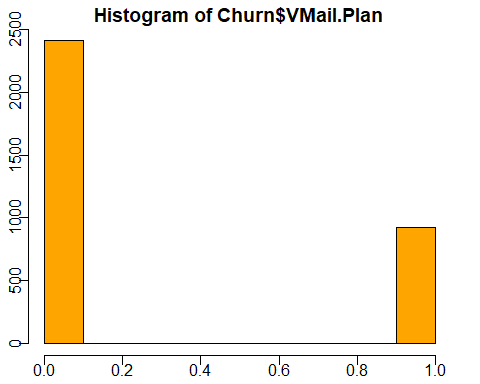
ggplot(data = Churn) + geom\_bar(mapping = aes(x = Churn), width = 0.1)



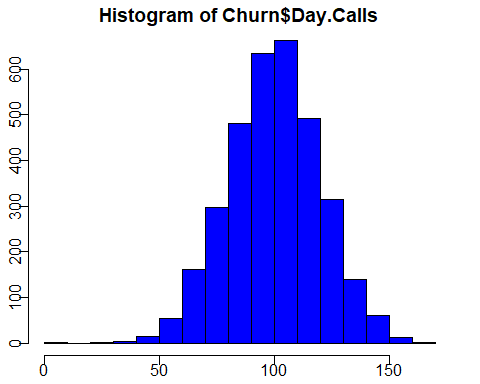
hist(Churn$`Int.l.Plan`, col = "blue")



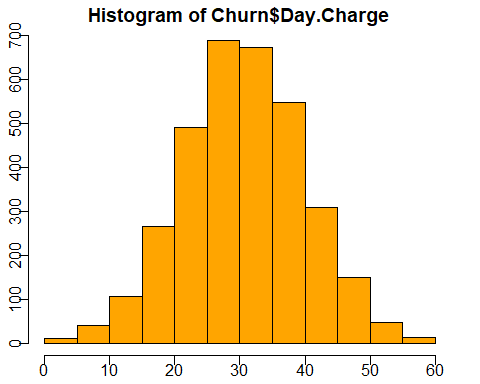
hist(Churn$`VMail.Plan`, col = "orange")



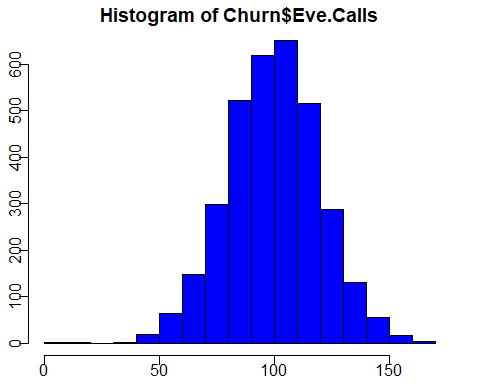
hist(Churn$`Day.Calls`, col = "blue")



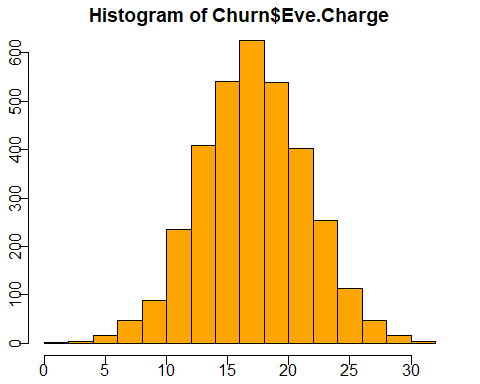
hist(Churn$`Day.Charge`, col = "orange")



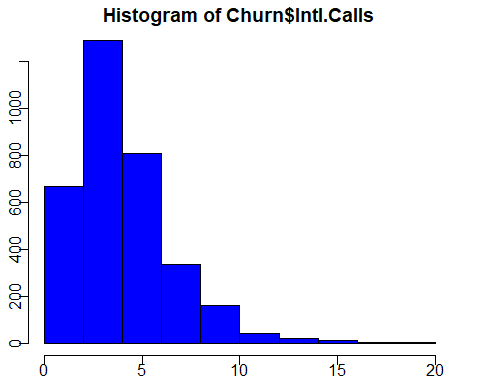
hist(Churn$`Eve.Calls`, col = "blue")



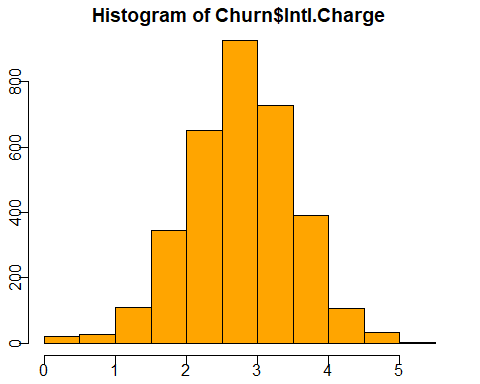
hist(Churn$`Eve.Charge`, col = "orange")



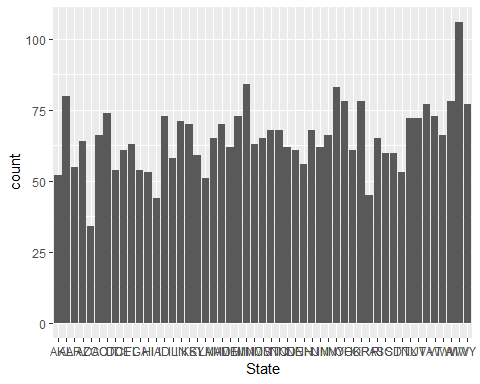
hist(Churn$`Intl.Calls`, col = "blue")



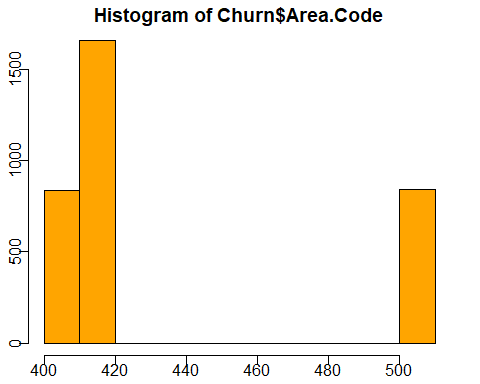
hist(Churn$`Intl.Charge`, col = "orange")



ggplot(data = Churn) + geom\_bar(mapping = aes(x = State))



hist(Churn$`Area.Code`, col = "orange")



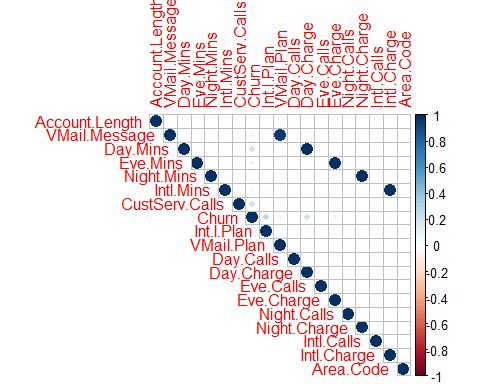
##correlations between variables  
cor(Churn[, c(-19,-21)])

## Account.Length VMail.Message Day.Mins Eve.Mins  
## Account.Length 1.000000000 -0.0046278243 0.0062160205 -0.006757142  
## VMail.Message -0.004627824 1.0000000000 0.0007782741 0.017562034  
## Day.Mins 0.006216021 0.0007782741 1.0000000000 0.007042511  
## Eve.Mins -0.006757142 0.0175620343 0.0070425110 1.000000000  
## Night.Mins -0.008955192 0.0076811359 0.0043233666 -0.012583678  
## Intl.Mins 0.009513902 0.0028561959 -0.0101545856 -0.011034714  
## CustServ.Calls -0.003795939 -0.0132625831 -0.0134231864 -0.012984553  
## Churn 0.016540742 -0.0897279698 0.2051508293 0.092795790  
## Int.l.Plan 0.024734655 0.0087454859 0.0493958241 0.019100013  
## VMail.Plan 0.002918409 0.9569266421 -0.0016840687 0.021545046  
## Day.Calls 0.038469882 -0.0095480677 0.0067504139 -0.021451408  
## Day.Charge 0.006214135 0.0007755235 0.9999999522 0.007049607  
## Eve.Calls 0.019259967 -0.0058643513 0.0157689932 -0.011430108  
## Eve.Charge -0.006745302 0.0175777801 0.0070290353 0.999999776  
## Night.Calls -0.013176275 0.0071230629 0.0229724555 0.007585643  
## Night.Charge -0.008959535 0.0076632904 0.0043003570 -0.012592806  
## Intl.Calls 0.020661428 0.0139573387 0.0080333570 0.002541292  
## Intl.Charge 0.009545675 0.0028836579 -0.0100919742 -0.011066621  
## Area.Code -0.012463497 -0.0019943701 -0.0082643662 0.003580395  
## Night.Mins Intl.Mins CustServ.Calls Churn  
## Account.Length -0.008955192 0.009513902 -0.003795939 0.016540742  
## VMail.Message 0.007681136 0.002856196 -0.013262583 -0.089727970  
## Day.Mins 0.004323367 -0.010154586 -0.013423186 0.205150829  
## Eve.Mins -0.012583678 -0.011034714 -0.012984553 0.092795790  
## Night.Mins 1.000000000 -0.015207297 -0.009287613 0.035492853  
## Intl.Mins -0.015207297 1.000000000 -0.009639680 0.068238776  
## CustServ.Calls -0.009287613 -0.009639680 1.000000000 0.208749999  
## Churn 0.035492853 0.068238776 0.208749999 1.000000000  
## Int.l.Plan -0.028904754 0.045870743 -0.024521956 0.259851847  
## VMail.Plan 0.006078657 -0.001317871 -0.017823944 -0.102148141  
## Day.Calls 0.022937845 0.021564794 -0.018941930 0.018459312  
## Day.Charge 0.004323879 -0.010156862 -0.013426969 0.205150743  
## Eve.Calls -0.002092768 0.008702881 0.002422575 0.009233132  
## Eve.Charge -0.012592020 -0.011042582 -0.012987407 0.092786039  
## Night.Calls 0.011203856 -0.013604996 -0.012801927 0.006141203  
## Night.Charge 0.999999215 -0.015213526 -0.009276954 0.035495556  
## Intl.Calls -0.012353432 0.032303884 -0.017560599 -0.052844336  
## Intl.Charge -0.015179849 0.999992742 -0.009674732 0.068258632  
## Area.Code -0.005824660 -0.018288168 0.027572226 0.006174233  
## Int.l.Plan VMail.Plan Day.Calls Day.Charge  
## Account.Length 0.024734655 0.0029184091 0.038469882 0.0062141347  
## VMail.Message 0.008745486 0.9569266421 -0.009548068 0.0007755235  
## Day.Mins 0.049395824 -0.0016840687 0.006750414 0.9999999522  
## Eve.Mins 0.019100013 0.0215450456 -0.021451408 0.0070496072  
## Night.Mins -0.028904754 0.0060786567 0.022937845 0.0043238794  
## Intl.Mins 0.045870743 -0.0013178710 0.021564794 -0.0101568616  
## CustServ.Calls -0.024521956 -0.0178239441 -0.018941930 -0.0134269694  
## Churn 0.259851847 -0.1021481407 0.018459312 0.2051507432  
## Int.l.Plan 1.000000000 0.0060063707 0.003754626 0.0493980611  
## VMail.Plan 0.006006371 1.0000000000 -0.011085902 -0.0016856290  
## Day.Calls 0.003754626 -0.0110859022 1.000000000 0.0067529620  
## Day.Charge 0.049398061 -0.0016856290 0.006752962 1.0000000000  
## Eve.Calls 0.006113577 -0.0064444437 0.006462114 0.0157692822  
## Eve.Charge 0.019106152 0.0215585006 -0.021449263 0.0070361315  
## Night.Calls 0.012451186 0.0155525275 -0.019556965 0.0229724195  
## Night.Charge -0.028912765 0.0060637086 0.022926638 0.0043008608  
## Intl.Calls 0.017366343 0.0076175947 0.004574268 0.0080315720  
## Intl.Charge 0.045780130 -0.0012762669 0.021666095 -0.0100942572  
## Area.Code 0.048550511 -0.0007470951 -0.009646044 -0.0082644411  
## Eve.Calls Eve.Charge Night.Calls Night.Charge  
## Account.Length 0.019259967 -0.006745302 -0.0131762751 -0.008959535  
## VMail.Message -0.005864351 0.017577780 0.0071230629 0.007663290  
## Day.Mins 0.015768993 0.007029035 0.0229724555 0.004300357  
## Eve.Mins -0.011430108 0.999999776 0.0075856431 -0.012592806  
## Night.Mins -0.002092768 -0.012592020 0.0112038563 0.999999215  
## Intl.Mins 0.008702881 -0.011042582 -0.0136049964 -0.015213526  
## CustServ.Calls 0.002422575 -0.012987407 -0.0128019273 -0.009276954  
## Churn 0.009233132 0.092786039 0.0061412030 0.035495556  
## Int.l.Plan 0.006113577 0.019106152 0.0124511861 -0.028912765  
## VMail.Plan -0.006444444 0.021558501 0.0155525275 0.006063709  
## Day.Calls 0.006462114 -0.021449263 -0.0195569654 0.022926638  
## Day.Charge 0.015769282 0.007036131 0.0229724195 0.004300861  
## Eve.Calls 1.000000000 -0.011422894 0.0077097055 -0.002055984  
## Eve.Charge -0.011422894 1.000000000 0.0075958430 -0.012601142  
## Night.Calls 0.007709706 0.007595843 1.0000000000 0.011187820  
## Night.Charge -0.002055984 -0.012601142 0.0111878197 1.000000000  
## Intl.Calls 0.017433692 0.002541458 0.0003045795 -0.012329215  
## Intl.Charge 0.008673858 -0.011074499 -0.0136301696 -0.015186139  
## Area.Code -0.011886271 0.003606690 0.0165223169 -0.005845376  
## Intl.Calls Intl.Charge Area.Code  
## Account.Length 0.0206614284 0.009545675 -0.0124634967  
## VMail.Message 0.0139573387 0.002883658 -0.0019943701  
## Day.Mins 0.0080333570 -0.010091974 -0.0082643662  
## Eve.Mins 0.0025412917 -0.011066621 0.0035803948  
## Night.Mins -0.0123534324 -0.015179849 -0.0058246598  
## Intl.Mins 0.0323038841 0.999992742 -0.0182881682  
## CustServ.Calls -0.0175605992 -0.009674732 0.0275722257  
## Churn -0.0528443358 0.068258632 0.0061742332  
## Int.l.Plan 0.0173663432 0.045780130 0.0485505114  
## VMail.Plan 0.0076175947 -0.001276267 -0.0007470951  
## Day.Calls 0.0045742682 0.021666095 -0.0096460440  
## Day.Charge 0.0080315720 -0.010094257 -0.0082644411  
## Eve.Calls 0.0174336921 0.008673858 -0.0118862708  
## Eve.Charge 0.0025414580 -0.011074499 0.0036066902  
## Night.Calls 0.0003045795 -0.013630170 0.0165223169  
## Night.Charge -0.0123292150 -0.015186139 -0.0058453757  
## Intl.Calls 1.0000000000 0.032372145 -0.0241785887  
## Intl.Charge 0.0323721453 1.000000000 -0.0183946960  
## Area.Code -0.0241785887 -0.018394696 1.0000000000

library(corrplot)

## corrplot 0.84 loaded

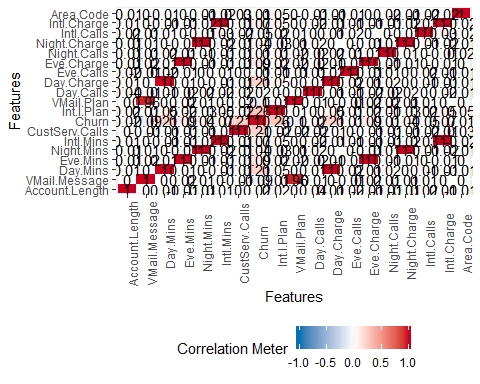
corrplot(cor(Churn[,c(-19,-21)]), type = "upper")



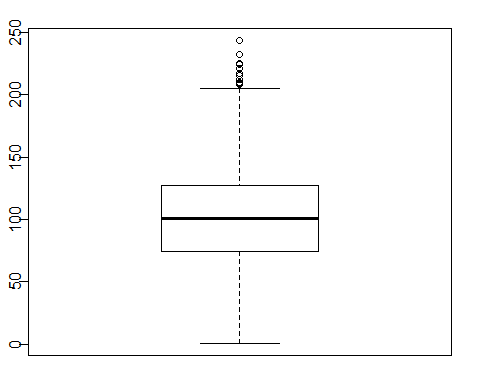
library(DataExplorer)

## Warning: package 'DataExplorer' was built under R version 3.5.3

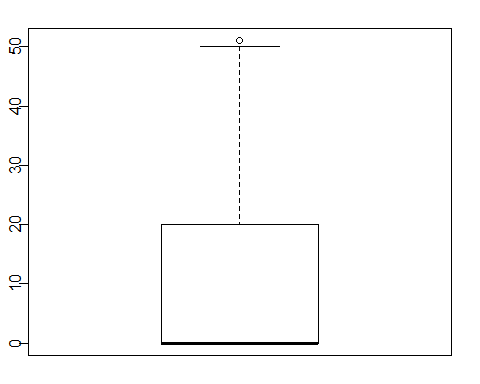
plot\_correlation(Churn, type = 'continuous')



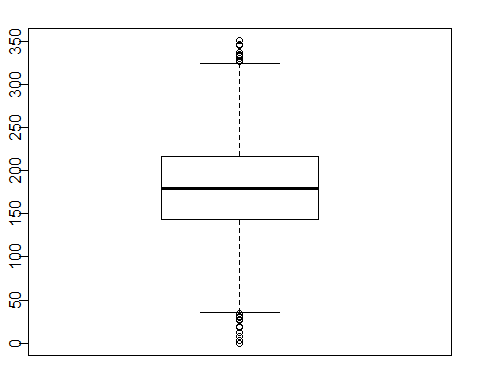
##Checking outliers  
boxplot(Churn$`Account.Length`)



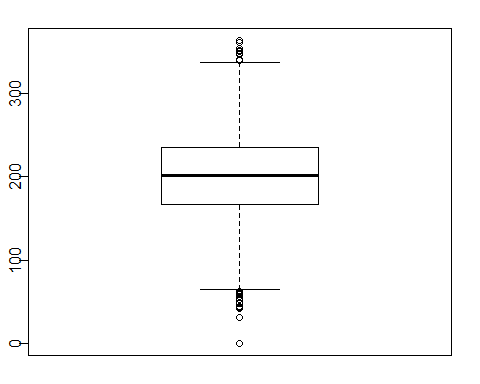
boxplot(Churn$`VMail.Message`)



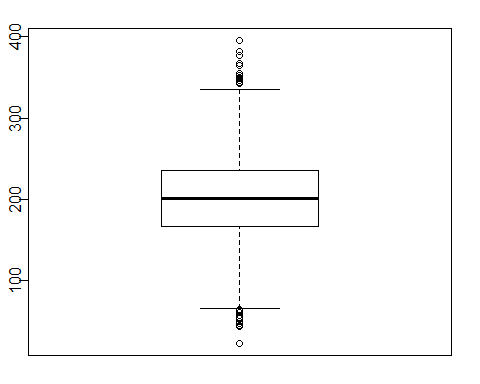
boxplot(Churn$`Day.Mins`)



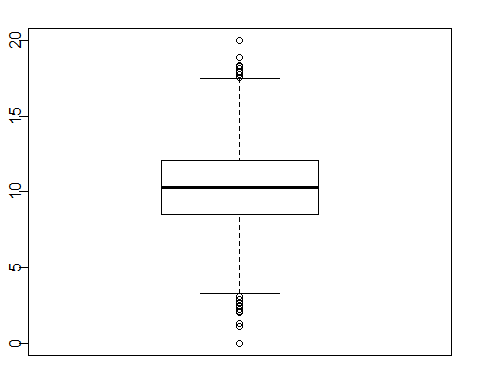
boxplot(Churn$`Eve.Mins`)



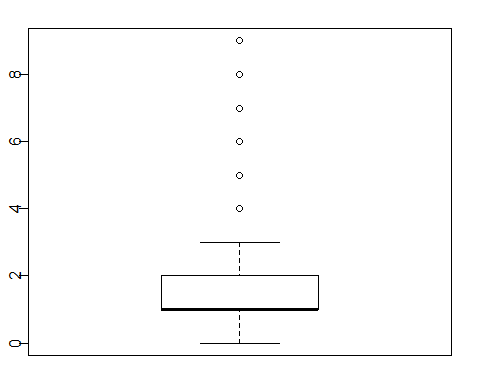
boxplot(Churn$`Night.Mins`)



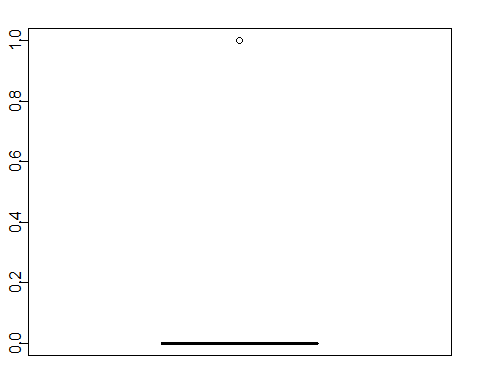
boxplot(Churn$`Intl.Mins`)



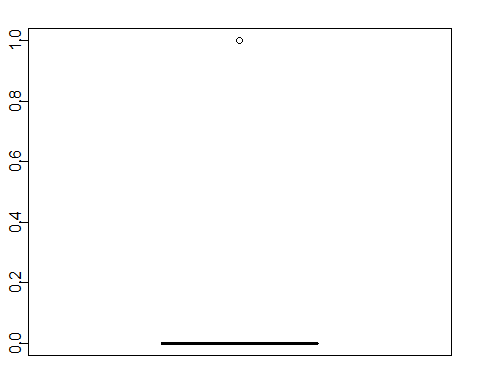
boxplot(Churn$`CustServ.Calls`)



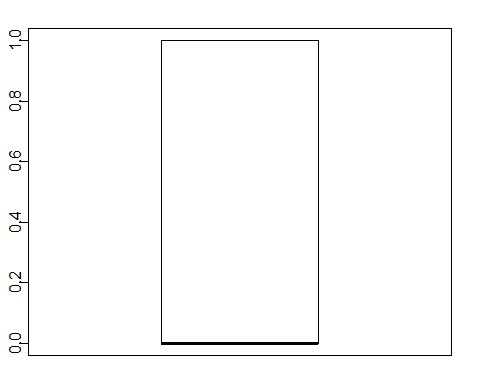
boxplot(Churn$Churn)



boxplot(Churn$`Int.l.Plan`)



boxplot(Churn$`VMail.Plan`)



###Model building without outliers treatmrent.  
  
## MOdel1: simple split in 80 ~ 2- for training and testing  
library(caTools)  
set.seed(1)  
split <- sample.split(Churn, SplitRatio = 0.80)  
Churn\_train <- subset(Churn, split == TRUE)

## Warning: Length of logical index must be 1 or 3333, not 21

Churn\_train<-Churn\_train[,c(-19, -21)] ###Remove phone number variable and state as it will make the model complex  
## Churn\_train$State<-as.factor(Churn\_train$State)  
Churn\_test <- subset(Churn, split == FALSE)

## Warning: Length of logical index must be 1 or 3333, not 21

table(Churn$Churn) #0 = 2850, 1 = 483

##   
## 0 1   
## 2850 483

table(Churn\_train$Churn) #0 = 1885, 1 = 338

##   
## 0 1   
## 2163 377

table(Churn\_test$Churn) #0 = 965, 1 = 145

##   
## 0 1   
## 687 106

##MODEL 1 : 80:20 split without outlier treatment  
glm\_model1.null<-glm(Churn~1,family=binomial(link="logit"),data= Churn\_train)  
summary(glm\_model1.null)

##   
## Call:  
## glm(formula = Churn ~ 1, family = binomial(link = "logit"), data = Churn\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.5669 -0.5669 -0.5669 -0.5669 1.9533   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.74701 0.05581 -31.3 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2133.4 on 2539 degrees of freedom  
## Residual deviance: 2133.4 on 2539 degrees of freedom  
## AIC: 2135.4  
##   
## Number of Fisher Scoring iterations: 4

glm\_model1.full<-glm(Churn ~., family=binomial(link="logit"), data=Churn\_train)  
summary(glm\_model1.full)

##   
## Call:  
## glm(formula = Churn ~ ., family = binomial(link = "logit"), data = Churn\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.2091 -0.5122 -0.3395 -0.1887 2.8898   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -8.978e+00 1.066e+00 -8.420 < 2e-16 \*\*\*  
## Account.Length 1.392e-03 1.590e-03 0.875 0.381527   
## VMail.Message 4.789e-02 2.135e-02 2.244 0.024843 \*   
## Day.Mins -9.130e-01 3.750e+00 -0.243 0.807645   
## Eve.Mins 3.258e+00 1.870e+00 1.743 0.081415 .   
## Night.Mins 1.603e-01 1.002e+00 0.160 0.872892   
## Intl.Mins -7.161e+00 6.031e+00 -1.187 0.235040   
## CustServ.Calls 5.103e-01 4.452e-02 11.461 < 2e-16 \*\*\*  
## Int.l.Plan 2.067e+00 1.667e-01 12.404 < 2e-16 \*\*\*  
## VMail.Plan -2.391e+00 6.902e-01 -3.465 0.000531 \*\*\*  
## Day.Calls 3.714e-03 3.138e-03 1.183 0.236719   
## Day.Charge 5.445e+00 2.206e+01 0.247 0.805033   
## Eve.Calls 2.325e-03 3.177e-03 0.732 0.464339   
## Eve.Charge -3.824e+01 2.200e+01 -1.738 0.082185 .   
## Night.Calls 2.269e-04 3.273e-03 0.069 0.944730   
## Night.Charge -3.478e+00 2.226e+01 -0.156 0.875839   
## Intl.Calls -9.729e-02 2.885e-02 -3.372 0.000745 \*\*\*  
## Intl.Charge 2.677e+01 2.233e+01 1.199 0.230671   
## Area.Code 5.370e-04 1.483e-03 0.362 0.717163   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2133.4 on 2539 degrees of freedom  
## Residual deviance: 1654.3 on 2521 degrees of freedom  
## AIC: 1692.3  
##   
## Number of Fisher Scoring iterations: 6

#to find the best model with lowest AIC  
step(glm\_model1.null, scope = list(upper=glm\_model1.full), direction="both", test="Chisq", data=Churn\_train)

## Start: AIC=2135.44  
## Churn ~ 1  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Int.l.Plan 1 1998.3 2002.3 135.121 < 2.2e-16 \*\*\*  
## + Day.Charge 1 2021.9 2025.9 111.557 < 2.2e-16 \*\*\*  
## + Day.Mins 1 2021.9 2025.9 111.556 < 2.2e-16 \*\*\*  
## + CustServ.Calls 1 2031.1 2035.1 102.294 < 2.2e-16 \*\*\*  
## + Eve.Mins 1 2102.6 2106.6 30.874 2.753e-08 \*\*\*  
## + Eve.Charge 1 2102.6 2106.6 30.861 2.772e-08 \*\*\*  
## + VMail.Plan 1 2105.4 2109.4 28.033 1.193e-07 \*\*\*  
## + VMail.Message 1 2112.9 2116.9 20.493 5.987e-06 \*\*\*  
## + Intl.Calls 1 2122.7 2126.7 10.724 0.001058 \*\*   
## + Intl.Charge 1 2124.6 2128.6 8.857 0.002920 \*\*   
## + Intl.Mins 1 2124.6 2128.6 8.841 0.002946 \*\*   
## + Night.Charge 1 2128.3 2132.3 5.124 0.023601 \*   
## + Night.Mins 1 2128.3 2132.3 5.123 0.023615 \*   
## <none> 2133.4 2135.4   
## + Day.Calls 1 2131.9 2135.9 1.511 0.218973   
## + Account.Length 1 2132.7 2136.7 0.727 0.393967   
## + Area.Code 1 2132.8 2136.8 0.625 0.429057   
## + Eve.Calls 1 2133.3 2137.3 0.160 0.688803   
## + Night.Calls 1 2133.3 2137.3 0.138 0.710740   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2002.32  
## Churn ~ Int.l.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + CustServ.Calls 1 1881.6 1887.6 116.701 < 2.2e-16 \*\*\*  
## + Day.Charge 1 1891.8 1897.8 106.543 < 2.2e-16 \*\*\*  
## + Day.Mins 1 1891.8 1897.8 106.543 < 2.2e-16 \*\*\*  
## + VMail.Plan 1 1965.2 1971.2 33.096 8.771e-09 \*\*\*  
## + Eve.Mins 1 1970.5 1976.5 27.864 1.301e-07 \*\*\*  
## + Eve.Charge 1 1970.5 1976.5 27.851 1.310e-07 \*\*\*  
## + VMail.Message 1 1973.4 1979.4 24.872 6.125e-07 \*\*\*  
## + Intl.Calls 1 1986.1 1992.1 12.212 0.0004748 \*\*\*  
## + Night.Charge 1 1990.3 1996.3 8.042 0.0045713 \*\*   
## + Night.Mins 1 1990.3 1996.3 8.040 0.0045764 \*\*   
## + Intl.Charge 1 1992.8 1998.8 5.469 0.0193514 \*   
## + Intl.Mins 1 1992.9 1998.9 5.449 0.0195821 \*   
## <none> 1998.3 2002.3   
## + Day.Calls 1 1997.0 2003.0 1.350 0.2452437   
## + Account.Length 1 1998.0 2004.0 0.285 0.5936836   
## + Eve.Calls 1 1998.2 2004.2 0.128 0.7208666   
## + Night.Calls 1 1998.3 2004.3 0.003 0.9549851   
## + Area.Code 1 1998.3 2004.3 0.000 0.9833689   
## - Int.l.Plan 1 2133.4 2135.4 135.121 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1887.61  
## Churn ~ Int.l.Plan + CustServ.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Day.Charge 1 1765.1 1773.1 116.534 < 2.2e-16 \*\*\*  
## + Day.Mins 1 1765.1 1773.1 116.533 < 2.2e-16 \*\*\*  
## + Eve.Mins 1 1848.0 1856.0 33.625 6.681e-09 \*\*\*  
## + Eve.Charge 1 1848.0 1856.0 33.609 6.738e-09 \*\*\*  
## + VMail.Plan 1 1849.8 1857.8 31.818 1.693e-08 \*\*\*  
## + VMail.Message 1 1857.3 1865.3 24.349 8.035e-07 \*\*\*  
## + Intl.Calls 1 1871.2 1879.2 10.431 0.001239 \*\*   
## + Night.Mins 1 1873.8 1881.8 7.765 0.005328 \*\*   
## + Night.Charge 1 1873.8 1881.8 7.764 0.005330 \*\*   
## + Intl.Charge 1 1874.9 1882.9 6.738 0.009435 \*\*   
## + Intl.Mins 1 1874.9 1882.9 6.716 0.009554 \*\*   
## <none> 1881.6 1887.6   
## + Day.Calls 1 1879.7 1887.7 1.945 0.163157   
## + Eve.Calls 1 1881.3 1889.3 0.355 0.551029   
## + Account.Length 1 1881.3 1889.3 0.292 0.589094   
## + Area.Code 1 1881.5 1889.5 0.088 0.766333   
## + Night.Calls 1 1881.6 1889.6 0.012 0.911157   
## - CustServ.Calls 1 1998.3 2002.3 116.701 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2031.1 2035.1 149.527 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1773.08  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Eve.Mins 1 1729.5 1739.5 35.549 2.487e-09 \*\*\*  
## + Eve.Charge 1 1729.5 1739.5 35.533 2.507e-09 \*\*\*  
## + VMail.Plan 1 1732.4 1742.4 32.690 1.081e-08 \*\*\*  
## + VMail.Message 1 1739.7 1749.7 25.408 4.640e-07 \*\*\*  
## + Intl.Calls 1 1753.8 1763.8 11.322 0.0007661 \*\*\*  
## + Night.Mins 1 1757.6 1767.6 7.492 0.0061974 \*\*   
## + Night.Charge 1 1757.6 1767.6 7.492 0.0061991 \*\*   
## + Intl.Charge 1 1758.9 1768.9 6.149 0.0131466 \*   
## + Intl.Mins 1 1759.0 1769.0 6.129 0.0132979 \*   
## <none> 1765.1 1773.1   
## + Day.Calls 1 1764.0 1774.0 1.081 0.2985360   
## + Eve.Calls 1 1764.8 1774.8 0.290 0.5905106   
## + Account.Length 1 1764.8 1774.8 0.226 0.6344662   
## + Night.Calls 1 1765.0 1775.0 0.053 0.8180655   
## + Day.Mins 1 1765.1 1775.1 0.012 0.9118915   
## + Area.Code 1 1765.1 1775.1 0.000 0.9843537   
## - Day.Charge 1 1881.6 1887.6 116.534 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1891.8 1897.8 126.692 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1912.0 1918.0 146.888 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1739.53  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge + Eve.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Plan 1 1694.9 1706.9 34.673 3.901e-09 \*\*\*  
## + VMail.Message 1 1702.4 1714.4 27.124 1.908e-07 \*\*\*  
## + Intl.Calls 1 1717.9 1729.9 11.615 0.0006544 \*\*\*  
## + Night.Charge 1 1721.3 1733.3 8.221 0.0041417 \*\*   
## + Night.Mins 1 1721.3 1733.3 8.221 0.0041418 \*\*   
## + Intl.Charge 1 1722.3 1734.3 7.205 0.0072713 \*\*   
## + Intl.Mins 1 1722.3 1734.3 7.182 0.0073629 \*\*   
## + Eve.Charge 1 1725.8 1737.8 3.751 0.0527873 .   
## <none> 1729.5 1739.5   
## + Day.Calls 1 1727.9 1739.9 1.643 0.1998623   
## + Account.Length 1 1729.1 1741.1 0.444 0.5052375   
## + Eve.Calls 1 1729.1 1741.1 0.442 0.5061619   
## + Night.Calls 1 1729.4 1741.4 0.113 0.7370246   
## + Day.Mins 1 1729.5 1741.5 0.005 0.9411431   
## + Area.Code 1 1729.5 1741.5 0.000 0.9822587   
## - Eve.Mins 1 1765.1 1773.1 35.549 2.487e-09 \*\*\*  
## - Day.Charge 1 1848.0 1856.0 118.458 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1863.4 1871.4 133.913 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1874.1 1882.1 144.539 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1706.86  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge + Eve.Mins +   
## VMail.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Calls 1 1683.8 1697.8 11.020 0.0009014 \*\*\*  
## + Night.Mins 1 1685.7 1699.7 9.138 0.0025031 \*\*   
## + Night.Charge 1 1685.7 1699.7 9.137 0.0025048 \*\*   
## + Intl.Charge 1 1687.1 1701.1 7.732 0.0054249 \*\*   
## + Intl.Mins 1 1687.2 1701.2 7.708 0.0054967 \*\*   
## + VMail.Message 1 1690.6 1704.6 4.261 0.0390068 \*   
## + Eve.Charge 1 1691.7 1705.7 3.196 0.0738006 .   
## <none> 1694.9 1706.9   
## + Day.Calls 1 1693.0 1707.0 1.810 0.1785229   
## + Account.Length 1 1694.3 1708.3 0.570 0.4504224   
## + Eve.Calls 1 1694.4 1708.4 0.457 0.4991319   
## + Night.Calls 1 1694.8 1708.8 0.029 0.8648760   
## + Day.Mins 1 1694.8 1708.8 0.011 0.9179731   
## + Area.Code 1 1694.8 1708.8 0.010 0.9191376   
## - VMail.Plan 1 1729.5 1739.5 34.673 3.901e-09 \*\*\*  
## - Eve.Mins 1 1732.4 1742.4 37.532 8.992e-10 \*\*\*  
## - Day.Charge 1 1814.4 1824.4 119.525 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1828.7 1838.7 133.869 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1844.9 1854.9 150.033 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1697.84  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge + Eve.Mins +   
## VMail.Plan + Intl.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Night.Mins 1 1675.0 1691.0 8.878 0.0028860 \*\*   
## + Night.Charge 1 1675.0 1691.0 8.878 0.0028869 \*\*   
## + Intl.Charge 1 1675.2 1691.2 8.661 0.0032510 \*\*   
## + Intl.Mins 1 1675.2 1691.2 8.633 0.0033007 \*\*   
## + VMail.Message 1 1679.5 1695.5 4.296 0.0382115 \*   
## + Eve.Charge 1 1680.7 1696.7 3.131 0.0768358 .   
## <none> 1683.8 1697.8   
## + Day.Calls 1 1682.2 1698.2 1.651 0.1988774   
## + Account.Length 1 1683.1 1699.1 0.708 0.4002429   
## + Eve.Calls 1 1683.3 1699.3 0.547 0.4597276   
## + Day.Mins 1 1683.8 1699.8 0.035 0.8521482   
## + Area.Code 1 1683.8 1699.8 0.021 0.8855284   
## + Night.Calls 1 1683.8 1699.8 0.010 0.9189817   
## - Intl.Calls 1 1694.9 1706.9 11.020 0.0009014 \*\*\*  
## - VMail.Plan 1 1717.9 1729.9 34.078 5.295e-09 \*\*\*  
## - Eve.Mins 1 1721.5 1733.5 37.616 8.614e-10 \*\*\*  
## - Day.Charge 1 1803.4 1815.4 119.570 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1817.1 1829.1 133.248 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1834.5 1846.5 150.713 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1690.96  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge + Eve.Mins +   
## VMail.Plan + Intl.Calls + Night.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Charge 1 1666.4 1684.4 8.583 0.003394 \*\*   
## + Intl.Mins 1 1666.4 1684.4 8.556 0.003444 \*\*   
## + VMail.Message 1 1670.5 1688.5 4.443 0.035050 \*   
## + Eve.Charge 1 1672.0 1690.0 3.006 0.082942 .   
## <none> 1675.0 1691.0   
## + Day.Calls 1 1673.4 1691.4 1.540 0.214620   
## + Account.Length 1 1674.2 1692.2 0.733 0.391823   
## + Eve.Calls 1 1674.5 1692.5 0.513 0.473792   
## + Day.Mins 1 1674.9 1692.9 0.047 0.828904   
## + Area.Code 1 1674.9 1692.9 0.023 0.880692   
## + Night.Calls 1 1674.9 1692.9 0.020 0.887277   
## + Night.Charge 1 1675.0 1693.0 0.005 0.944534   
## - Night.Mins 1 1683.8 1697.8 8.878 0.002886 \*\*   
## - Intl.Calls 1 1685.7 1699.7 10.760 0.001037 \*\*   
## - VMail.Plan 1 1709.9 1723.9 34.977 3.337e-09 \*\*\*  
## - Eve.Mins 1 1713.6 1727.6 38.640 5.095e-10 \*\*\*  
## - Day.Charge 1 1794.4 1808.4 119.404 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1808.7 1822.7 133.726 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1829.2 1843.2 154.258 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1684.38  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge + Eve.Mins +   
## VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Message 1 1661.7 1681.7 4.711 0.0299631 \*   
## + Eve.Charge 1 1663.8 1683.8 2.603 0.1066287   
## <none> 1666.4 1684.4   
## + Day.Calls 1 1664.9 1684.9 1.486 0.2229084   
## + Intl.Mins 1 1665.0 1685.0 1.405 0.2358998   
## + Account.Length 1 1665.6 1685.6 0.739 0.3899965   
## + Eve.Calls 1 1665.8 1685.8 0.532 0.4656422   
## + Area.Code 1 1666.3 1686.3 0.064 0.7997252   
## + Day.Mins 1 1666.3 1686.3 0.045 0.8314940   
## + Night.Charge 1 1666.4 1686.4 0.012 0.9131135   
## + Night.Calls 1 1666.4 1686.4 0.005 0.9413910   
## - Intl.Charge 1 1675.0 1691.0 8.583 0.0033937 \*\*   
## - Night.Mins 1 1675.2 1691.2 8.800 0.0030124 \*\*   
## - Intl.Calls 1 1678.0 1694.0 11.654 0.0006405 \*\*\*  
## - VMail.Plan 1 1701.6 1717.6 35.194 2.985e-09 \*\*\*  
## - Eve.Mins 1 1706.4 1722.4 40.034 2.496e-10 \*\*\*  
## - Day.Charge 1 1786.2 1802.2 119.835 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1801.8 1817.8 135.437 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1816.9 1832.9 150.481 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1681.67  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge + Eve.Mins +   
## VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge + VMail.Message  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Eve.Charge 1 1658.8 1680.8 2.908 0.0881400 .   
## <none> 1661.7 1681.7   
## + Intl.Mins 1 1660.2 1682.2 1.490 0.2222875   
## + Day.Calls 1 1660.2 1682.2 1.474 0.2246640   
## + Account.Length 1 1660.9 1682.9 0.727 0.3937036   
## + Eve.Calls 1 1661.1 1683.1 0.579 0.4468301   
## + Day.Mins 1 1661.6 1683.6 0.053 0.8172235   
## + Area.Code 1 1661.6 1683.6 0.051 0.8221383   
## + Night.Charge 1 1661.6 1683.6 0.025 0.8750740   
## + Night.Calls 1 1661.7 1683.7 0.000 0.9872236   
## - VMail.Message 1 1666.4 1684.4 4.711 0.0299631 \*   
## - Intl.Charge 1 1670.5 1688.5 8.851 0.0029286 \*\*   
## - Night.Mins 1 1670.7 1688.7 8.992 0.0027118 \*\*   
## - Intl.Calls 1 1673.4 1691.4 11.751 0.0006081 \*\*\*  
## - VMail.Plan 1 1674.3 1692.3 12.604 0.0003850 \*\*\*  
## - Eve.Mins 1 1701.8 1719.8 40.086 2.43e-10 \*\*\*  
## - Day.Charge 1 1781.7 1799.7 120.050 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1796.6 1814.6 134.946 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1811.6 1829.6 149.955 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1680.76  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge + Eve.Mins +   
## VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge + VMail.Message +   
## Eve.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 1658.8 1680.8   
## + Day.Calls 1 1657.3 1681.3 1.496 0.2212401   
## + Intl.Mins 1 1657.3 1681.3 1.494 0.2215277   
## - Eve.Charge 1 1661.7 1681.7 2.908 0.0881400 .   
## - Eve.Mins 1 1661.7 1681.7 2.923 0.0873296 .   
## + Account.Length 1 1658.0 1682.0 0.794 0.3728274   
## + Eve.Calls 1 1658.2 1682.2 0.547 0.4594430   
## + Area.Code 1 1658.7 1682.7 0.090 0.7641064   
## + Day.Mins 1 1658.7 1682.7 0.063 0.8018929   
## + Night.Charge 1 1658.7 1682.7 0.024 0.8760610   
## + Night.Calls 1 1658.8 1682.8 0.000 0.9977127   
## - VMail.Message 1 1663.8 1683.8 5.016 0.0251150 \*   
## - Intl.Charge 1 1667.2 1687.2 8.440 0.0036714 \*\*   
## - Night.Mins 1 1667.6 1687.6 8.851 0.0029297 \*\*   
## - Intl.Calls 1 1670.5 1690.5 11.716 0.0006196 \*\*\*  
## - VMail.Plan 1 1671.8 1691.8 13.012 0.0003095 \*\*\*  
## - Day.Charge 1 1778.2 1798.2 119.450 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1794.5 1814.5 135.711 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1809.0 1829.0 150.205 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Call: glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge +   
## Eve.Mins + VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge +   
## VMail.Message + Eve.Charge, family = binomial(link = "logit"),   
## data = Churn\_train)  
##   
## Coefficients:  
## (Intercept) Int.l.Plan CustServ.Calls Day.Charge   
## -7.951285 2.067146 0.507400 0.074615   
## Eve.Mins VMail.Plan Intl.Calls Night.Mins   
## 3.181745 -2.363100 -0.095363 0.003784   
## Intl.Charge VMail.Message Eve.Charge   
## 0.248350 0.047322 -37.336479   
##   
## Degrees of Freedom: 2539 Total (i.e. Null); 2529 Residual  
## Null Deviance: 2133   
## Residual Deviance: 1659 AIC: 1681

#final model 1 training:  
glm\_model1 <- glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge +   
 Eve.Mins + VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge +   
 VMail.Message + Eve.Charge, family = binomial(link = "logit"),   
 data = Churn\_train)  
  
summary(glm\_model1)

##   
## Call:  
## glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge +   
## Eve.Mins + VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge +   
## VMail.Message + Eve.Charge, family = binomial(link = "logit"),   
## data = Churn\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.1989 -0.5207 -0.3427 -0.1912 2.8986   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -7.951285 0.583028 -13.638 < 2e-16 \*\*\*  
## Int.l.Plan 2.067146 0.166065 12.448 < 2e-16 \*\*\*  
## CustServ.Calls 0.507400 0.044288 11.457 < 2e-16 \*\*\*  
## Day.Charge 0.074615 0.007176 10.398 < 2e-16 \*\*\*  
## Eve.Mins 3.181745 1.864004 1.707 0.087833 .   
## VMail.Plan -2.363100 0.688112 -3.434 0.000594 \*\*\*  
## Intl.Calls -0.095363 0.028684 -3.325 0.000886 \*\*\*  
## Night.Mins 0.003784 0.001276 2.967 0.003008 \*\*   
## Intl.Charge 0.248350 0.086004 2.888 0.003881 \*\*   
## VMail.Message 0.047322 0.021304 2.221 0.026332 \*   
## Eve.Charge -37.336479 21.929129 -1.703 0.088643 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2133.4 on 2539 degrees of freedom  
## Residual deviance: 1658.8 on 2529 degrees of freedom  
## AIC: 1680.8  
##   
## Number of Fisher Scoring iterations: 6

## Model2: Split the data into 80~ 20 having equal proportion of '0' & '1'  
data\_1<-Churn[Churn$Churn==1,]  
set.seed(1)  
nrow(data\_1)

## [1] 483

ind\_1<-sample(1:nrow(data\_1),0.8\*nrow(data\_1))   
  
data\_0<-Churn[Churn$Churn==0,]  
set.seed(1)  
nrow(data\_0)

## [1] 2850

ind\_0<-sample(1:nrow(data\_0),0.8\*nrow(data\_0))  
  
Churn\_train2<-Churn[c(ind\_1,ind\_0) , -c(19, 21) ]  
str(Churn\_train2)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 2666 obs. of 19 variables:  
## $ Account.Length: num 92 70 83 58 73 49 85 128 113 49 ...  
## $ VMail.Message : num 0 0 0 29 0 0 0 0 0 0 ...  
## $ Day.Mins : num 91.7 232.1 221.4 240.4 160.1 ...  
## $ Eve.Mins : num 194 292 232 119 213 ...  
## $ Night.Mins : num 175 201 122 164 174 ...  
## $ Intl.Mins : num 9.2 0 9.8 11.2 13 12.4 12.5 18.9 6.6 11.1 ...  
## $ CustServ.Calls: num 2 3 3 1 0 2 0 0 4 1 ...  
## $ Churn : num 0 0 0 0 0 0 0 0 0 0 ...  
## $ Int.l.Plan : num 0 0 0 0 0 0 0 0 0 0 ...  
## $ VMail.Plan : num 0 0 0 1 0 0 0 0 0 0 ...  
## $ Day.Calls : num 90 122 103 80 110 69 107 99 129 117 ...  
## $ Day.Charge : num 15.6 39.5 37.6 40.9 27.2 ...  
## $ Eve.Calls : num 123 112 103 91 72 92 124 107 117 109 ...  
## $ Eve.Charge : num 16.5 24.9 19.7 10.1 18.1 ...  
## $ Night.Calls : num 86 112 100 108 72 153 76 111 91 90 ...  
## $ Night.Charge : num 7.88 9.05 5.51 7.39 7.83 ...  
## $ Intl.Calls : num 4 0 5 3 4 7 5 2 9 1 ...  
## $ Intl.Charge : num 2.48 0 2.65 3.02 3.51 3.35 3.38 5.1 1.78 3 ...  
## $ Area.Code : num 408 408 415 510 415 510 408 408 415 510 ...

table(Churn\_train2$Churn)

##   
## 0 1   
## 2296 370

Churn\_test2<-Churn[-c(ind\_1,ind\_0),]  
table(Churn\_test2$Churn)

##   
## 0 1   
## 828 156

glm\_model2.null<-glm(Churn~1,family=binomial(link="logit"),data= Churn\_train2)  
summary(glm\_model1.null)

##   
## Call:  
## glm(formula = Churn ~ 1, family = binomial(link = "logit"), data = Churn\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.5669 -0.5669 -0.5669 -0.5669 1.9533   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.74701 0.05581 -31.3 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2133.4 on 2539 degrees of freedom  
## Residual deviance: 2133.4 on 2539 degrees of freedom  
## AIC: 2135.4  
##   
## Number of Fisher Scoring iterations: 4

glm\_model2.full<-glm(Churn ~., family=binomial(link="logit"), data=Churn\_train2)  
summary(glm\_model1.full)

##   
## Call:  
## glm(formula = Churn ~ ., family = binomial(link = "logit"), data = Churn\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.2091 -0.5122 -0.3395 -0.1887 2.8898   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -8.978e+00 1.066e+00 -8.420 < 2e-16 \*\*\*  
## Account.Length 1.392e-03 1.590e-03 0.875 0.381527   
## VMail.Message 4.789e-02 2.135e-02 2.244 0.024843 \*   
## Day.Mins -9.130e-01 3.750e+00 -0.243 0.807645   
## Eve.Mins 3.258e+00 1.870e+00 1.743 0.081415 .   
## Night.Mins 1.603e-01 1.002e+00 0.160 0.872892   
## Intl.Mins -7.161e+00 6.031e+00 -1.187 0.235040   
## CustServ.Calls 5.103e-01 4.452e-02 11.461 < 2e-16 \*\*\*  
## Int.l.Plan 2.067e+00 1.667e-01 12.404 < 2e-16 \*\*\*  
## VMail.Plan -2.391e+00 6.902e-01 -3.465 0.000531 \*\*\*  
## Day.Calls 3.714e-03 3.138e-03 1.183 0.236719   
## Day.Charge 5.445e+00 2.206e+01 0.247 0.805033   
## Eve.Calls 2.325e-03 3.177e-03 0.732 0.464339   
## Eve.Charge -3.824e+01 2.200e+01 -1.738 0.082185 .   
## Night.Calls 2.269e-04 3.273e-03 0.069 0.944730   
## Night.Charge -3.478e+00 2.226e+01 -0.156 0.875839   
## Intl.Calls -9.729e-02 2.885e-02 -3.372 0.000745 \*\*\*  
## Intl.Charge 2.677e+01 2.233e+01 1.199 0.230671   
## Area.Code 5.370e-04 1.483e-03 0.362 0.717163   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2133.4 on 2539 degrees of freedom  
## Residual deviance: 1654.3 on 2521 degrees of freedom  
## AIC: 1692.3  
##   
## Number of Fisher Scoring iterations: 6

#to find the best model with lowest AIC  
step(glm\_model2.null, scope = list(upper=glm\_model1.full), direction="both", test="Chisq", data=Churn\_train)

## Start: AIC=2149.47  
## Churn ~ 1  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Int.l.Plan 1 2013.9 2017.9 133.597 < 2.2e-16 \*\*\*  
## + Day.Mins 1 2021.3 2025.3 126.128 < 2.2e-16 \*\*\*  
## + Day.Charge 1 2021.3 2025.3 126.120 < 2.2e-16 \*\*\*  
## + CustServ.Calls 1 2074.5 2078.5 72.930 < 2.2e-16 \*\*\*  
## + Eve.Mins 1 2121.1 2125.1 26.418 2.749e-07 \*\*\*  
## + Eve.Charge 1 2121.1 2125.1 26.404 2.769e-07 \*\*\*  
## + VMail.Plan 1 2125.4 2129.4 22.017 2.703e-06 \*\*\*  
## + VMail.Message 1 2130.9 2134.9 16.569 4.690e-05 \*\*\*  
## + Intl.Charge 1 2135.7 2139.7 11.768 0.0006025 \*\*\*  
## + Intl.Mins 1 2135.7 2139.7 11.732 0.0006142 \*\*\*  
## + Intl.Calls 1 2137.2 2141.2 10.322 0.0013145 \*\*   
## + Night.Charge 1 2144.1 2148.1 3.360 0.0667807 .   
## + Night.Mins 1 2144.1 2148.1 3.359 0.0668545 .   
## <none> 2147.5 2149.5   
## + Night.Calls 1 2146.6 2150.6 0.919 0.3376301   
## + Day.Calls 1 2146.7 2150.7 0.792 0.3733855   
## + Eve.Calls 1 2146.7 2150.7 0.728 0.3936441   
## + Account.Length 1 2146.8 2150.8 0.631 0.4268465   
## + Area.Code 1 2147.4 2151.4 0.017 0.8951050   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2017.87  
## Churn ~ Int.l.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Day.Mins 1 1900.3 1906.3 113.551 < 2.2e-16 \*\*\*  
## + Day.Charge 1 1900.3 1906.3 113.541 < 2.2e-16 \*\*\*  
## + CustServ.Calls 1 1923.2 1929.2 90.674 < 2.2e-16 \*\*\*  
## + VMail.Plan 1 1988.3 1994.3 25.554 4.301e-07 \*\*\*  
## + Eve.Mins 1 1989.7 1995.7 24.153 8.897e-07 \*\*\*  
## + Eve.Charge 1 1989.7 1995.7 24.139 8.961e-07 \*\*\*  
## + VMail.Message 1 1993.5 1999.5 20.359 6.419e-06 \*\*\*  
## + Intl.Calls 1 2002.5 2008.5 11.331 0.0007621 \*\*\*  
## + Intl.Charge 1 2005.2 2011.2 8.648 0.0032739 \*\*   
## + Intl.Mins 1 2005.3 2011.3 8.604 0.0033545 \*\*   
## + Night.Charge 1 2008.4 2014.4 5.478 0.0192626 \*   
## + Night.Mins 1 2008.4 2014.4 5.475 0.0192887 \*   
## <none> 2013.9 2017.9   
## + Area.Code 1 2013.3 2019.3 0.584 0.4449075   
## + Night.Calls 1 2013.4 2019.4 0.515 0.4730714   
## + Day.Calls 1 2013.4 2019.4 0.443 0.5056162   
## + Eve.Calls 1 2013.4 2019.4 0.434 0.5099458   
## + Account.Length 1 2013.6 2019.6 0.303 0.5820394   
## - Int.l.Plan 1 2147.5 2149.5 133.597 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1906.32  
## Churn ~ Int.l.Plan + Day.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + CustServ.Calls 1 1798.6 1806.6 101.694 < 2.2e-16 \*\*\*  
## + VMail.Plan 1 1873.1 1881.1 27.228 1.808e-07 \*\*\*  
## + Eve.Mins 1 1873.3 1881.3 26.976 2.059e-07 \*\*\*  
## + Eve.Charge 1 1873.4 1881.4 26.963 2.074e-07 \*\*\*  
## + VMail.Message 1 1878.2 1886.2 22.162 2.506e-06 \*\*\*  
## + Intl.Calls 1 1886.6 1894.6 13.685 0.0002162 \*\*\*  
## + Intl.Charge 1 1891.4 1899.4 8.923 0.0028160 \*\*   
## + Intl.Mins 1 1891.4 1899.4 8.882 0.0028794 \*\*   
## + Night.Charge 1 1894.5 1902.5 5.847 0.0156018 \*   
## + Night.Mins 1 1894.5 1902.5 5.844 0.0156264 \*   
## <none> 1900.3 1906.3   
## + Day.Charge 1 1898.5 1906.5 1.785 0.1815646   
## + Night.Calls 1 1899.9 1907.9 0.437 0.5084348   
## + Area.Code 1 1899.9 1907.9 0.435 0.5097362   
## + Account.Length 1 1900.0 1908.0 0.309 0.5784034   
## + Day.Calls 1 1900.1 1908.1 0.237 0.6264213   
## + Eve.Calls 1 1900.2 1908.2 0.078 0.7804789   
## - Day.Mins 1 2013.9 2017.9 113.551 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2021.3 2025.3 121.020 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1806.63  
## Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Eve.Mins 1 1767.1 1777.1 31.547 1.946e-08 \*\*\*  
## + Eve.Charge 1 1767.1 1777.1 31.534 1.959e-08 \*\*\*  
## + VMail.Plan 1 1771.0 1781.0 27.596 1.495e-07 \*\*\*  
## + VMail.Message 1 1776.5 1786.5 22.115 2.567e-06 \*\*\*  
## + Intl.Calls 1 1784.5 1794.5 14.182 0.000166 \*\*\*  
## + Intl.Charge 1 1788.3 1798.3 10.365 0.001284 \*\*   
## + Intl.Mins 1 1788.3 1798.3 10.320 0.001316 \*\*   
## + Night.Charge 1 1791.5 1801.5 7.074 0.007822 \*\*   
## + Night.Mins 1 1791.6 1801.6 7.071 0.007836 \*\*   
## <none> 1798.6 1806.6   
## + Day.Charge 1 1797.0 1807.0 1.680 0.194981   
## + Area.Code 1 1797.7 1807.7 0.913 0.339367   
## + Account.Length 1 1798.4 1808.4 0.209 0.647335   
## + Day.Calls 1 1798.5 1808.5 0.161 0.688312   
## + Night.Calls 1 1798.6 1808.6 0.039 0.842766   
## + Eve.Calls 1 1798.6 1808.6 0.003 0.958129   
## - CustServ.Calls 1 1900.3 1906.3 101.694 < 2.2e-16 \*\*\*  
## - Day.Mins 1 1923.2 1929.2 124.570 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1939.0 1945.0 140.394 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1777.08  
## Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls + Eve.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Plan 1 1738.5 1750.5 28.570 9.035e-08 \*\*\*  
## + VMail.Message 1 1744.0 1756.0 23.122 1.520e-06 \*\*\*  
## + Intl.Calls 1 1752.9 1764.9 14.215 0.0001631 \*\*\*  
## + Intl.Charge 1 1754.8 1766.8 12.319 0.0004484 \*\*\*  
## + Intl.Mins 1 1754.8 1766.8 12.271 0.0004601 \*\*\*  
## + Night.Charge 1 1758.6 1770.6 8.508 0.0035359 \*\*   
## + Night.Mins 1 1758.6 1770.6 8.505 0.0035409 \*\*   
## + Eve.Charge 1 1764.0 1776.0 3.043 0.0810718 .   
## + Day.Charge 1 1765.1 1777.1 2.004 0.1569196   
## <none> 1767.1 1777.1   
## + Area.Code 1 1766.1 1778.1 0.950 0.3297865   
## + Account.Length 1 1766.7 1778.7 0.363 0.5470498   
## + Day.Calls 1 1767.0 1779.0 0.117 0.7326315   
## + Eve.Calls 1 1767.1 1779.1 0.013 0.9082359   
## + Night.Calls 1 1767.1 1779.1 0.012 0.9116943   
## - Eve.Mins 1 1798.6 1806.6 31.547 1.946e-08 \*\*\*  
## - CustServ.Calls 1 1873.3 1881.3 106.265 < 2.2e-16 \*\*\*  
## - Day.Mins 1 1895.6 1903.6 128.548 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1907.0 1915.0 139.964 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1750.51  
## Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls + Eve.Mins + VMail.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Charge 1 1724.6 1738.6 13.923 0.0001904 \*\*\*  
## + Intl.Mins 1 1724.6 1738.6 13.870 0.0001959 \*\*\*  
## + Intl.Calls 1 1725.4 1739.4 13.114 0.0002932 \*\*\*  
## + Night.Charge 1 1729.2 1743.2 9.304 0.0022863 \*\*   
## + Night.Mins 1 1729.2 1743.2 9.302 0.0022896 \*\*   
## + Eve.Charge 1 1736.2 1750.2 2.362 0.1243291   
## + VMail.Message 1 1736.4 1750.4 2.151 0.1424917   
## <none> 1738.5 1750.5   
## + Day.Charge 1 1736.8 1750.8 1.729 0.1885679   
## + Area.Code 1 1737.7 1751.7 0.822 0.3645257   
## + Account.Length 1 1738.1 1752.1 0.423 0.5154419   
## + Day.Calls 1 1738.3 1752.3 0.172 0.6786795   
## + Night.Calls 1 1738.5 1752.5 0.047 0.8279372   
## + Eve.Calls 1 1738.5 1752.5 0.039 0.8437749   
## - VMail.Plan 1 1767.1 1777.1 28.570 9.035e-08 \*\*\*  
## - Eve.Mins 1 1771.0 1781.0 32.522 1.179e-08 \*\*\*  
## - CustServ.Calls 1 1845.8 1855.8 107.283 < 2.2e-16 \*\*\*  
## - Day.Mins 1 1869.4 1879.4 130.876 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1882.0 1892.0 143.491 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1738.59  
## Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls + Eve.Mins + VMail.Plan +   
## Intl.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Calls 1 1710.2 1726.2 14.431 0.0001454 \*\*\*  
## + Night.Charge 1 1714.3 1730.3 10.253 0.0013649 \*\*   
## + Night.Mins 1 1714.3 1730.3 10.250 0.0013668 \*\*   
## + Intl.Mins 1 1721.2 1737.2 3.409 0.0648600 .   
## <none> 1724.6 1738.6   
## + Eve.Charge 1 1722.6 1738.6 1.990 0.1583347   
## + VMail.Message 1 1722.6 1738.6 1.944 0.1632189   
## + Day.Charge 1 1723.0 1739.0 1.536 0.2151901   
## + Area.Code 1 1723.9 1739.9 0.712 0.3986630   
## + Account.Length 1 1724.2 1740.2 0.390 0.5322627   
## + Eve.Calls 1 1724.5 1740.5 0.102 0.7490798   
## + Day.Calls 1 1724.5 1740.5 0.094 0.7591467   
## + Night.Calls 1 1724.5 1740.5 0.055 0.8152951   
## - Intl.Charge 1 1738.5 1750.5 13.923 0.0001904 \*\*\*  
## - VMail.Plan 1 1754.8 1766.8 30.175 3.948e-08 \*\*\*  
## - Eve.Mins 1 1759.5 1771.5 34.949 3.384e-09 \*\*\*  
## - CustServ.Calls 1 1834.5 1846.5 109.870 < 2.2e-16 \*\*\*  
## - Day.Mins 1 1857.0 1869.0 132.407 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1863.5 1875.5 138.960 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1726.16  
## Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls + Eve.Mins + VMail.Plan +   
## Intl.Charge + Intl.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Night.Charge 1 1699.0 1717.0 11.136 0.0008467 \*\*\*  
## + Night.Mins 1 1699.0 1717.0 11.133 0.0008482 \*\*\*  
## + Intl.Mins 1 1706.0 1724.0 4.109 0.0426509 \*   
## + VMail.Message 1 1707.9 1725.9 2.227 0.1355901   
## + Eve.Charge 1 1708.0 1726.0 2.128 0.1445900   
## <none> 1710.2 1726.2   
## + Day.Charge 1 1708.9 1726.9 1.226 0.2682247   
## + Area.Code 1 1709.3 1727.3 0.875 0.3495996   
## + Account.Length 1 1709.7 1727.7 0.506 0.4770370   
## + Day.Calls 1 1710.0 1728.0 0.135 0.7130693   
## + Eve.Calls 1 1710.0 1728.0 0.116 0.7332281   
## + Night.Calls 1 1710.1 1728.1 0.098 0.7544056   
## - Intl.Calls 1 1724.6 1738.6 14.431 0.0001454 \*\*\*  
## - Intl.Charge 1 1725.4 1739.4 15.241 9.463e-05 \*\*\*  
## - VMail.Plan 1 1739.1 1753.1 28.957 7.401e-08 \*\*\*  
## - Eve.Mins 1 1745.0 1759.0 34.869 3.526e-09 \*\*\*  
## - CustServ.Calls 1 1821.2 1835.2 111.052 < 2.2e-16 \*\*\*  
## - Day.Mins 1 1846.0 1860.0 135.813 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1850.1 1864.1 139.913 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1717.02  
## Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls + Eve.Mins + VMail.Plan +   
## Intl.Charge + Intl.Calls + Night.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Mins 1 1695.1 1715.1 3.908 0.0480693 \*   
## + VMail.Message 1 1696.7 1716.7 2.356 0.1247813   
## <none> 1699.0 1717.0   
## + Eve.Charge 1 1697.3 1717.3 1.696 0.1928593   
## + Day.Charge 1 1697.9 1717.9 1.148 0.2839523   
## + Area.Code 1 1698.1 1718.1 0.919 0.3376859   
## + Account.Length 1 1698.5 1718.5 0.545 0.4602202   
## + Day.Calls 1 1698.9 1718.9 0.155 0.6941337   
## + Night.Mins 1 1698.9 1718.9 0.154 0.6951263   
## + Eve.Calls 1 1698.9 1718.9 0.130 0.7186008   
## + Night.Calls 1 1698.9 1718.9 0.080 0.7768654   
## - Night.Charge 1 1710.2 1726.2 11.136 0.0008467 \*\*\*  
## - Intl.Calls 1 1714.3 1730.3 15.314 9.102e-05 \*\*\*  
## - Intl.Charge 1 1715.3 1731.3 16.304 5.395e-05 \*\*\*  
## - VMail.Plan 1 1729.0 1745.0 29.996 4.329e-08 \*\*\*  
## - Eve.Mins 1 1735.8 1751.8 36.806 1.305e-09 \*\*\*  
## - CustServ.Calls 1 1812.8 1828.8 113.826 < 2.2e-16 \*\*\*  
## - Day.Mins 1 1836.5 1852.5 137.498 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1842.5 1858.5 143.443 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1715.11  
## Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls + Eve.Mins + VMail.Plan +   
## Intl.Charge + Intl.Calls + Night.Charge + Intl.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Message 1 1692.7 1714.7 2.430 0.119051   
## <none> 1695.1 1715.1   
## + Eve.Charge 1 1693.4 1715.4 1.689 0.193707   
## + Day.Charge 1 1693.9 1715.9 1.256 0.262408   
## + Area.Code 1 1694.3 1716.3 0.806 0.369405   
## + Account.Length 1 1694.6 1716.6 0.525 0.468614   
## + Day.Calls 1 1694.9 1716.9 0.190 0.662713   
## + Night.Mins 1 1695.0 1717.0 0.142 0.706676   
## - Intl.Mins 1 1699.0 1717.0 3.908 0.048069 \*   
## + Eve.Calls 1 1695.0 1717.0 0.084 0.772052   
## + Night.Calls 1 1695.0 1717.0 0.079 0.778226   
## - Intl.Charge 1 1699.1 1717.1 3.969 0.046344 \*   
## - Night.Charge 1 1706.0 1724.0 10.934 0.000944 \*\*\*  
## - Intl.Calls 1 1711.1 1729.1 16.027 6.245e-05 \*\*\*  
## - VMail.Plan 1 1725.4 1743.4 30.269 3.760e-08 \*\*\*  
## - Eve.Mins 1 1731.9 1749.9 36.745 1.346e-09 \*\*\*  
## - CustServ.Calls 1 1809.1 1827.1 113.977 < 2.2e-16 \*\*\*  
## - Day.Mins 1 1831.4 1849.4 136.305 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1840.7 1858.7 145.587 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1714.68  
## Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls + Eve.Mins + VMail.Plan +   
## Intl.Charge + Intl.Calls + Night.Charge + Intl.Mins + VMail.Message  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 1692.7 1714.7   
## + Eve.Charge 1 1690.8 1714.8 1.831 0.1760366   
## - VMail.Message 1 1695.1 1715.1 2.430 0.1190511   
## + Day.Charge 1 1691.4 1715.4 1.242 0.2650001   
## + Area.Code 1 1692.0 1716.0 0.709 0.3997325   
## + Account.Length 1 1692.2 1716.2 0.529 0.4669634   
## + Day.Calls 1 1692.5 1716.5 0.201 0.6542856   
## + Night.Mins 1 1692.5 1716.5 0.159 0.6905060   
## + Night.Calls 1 1692.6 1716.6 0.100 0.7517426   
## + Eve.Calls 1 1692.6 1716.6 0.096 0.7569781   
## - Intl.Mins 1 1696.7 1716.7 3.981 0.0460144 \*   
## - Intl.Charge 1 1696.7 1716.7 4.043 0.0443545 \*   
## - VMail.Plan 1 1701.1 1721.1 8.451 0.0036488 \*\*   
## - Night.Charge 1 1703.7 1723.7 11.047 0.0008881 \*\*\*  
## - Intl.Calls 1 1709.0 1729.0 16.340 5.292e-05 \*\*\*  
## - Eve.Mins 1 1729.3 1749.3 36.618 1.437e-09 \*\*\*  
## - CustServ.Calls 1 1807.3 1827.3 114.575 < 2.2e-16 \*\*\*  
## - Day.Mins 1 1828.7 1848.7 135.974 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1837.4 1857.4 144.692 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Call: glm(formula = Churn ~ Int.l.Plan + Day.Mins + CustServ.Calls +   
## Eve.Mins + VMail.Plan + Intl.Charge + Intl.Calls + Night.Charge +   
## Intl.Mins + VMail.Message, family = binomial(link = "logit"),   
## data = Churn\_train2)  
##   
## Coefficients:  
## (Intercept) Int.l.Plan Day.Mins CustServ.Calls   
## -8.439786 2.023139 0.013989 0.485545   
## Eve.Mins VMail.Plan Intl.Charge Intl.Calls   
## 0.007611 -1.772952 44.482967 -0.110960   
## Night.Charge Intl.Mins VMail.Message   
## 0.092228 -11.919332 0.031009   
##   
## Degrees of Freedom: 2665 Total (i.e. Null); 2655 Residual  
## Null Deviance: 2147   
## Residual Deviance: 1693 AIC: 1715

#final model 2 training:  
glm\_model2 <- glm(formula = Churn ~ `Intl.Mins` + `Intl.Charge` + `VMail.Plan` + `Night.Charge` + `Intl.Calls` +  
 `Eve.Mins` + `CustServ.Calls` + `Day.Mins` + `Int.l.Plan`, family = binomial(link = "logit"),   
 data = Churn\_train)  
summary(glm\_model2)

##   
## Call:  
## glm(formula = Churn ~ Intl.Mins + Intl.Charge + VMail.Plan +   
## Night.Charge + Intl.Calls + Eve.Mins + CustServ.Calls + Day.Mins +   
## Int.l.Plan, family = binomial(link = "logit"), data = Churn\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.1208 -0.5230 -0.3415 -0.1978 2.8933   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -7.952806 0.580263 -13.706 < 2e-16 \*\*\*  
## Intl.Mins -7.113384 6.005623 -1.184 0.236233   
## Intl.Charge 26.593394 22.241582 1.196 0.231829   
## VMail.Plan -0.923828 0.165371 -5.586 2.32e-08 \*\*\*  
## Night.Charge 0.083304 0.028292 2.944 0.003236 \*\*   
## Intl.Calls -0.095846 0.028627 -3.348 0.000814 \*\*\*  
## Eve.Mins 0.008112 0.001306 6.212 5.25e-10 \*\*\*  
## CustServ.Calls 0.506207 0.044098 11.479 < 2e-16 \*\*\*  
## Day.Mins 0.012695 0.001219 10.413 < 2e-16 \*\*\*  
## Int.l.Plan 2.073378 0.165807 12.505 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2133.4 on 2539 degrees of freedom  
## Residual deviance: 1665.0 on 2530 degrees of freedom  
## AIC: 1685  
##   
## Number of Fisher Scoring iterations: 5

##Model3: undresampling with 483 'Zeros' and 483 'Ones'  
library(ROSE)

## Warning: package 'ROSE' was built under R version 3.5.3

## Loaded ROSE 0.0-3

data\_483 <- ovun.sample(Churn ~ ., data = Churn, method = "under", N = 966, seed = 1)$data  
data\_483 <- data\_483[, -c(19, 21)]  
table(data\_483$Churn)

##   
## 0 1   
## 483 483

glm\_model483.null<-glm(Churn ~ 1,family=binomial(link="logit"),data=data\_483)  
glm\_model483.full<-glm(Churn ~ .,family=binomial(link="logit"),data=data\_483)  
  
#to find the best model with lowest AIC  
step(glm\_model483.null, scope = list(upper=glm\_model483.full), direction="both", test="Chisq", data=data\_483)

## Start: AIC=1341.16  
## Churn ~ 1  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Int.l.Plan 1 1255.0 1259.0 84.133 < 2.2e-16 \*\*\*  
## + Day.Mins 1 1282.0 1286.0 57.134 4.072e-14 \*\*\*  
## + Day.Charge 1 1282.0 1286.0 57.128 4.084e-14 \*\*\*  
## + CustServ.Calls 1 1283.1 1287.1 56.020 7.175e-14 \*\*\*  
## + VMail.Plan 1 1315.8 1319.8 23.391 1.322e-06 \*\*\*  
## + VMail.Message 1 1322.3 1326.3 16.872 3.999e-05 \*\*\*  
## + Eve.Mins 1 1322.5 1326.5 16.616 4.577e-05 \*\*\*  
## + Eve.Charge 1 1322.5 1326.5 16.609 4.593e-05 \*\*\*  
## + Intl.Charge 1 1331.4 1335.4 7.798 0.005230 \*\*   
## + Intl.Mins 1 1331.4 1335.4 7.774 0.005302 \*\*   
## + Day.Calls 1 1333.8 1337.8 5.402 0.020118 \*   
## + Intl.Calls 1 1335.8 1339.8 3.396 0.065373 .   
## <none> 1339.2 1341.2   
## + Account.Length 1 1337.4 1341.4 1.805 0.179110   
## + Night.Charge 1 1337.5 1341.5 1.681 0.194792   
## + Night.Mins 1 1337.5 1341.5 1.680 0.194964   
## + Area.Code 1 1338.8 1342.8 0.392 0.531512   
## + Eve.Calls 1 1339.1 1343.1 0.047 0.829169   
## + Night.Calls 1 1339.2 1343.2 0.014 0.905270   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1259.03  
## Churn ~ Int.l.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + CustServ.Calls 1 1181.6 1187.6 73.443 < 2.2e-16 \*\*\*  
## + Day.Mins 1 1194.2 1200.2 60.803 6.307e-15 \*\*\*  
## + Day.Charge 1 1194.2 1200.2 60.798 6.326e-15 \*\*\*  
## + VMail.Plan 1 1225.2 1231.2 29.869 4.622e-08 \*\*\*  
## + VMail.Message 1 1232.4 1238.4 22.602 1.992e-06 \*\*\*  
## + Eve.Mins 1 1238.5 1244.5 16.527 4.797e-05 \*\*\*  
## + Eve.Charge 1 1238.5 1244.5 16.518 4.819e-05 \*\*\*  
## + Day.Calls 1 1248.6 1254.6 6.391 0.01147 \*   
## + Night.Charge 1 1250.9 1256.9 4.167 0.04122 \*   
## + Night.Mins 1 1250.9 1256.9 4.164 0.04129 \*   
## + Intl.Calls 1 1251.9 1257.9 3.090 0.07877 .   
## + Intl.Charge 1 1252.5 1258.5 2.513 0.11295   
## + Intl.Mins 1 1252.5 1258.5 2.493 0.11438   
## <none> 1255.0 1259.0   
## + Account.Length 1 1254.0 1260.0 1.004 0.31643   
## + Eve.Calls 1 1254.9 1260.9 0.130 0.71873   
## + Night.Calls 1 1255.0 1261.0 0.031 0.86098   
## + Area.Code 1 1255.0 1261.0 0.007 0.93380   
## - Int.l.Plan 1 1339.2 1341.2 84.133 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1187.58  
## Churn ~ Int.l.Plan + CustServ.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Day.Mins 1 1077.6 1085.6 103.942 < 2.2e-16 \*\*\*  
## + Day.Charge 1 1077.7 1085.7 103.934 < 2.2e-16 \*\*\*  
## + VMail.Plan 1 1150.1 1158.1 31.441 2.056e-08 \*\*\*  
## + Eve.Mins 1 1154.4 1162.4 27.198 1.837e-07 \*\*\*  
## + Eve.Charge 1 1154.4 1162.4 27.187 1.847e-07 \*\*\*  
## + VMail.Message 1 1156.8 1164.8 24.817 6.304e-07 \*\*\*  
## + Night.Charge 1 1174.8 1182.8 6.807 0.009082 \*\*   
## + Night.Mins 1 1174.8 1182.8 6.805 0.009092 \*\*   
## + Day.Calls 1 1174.9 1182.9 6.641 0.009965 \*\*   
## + Intl.Charge 1 1177.8 1185.8 3.775 0.052030 .   
## + Intl.Mins 1 1177.8 1185.8 3.747 0.052916 .   
## + Intl.Calls 1 1179.2 1187.2 2.427 0.119267   
## <none> 1181.6 1187.6   
## + Account.Length 1 1180.6 1188.6 0.976 0.323239   
## + Area.Code 1 1181.3 1189.3 0.236 0.626952   
## + Night.Calls 1 1181.3 1189.3 0.233 0.629603   
## + Eve.Calls 1 1181.5 1189.5 0.084 0.771912   
## - CustServ.Calls 1 1255.0 1259.0 73.443 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1283.1 1287.1 101.557 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1085.64  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Plan 1 1057.5 1067.5 20.095 7.370e-06 \*\*\*  
## + Eve.Mins 1 1058.5 1068.5 19.138 1.216e-05 \*\*\*  
## + Eve.Charge 1 1058.5 1068.5 19.130 1.221e-05 \*\*\*  
## + VMail.Message 1 1062.7 1072.7 14.941 0.0001109 \*\*\*  
## + Night.Charge 1 1071.2 1081.2 6.469 0.0109794 \*   
## + Night.Mins 1 1071.2 1081.2 6.468 0.0109826 \*   
## + Intl.Charge 1 1072.5 1082.5 5.197 0.0226261 \*   
## + Intl.Mins 1 1072.5 1082.5 5.169 0.0229971 \*   
## + Intl.Calls 1 1072.6 1082.6 5.065 0.0244071 \*   
## + Day.Calls 1 1073.5 1083.5 4.185 0.0407873 \*   
## + Day.Charge 1 1075.6 1085.6 2.074 0.1497905   
## <none> 1077.6 1085.6   
## + Account.Length 1 1077.0 1087.0 0.633 0.4263047   
## + Area.Code 1 1077.2 1087.2 0.417 0.5182345   
## + Eve.Calls 1 1077.5 1087.5 0.136 0.7120295   
## + Night.Calls 1 1077.6 1087.6 0.016 0.9000326   
## - Day.Mins 1 1181.6 1187.6 103.942 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1192.1 1198.1 114.445 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1194.2 1200.2 116.582 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1067.55  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Eve.Mins 1 1037.9 1049.9 19.687 9.124e-06 \*\*\*  
## + Eve.Charge 1 1037.9 1049.9 19.681 9.151e-06 \*\*\*  
## + Night.Mins 1 1051.6 1063.6 5.976 0.01450 \*   
## + Night.Charge 1 1051.6 1063.6 5.975 0.01451 \*   
## + Intl.Charge 1 1051.7 1063.7 5.818 0.01586 \*   
## + Intl.Mins 1 1051.8 1063.8 5.789 0.01612 \*   
## + Day.Calls 1 1053.1 1065.1 4.421 0.03551 \*   
## + Intl.Calls 1 1053.8 1065.8 3.745 0.05298 .   
## + VMail.Message 1 1054.5 1066.5 2.995 0.08351 .   
## + Day.Charge 1 1055.3 1067.3 2.262 0.13258   
## <none> 1057.5 1067.5   
## + Account.Length 1 1056.9 1068.9 0.634 0.42590   
## + Area.Code 1 1057.2 1069.2 0.366 0.54494   
## + Night.Calls 1 1057.5 1069.5 0.055 0.81497   
## + Eve.Calls 1 1057.5 1069.5 0.014 0.90513   
## - VMail.Plan 1 1077.6 1085.6 20.095 7.370e-06 \*\*\*  
## - Day.Mins 1 1150.1 1158.1 92.595 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1173.2 1181.2 115.683 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1177.5 1185.5 119.948 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1049.86  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Charge 1 1031.7 1045.7 6.193 0.01283 \*   
## + Intl.Mins 1 1031.7 1045.7 6.162 0.01305 \*   
## + Night.Charge 1 1032.6 1046.6 5.258 0.02185 \*   
## + Night.Mins 1 1032.6 1046.6 5.258 0.02185 \*   
## + Day.Calls 1 1033.8 1047.8 4.053 0.04410 \*   
## + VMail.Message 1 1034.3 1048.3 3.549 0.05957 .   
## + Intl.Calls 1 1034.5 1048.5 3.403 0.06509 .   
## + Day.Charge 1 1035.7 1049.7 2.205 0.13758   
## <none> 1037.9 1049.9   
## + Eve.Charge 1 1037.0 1051.0 0.867 0.35190   
## + Account.Length 1 1037.2 1051.2 0.643 0.42267   
## + Area.Code 1 1037.5 1051.5 0.351 0.55334   
## + Eve.Calls 1 1037.9 1051.9 0.001 0.98136   
## + Night.Calls 1 1037.9 1051.9 0.000 0.99320   
## - Eve.Mins 1 1057.5 1067.5 19.687 9.124e-06 \*\*\*  
## - VMail.Plan 1 1058.5 1068.5 20.644 5.532e-06 \*\*\*  
## - Day.Mins 1 1123.0 1133.0 85.113 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1159.9 1169.9 122.062 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1162.9 1172.9 125.066 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1045.67  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Night.Mins 1 1026.1 1042.1 5.534 0.01865 \*   
## + Night.Charge 1 1026.1 1042.1 5.534 0.01865 \*   
## + Intl.Calls 1 1027.2 1043.2 4.501 0.03387 \*   
## + Day.Calls 1 1027.6 1043.6 4.097 0.04296 \*   
## + VMail.Message 1 1028.3 1044.3 3.357 0.06692 .   
## + Intl.Mins 1 1029.0 1045.0 2.623 0.10530   
## + Day.Charge 1 1029.4 1045.4 2.293 0.12998   
## <none> 1031.7 1045.7   
## + Eve.Charge 1 1030.7 1046.7 1.018 0.31303   
## + Account.Length 1 1031.0 1047.0 0.656 0.41800   
## + Area.Code 1 1031.3 1047.3 0.375 0.54009   
## + Night.Calls 1 1031.7 1047.7 0.003 0.95552   
## + Eve.Calls 1 1031.7 1047.7 0.000 0.99246   
## - Intl.Charge 1 1037.9 1049.9 6.193 0.01283 \*   
## - Eve.Mins 1 1051.7 1063.7 20.061 7.500e-06 \*\*\*  
## - VMail.Plan 1 1053.0 1065.0 21.340 3.846e-06 \*\*\*  
## - Day.Mins 1 1118.3 1130.3 86.606 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1145.4 1157.4 113.708 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1159.5 1171.5 127.822 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1042.13  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Calls 1 1021.2 1039.2 4.879 0.02718 \*   
## + Day.Calls 1 1022.0 1040.0 4.148 0.04169 \*   
## + VMail.Message 1 1022.6 1040.6 3.548 0.05962 .   
## + Intl.Mins 1 1023.4 1041.4 2.692 0.10086   
## + Day.Charge 1 1023.9 1041.9 2.202 0.13781   
## <none> 1026.1 1042.1   
## + Eve.Charge 1 1025.2 1043.2 0.949 0.32986   
## + Account.Length 1 1025.4 1043.4 0.727 0.39377   
## + Area.Code 1 1025.7 1043.7 0.443 0.50575   
## + Night.Calls 1 1026.1 1044.1 0.032 0.85912   
## + Night.Charge 1 1026.1 1044.1 0.007 0.93328   
## + Eve.Calls 1 1026.1 1044.1 0.000 0.98381   
## - Night.Mins 1 1031.7 1045.7 5.534 0.01865 \*   
## - Intl.Charge 1 1032.6 1046.6 6.470 0.01097 \*   
## - Eve.Mins 1 1045.4 1059.4 19.291 1.122e-05 \*\*\*  
## - VMail.Plan 1 1047.0 1061.0 20.826 5.028e-06 \*\*\*  
## - Day.Mins 1 1112.4 1126.4 86.253 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1143.0 1157.0 116.900 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1156.9 1170.9 130.756 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1039.25  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Mins + Intl.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Message 1 1017.5 1037.5 3.803 0.05117 .   
## + Day.Calls 1 1017.5 1037.5 3.769 0.05221 .   
## + Intl.Mins 1 1018.4 1038.4 2.846 0.09159 .   
## + Day.Charge 1 1019.1 1039.1 2.141 0.14340   
## <none> 1021.2 1039.2   
## + Eve.Charge 1 1020.4 1040.4 0.836 0.36063   
## + Account.Length 1 1020.4 1040.4 0.823 0.36434   
## + Area.Code 1 1020.8 1040.8 0.509 0.47554   
## + Night.Charge 1 1021.2 1041.2 0.027 0.86835   
## + Night.Calls 1 1021.2 1041.2 0.009 0.92624   
## + Eve.Calls 1 1021.2 1041.2 0.003 0.96001   
## - Intl.Calls 1 1026.1 1042.1 4.879 0.02718 \*   
## - Night.Mins 1 1027.2 1043.2 5.912 0.01504 \*   
## - Intl.Charge 1 1028.9 1044.9 7.684 0.00557 \*\*   
## - Eve.Mins 1 1040.0 1056.0 18.791 1.458e-05 \*\*\*  
## - VMail.Plan 1 1040.5 1056.5 19.212 1.170e-05 \*\*\*  
## - Day.Mins 1 1110.5 1126.5 89.205 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1136.7 1152.7 115.467 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1152.1 1168.1 130.828 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1037.45  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Mins + Intl.Calls + VMail.Message  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Day.Calls 1 1013.7 1035.7 3.737 0.053213 .   
## + Intl.Mins 1 1014.6 1036.6 2.824 0.092869 .   
## + Day.Charge 1 1015.1 1037.1 2.370 0.123698   
## <none> 1017.5 1037.5   
## + Eve.Charge 1 1016.5 1038.5 0.933 0.334165   
## + Account.Length 1 1016.5 1038.5 0.904 0.341733   
## + Area.Code 1 1016.9 1038.9 0.560 0.454437   
## - VMail.Message 1 1021.2 1039.2 3.803 0.051170 .   
## + Eve.Calls 1 1017.4 1039.4 0.012 0.911555   
## + Night.Charge 1 1017.5 1039.5 0.001 0.972223   
## + Night.Calls 1 1017.5 1039.5 0.000 0.993629   
## - Intl.Calls 1 1022.6 1040.6 5.134 0.023461 \*   
## - Night.Mins 1 1023.6 1041.6 6.168 0.013007 \*   
## - Intl.Charge 1 1025.0 1043.0 7.512 0.006131 \*\*   
## - VMail.Plan 1 1026.7 1044.7 9.223 0.002389 \*\*   
## - Eve.Mins 1 1036.8 1054.8 19.307 1.113e-05 \*\*\*  
## - Day.Mins 1 1106.9 1124.9 89.470 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1131.3 1149.3 113.833 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1146.5 1164.5 129.041 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1035.71  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Mins + Intl.Calls + VMail.Message +   
## Day.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Mins 1 1011.1 1035.1 2.652 0.103412   
## + Day.Charge 1 1011.4 1035.4 2.303 0.129116   
## <none> 1013.7 1035.7   
## + Eve.Charge 1 1012.8 1036.8 0.948 0.330222   
## + Account.Length 1 1013.0 1037.0 0.724 0.394950   
## + Area.Code 1 1013.2 1037.2 0.487 0.485315   
## - Day.Calls 1 1017.5 1037.5 3.737 0.053213 .   
## - VMail.Message 1 1017.5 1037.5 3.771 0.052149 .   
## + Eve.Calls 1 1013.7 1037.7 0.019 0.890397   
## + Night.Charge 1 1013.7 1037.7 0.000 0.995557   
## + Night.Calls 1 1013.7 1037.7 0.000 0.996504   
## - Intl.Calls 1 1018.5 1038.5 4.760 0.029131 \*   
## - Night.Mins 1 1019.9 1039.9 6.221 0.012623 \*   
## - Intl.Charge 1 1021.1 1041.2 7.431 0.006410 \*\*   
## - VMail.Plan 1 1022.9 1042.9 9.187 0.002438 \*\*   
## - Eve.Mins 1 1032.8 1052.8 19.041 1.279e-05 \*\*\*  
## - Day.Mins 1 1100.5 1120.5 86.794 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1128.1 1148.1 114.357 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1142.5 1162.5 128.769 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1035.06  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Mins + Intl.Calls + VMail.Message +   
## Day.Calls + Intl.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Day.Charge 1 1008.8 1034.8 2.308 0.12871   
## <none> 1011.1 1035.1   
## - Intl.Mins 1 1013.7 1035.7 2.652 0.10341   
## - Intl.Charge 1 1013.8 1035.8 2.686 0.10122   
## + Eve.Charge 1 1010.2 1036.2 0.903 0.34189   
## + Account.Length 1 1010.3 1036.3 0.774 0.37911   
## - Day.Calls 1 1014.6 1036.6 3.565 0.05900 .   
## + Area.Code 1 1010.7 1036.7 0.406 0.52399   
## - VMail.Message 1 1014.8 1036.8 3.744 0.05298 .   
## + Eve.Calls 1 1011.0 1037.0 0.026 0.87094   
## + Night.Charge 1 1011.1 1037.1 0.002 0.96891   
## + Night.Calls 1 1011.1 1037.1 0.000 0.99677   
## - Intl.Calls 1 1016.0 1038.0 4.905 0.02678 \*   
## - Night.Mins 1 1017.4 1039.4 6.295 0.01211 \*   
## - VMail.Plan 1 1020.2 1042.2 9.134 0.00251 \*\*   
## - Eve.Mins 1 1030.1 1052.1 19.062 1.265e-05 \*\*\*  
## - Day.Mins 1 1097.1 1119.1 85.995 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1126.8 1148.8 115.744 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1140.5 1162.5 129.421 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1034.75  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Mins + Intl.Calls + VMail.Message +   
## Day.Calls + Intl.Mins + Day.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 1008.8 1034.8   
## - Day.Charge 1 1011.1 1035.1 2.308 0.128709   
## - Day.Mins 1 1011.1 1035.1 2.316 0.128043   
## - Intl.Mins 1 1011.4 1035.4 2.657 0.103094   
## - Intl.Charge 1 1011.5 1035.5 2.691 0.100896   
## + Eve.Charge 1 1007.9 1035.9 0.895 0.344030   
## + Account.Length 1 1008.0 1036.0 0.776 0.378251   
## - Day.Calls 1 1012.2 1036.2 3.495 0.061540 .   
## + Area.Code 1 1008.4 1036.4 0.356 0.551014   
## + Eve.Calls 1 1008.7 1036.7 0.021 0.884949   
## - VMail.Message 1 1012.8 1036.8 3.991 0.045743 \*   
## + Night.Charge 1 1008.8 1036.8 0.008 0.927664   
## + Night.Calls 1 1008.8 1036.8 0.005 0.945259   
## - Intl.Calls 1 1013.6 1037.7 4.896 0.026916 \*   
## - Night.Mins 1 1015.0 1039.0 6.214 0.012674 \*   
## - VMail.Plan 1 1018.3 1042.3 9.537 0.002014 \*\*   
## - Eve.Mins 1 1027.7 1051.7 18.964 1.332e-05 \*\*\*  
## - Int.l.Plan 1 1124.8 1148.8 116.094 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1137.8 1161.8 128.994 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Call: glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins +   
## VMail.Plan + Eve.Mins + Intl.Charge + Night.Mins + Intl.Calls +   
## VMail.Message + Day.Calls + Intl.Mins + Day.Charge, family = binomial(link = "logit"),   
## data = data\_483)  
##   
## Coefficients:  
## (Intercept) Int.l.Plan CustServ.Calls Day.Mins   
## -6.968886 2.324890 0.585793 6.977590   
## VMail.Plan Eve.Mins Intl.Charge Night.Mins   
## -2.229553 0.006642 44.295194 0.004021   
## Intl.Calls VMail.Message Day.Calls Intl.Mins   
## -0.070347 0.046814 0.007004 -11.883085   
## Day.Charge   
## -40.973533   
##   
## Degrees of Freedom: 965 Total (i.e. Null); 953 Residual  
## Null Deviance: 1339   
## Residual Deviance: 1009 AIC: 1035

#final model 3 training:  
glm\_model3<- glm(formula = Churn ~ `Day.Calls` + `VMail.Message` + `Intl.Calls` + `Night.Mins` + `VMail.Plan` +  
 `Eve.Mins` + `Int.l.Plan` + `CustServ.Calls`, family = binomial(link = "logit"),   
 data = data\_483)  
summary(glm\_model3)

##   
## Call:  
## glm(formula = Churn ~ Day.Calls + VMail.Message + Intl.Calls +   
## Night.Mins + VMail.Plan + Eve.Mins + Int.l.Plan + CustServ.Calls,   
## family = binomial(link = "logit"), data = data\_483)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.60392 -0.96323 -0.08139 0.97784 2.47548   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.956845 0.614715 -6.437 1.22e-10 \*\*\*  
## Day.Calls 0.008840 0.003530 2.505 0.01226 \*   
## VMail.Message 0.042604 0.022620 1.883 0.05964 .   
## Intl.Calls -0.034668 0.030026 -1.155 0.24826   
## Night.Mins 0.003692 0.001509 2.446 0.01443 \*   
## VMail.Plan -2.230831 0.707462 -3.153 0.00161 \*\*   
## Eve.Mins 0.007440 0.001476 5.042 4.61e-07 \*\*\*  
## Int.l.Plan 2.174433 0.230956 9.415 < 2e-16 \*\*\*  
## CustServ.Calls 0.444131 0.051495 8.625 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 1339.2 on 965 degrees of freedom  
## Residual deviance: 1106.1 on 957 degrees of freedom  
## AIC: 1124.1  
##   
## Number of Fisher Scoring iterations: 4

# Model 4-------------------------------------------------------------------------  
# undersampling with 1449 zeros and 483 ones  
data\_1449 <- ovun.sample(Churn ~ ., data = Churn, method = "under", N = 1932, seed = 1)$data  
data\_1449 <- data\_1449[, -c(19, 21)]  
table(data\_1449$Churn)

##   
## 0 1   
## 1449 483

glm\_model1449.null<-glm(Churn ~ 1,family=binomial(link="logit"),data=data\_1449)  
glm\_model1449.full<-glm(Churn ~ .,family=binomial(link="logit"),data=data\_1449)  
  
#to find the best model with lowest AIC  
step(glm\_model1449.null, scope = list(upper=glm\_model1449.full), direction="both", test="Chisq", data=data\_1449)

## Start: AIC=2174.86  
## Churn ~ 1  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Int.l.Plan 1 2033.2 2037.2 139.692 < 2.2e-16 \*\*\*  
## + Day.Mins 1 2063.1 2067.1 109.806 < 2.2e-16 \*\*\*  
## + Day.Charge 1 2063.1 2067.1 109.802 < 2.2e-16 \*\*\*  
## + CustServ.Calls 1 2069.1 2073.1 103.782 < 2.2e-16 \*\*\*  
## + Eve.Mins 1 2141.4 2145.4 31.502 1.992e-08 \*\*\*  
## + Eve.Charge 1 2141.4 2145.4 31.492 2.003e-08 \*\*\*  
## + VMail.Plan 1 2144.8 2148.8 28.079 1.165e-07 \*\*\*  
## + VMail.Message 1 2151.1 2155.1 21.759 3.092e-06 \*\*\*  
## + Intl.Charge 1 2159.5 2163.5 13.369 0.0002558 \*\*\*  
## + Intl.Mins 1 2159.5 2163.5 13.352 0.0002582 \*\*\*  
## + Intl.Calls 1 2165.2 2169.2 7.700 0.0055222 \*\*   
## + Night.Charge 1 2168.9 2172.9 3.909 0.0480334 \*   
## + Night.Mins 1 2169.0 2173.0 3.907 0.0480720 \*   
## <none> 2172.9 2174.9   
## + Day.Calls 1 2172.0 2176.0 0.871 0.3505454   
## + Eve.Calls 1 2172.2 2176.2 0.637 0.4248393   
## + Area.Code 1 2172.4 2176.4 0.483 0.4871900   
## + Account.Length 1 2172.6 2176.6 0.242 0.6225569   
## + Night.Calls 1 2172.8 2176.8 0.048 0.8258419   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2037.17  
## Churn ~ Int.l.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + CustServ.Calls 1 1910.9 1916.9 122.278 < 2.2e-16 \*\*\*  
## + Day.Mins 1 1928.6 1934.6 104.605 < 2.2e-16 \*\*\*  
## + Day.Charge 1 1928.6 1934.6 104.601 < 2.2e-16 \*\*\*  
## + VMail.Plan 1 1997.9 2003.9 35.304 2.820e-09 \*\*\*  
## + Eve.Mins 1 2002.8 2008.8 30.416 3.486e-08 \*\*\*  
## + Eve.Charge 1 2002.8 2008.8 30.405 3.506e-08 \*\*\*  
## + VMail.Message 1 2005.2 2011.2 27.994 1.217e-07 \*\*\*  
## + Intl.Charge 1 2025.4 2031.4 7.748 0.005377 \*\*   
## + Intl.Mins 1 2025.4 2031.4 7.726 0.005442 \*\*   
## + Intl.Calls 1 2026.2 2032.2 6.978 0.008251 \*\*   
## + Night.Charge 1 2027.3 2033.3 5.915 0.015010 \*   
## + Night.Mins 1 2027.3 2033.3 5.913 0.015027 \*   
## <none> 2033.2 2037.2   
## + Day.Calls 1 2032.2 2038.2 0.975 0.323486   
## + Eve.Calls 1 2032.5 2038.5 0.726 0.394206   
## + Night.Calls 1 2033.1 2039.1 0.053 0.818105   
## + Account.Length 1 2033.2 2039.2 0.026 0.872409   
## + Area.Code 1 2033.2 2039.2 0.011 0.914734   
## - Int.l.Plan 1 2172.9 2174.9 139.692 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1916.89  
## Churn ~ Int.l.Plan + CustServ.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Day.Mins 1 1778.3 1786.3 132.576 < 2.2e-16 \*\*\*  
## + Day.Charge 1 1778.3 1786.3 132.571 < 2.2e-16 \*\*\*  
## + Eve.Mins 1 1873.2 1881.2 37.696 8.269e-10 \*\*\*  
## + Eve.Charge 1 1873.2 1881.2 37.683 8.324e-10 \*\*\*  
## + VMail.Plan 1 1873.6 1881.6 37.295 1.016e-09 \*\*\*  
## + VMail.Message 1 1880.9 1888.9 30.010 4.298e-08 \*\*\*  
## + Intl.Charge 1 1902.1 1910.1 8.813 0.002992 \*\*   
## + Intl.Mins 1 1902.1 1910.1 8.790 0.003029 \*\*   
## + Night.Charge 1 1902.3 1910.3 8.571 0.003416 \*\*   
## + Night.Mins 1 1902.3 1910.3 8.569 0.003419 \*\*   
## + Intl.Calls 1 1905.4 1913.4 5.462 0.019429 \*   
## <none> 1910.9 1916.9   
## + Day.Calls 1 1909.6 1917.6 1.334 0.248081   
## + Eve.Calls 1 1910.2 1918.2 0.639 0.423938   
## + Night.Calls 1 1910.6 1918.6 0.282 0.595171   
## + Account.Length 1 1910.9 1918.9 0.023 0.879157   
## + Area.Code 1 1910.9 1918.9 0.021 0.883510   
## - CustServ.Calls 1 2033.2 2037.2 122.278 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2069.1 2073.1 158.189 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1786.32  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Plan 1 1743.4 1753.4 34.960 3.366e-09 \*\*\*  
## + Eve.Mins 1 1747.6 1757.6 30.682 3.039e-08 \*\*\*  
## + Eve.Charge 1 1747.6 1757.6 30.672 3.055e-08 \*\*\*  
## + VMail.Message 1 1749.8 1759.8 28.530 9.224e-08 \*\*\*  
## + Intl.Charge 1 1767.6 1777.6 10.681 0.001082 \*\*   
## + Intl.Mins 1 1767.7 1777.7 10.664 0.001092 \*\*   
## + Night.Charge 1 1769.8 1779.8 8.552 0.003452 \*\*   
## + Night.Mins 1 1769.8 1779.8 8.550 0.003456 \*\*   
## + Intl.Calls 1 1771.4 1781.4 6.927 0.008492 \*\*   
## <none> 1778.3 1786.3   
## + Day.Calls 1 1777.2 1787.2 1.168 0.279797   
## + Day.Charge 1 1777.9 1787.9 0.433 0.510298   
## + Account.Length 1 1778.1 1788.1 0.190 0.662958   
## + Eve.Calls 1 1778.2 1788.2 0.144 0.703959   
## + Night.Calls 1 1778.2 1788.2 0.091 0.763450   
## + Area.Code 1 1778.3 1788.3 0.006 0.939091   
## - Day.Mins 1 1910.9 1916.9 132.576 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1928.6 1934.6 150.248 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1933.9 1939.9 155.546 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1753.36  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Eve.Mins 1 1710.7 1722.7 32.628 1.116e-08 \*\*\*  
## + Eve.Charge 1 1710.7 1722.7 32.619 1.121e-08 \*\*\*  
## + Intl.Charge 1 1731.7 1743.7 11.673 0.0006340 \*\*\*  
## + Intl.Mins 1 1731.7 1743.7 11.656 0.0006398 \*\*\*  
## + Night.Charge 1 1735.1 1747.1 8.290 0.0039872 \*\*   
## + Night.Mins 1 1735.1 1747.1 8.289 0.0039891 \*\*   
## + Intl.Calls 1 1737.5 1749.5 5.911 0.0150447 \*   
## + VMail.Message 1 1740.7 1752.7 2.673 0.1020427   
## <none> 1743.4 1753.4   
## + Day.Calls 1 1742.3 1754.3 1.051 0.3051707   
## + Eve.Calls 1 1743.1 1755.1 0.222 0.6372750   
## + Day.Charge 1 1743.2 1755.2 0.170 0.6803958   
## + Account.Length 1 1743.2 1755.2 0.119 0.7306552   
## + Night.Calls 1 1743.2 1755.2 0.109 0.7407283   
## + Area.Code 1 1743.3 1755.3 0.008 0.9274778   
## - VMail.Plan 1 1778.3 1786.3 34.960 3.366e-09 \*\*\*  
## - Day.Mins 1 1873.6 1881.6 130.241 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1896.8 1904.8 153.407 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1906.4 1914.4 163.051 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1722.73  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Charge 1 1697.2 1711.2 13.543 0.0002331 \*\*\*  
## + Intl.Mins 1 1697.2 1711.2 13.525 0.0002354 \*\*\*  
## + Night.Charge 1 1701.1 1715.1 9.675 0.0018683 \*\*   
## + Night.Mins 1 1701.1 1715.1 9.673 0.0018697 \*\*   
## + Intl.Calls 1 1704.6 1718.6 6.169 0.0129974 \*   
## + VMail.Message 1 1707.5 1721.5 3.281 0.0700689 .   
## <none> 1710.7 1722.7   
## + Eve.Charge 1 1709.5 1723.5 1.180 0.2774183   
## + Day.Calls 1 1709.6 1723.6 1.138 0.2860271   
## + Eve.Calls 1 1710.4 1724.4 0.342 0.5584091   
## + Day.Charge 1 1710.5 1724.5 0.254 0.6143040   
## + Account.Length 1 1710.7 1724.7 0.068 0.7949460   
## + Night.Calls 1 1710.7 1724.7 0.058 0.8097638   
## + Area.Code 1 1710.7 1724.7 0.007 0.9312943   
## - Eve.Mins 1 1743.4 1753.4 32.628 1.116e-08 \*\*\*  
## - VMail.Plan 1 1747.6 1757.6 36.905 1.240e-09 \*\*\*  
## - Day.Mins 1 1833.8 1843.8 123.033 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1871.2 1881.2 160.516 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1875.3 1885.3 164.557 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1711.19  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Night.Charge 1 1687.6 1703.6 9.607 0.0019381 \*\*   
## + Night.Mins 1 1687.6 1703.6 9.606 0.0019391 \*\*   
## + Intl.Calls 1 1689.5 1705.5 7.656 0.0056581 \*\*   
## + VMail.Message 1 1693.9 1709.9 3.255 0.0712074 .   
## <none> 1697.2 1711.2   
## + Day.Calls 1 1696.1 1712.1 1.101 0.2940275   
## + Eve.Charge 1 1696.2 1712.2 0.944 0.3313238   
## + Intl.Mins 1 1696.8 1712.8 0.426 0.5138852   
## + Eve.Calls 1 1696.8 1712.8 0.338 0.5610636   
## + Day.Charge 1 1696.9 1712.9 0.253 0.6151671   
## + Night.Calls 1 1697.1 1713.1 0.092 0.7620901   
## + Account.Length 1 1697.1 1713.1 0.052 0.8188616   
## + Area.Code 1 1697.2 1713.2 0.024 0.8777733   
## - Intl.Charge 1 1710.7 1722.7 13.543 0.0002331 \*\*\*  
## - Eve.Mins 1 1731.7 1743.7 34.498 4.268e-09 \*\*\*  
## - VMail.Plan 1 1735.2 1747.2 38.022 6.996e-10 \*\*\*  
## - Day.Mins 1 1823.1 1835.1 125.869 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1854.4 1866.4 157.199 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1860.5 1872.5 163.296 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1703.58  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Calls 1 1679.8 1697.8 7.764 0.0053288 \*\*   
## + VMail.Message 1 1684.0 1702.0 3.556 0.0593354 .   
## <none> 1687.6 1703.6   
## + Day.Calls 1 1686.6 1704.6 0.962 0.3266421   
## + Eve.Charge 1 1686.8 1704.8 0.780 0.3771343   
## + Intl.Mins 1 1687.2 1705.2 0.402 0.5258131   
## + Eve.Calls 1 1687.2 1705.2 0.384 0.5353090   
## + Day.Charge 1 1687.4 1705.4 0.145 0.7034060   
## + Night.Calls 1 1687.5 1705.5 0.074 0.7858911   
## + Account.Length 1 1687.5 1705.5 0.045 0.8322633   
## + Area.Code 1 1687.6 1705.6 0.014 0.9057736   
## + Night.Mins 1 1687.6 1705.6 0.012 0.9136144   
## - Night.Charge 1 1697.2 1711.2 9.607 0.0019381 \*\*   
## - Intl.Charge 1 1701.1 1715.1 13.476 0.0002417 \*\*\*  
## - Eve.Mins 1 1723.4 1737.4 35.806 2.180e-09 \*\*\*  
## - VMail.Plan 1 1725.4 1739.4 37.818 7.767e-10 \*\*\*  
## - Day.Mins 1 1812.9 1826.9 125.306 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1849.1 1863.1 161.513 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1855.0 1869.0 167.406 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1697.82  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Charge + Intl.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Message 1 1676.1 1696.1 3.707 0.0542004 .   
## <none> 1679.8 1697.8   
## + Day.Calls 1 1678.9 1698.9 0.871 0.3507941   
## + Eve.Charge 1 1679.1 1699.1 0.760 0.3833916   
## + Intl.Mins 1 1679.2 1699.2 0.568 0.4512264   
## + Eve.Calls 1 1679.4 1699.4 0.401 0.5266732   
## + Day.Charge 1 1679.7 1699.7 0.143 0.7057176   
## + Night.Calls 1 1679.7 1699.7 0.126 0.7220908   
## + Night.Mins 1 1679.8 1699.8 0.026 0.8710627   
## + Account.Length 1 1679.8 1699.8 0.023 0.8782336   
## + Area.Code 1 1679.8 1699.8 0.007 0.9331965   
## - Intl.Calls 1 1687.6 1703.6 7.764 0.0053288 \*\*   
## - Night.Charge 1 1689.5 1705.5 9.715 0.0018272 \*\*   
## - Intl.Charge 1 1694.8 1710.8 14.949 0.0001105 \*\*\*  
## - Eve.Mins 1 1715.8 1731.8 35.947 2.027e-09 \*\*\*  
## - VMail.Plan 1 1716.2 1732.2 36.338 1.659e-09 \*\*\*  
## - Day.Mins 1 1807.0 1823.0 127.152 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1839.5 1855.5 159.689 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1846.3 1862.3 166.485 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1696.11  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Night.Charge + Intl.Calls + VMail.Message  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 1676.1 1696.1   
## + Day.Calls 1 1675.2 1697.2 0.897 0.3435830   
## + Eve.Charge 1 1675.2 1697.2 0.873 0.3502541   
## + Intl.Mins 1 1675.5 1697.5 0.641 0.4234209   
## + Eve.Calls 1 1675.6 1697.6 0.478 0.4894586   
## - VMail.Message 1 1679.8 1697.8 3.707 0.0542004 .   
## + Day.Charge 1 1675.9 1697.9 0.195 0.6585167   
## + Night.Calls 1 1675.9 1697.9 0.193 0.6606894   
## + Night.Mins 1 1676.1 1698.1 0.047 0.8283355   
## + Account.Length 1 1676.1 1698.1 0.012 0.9126273   
## + Area.Code 1 1676.1 1698.1 0.006 0.9373578   
## - Intl.Calls 1 1684.0 1702.0 7.915 0.0049027 \*\*   
## - Night.Charge 1 1686.2 1704.2 10.074 0.0015036 \*\*   
## - VMail.Plan 1 1687.2 1705.2 11.069 0.0008777 \*\*\*  
## - Intl.Charge 1 1691.1 1709.1 14.998 0.0001076 \*\*\*  
## - Eve.Mins 1 1712.7 1730.7 36.627 1.431e-09 \*\*\*  
## - Day.Mins 1 1802.5 1820.5 126.429 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 1835.2 1853.2 159.119 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 1842.4 1860.4 166.250 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Call: glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins +   
## VMail.Plan + Eve.Mins + Intl.Charge + Night.Charge + Intl.Calls +   
## VMail.Message, family = binomial(link = "logit"), data = data\_1449)  
##   
## Coefficients:  
## (Intercept) Int.l.Plan CustServ.Calls Day.Mins   
## -7.447007 2.088159 0.540396 0.012215   
## VMail.Plan Eve.Mins Intl.Charge Night.Charge   
## -2.087791 0.007391 0.313244 0.087392   
## Intl.Calls VMail.Message   
## -0.072462 0.039058   
##   
## Degrees of Freedom: 1931 Total (i.e. Null); 1922 Residual  
## Null Deviance: 2173   
## Residual Deviance: 1676 AIC: 1696

#final model 4 training:  
glm\_model4<- glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins +   
 VMail.Plan + Eve.Mins + Intl.Charge + Night.Charge + Intl.Calls +   
 VMail.Message, family = binomial(link = "logit"),   
 data = data\_1449)  
summary(glm\_model4)

##   
## Call:  
## glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins +   
## VMail.Plan + Eve.Mins + Intl.Charge + Night.Charge + Intl.Calls +   
## VMail.Message, family = binomial(link = "logit"), data = data\_1449)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.42267 -0.65998 -0.42383 -0.03375 3.03513   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -7.447007 0.556513 -13.382 < 2e-16 \*\*\*  
## Int.l.Plan 2.088159 0.170924 12.217 < 2e-16 \*\*\*  
## CustServ.Calls 0.540396 0.044005 12.280 < 2e-16 \*\*\*  
## Day.Mins 0.012215 0.001148 10.640 < 2e-16 \*\*\*  
## VMail.Plan -2.087791 0.645284 -3.235 0.001214 \*\*   
## Eve.Mins 0.007391 0.001242 5.949 2.71e-09 \*\*\*  
## Intl.Charge 0.313244 0.081715 3.833 0.000126 \*\*\*  
## Night.Charge 0.087392 0.027655 3.160 0.001577 \*\*   
## Intl.Calls -0.072462 0.026303 -2.755 0.005870 \*\*   
## VMail.Message 0.039058 0.020338 1.920 0.054804 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2172.9 on 1931 degrees of freedom  
## Residual deviance: 1676.1 on 1922 degrees of freedom  
## AIC: 1696.1  
##   
## Number of Fisher Scoring iterations: 5

# Model 5--------------------------------------------------------------------------  
# undersampling with 2415 zeros and 483 ones  
data\_2415 <- ovun.sample(Churn ~ ., data = Churn, method = "under", N = 2898, seed = 1)$data  
data\_2415 <- data\_2415[, -c(19, 21)]  
table(data\_2415$Churn)

##   
## 0 1   
## 2415 483

glm\_model2415.null<-glm(Churn ~ 1,family=binomial(link="logit"),data=data\_2415)  
glm\_model2415.full<-glm(Churn ~ .,family=binomial(link="logit"),data=data\_2415)  
  
#to find the best model with lowest AIC  
step(glm\_model2415.null, scope = list(upper=glm\_model2415.full), direction="both", test="Chisq", data=data\_2415)

## Start: AIC=2613.45  
## Churn ~ 1  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Int.l.Plan 1 2451.5 2455.5 159.919 < 2.2e-16 \*\*\*  
## + Day.Mins 1 2477.4 2481.4 134.076 < 2.2e-16 \*\*\*  
## + Day.Charge 1 2477.4 2481.4 134.073 < 2.2e-16 \*\*\*  
## + CustServ.Calls 1 2488.7 2492.7 122.762 < 2.2e-16 \*\*\*  
## + VMail.Plan 1 2573.8 2577.8 37.626 8.569e-10 \*\*\*  
## + VMail.Message 1 2582.2 2586.2 29.217 6.472e-08 \*\*\*  
## + Eve.Mins 1 2582.9 2586.9 28.559 9.087e-08 \*\*\*  
## + Eve.Charge 1 2582.9 2586.9 28.552 9.123e-08 \*\*\*  
## + Intl.Charge 1 2597.1 2601.1 14.371 0.0001501 \*\*\*  
## + Intl.Mins 1 2597.1 2601.1 14.358 0.0001511 \*\*\*  
## + Intl.Calls 1 2602.8 2606.8 8.693 0.0031936 \*\*   
## + Night.Charge 1 2607.1 2611.1 4.350 0.0369983 \*   
## + Night.Mins 1 2607.1 2611.1 4.348 0.0370443 \*   
## <none> 2611.4 2613.4   
## + Day.Calls 1 2610.3 2614.3 1.157 0.2819849   
## + Account.Length 1 2610.5 2614.5 0.912 0.3395514   
## + Area.Code 1 2611.2 2615.2 0.267 0.6055198   
## + Eve.Calls 1 2611.2 2615.2 0.241 0.6237777   
## + Night.Calls 1 2611.3 2615.3 0.185 0.6673233   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2455.53  
## Churn ~ Int.l.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + CustServ.Calls 1 2309.2 2315.2 142.306 < 2.2e-16 \*\*\*  
## + Day.Mins 1 2325.6 2331.6 125.900 < 2.2e-16 \*\*\*  
## + Day.Charge 1 2325.6 2331.6 125.896 < 2.2e-16 \*\*\*  
## + VMail.Plan 1 2410.2 2416.2 41.358 1.268e-10 \*\*\*  
## + VMail.Message 1 2418.7 2424.7 32.862 9.894e-09 \*\*\*  
## + Eve.Mins 1 2424.9 2430.9 26.646 2.444e-07 \*\*\*  
## + Eve.Charge 1 2424.9 2430.9 26.637 2.455e-07 \*\*\*  
## + Intl.Calls 1 2440.9 2446.9 10.668 0.001090 \*\*   
## + Intl.Charge 1 2441.3 2447.3 10.208 0.001398 \*\*   
## + Intl.Mins 1 2441.3 2447.3 10.192 0.001410 \*\*   
## + Night.Charge 1 2445.0 2451.0 6.507 0.010746 \*   
## + Night.Mins 1 2445.0 2451.0 6.504 0.010765 \*   
## <none> 2451.5 2455.5   
## + Day.Calls 1 2450.3 2456.3 1.258 0.262076   
## + Account.Length 1 2451.0 2457.0 0.496 0.481136   
## + Eve.Calls 1 2451.3 2457.3 0.199 0.655220   
## + Area.Code 1 2451.4 2457.4 0.097 0.755608   
## + Night.Calls 1 2451.5 2457.5 0.063 0.801292   
## - Int.l.Plan 1 2611.4 2613.4 159.919 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2315.23  
## Churn ~ Int.l.Plan + CustServ.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Day.Mins 1 2166.6 2174.6 142.613 < 2.2e-16 \*\*\*  
## + Day.Charge 1 2166.6 2174.6 142.610 < 2.2e-16 \*\*\*  
## + VMail.Plan 1 2268.8 2276.8 40.444 2.023e-10 \*\*\*  
## + Eve.Mins 1 2276.1 2284.1 33.178 8.409e-09 \*\*\*  
## + Eve.Charge 1 2276.1 2284.1 33.169 8.449e-09 \*\*\*  
## + VMail.Message 1 2277.1 2285.1 32.162 1.419e-08 \*\*\*  
## + Intl.Charge 1 2297.6 2305.6 11.594 0.0006617 \*\*\*  
## + Intl.Mins 1 2297.7 2305.7 11.579 0.0006670 \*\*\*  
## + Intl.Calls 1 2300.1 2308.1 9.125 0.0025212 \*\*   
## + Night.Charge 1 2301.3 2309.3 7.967 0.0047651 \*\*   
## + Night.Mins 1 2301.3 2309.3 7.964 0.0047724 \*\*   
## <none> 2309.2 2315.2   
## + Day.Calls 1 2307.5 2315.5 1.684 0.1943808   
## + Area.Code 1 2308.8 2316.8 0.401 0.5266158   
## + Account.Length 1 2308.9 2316.9 0.369 0.5433209   
## + Night.Calls 1 2309.1 2317.1 0.174 0.6765046   
## + Eve.Calls 1 2309.1 2317.1 0.171 0.6793292   
## - CustServ.Calls 1 2451.5 2455.5 142.306 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2488.7 2492.7 179.463 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2174.61  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Plan 1 2125.4 2135.4 41.184 1.386e-10 \*\*\*  
## + VMail.Message 1 2133.4 2143.4 33.207 8.287e-09 \*\*\*  
## + Eve.Mins 1 2134.4 2144.4 32.237 1.365e-08 \*\*\*  
## + Eve.Charge 1 2134.4 2144.4 32.230 1.369e-08 \*\*\*  
## + Intl.Charge 1 2153.6 2163.6 13.001 0.0003114 \*\*\*  
## + Intl.Mins 1 2153.6 2163.6 12.989 0.0003133 \*\*\*  
## + Intl.Calls 1 2155.9 2165.9 10.662 0.0010938 \*\*   
## + Night.Charge 1 2158.1 2168.1 8.491 0.0035697 \*\*   
## + Night.Mins 1 2158.1 2168.1 8.487 0.0035764 \*\*   
## <none> 2166.6 2174.6   
## + Day.Calls 1 2165.2 2175.2 1.452 0.2281374   
## + Day.Charge 1 2166.4 2176.4 0.188 0.6643343   
## + Area.Code 1 2166.4 2176.4 0.174 0.6769386   
## + Account.Length 1 2166.5 2176.5 0.077 0.7817475   
## + Eve.Calls 1 2166.6 2176.6 0.003 0.9588100   
## + Night.Calls 1 2166.6 2176.6 0.002 0.9653554   
## - Day.Mins 1 2309.2 2315.2 142.613 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 2325.6 2331.6 159.019 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2340.7 2346.7 174.071 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2135.43  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Eve.Mins 1 2090.5 2102.5 34.909 3.456e-09 \*\*\*  
## + Eve.Charge 1 2090.5 2102.5 34.903 3.466e-09 \*\*\*  
## + Intl.Charge 1 2112.1 2124.1 13.385 0.0002537 \*\*\*  
## + Intl.Mins 1 2112.1 2124.1 13.369 0.0002558 \*\*\*  
## + Intl.Calls 1 2115.2 2127.2 10.228 0.0013831 \*\*   
## + Night.Charge 1 2116.4 2128.4 9.042 0.0026380 \*\*   
## + Night.Mins 1 2116.4 2128.4 9.039 0.0026423 \*\*   
## + VMail.Message 1 2122.3 2134.3 3.154 0.0757380 .   
## <none> 2125.4 2135.4   
## + Day.Calls 1 2124.1 2136.1 1.373 0.2412913   
## + Area.Code 1 2125.3 2137.3 0.122 0.7269086   
## + Account.Length 1 2125.3 2137.3 0.096 0.7560946   
## + Day.Charge 1 2125.3 2137.3 0.089 0.7651415   
## + Night.Calls 1 2125.4 2137.4 0.005 0.9443770   
## + Eve.Calls 1 2125.4 2137.4 0.000 0.9941120   
## - VMail.Plan 1 2166.6 2174.6 41.184 1.386e-10 \*\*\*  
## - Day.Mins 1 2268.8 2276.8 143.352 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 2284.3 2292.3 158.850 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2303.6 2311.6 178.172 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2102.52  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Charge 1 2075.2 2089.2 15.351 8.927e-05 \*\*\*  
## + Intl.Mins 1 2075.2 2089.2 15.334 9.008e-05 \*\*\*  
## + Night.Charge 1 2080.1 2094.1 10.435 0.001237 \*\*   
## + Night.Mins 1 2080.1 2094.1 10.431 0.001239 \*\*   
## + Intl.Calls 1 2080.2 2094.2 10.331 0.001308 \*\*   
## + VMail.Message 1 2087.1 2101.1 3.391 0.065546 .   
## <none> 2090.5 2102.5   
## + Day.Calls 1 2088.7 2102.7 1.820 0.177257   
## + Eve.Charge 1 2090.0 2104.0 0.482 0.487550   
## + Account.Length 1 2090.3 2104.3 0.197 0.656841   
## + Area.Code 1 2090.3 2104.3 0.195 0.658467   
## + Day.Charge 1 2090.4 2104.4 0.142 0.705993   
## + Eve.Calls 1 2090.5 2104.5 0.026 0.871856   
## + Night.Calls 1 2090.5 2104.5 0.017 0.896258   
## - Eve.Mins 1 2125.4 2135.4 34.909 3.456e-09 \*\*\*  
## - VMail.Plan 1 2134.4 2144.4 43.855 3.535e-11 \*\*\*  
## - Day.Mins 1 2233.0 2243.0 142.519 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 2256.6 2266.6 166.118 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2268.7 2278.7 178.215 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2089.17  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Intl.Calls 1 2063.2 2079.2 11.942 0.0005487 \*\*\*  
## + Night.Charge 1 2064.6 2080.6 10.539 0.0011686 \*\*   
## + Night.Mins 1 2064.6 2080.6 10.536 0.0011706 \*\*   
## + VMail.Message 1 2071.7 2087.7 3.498 0.0614586 .   
## <none> 2075.2 2089.2   
## + Day.Calls 1 2073.6 2089.6 1.607 0.2049836   
## + Eve.Charge 1 2074.7 2090.7 0.428 0.5129056   
## + Intl.Mins 1 2074.8 2090.8 0.317 0.5736265   
## + Account.Length 1 2074.9 2090.9 0.229 0.6320688   
## + Area.Code 1 2075.0 2091.0 0.169 0.6806015   
## + Day.Charge 1 2075.0 2091.0 0.146 0.7027228   
## + Night.Calls 1 2075.1 2091.1 0.033 0.8568914   
## + Eve.Calls 1 2075.2 2091.2 0.024 0.8775339   
## - Intl.Charge 1 2090.5 2102.5 15.351 8.927e-05 \*\*\*  
## - Eve.Mins 1 2112.1 2124.1 36.875 1.260e-09 \*\*\*  
## - VMail.Plan 1 2119.6 2131.6 44.473 2.579e-11 \*\*\*  
## - Day.Mins 1 2219.9 2231.9 144.713 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 2244.0 2256.0 168.862 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2248.5 2260.5 173.367 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2079.23  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Intl.Calls  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + Night.Charge 1 2052.7 2070.7 10.560 0.0011554 \*\*   
## + Night.Mins 1 2052.7 2070.7 10.557 0.0011575 \*\*   
## + VMail.Message 1 2059.7 2077.7 3.537 0.0600315 .   
## <none> 2063.2 2079.2   
## + Day.Calls 1 2061.9 2079.9 1.363 0.2430717   
## + Intl.Mins 1 2062.8 2080.8 0.446 0.5043230   
## + Eve.Charge 1 2062.8 2080.8 0.428 0.5129127   
## + Account.Length 1 2062.9 2080.9 0.315 0.5747995   
## + Area.Code 1 2063.1 2081.1 0.157 0.6918276   
## + Day.Charge 1 2063.1 2081.1 0.136 0.7122452   
## + Night.Calls 1 2063.2 2081.2 0.071 0.7895891   
## + Eve.Calls 1 2063.2 2081.2 0.050 0.8231240   
## - Intl.Calls 1 2075.2 2089.2 11.942 0.0005487 \*\*\*  
## - Intl.Charge 1 2080.2 2094.2 16.962 3.813e-05 \*\*\*  
## - Eve.Mins 1 2100.3 2114.3 37.073 1.138e-09 \*\*\*  
## - VMail.Plan 1 2106.9 2120.9 43.697 3.833e-11 \*\*\*  
## - Day.Mins 1 2209.7 2223.7 146.481 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 2231.4 2245.4 168.181 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2238.5 2252.5 175.292 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2070.67  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Intl.Calls + Night.Charge  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + VMail.Message 1 2048.7 2068.7 3.966 0.0464322 \*   
## <none> 2052.7 2070.7   
## + Day.Calls 1 2051.4 2071.4 1.256 0.2623395   
## + Intl.Mins 1 2052.2 2072.2 0.484 0.4867141   
## + Eve.Charge 1 2052.3 2072.3 0.368 0.5442443   
## + Account.Length 1 2052.3 2072.3 0.315 0.5747213   
## + Night.Mins 1 2052.5 2072.5 0.163 0.6864984   
## + Area.Code 1 2052.5 2072.5 0.152 0.6963759   
## + Day.Charge 1 2052.6 2072.6 0.110 0.7403090   
## + Eve.Calls 1 2052.6 2072.6 0.054 0.8163929   
## + Night.Calls 1 2052.6 2072.6 0.049 0.8240372   
## - Night.Charge 1 2063.2 2079.2 10.560 0.0011554 \*\*   
## - Intl.Calls 1 2064.6 2080.6 11.963 0.0005425 \*\*\*  
## - Intl.Charge 1 2069.7 2085.7 17.015 3.708e-05 \*\*\*  
## - Eve.Mins 1 2091.1 2107.1 38.443 5.638e-10 \*\*\*  
## - VMail.Plan 1 2097.1 2113.1 44.423 2.646e-11 \*\*\*  
## - Day.Mins 1 2199.8 2215.8 147.138 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 2223.7 2239.7 171.032 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2232.0 2248.0 179.291 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=2068.7  
## Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins + VMail.Plan +   
## Eve.Mins + Intl.Charge + Intl.Calls + Night.Charge + VMail.Message  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 2048.7 2068.7   
## + Day.Calls 1 2047.5 2069.5 1.199 0.2734709   
## + Intl.Mins 1 2048.1 2070.1 0.560 0.4541730   
## + Eve.Charge 1 2048.2 2070.2 0.482 0.4873466   
## + Account.Length 1 2048.4 2070.4 0.336 0.5620661   
## + Night.Mins 1 2048.5 2070.5 0.190 0.6625520   
## + Area.Code 1 2048.5 2070.5 0.175 0.6753393   
## + Day.Charge 1 2048.6 2070.6 0.126 0.7224256   
## + Night.Calls 1 2048.6 2070.6 0.093 0.7598266   
## + Eve.Calls 1 2048.6 2070.6 0.071 0.7902774   
## - VMail.Message 1 2052.7 2070.7 3.966 0.0464322 \*   
## - Night.Charge 1 2059.7 2077.7 10.990 0.0009162 \*\*\*  
## - Intl.Calls 1 2060.8 2078.8 12.049 0.0005183 \*\*\*  
## - VMail.Plan 1 2061.7 2079.7 12.998 0.0003117 \*\*\*  
## - Intl.Charge 1 2065.9 2083.9 17.196 3.371e-05 \*\*\*  
## - Eve.Mins 1 2087.4 2105.4 38.739 4.845e-10 \*\*\*  
## - Day.Mins 1 2195.8 2213.8 147.084 < 2.2e-16 \*\*\*  
## - CustServ.Calls 1 2219.9 2237.9 171.250 < 2.2e-16 \*\*\*  
## - Int.l.Plan 1 2227.1 2245.1 178.412 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Call: glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins +   
## VMail.Plan + Eve.Mins + Intl.Charge + Intl.Calls + Night.Charge +   
## VMail.Message, family = binomial(link = "logit"), data = data\_2415)  
##   
## Coefficients:  
## (Intercept) Int.l.Plan CustServ.Calls Day.Mins   
## -7.802871 2.013902 0.508444 0.012568   
## VMail.Plan Eve.Mins Intl.Charge Intl.Calls   
## -2.015300 0.007144 0.312204 -0.085245   
## Night.Charge VMail.Message   
## 0.083254 0.036275   
##   
## Degrees of Freedom: 2897 Total (i.e. Null); 2888 Residual  
## Null Deviance: 2611   
## Residual Deviance: 2049 AIC: 2069

#final model 5 training:  
glm\_model5<- glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins +   
 VMail.Plan + Eve.Mins + Intl.Charge + Intl.Calls + Night.Charge +   
 VMail.Message, family = binomial(link = "logit"),   
 data = data\_2415)  
summary(glm\_model5)

##   
## Call:  
## glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Mins +   
## VMail.Plan + Eve.Mins + Intl.Charge + Intl.Calls + Night.Charge +   
## VMail.Message, family = binomial(link = "logit"), data = data\_2415)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.1646 -0.5500 -0.3630 -0.2022 3.1780   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -7.802871 0.520417 -14.994 < 2e-16 \*\*\*  
## Int.l.Plan 2.013902 0.149295 13.489 < 2e-16 \*\*\*  
## CustServ.Calls 0.508444 0.039713 12.803 < 2e-16 \*\*\*  
## Day.Mins 0.012568 0.001089 11.536 < 2e-16 \*\*\*  
## VMail.Plan -2.015300 0.582947 -3.457 0.000546 \*\*\*  
## Eve.Mins 0.007144 0.001164 6.138 8.34e-10 \*\*\*  
## Intl.Charge 0.312204 0.076132 4.101 4.12e-05 \*\*\*  
## Intl.Calls -0.085245 0.025217 -3.380 0.000724 \*\*\*  
## Night.Charge 0.083254 0.025192 3.305 0.000951 \*\*\*  
## VMail.Message 0.036275 0.018322 1.980 0.047726 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2611.5 on 2897 degrees of freedom  
## Residual deviance: 2048.7 on 2888 degrees of freedom  
## AIC: 2068.7  
##   
## Number of Fisher Scoring iterations: 5

##TESTING THE MODEL AND FINDING ACCURACY, ROC and AUC  
  
#Check dispresion of model  
summary(glm\_model1)$deviance / summary(glm\_model1)$df.residual

## [1] 0.6558949

#=0.6669978, model is not overfitted.  
summary(glm\_model2)$deviance / summary(glm\_model2)$df.residual

## [1] 0.6580929

summary(glm\_model3)$deviance / summary(glm\_model3)$df.residual

## [1] 1.155792

summary(glm\_model4)$deviance / summary(glm\_model4)$df.residual

## [1] 0.8720648

summary(glm\_model5)$deviance / summary(glm\_model5)$df.residual

## [1] 0.7093847

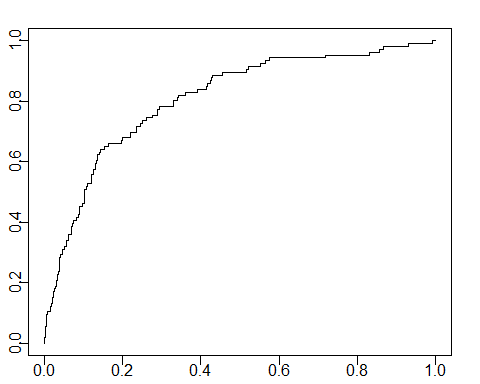
# Finding the best model  
library(ROCR)

## Loading required package: gplots

##   
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':  
##   
## lowess

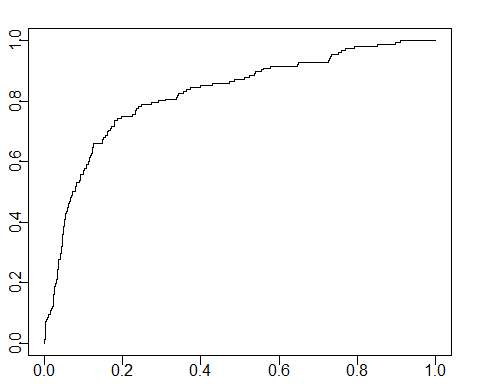
pred1 <- predict(glm\_model1,Churn\_test,type="response")  
pr1<-prediction(pred1, Churn\_test$Churn)  
#ROC curve  
prf1 <- performance(pr1, measure = "tpr", x.measure = "fpr")  
plot(prf1)



#Area under the curve  
auc1 <- performance(pr1, measure = "auc")  
auc1 <- auc1@y.values[[1]]  
auc1 # model1

## [1] 0.8083958

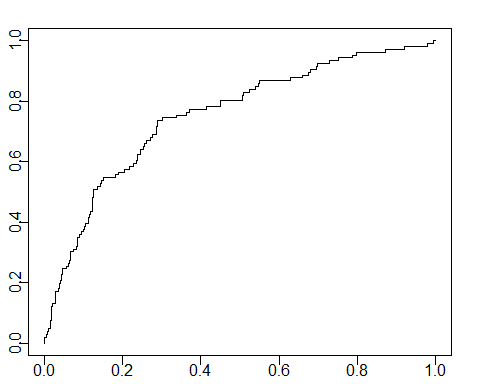
pred2 <- predict(glm\_model2,Churn\_test2,type="response")  
pr2<-prediction(pred2, Churn\_test2$Churn)  
#ROC curve  
prf2 <- performance(pr2, measure = "tpr", x.measure = "fpr")  
plot(prf2)



#Area under the curve  
auc2 <- performance(pr2, measure = "auc")  
auc2 <- auc2@y.values[[1]]  
auc2 # model2

## [1] 0.8234238

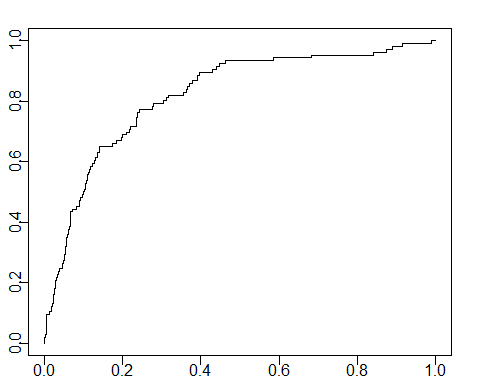
pred3 <- predict(glm\_model3,Churn\_test,type="response")  
pr3<-prediction(pred3, Churn\_test$Churn)  
prf3 <- performance(pr3, measure = "tpr", x.measure = "fpr")  
plot(prf3)



auc3 <- performance(pr3, measure = "auc")  
auc3 <- auc3@y.values[[1]]  
auc3 # model3

## [1] 0.7533575

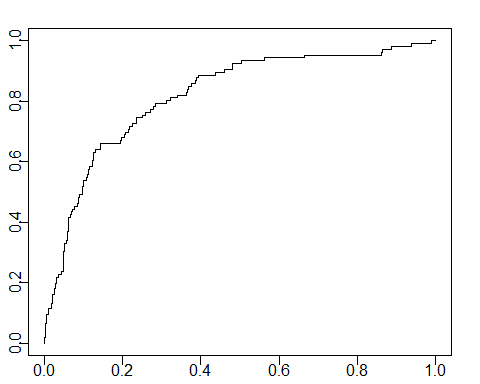
pred4 <- predict(glm\_model4,Churn\_test,type="response")  
pr4<-prediction(pred4, Churn\_test$Churn)  
prf4 <- performance(pr4, measure = "tpr", x.measure = "fpr")  
plot(prf4)



auc4 <- performance(pr4, measure = "auc")  
auc4 <- auc4@y.values[[1]]  
auc4 # model4

## [1] 0.8205075

pred5 <- predict(glm\_model5,Churn\_test,type="response")  
pr5<-prediction(pred5, Churn\_test$Churn)  
prf5 <- performance(pr5, measure = "tpr", x.measure = "fpr")  
plot(prf5)



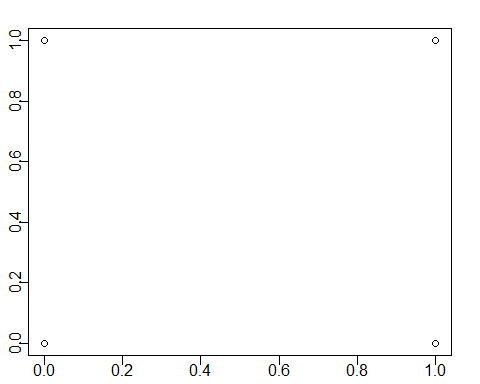
auc5 <- performance(pr5, measure = "auc")  
auc5 <- auc5@y.values[[1]]  
auc5 # model5

## [1] 0.8190382

sort(as.data.frame(table(auc1, auc2, auc3, auc4, auc5)))

## auc3 auc1 auc5 auc4  
## 1 0.753357501853836 0.808395814451674 0.819038202740929 0.820507538930543  
## auc2 Freq  
## 1 0.823423758206363 1

#Accuracy  
library(rpart)  
pred1 <- predict(glm\_model1,Churn\_test,type="response")  
pred1 <- ifelse(pred1 > 0.5,1,0)  
plot(Churn\_test$Churn~pred1)



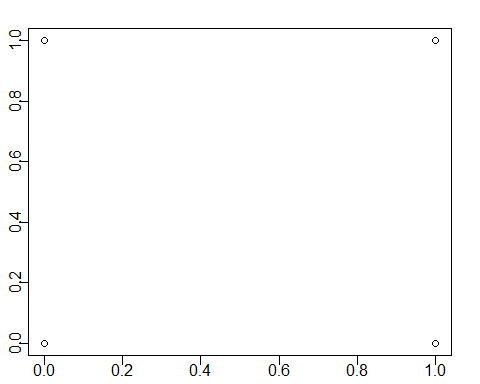
cfmatrix1=table(Churn\_test$Churn,pred1)  
cfmatrix1

## pred1  
## 0 1  
## 0 669 18  
## 1 88 18

Accuracy\_model1<-((cfmatrix1[1]+cfmatrix1[2,2])/  
 (cfmatrix1[1,1]+cfmatrix1[1,2]+cfmatrix1[2,1]+cfmatrix1[2,2]))  
Accuracy\_model1 # model1

## [1] 0.8663304

pred2 <- predict(glm\_model2,Churn\_test2,type="response")  
pred2 <- ifelse(pred2 > 0.5,1,0)  
plot(Churn\_test2$Churn~pred2)



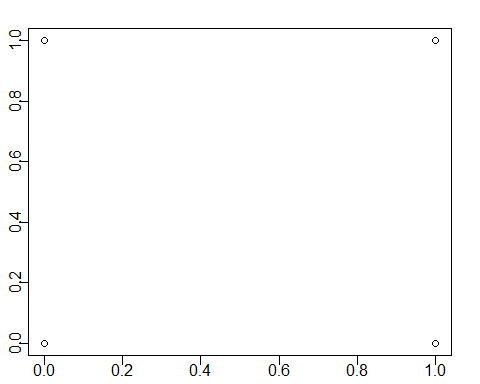
cfmatrix2=table(Churn\_test2$Churn,pred2)  
cfmatrix2

## pred2  
## 0 1  
## 0 801 27  
## 1 120 36

Accuracy\_model2<-((cfmatrix2[1]+cfmatrix2[2,2])/  
 (cfmatrix2[1,1]+cfmatrix2[1,2]+cfmatrix2[2,1]+cfmatrix2[2,2]))  
Accuracy\_model2 # model2

## [1] 0.8506098

pred3 <- predict(glm\_model3,Churn\_test,type="response")  
pred3 <- ifelse(pred3 > 0.5,1,0)  
plot(Churn\_test$Churn~pred3)



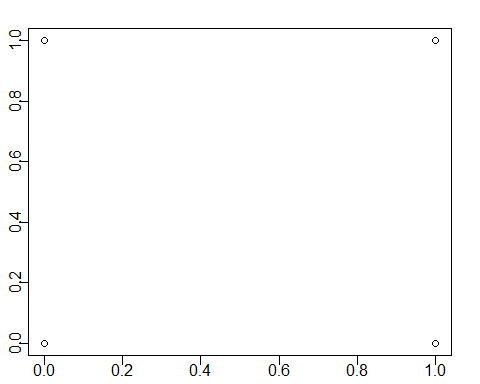
cfmatrix3=table(Churn\_test$Churn,pred3)  
cfmatrix3

## pred3  
## 0 1  
## 0 500 187  
## 1 34 72

Accuracy\_model3<-((cfmatrix3[1]+cfmatrix3[2,2])/  
 (cfmatrix3[1,1]+cfmatrix3[1,2]+cfmatrix3[2,1]+cfmatrix3[2,2]))  
Accuracy\_model3 # model3

## [1] 0.7213115

pred4 <- predict(glm\_model4,Churn\_test,type="response")  
pred4 <- ifelse(pred4 > 0.5,1,0)  
plot(Churn\_test$Churn~pred4)



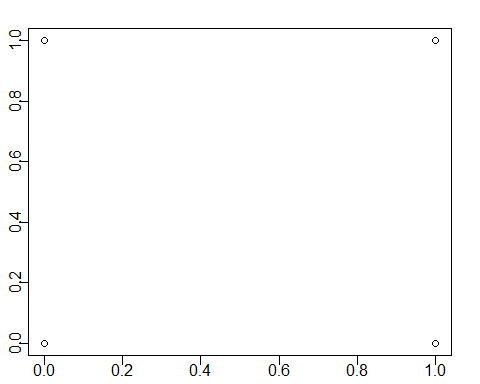
cfmatrix4=table(Churn\_test$Churn,pred4)  
cfmatrix4

## pred4  
## 0 1  
## 0 644 43  
## 1 67 39

Accuracy\_model4<-((cfmatrix4[1]+cfmatrix4[2,2])/  
 (cfmatrix4[1,1]+cfmatrix4[1,2]+cfmatrix4[2,1]+cfmatrix4[2,2]))  
Accuracy\_model4 # model4

## [1] 0.8612863

pred5 <- predict(glm\_model5,Churn\_test,type="response")  
pred5 <- ifelse(pred5 > 0.5,1,0)  
plot(Churn\_test$Churn~pred5)



cfmatrix5=table(Churn\_test$Churn,pred5)  
cfmatrix5

## pred5  
## 0 1  
## 0 668 19  
## 1 86 20

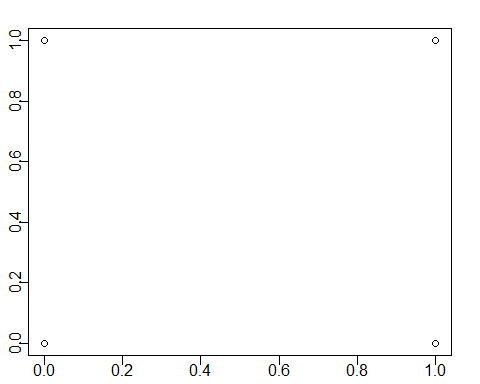
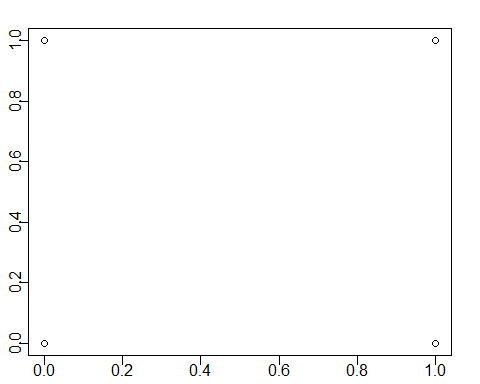
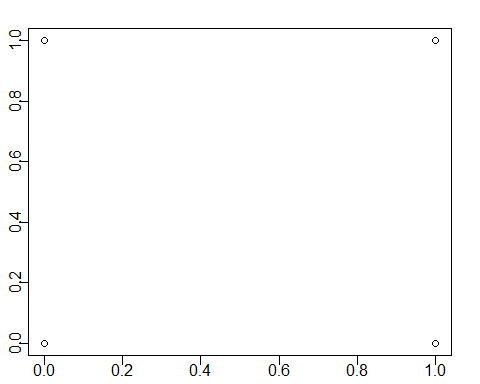
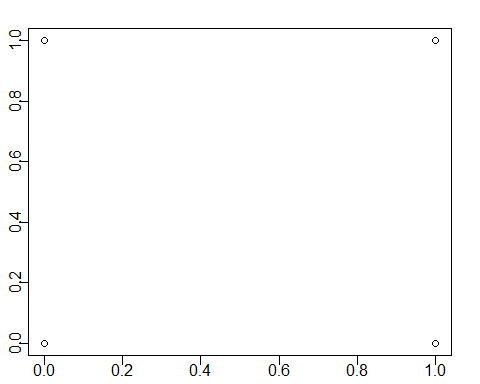
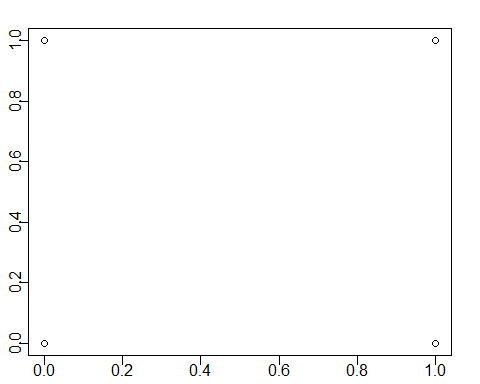
Accuracy\_model5<-((cfmatrix5[1]+cfmatrix5[2,2])/  
 (cfmatrix5[1,1]+cfmatrix5[1,2]+cfmatrix5[2,1]+cfmatrix5[2,2]))  
Accuracy\_model5 # model5

## [1] 0.8675914

sort(as.data.frame(table(Accuracy\_model1, Accuracy\_model2, Accuracy\_model3, Accuracy\_model4,   
 Accuracy\_model5)))

## Accuracy\_model3 Accuracy\_model2 Accuracy\_model4 Accuracy\_model1  
## 1 0.721311475409836 0.850609756097561 0.861286254728878 0.866330390920555  
## Accuracy\_model5 Freq  
## 1 0.867591424968474 1

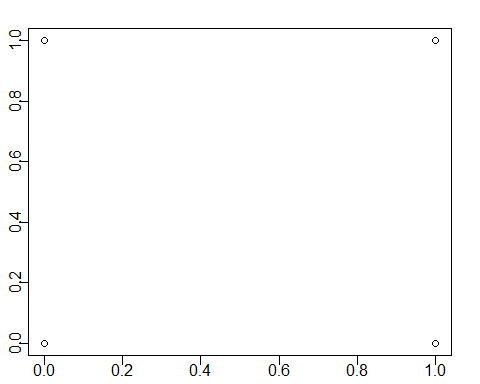
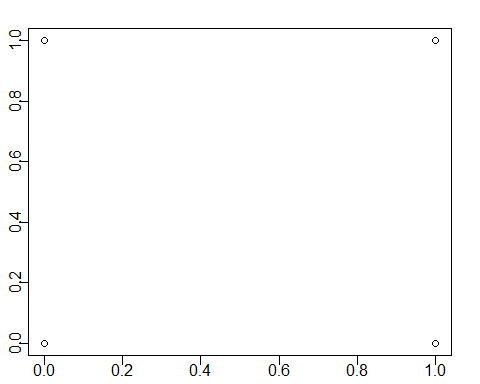
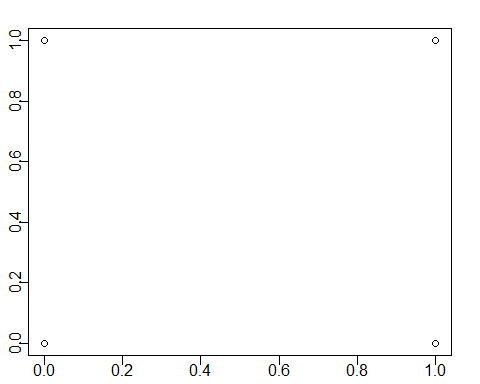
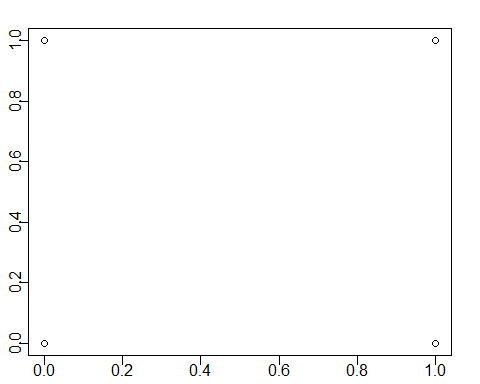
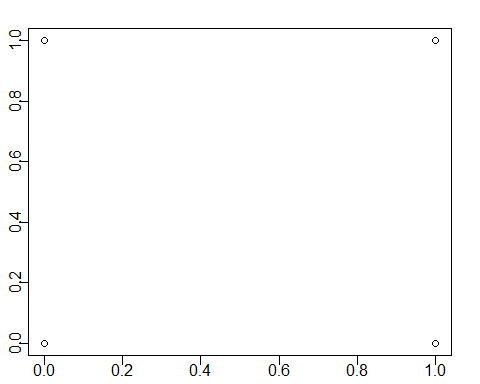
# Model 5 has greater accuracy   
  
  
  
##from the above results we can conclude that we will use model as the best fit.  
  
# Adjusting the probability limit in prediction  
#we create a function to test different probability limit  
Accuracyfunction<- function(x){  
 glm\_model1 <- glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge +   
 Eve.Mins + VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge +   
 VMail.Message + Eve.Charge, family = binomial(link = "logit"),   
 data = Churn\_train)  
 pred1 <- predict(glm\_model1,Churn\_test,type="response")  
 pred1 <- ifelse(pred1 > x ,1,0)  
 plot(Churn\_test$Churn~pred1)  
 cfmatrix1=table(Churn\_test$Churn,pred1)  
 cfmatrix1  
 Accuracy\_model1<-((cfmatrix1[1]+cfmatrix1[2,2])/  
 (cfmatrix1[1,1]+cfmatrix1[1,2]+cfmatrix1[2,1]+cfmatrix1[2,2]))  
 paste(Accuracy = Accuracy\_model1)}  
  
x<- c(Accuracyfunction(0.3),  
 Accuracyfunction(0.4),  
 Accuracyfunction(0.5),  
 Accuracyfunction(0.6),  
 Accuracyfunction(0.7))



as.data.frame(x)

## x  
## 1 0.849936948297604  
## 2 0.867591424968474  
## 3 0.866330390920555  
## 4 0.868852459016393  
## 5 0.868852459016393

# threshhold probability is between 0.4 to 0.6  
x<- c(Accuracyfunction(0.406),  
 Accuracyfunction(0.408),  
 Accuracyfunction(0.500),  
 Accuracyfunction(0.502),  
 Accuracyfunction(0.504),  
 Accuracyfunction(0.506),  
 Accuracyfunction(0.508))



as.data.frame(x)

## x  
## 1 0.868852459016393  
## 2 0.867591424968474  
## 3 0.866330390920555  
## 4 0.866330390920555  
## 5 0.868852459016393  
## 6 0.868852459016393  
## 7 0.867591424968474

# Hence Threshhold Probability = 0.506  
  
#============================================================================================  
# CONCLUSION #  
#============================================================================================  
  
  
# Thus the final Model is as below with threshhold probability = 0.506  
  
glm\_model\_final <- glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge +   
 Eve.Mins + VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge +   
 VMail.Message + Eve.Charge, family = binomial(link = "logit"),   
 data = Churn\_train)  
  
summary(glm\_model\_final)

##   
## Call:  
## glm(formula = Churn ~ Int.l.Plan + CustServ.Calls + Day.Charge +   
## Eve.Mins + VMail.Plan + Intl.Calls + Night.Mins + Intl.Charge +   
## VMail.Message + Eve.Charge, family = binomial(link = "logit"),   
## data = Churn\_train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.1989 -0.5207 -0.3427 -0.1912 2.8986   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -7.951285 0.583028 -13.638 < 2e-16 \*\*\*  
## Int.l.Plan 2.067146 0.166065 12.448 < 2e-16 \*\*\*  
## CustServ.Calls 0.507400 0.044288 11.457 < 2e-16 \*\*\*  
## Day.Charge 0.074615 0.007176 10.398 < 2e-16 \*\*\*  
## Eve.Mins 3.181745 1.864004 1.707 0.087833 .   
## VMail.Plan -2.363100 0.688112 -3.434 0.000594 \*\*\*  
## Intl.Calls -0.095363 0.028684 -3.325 0.000886 \*\*\*  
## Night.Mins 0.003784 0.001276 2.967 0.003008 \*\*   
## Intl.Charge 0.248350 0.086004 2.888 0.003881 \*\*   
## VMail.Message 0.047322 0.021304 2.221 0.026332 \*   
## Eve.Charge -37.336479 21.929129 -1.703 0.088643 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 2133.4 on 2539 degrees of freedom  
## Residual deviance: 1658.8 on 2529 degrees of freedom  
## AIC: 1680.8  
##   
## Number of Fisher Scoring iterations: 6

pred\_final <- predict(glm\_model\_final,Churn\_test,type="response")  
pred\_final <- ifelse(pred\_final > 0.506 ,1,0)  
table(pred\_final)

## pred\_final  
## 0 1   
## 759 34