# Create an RF model for predicting mineral or rock

# We will need these libraries ... install.packages("    "), before running

library(caret)

library(pROC)

# Read in the dataset

data<-read.csv("~/winequality-w.csv")

# Look at the column names

names(data)

#data1<- na.omit(data1)

# Split the dataset into a training and test dataset using a split of 75% training

# and 25% test

set.seed(42)

nrow(data)

trainingObs<-sample(nrow(data),0.75\*nrow(data),replace=FALSE)

# Create the training dataset

trainingDS<-data[trainingObs,]

View(trainingDS)

# Create the test dataset

testDS<-data[-trainingObs,]

View(testDS)

write.csv(testDS,file="white\_test.csv")

#smote training data

table(trainingDS$BQ2\_quality)

smoted\_white<-SMOTE(BQ2\_quality ~ ., trainingDS, perc.over=330, perc.under= 135)

table(smoted\_white$BQ2\_quality)

write.csv(smoted\_white, file="white\_train\_smote.csv")