wine quality prediction\_full attributes

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Import to R the following file:

winequality <- read.csv(file = "http://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/winequality-white.csv", header = TRUE, sep = ";")

Check data characteristics. Is there missing data?

str(winequality)

## 'data.frame': 4898 obs. of 12 variables:  
## $ fixed.acidity : num 7 6.3 8.1 7.2 7.2 8.1 6.2 7 6.3 8.1 ...  
## $ volatile.acidity : num 0.27 0.3 0.28 0.23 0.23 0.28 0.32 0.27 0.3 0.22 ...  
## $ citric.acid : num 0.36 0.34 0.4 0.32 0.32 0.4 0.16 0.36 0.34 0.43 ...  
## $ residual.sugar : num 20.7 1.6 6.9 8.5 8.5 6.9 7 20.7 1.6 1.5 ...  
## $ chlorides : num 0.045 0.049 0.05 0.058 0.058 0.05 0.045 0.045 0.049 0.044 ...  
## $ free.sulfur.dioxide : num 45 14 30 47 47 30 30 45 14 28 ...  
## $ total.sulfur.dioxide: num 170 132 97 186 186 97 136 170 132 129 ...  
## $ density : num 1.001 0.994 0.995 0.996 0.996 ...  
## $ pH : num 3 3.3 3.26 3.19 3.19 3.26 3.18 3 3.3 3.22 ...  
## $ sulphates : num 0.45 0.49 0.44 0.4 0.4 0.44 0.47 0.45 0.49 0.45 ...  
## $ alcohol : num 8.8 9.5 10.1 9.9 9.9 10.1 9.6 8.8 9.5 11 ...  
## $ quality : int 6 6 6 6 6 6 6 6 6 6 ...

summary(winequality)

## fixed.acidity volatile.acidity citric.acid residual.sugar   
## Min. : 3.800 Min. :0.0800 Min. :0.0000 Min. : 0.600   
## 1st Qu.: 6.300 1st Qu.:0.2100 1st Qu.:0.2700 1st Qu.: 1.700   
## Median : 6.800 Median :0.2600 Median :0.3200 Median : 5.200   
## Mean : 6.855 Mean :0.2782 Mean :0.3342 Mean : 6.391   
## 3rd Qu.: 7.300 3rd Qu.:0.3200 3rd Qu.:0.3900 3rd Qu.: 9.900   
## Max. :14.200 Max. :1.1000 Max. :1.6600 Max. :65.800   
## chlorides free.sulfur.dioxide total.sulfur.dioxide  
## Min. :0.00900 Min. : 2.00 Min. : 9.0   
## 1st Qu.:0.03600 1st Qu.: 23.00 1st Qu.:108.0   
## Median :0.04300 Median : 34.00 Median :134.0   
## Mean :0.04577 Mean : 35.31 Mean :138.4   
## 3rd Qu.:0.05000 3rd Qu.: 46.00 3rd Qu.:167.0   
## Max. :0.34600 Max. :289.00 Max. :440.0   
## density pH sulphates alcohol   
## Min. :0.9871 Min. :2.720 Min. :0.2200 Min. : 8.00   
## 1st Qu.:0.9917 1st Qu.:3.090 1st Qu.:0.4100 1st Qu.: 9.50   
## Median :0.9937 Median :3.180 Median :0.4700 Median :10.40   
## Mean :0.9940 Mean :3.188 Mean :0.4898 Mean :10.51   
## 3rd Qu.:0.9961 3rd Qu.:3.280 3rd Qu.:0.5500 3rd Qu.:11.40   
## Max. :1.0390 Max. :3.820 Max. :1.0800 Max. :14.20   
## quality   
## Min. :3.000   
## 1st Qu.:5.000   
## Median :6.000   
## Mean :5.878   
## 3rd Qu.:6.000   
## Max. :9.000

sum(is.na(winequality) == TRUE)

## [1] 0

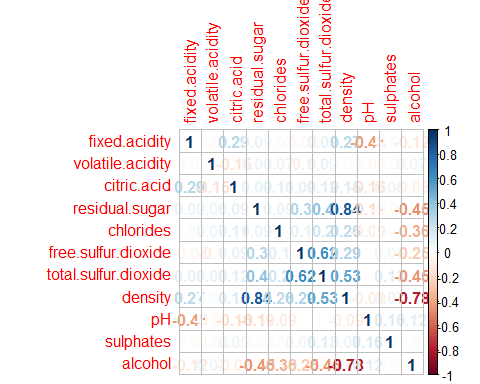
There is no missing data

The correlation between the attributes other than wine quality?

wine <- winequality[,-12]  
wine\_cor <- cor(wine)  
#install.packages("corrplot")  
library(corrplot)

## corrplot 0.84 loaded

corrplot(wine\_cor, method = "number")

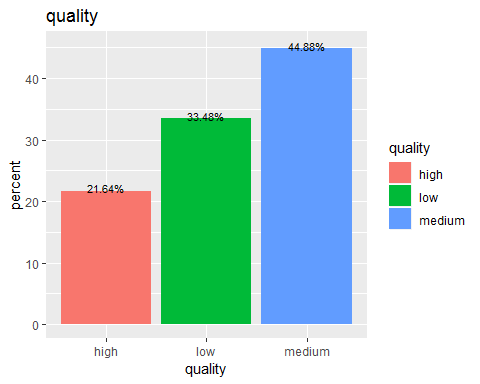
 Taking a look at the correlation coefficients r for the predictor variables, we see that density is strongly correlated with residual.sugar (r=0.84) and alcohol (r=???0.78), and moderately correlated with total.sulfur.dioxide (r=0.53). free.sulfur.dioxide and total.sulfur.dioxide are also moderately correlated with each other (r=0.62).

Reduce the levels of rating for quality to three levels as high, medium and low

winequality$quality[which(winequality$quality %in% c(3,4,5))] = "low"  
winequality$quality[which(winequality$quality %in% c(6))] = "medium"  
winequality$quality[which(winequality$quality %in% c(7,8,9))] = "high"  
winequality$quality <- as.factor(winequality$quality)

Graph the frequency distribution of wine quality.

#install.packages("ggplot2")  
library(ggplot2)  
ggplot(winequality, aes(x = quality, fill = quality)) +  
 geom\_bar(aes(y = prop.table(..count..) \* 100), position='dodge') +  
 geom\_text(aes(y = prop.table(..count..) \* 100 + 0.5,   
 label = paste0(round(prop.table(..count..) \* 100, 2), '%')),   
 stat = 'count',   
 position = position\_dodge(.9),   
 size = 3) +   
 labs(x = 'quality', y = 'percent', fill = 'quality', title = 'quality')

 Normalize the data set.

normalize <- function(x) {  
 return ((x - min(x)) / (max(x) - min(x))) }  
winedata <- as.data.frame(lapply(wine, normalize))

Divide the data to training and testing groups.

split\_data <- sample(nrow(winedata), floor(nrow(winedata)\*0.7))  
winedata\_train <- winedata[split\_data,]  
winedata\_test <- winedata[-split\_data,]  
  
winedata\_train\_labels <- winequality[split\_data, 12]  
winedata\_test\_labels <- winequality[-split\_data, 12]

Use the KNN algorithm to predict the quality of wine using its attributes.

#install.packages("class")  
#install.packages("gmodels")  
library(class)  
library(gmodels)  
winedata\_test\_pred <- knn(train = winedata\_train, test = winedata\_test, cl = winedata\_train\_labels, k=10, prob = TRUE)  
winedata\_test\_pred

## [1] medium low low medium medium medium medium medium medium  
## [10] low medium low medium low medium medium medium medium  
## [19] low high low high low medium medium low low   
## [28] medium low medium medium medium medium low low high   
## [37] low medium low medium medium low low medium medium  
## [46] medium medium low medium low low low high low   
## [55] low medium medium medium medium medium medium medium medium  
## [64] medium low medium high medium medium low high high   
## [73] medium low medium medium high medium medium medium medium  
## [82] medium low low low low medium medium medium medium  
## [91] medium high medium medium low high high low medium  
## [100] low low low low medium medium high low high   
## [109] low low medium medium high low medium medium low   
## [118] high high low medium medium low medium low high   
## [127] medium medium medium medium low low high low medium  
## [136] medium high medium low low low medium medium low   
## [145] low low medium high low medium medium medium high   
## [154] low medium medium high high low medium high medium  
## [163] medium medium medium low low medium medium medium medium  
## [172] low low medium low medium low medium low high   
## [181] medium medium high high low medium low high medium  
## [190] low low medium medium low low medium medium low   
## [199] low medium medium low medium low low medium low   
## [208] low high low high high high medium low medium  
## [217] low low medium medium medium medium medium medium low   
## [226] medium medium medium medium medium medium medium medium high   
## [235] low medium medium medium medium low medium medium medium  
## [244] low low medium medium low low low low medium  
## [253] high medium medium high medium high medium medium medium  
## [262] medium medium low low medium medium medium low high   
## [271] medium medium medium low low low medium medium low   
## [280] low medium medium low medium medium low high high   
## [289] medium low medium low medium high medium low low   
## [298] medium medium medium medium medium medium medium low low   
## [307] low low medium medium medium low low low medium  
## [316] low medium medium medium low low low medium medium  
## [325] low low medium low medium medium high medium high   
## [334] medium low low low low high low medium low   
## [343] low low medium low low low medium low low   
## [352] medium medium medium low low low low low medium  
## [361] medium low high low low medium high medium low   
## [370] low medium high high high high high medium medium  
## [379] low medium medium medium medium medium low medium high   
## [388] low medium low low low medium medium medium high   
## [397] medium low low low low medium high high medium  
## [406] low medium low medium medium low high high high   
## [415] medium medium low low medium medium low medium medium  
## [424] medium high medium medium medium medium high medium medium  
## [433] medium low medium medium high low medium medium medium  
## [442] medium medium medium medium high medium medium high medium  
## [451] low low medium medium medium medium medium low medium  
## [460] low high medium low medium low low medium medium  
## [469] medium high low low high high low medium medium  
## [478] medium low medium high medium medium low low low   
## [487] medium medium medium low medium low medium low low   
## [496] medium low high medium medium low low medium medium  
## [505] low high high low low medium medium medium medium  
## [514] medium low medium medium medium medium high high low   
## [523] medium low medium low high low medium medium medium  
## [532] low low medium low high high medium medium medium  
## [541] medium medium medium low medium high high low high   
## [550] medium medium low medium medium medium low medium medium  
## [559] medium medium medium low medium medium medium low medium  
## [568] medium low low low medium low low low low   
## [577] high medium medium low low high high low low   
## [586] low low medium medium medium low low medium low   
## [595] low medium low medium high medium low medium medium  
## [604] high high low low low medium low medium medium  
## [613] high high medium medium high medium medium medium medium  
## [622] medium low low medium low low low high low   
## [631] high high medium medium medium low medium medium medium  
## [640] low medium low low medium medium low medium medium  
## [649] low low medium medium low medium medium medium medium  
## [658] medium high high low high low medium high medium  
## [667] medium low low low medium medium medium low medium  
## [676] medium low medium high low medium medium low low   
## [685] low medium high medium medium medium medium low medium  
## [694] medium low low low low low low medium high   
## [703] high low medium medium medium medium medium medium medium  
## [712] low low low high high medium medium medium low   
## [721] low medium medium low low low high medium medium  
## [730] high low medium high low low medium low medium  
## [739] low medium high high low low low medium medium  
## [748] low low medium low low low medium medium medium  
## [757] high medium medium low low medium low high medium  
## [766] medium low medium medium low high low high medium  
## [775] medium low low high high high low medium high   
## [784] medium medium low medium medium medium medium high medium  
## [793] medium low low high high medium medium medium medium  
## [802] low medium medium low medium medium high medium low   
## [811] medium low low medium medium low medium low medium  
## [820] medium medium medium medium medium medium low medium medium  
## [829] medium medium high high medium high medium medium high   
## [838] medium medium low high medium medium low medium medium  
## [847] medium low medium medium medium medium medium medium low   
## [856] medium medium high medium medium medium medium high high   
## [865] high high medium medium medium high medium medium medium  
## [874] low high medium low high high medium high medium  
## [883] medium medium medium medium low medium medium medium medium  
## [892] medium medium medium medium medium medium medium high high   
## [901] high high high high high medium high medium low   
## [910] medium medium high medium medium medium medium medium high   
## [919] medium low high high medium medium high medium medium  
## [928] low high low high medium medium medium medium high   
## [937] medium medium medium medium medium medium medium medium medium  
## [946] medium medium high low medium low medium low low   
## [955] low medium high high high medium high medium medium  
## [964] medium medium high high low medium low medium medium  
## [973] low high high medium low medium low medium medium  
## [982] medium medium high low medium medium medium high medium  
## [991] low low medium medium high high medium medium medium  
## [1000] medium medium medium low low low medium medium medium  
## [1009] medium medium medium medium medium medium medium high low   
## [1018] high high high low medium medium medium medium medium  
## [1027] medium high medium low medium medium medium medium high   
## [1036] medium high low high medium medium medium high medium  
## [1045] low high high low medium medium medium medium low   
## [1054] low low low medium medium medium high medium low   
## [1063] medium medium medium medium low high medium medium medium  
## [1072] medium medium high low medium medium low medium medium  
## [1081] medium low medium medium low medium medium medium medium  
## [1090] low low medium medium medium medium low medium high   
## [1099] high high medium high medium low high medium medium  
## [1108] low high medium medium medium medium medium medium high   
## [1117] high low medium low medium low low high high   
## [1126] high low medium high low low low medium medium  
## [1135] low medium medium low medium medium medium low low   
## [1144] medium low low low low medium high low high   
## [1153] high low low low medium medium low medium medium  
## [1162] medium medium high medium low medium medium high low   
## [1171] high high high high low low medium medium low   
## [1180] medium medium high low medium low high high low   
## [1189] low high low medium low medium medium medium medium  
## [1198] medium high medium medium medium medium low low medium  
## [1207] high high high high high high medium high medium  
## [1216] medium low low high medium high medium low medium  
## [1225] medium medium medium medium low medium low low medium  
## [1234] medium high medium low high high low medium medium  
## [1243] medium medium medium medium medium low high low medium  
## [1252] medium low low low low medium medium high medium  
## [1261] low medium low medium low low medium high low   
## [1270] low low low low high high low medium medium  
## [1279] medium high medium medium low medium medium medium low   
## [1288] low low medium high medium high low high low   
## [1297] low high low high medium high low high high   
## [1306] high medium medium medium high medium medium medium medium  
## [1315] low medium medium medium low medium low high medium  
## [1324] low medium high medium medium high medium low low   
## [1333] low high low high medium high medium medium medium  
## [1342] high medium low medium medium medium low low medium  
## [1351] medium medium medium medium low low medium medium high   
## [1360] medium high medium medium medium high high high medium  
## [1369] high medium high low medium medium medium medium low   
## [1378] low low high high medium medium high medium low   
## [1387] medium medium medium medium low medium high medium high   
## [1396] high medium medium low high medium medium medium medium  
## [1405] medium high medium medium high medium low medium medium  
## [1414] medium high medium medium medium low medium low low   
## [1423] high medium medium medium low medium medium low low   
## [1432] high medium high low high medium high high medium  
## [1441] low medium medium high high medium high medium medium  
## [1450] high high high medium high low high low medium  
## [1459] medium high medium medium medium medium high medium medium  
## [1468] medium low high   
## attr(,"prob")  
## [1] 0.5000000 0.6000000 0.7000000 0.5000000 0.5000000 0.7000000  
## [7] 0.5000000 0.6000000 0.5000000 0.7000000 0.6000000 0.7500000  
## [13] 0.6000000 0.4000000 0.5000000 0.6000000 0.7000000 0.9000000  
## [19] 0.6000000 0.6000000 0.6000000 0.6000000 0.5000000 0.7000000  
## [25] 0.6000000 0.7000000 0.7000000 0.7000000 0.7000000 0.5000000  
## [31] 0.6363636 0.5000000 0.5000000 0.6000000 0.6000000 0.6000000  
## [37] 0.8000000 0.8000000 0.6000000 0.7000000 0.8000000 0.5000000  
## [43] 0.7000000 0.8000000 0.5000000 0.4545455 0.8000000 0.6000000  
## [49] 0.5454545 0.9000000 0.6000000 0.8000000 0.6000000 0.8000000  
## [55] 0.9000000 0.7000000 0.5000000 0.5000000 0.5000000 0.5000000  
## [61] 0.5454545 0.6000000 0.5454545 0.5454545 0.7272727 0.8000000  
## [67] 0.7142857 0.4000000 0.5000000 0.7272727 0.6000000 0.4000000  
## [73] 0.5000000 0.7000000 0.6000000 0.4545455 0.5000000 0.5000000  
## [79] 0.7000000 0.6000000 0.7000000 0.5000000 0.9000000 0.8000000  
## [85] 0.5384615 0.5000000 0.6000000 0.6000000 0.5000000 0.5454545  
## [91] 0.5000000 0.5454545 0.5454545 0.7000000 0.7000000 0.6000000  
## [97] 0.7272727 0.6000000 0.5000000 1.0000000 0.7000000 0.5000000  
## [103] 0.7272727 0.7000000 0.6000000 0.9000000 0.7000000 0.5833333  
## [109] 0.6000000 0.8181818 0.6000000 0.5000000 0.5000000 0.6000000  
## [115] 0.6000000 0.6000000 0.6000000 0.6000000 0.6000000 0.9000000  
## [121] 0.5000000 0.6000000 1.0000000 0.5000000 0.6000000 0.6363636  
## [127] 0.6000000 1.0000000 0.7000000 1.0000000 0.7000000 0.5000000  
## [133] 0.6000000 0.5000000 0.5000000 0.7000000 0.7272727 0.6000000  
## [139] 0.6000000 0.7272727 0.6000000 0.6000000 0.6000000 0.6000000  
## [145] 1.0000000 0.7000000 0.8000000 0.4000000 0.9000000 0.7000000  
## [151] 0.3636364 0.6000000 0.8000000 0.9000000 0.8181818 0.8181818  
## [157] 0.6000000 0.4000000 0.5000000 0.5000000 0.6000000 0.6000000  
## [163] 0.5454545 0.4545455 0.4545455 0.7000000 0.5000000 0.5000000  
## [169] 0.4000000 0.5000000 0.4545455 0.7000000 0.7000000 0.5454545  
## [175] 0.5000000 0.5000000 0.9000000 0.4000000 0.7000000 0.8000000  
## [181] 0.4545455 0.6000000 0.7000000 0.6000000 0.7000000 0.7272727  
## [187] 0.6000000 0.7000000 0.6363636 0.6363636 0.8000000 0.5000000  
## [193] 0.5000000 0.7000000 0.8000000 0.9000000 0.9000000 0.7000000  
## [199] 0.9000000 0.4000000 0.5000000 0.5000000 0.6000000 0.8000000  
## [205] 0.9000000 0.8000000 0.5000000 0.6000000 0.5000000 0.8000000  
## [211] 0.5000000 0.6000000 0.5000000 0.5000000 0.6000000 0.5000000  
## [217] 0.8000000 0.7272727 0.7000000 0.6363636 0.6000000 0.5000000  
## [223] 0.5000000 0.6000000 0.5454545 0.4000000 0.6000000 0.6000000  
## [229] 0.5000000 0.8000000 0.7000000 0.9000000 0.7272727 0.6000000  
## [235] 0.5000000 0.4000000 0.6000000 0.7000000 0.7000000 0.5000000  
## [241] 0.7000000 0.7000000 0.7000000 0.7000000 0.5000000 0.6000000  
## [247] 0.6000000 0.8181818 0.9000000 0.9000000 1.0000000 1.0000000  
## [253] 0.6000000 0.4000000 0.6000000 0.8000000 0.7000000 0.6000000  
## [259] 0.5000000 0.9090909 0.9090909 0.5000000 0.5000000 0.7000000  
## [265] 0.7272727 1.0000000 0.7000000 0.5000000 0.5000000 0.6000000  
## [271] 0.4000000 0.8000000 0.8000000 0.6000000 0.6000000 0.7000000  
## [277] 0.5000000 0.6363636 0.9000000 0.5000000 0.8000000 0.5000000  
## [283] 0.8000000 0.7000000 0.6000000 0.9000000 0.6000000 0.9000000  
## [289] 0.6000000 0.8000000 0.5000000 0.7000000 0.6000000 0.5000000  
## [295] 0.8000000 0.7000000 0.4000000 0.3636364 0.8000000 0.4000000  
## [301] 0.4000000 0.6000000 0.8000000 0.6000000 1.0000000 0.5000000  
## [307] 0.4000000 0.6000000 0.6000000 0.6000000 0.6000000 1.0000000  
## [313] 1.0000000 0.6000000 0.6000000 0.7000000 0.4000000 0.5454545  
## [319] 0.4000000 0.8000000 0.6000000 0.6000000 0.6000000 0.4615385  
## [325] 0.6000000 0.8181818 0.6666667 0.5000000 0.4000000 0.7000000  
## [331] 0.7272727 0.8000000 0.4000000 0.8000000 0.7000000 0.7000000  
## [337] 0.6000000 1.0000000 0.4000000 0.6000000 0.5000000 0.7000000  
## [343] 0.7000000 0.6000000 0.4000000 0.6000000 0.7000000 0.7000000  
## [349] 0.5000000 0.7000000 0.8000000 0.6000000 0.4000000 0.6000000  
## [355] 0.7000000 0.6000000 0.5000000 1.0000000 0.8000000 0.7000000  
## [361] 0.6000000 0.5000000 0.8000000 0.7000000 0.7000000 0.7000000  
## [367] 0.5000000 0.6363636 1.0000000 0.5000000 0.6000000 0.5000000  
## [373] 0.7000000 0.5000000 0.9000000 0.5000000 0.6000000 0.5000000  
## [379] 0.6000000 0.5000000 0.7272727 0.5000000 0.6000000 0.6000000  
## [385] 0.7000000 0.6000000 0.6000000 0.7000000 0.5000000 0.7000000  
## [391] 0.9000000 0.6000000 0.7000000 0.7000000 0.9000000 0.8000000  
## [397] 0.5000000 0.7000000 0.7000000 0.6000000 0.6000000 0.5000000  
## [403] 0.5000000 0.6000000 0.9000000 0.8181818 0.5000000 0.6000000  
## [409] 0.7000000 0.7000000 0.4000000 0.7000000 0.7000000 0.6000000  
## [415] 0.6363636 0.6000000 0.7000000 0.6000000 0.6000000 0.5000000  
## [421] 0.8000000 0.4000000 0.8000000 0.9000000 0.6000000 0.8000000  
## [427] 0.7000000 0.7000000 0.5000000 0.6000000 0.7000000 0.4545455  
## [433] 0.7000000 0.5000000 0.6000000 0.7000000 0.5000000 0.5454545  
## [439] 0.5000000 0.5000000 0.6000000 0.5000000 0.7000000 0.4000000  
## [445] 0.7000000 0.5000000 0.8000000 0.5000000 0.6000000 0.8000000  
## [451] 0.6000000 0.6000000 0.7000000 0.7000000 0.5000000 0.4000000  
## [457] 0.5000000 0.7000000 0.7000000 0.4000000 0.5454545 0.6000000  
## [463] 0.9000000 0.7000000 0.7000000 1.0000000 0.5000000 0.7000000  
## [469] 0.7000000 0.9000000 0.6000000 0.8000000 0.5000000 0.5000000  
## [475] 0.7000000 0.4000000 0.6363636 0.7000000 0.6000000 0.5000000  
## [481] 0.4000000 0.7000000 0.6363636 0.4000000 0.7000000 0.7272727  
## [487] 0.7000000 0.5454545 0.6000000 0.7000000 0.8000000 0.6000000  
## [493] 0.4000000 0.5000000 0.5000000 0.4000000 0.4000000 0.6000000  
## [499] 0.5000000 0.8000000 0.9000000 0.7000000 0.6000000 0.5000000  
## [505] 0.6363636 0.6000000 0.5000000 0.5000000 0.9000000 0.7000000  
## [511] 0.7000000 0.4000000 0.6000000 0.7000000 0.7000000 0.6000000  
## [517] 0.5000000 0.5000000 0.6000000 0.7000000 0.6000000 0.7000000  
## [523] 0.6363636 0.6000000 0.5000000 0.7000000 0.8000000 0.6000000  
## [529] 0.6000000 0.7000000 0.5000000 0.8181818 0.5000000 0.6000000  
## [535] 0.8000000 0.9000000 0.5000000 0.7000000 0.9000000 0.5000000  
## [541] 0.5000000 0.5000000 0.7000000 0.5000000 0.5000000 0.6000000  
## [547] 0.4000000 0.6000000 0.6000000 0.5000000 0.4000000 0.8000000  
## [553] 0.4000000 0.9000000 0.5000000 0.7500000 0.6000000 0.6000000  
## [559] 0.6000000 0.5000000 0.6000000 0.5000000 0.6000000 0.7000000  
## [565] 0.9000000 0.6000000 0.6000000 0.7000000 0.9000000 0.9000000  
## [571] 0.5000000 0.6000000 0.8000000 0.8000000 0.8181818 0.8000000  
## [577] 0.9000000 0.3636364 0.5000000 0.9000000 0.5454545 0.6000000  
## [583] 0.8000000 0.4000000 0.5000000 0.8000000 0.5000000 0.8000000  
## [589] 0.5000000 0.7000000 0.9000000 0.8000000 0.7000000 0.7000000  
## [595] 0.8000000 0.5000000 0.7000000 0.5000000 0.5454545 0.7272727  
## [601] 0.8000000 0.4000000 0.5000000 0.5000000 0.5000000 0.7000000  
## [607] 0.7000000 0.5000000 0.6000000 0.6363636 0.8000000 0.4000000  
## [613] 0.6000000 0.6000000 0.6000000 0.6000000 0.7000000 0.5000000  
## [619] 0.6000000 0.6000000 0.6000000 0.6000000 0.6000000 0.6000000  
## [625] 0.5454545 0.6000000 0.8181818 0.8181818 0.6000000 0.4000000  
## [631] 0.5000000 0.7000000 0.5000000 0.7000000 0.7000000 0.4000000  
## [637] 0.6000000 0.6000000 0.7272727 0.7000000 0.6000000 0.8000000  
## [643] 0.6000000 0.5000000 0.7000000 0.8000000 0.6000000 0.5000000  
## [649] 0.8000000 0.7272727 0.5000000 0.8181818 0.7000000 0.5000000  
## [655] 0.5454545 0.6000000 0.6000000 0.6000000 1.0000000 1.0000000  
## [661] 1.0000000 1.0000000 0.7000000 0.5000000 0.8000000 0.7000000  
## [667] 0.6000000 0.7000000 1.0000000 0.5000000 0.6363636 0.4000000  
## [673] 0.6000000 0.4000000 0.6000000 0.5000000 0.8000000 0.6000000  
## [679] 0.5000000 0.4000000 0.4000000 0.6363636 0.6000000 0.5833333  
## [685] 1.0000000 0.6000000 0.5000000 0.6000000 0.6000000 0.4000000  
## [691] 0.5454545 0.6000000 0.5454545 0.6000000 0.6000000 0.6000000  
## [697] 0.5000000 1.0000000 0.8000000 0.5000000 0.3636364 0.8000000  
## [703] 0.4000000 0.9000000 0.6000000 0.7000000 0.6000000 0.7000000  
## [709] 0.8000000 0.5000000 0.5000000 0.5000000 0.7272727 0.9000000  
## [715] 0.9090909 0.7000000 0.7000000 0.6000000 0.6666667 0.9000000  
## [721] 0.6363636 0.8000000 0.6000000 0.7000000 0.8000000 0.5454545  
## [727] 0.6000000 0.7000000 0.5000000 0.5454545 0.9000000 0.8000000  
## [733] 0.5000000 0.8000000 0.8000000 0.5000000 1.0000000 0.4000000  
## [739] 1.0000000 0.6000000 0.4000000 0.4000000 0.9000000 0.6000000  
## [745] 0.6000000 0.6000000 0.6000000 0.8000000 0.8000000 0.8000000  
## [751] 0.5454545 0.9000000 0.7000000 0.6000000 0.6000000 0.6000000  
## [757] 0.8000000 0.5000000 0.9000000 0.9000000 0.9000000 0.8461538  
## [763] 0.7000000 0.6000000 0.4545455 0.8461538 0.9000000 0.5000000  
## [769] 1.0000000 1.0000000 0.5000000 0.7000000 0.5000000 0.6000000  
## [775] 0.6000000 0.5000000 0.5000000 0.5833333 0.5833333 0.5833333  
## [781] 0.8000000 0.7000000 1.0000000 0.7000000 0.5000000 0.7000000  
## [787] 0.7000000 0.8000000 0.8000000 0.7000000 0.5000000 0.6000000  
## [793] 0.6363636 0.7000000 0.7000000 0.6000000 0.7000000 0.6000000  
## [799] 0.7000000 0.7000000 0.5000000 0.6000000 0.6000000 0.6000000  
## [805] 0.5000000 0.6000000 0.6000000 0.5000000 0.5000000 0.6000000  
## [811] 0.6000000 0.5000000 0.5000000 0.5000000 0.6000000 0.5000000  
## [817] 0.5000000 0.8000000 0.8000000 0.6363636 0.7000000 0.7000000  
## [823] 0.8000000 0.5454545 0.7000000 0.8000000 0.4545455 0.6000000  
## [829] 0.6000000 0.8000000 0.5000000 0.5000000 0.5000000 0.5000000  
## [835] 0.8000000 0.8000000 0.8000000 0.7000000 0.5000000 1.0000000  
## [841] 0.5000000 0.5000000 0.5000000 0.5000000 0.4000000 0.6363636  
## [847] 0.6000000 0.7000000 0.6000000 0.8000000 0.6363636 0.6000000  
## [853] 0.8000000 0.4000000 0.9000000 0.8000000 0.5454545 0.5000000  
## [859] 0.8000000 0.7000000 0.8000000 0.7000000 0.8000000 0.5000000  
## [865] 0.8000000 0.9000000 0.7000000 0.6000000 0.6000000 0.7272727  
## [871] 0.5000000 0.5000000 0.4545455 0.8000000 0.7000000 0.9000000  
## [877] 0.6000000 0.5000000 0.5000000 0.8181818 0.9090909 0.8000000  
## [883] 0.7000000 0.7000000 0.5000000 0.4545455 1.0000000 0.4545455  
## [889] 0.6000000 0.7000000 0.6000000 0.4545455 0.5000000 0.7000000  
## [895] 0.4000000 0.7000000 0.5000000 0.6000000 0.5454545 0.7000000  
## [901] 0.8000000 0.7000000 0.6000000 0.9000000 0.8000000 0.8000000  
## [907] 0.9000000 0.5454545 0.4000000 0.5000000 0.8000000 0.6000000  
## [913] 0.6000000 0.8000000 0.8000000 0.8000000 0.5000000 0.5000000  
## [919] 0.5454545 0.6000000 0.4000000 0.7000000 0.4000000 0.7000000  
## [925] 0.7000000 0.6000000 0.5000000 0.5454545 0.5000000 0.5000000  
## [931] 0.8000000 0.8000000 0.6363636 0.6000000 0.5000000 0.6000000  
## [937] 0.4000000 0.8181818 0.8000000 0.5000000 0.7000000 0.5000000  
## [943] 0.8000000 0.5000000 0.8000000 0.6000000 0.6000000 0.5454545  
## [949] 0.6000000 0.4545455 0.6000000 0.4000000 0.6000000 0.6000000  
## [955] 0.6000000 0.8181818 0.6000000 0.8000000 0.8000000 0.6000000  
## [961] 0.6000000 0.7000000 0.5000000 0.6000000 0.5000000 0.6000000  
## [967] 0.5454545 0.6000000 0.5000000 0.9000000 0.8000000 0.5000000  
## [973] 0.5000000 0.7000000 0.6363636 0.4000000 0.8000000 0.6000000  
## [979] 0.6000000 0.7000000 0.5000000 0.5000000 0.5000000 0.5000000  
## [985] 0.4000000 0.5000000 0.7000000 0.6000000 0.7000000 0.6000000  
## [991] 0.6000000 0.7000000 0.7000000 0.8000000 0.9000000 0.7000000  
## [997] 0.6000000 0.6000000 0.6000000 0.7000000 0.6000000 1.0000000  
## [1003] 0.5000000 0.6000000 0.5000000 0.8000000 0.8000000 0.4000000  
## [1009] 0.5000000 0.5000000 0.5000000 0.8000000 0.6000000 0.9000000  
## [1015] 0.5000000 0.5000000 0.8000000 0.5000000 0.9000000 0.7272727  
## [1021] 0.6000000 0.6000000 0.5454545 0.4000000 0.6000000 0.6000000  
## [1027] 0.4000000 0.8000000 0.8000000 0.9000000 0.5000000 0.8000000  
## [1033] 0.8000000 0.6000000 0.7000000 0.5000000 0.8000000 0.8000000  
## [1039] 0.8000000 0.8000000 0.5000000 0.4000000 0.7000000 0.5000000  
## [1045] 0.5000000 0.7000000 0.7000000 0.5000000 0.8000000 0.6000000  
## [1051] 0.6000000 0.6000000 0.7272727 0.6000000 0.7272727 0.8000000  
## [1057] 0.6000000 0.8000000 0.5000000 0.8333333 0.5000000 0.5000000  
## [1063] 0.7000000 0.7272727 0.7000000 1.0000000 0.9000000 0.6000000  
## [1069] 0.6000000 0.6000000 0.6000000 0.8000000 0.5000000 0.4000000  
## [1075] 0.5454545 0.5000000 0.4000000 0.6000000 0.8000000 0.8000000  
## [1081] 0.9000000 0.7000000 0.9000000 0.5000000 0.8181818 0.6000000  
## [1087] 0.6000000 0.6000000 0.5000000 0.8000000 0.8000000 0.7000000  
## [1093] 0.7272727 0.6000000 0.5000000 0.7000000 0.6000000 0.8000000  
## [1099] 0.6000000 0.7000000 0.6000000 0.7000000 0.8000000 0.6000000  
## [1105] 0.8000000 0.7000000 0.6000000 0.5000000 0.6000000 0.6000000  
## [1111] 0.6000000 0.6000000 0.5000000 0.7500000 0.7500000 0.5000000  
## [1117] 0.4000000 0.8000000 0.5384615 0.6000000 0.6000000 0.5000000  
## [1123] 0.5000000 1.0000000 1.0000000 0.5000000 0.6000000 0.5000000  
## [1129] 1.0000000 0.6363636 0.6363636 0.6363636 0.6000000 1.0000000  
## [1135] 1.0000000 0.8000000 0.6000000 0.5000000 0.6000000 0.6000000  
## [1141] 0.6000000 0.5000000 0.6000000 0.6000000 0.5000000 0.5000000  
## [1147] 0.6000000 0.5000000 0.5000000 0.5000000 0.6000000 0.6000000  
## [1153] 0.7000000 0.6000000 0.7000000 0.8000000 0.5000000 0.7000000  
## [1159] 0.6000000 0.7000000 0.8000000 0.7000000 1.0000000 0.6000000  
## [1165] 0.7000000 0.6000000 0.5000000 0.7000000 0.6000000 0.6000000  
## [1171] 0.4000000 0.8000000 0.6000000 0.6000000 0.8000000 0.7000000  
## [1177] 0.9000000 0.6000000 0.7000000 0.9000000 0.5000000 0.7000000  
## [1183] 0.7272727 0.7000000 0.9166667 0.5000000 0.6000000 0.7272727  
## [1189] 0.7272727 0.9000000 0.4000000 0.7000000 0.9000000 0.6000000  
## [1195] 0.6000000 0.7000000 0.5000000 0.8000000 0.5000000 0.8000000  
## [1201] 0.6000000 0.6000000 0.7000000 0.5000000 0.5000000 0.5000000  
## [1207] 1.0000000 1.0000000 1.0000000 0.8125000 0.5000000 1.0000000  
## [1213] 0.6000000 1.0000000 0.6000000 0.5000000 0.8000000 0.5000000  
## [1219] 0.5454545 0.5000000 0.5454545 0.5000000 1.0000000 0.9000000  
## [1225] 0.8000000 0.9000000 0.7000000 0.9000000 0.6000000 0.9000000  
## [1231] 0.8000000 0.5000000 0.7000000 0.5000000 0.8000000 0.8181818  
## [1237] 0.6363636 0.7000000 0.8181818 1.0000000 0.9000000 0.7000000  
## [1243] 0.7000000 0.8000000 0.6000000 0.5000000 0.8181818 0.7000000  
## [1249] 1.0000000 0.4000000 0.6000000 0.7000000 0.6000000 0.5000000  
## [1255] 0.4000000 0.4000000 0.6000000 0.9000000 0.7000000 0.9000000  
## [1261] 0.8000000 0.9000000 0.7000000 0.6000000 0.6000000 0.6000000  
## [1267] 0.7000000 0.8000000 0.5000000 0.6000000 0.5454545 0.8000000  
## [1273] 0.8000000 0.6000000 0.6363636 0.8000000 1.0000000 1.0000000  
## [1279] 1.0000000 0.6000000 0.9000000 1.0000000 0.8000000 0.6000000  
## [1285] 0.7000000 0.5000000 0.7000000 0.7000000 0.7272727 0.8000000  
## [1291] 0.6000000 0.6363636 0.8000000 0.7000000 0.5000000 0.7272727  
## [1297] 0.4000000 0.4000000 0.7000000 1.0000000 0.5000000 0.7000000  
## [1303] 0.6000000 0.8181818 0.6000000 0.4000000 0.7000000 0.7000000  
## [1309] 0.6000000 0.6363636 0.6000000 0.8000000 0.4545455 0.6666667  
## [1315] 0.7000000 0.7000000 0.7000000 0.7000000 0.7000000 0.5000000  
## [1321] 0.5000000 0.6000000 0.4000000 0.7000000 0.8000000 0.6000000  
## [1327] 0.6363636 0.7000000 0.7000000 0.7000000 0.8181818 0.8181818  
## [1333] 0.7272727 0.5000000 0.6000000 0.6000000 0.6000000 0.5454545  
## [1339] 0.7000000 0.8000000 0.4000000 0.5000000 0.5454545 0.9000000  
## [1345] 0.7000000 0.8000000 0.5000000 0.4000000 0.7000000 0.7000000  
## [1351] 0.6000000 0.6000000 0.6000000 0.8000000 0.5454545 0.5454545  
## [1357] 0.6000000 0.8000000 0.5000000 0.5000000 0.5833333 0.6000000  
## [1363] 0.6000000 0.5000000 0.6363636 0.7000000 0.7272727 0.5000000  
## [1369] 0.5000000 0.5000000 0.9000000 0.7000000 0.9000000 0.5384615  
## [1375] 0.6000000 0.6000000 0.9000000 0.5000000 1.0000000 0.9000000  
## [1381] 0.5000000 0.7000000 0.5000000 0.6000000 0.6000000 0.6363636  
## [1387] 0.8000000 0.7000000 0.7000000 0.5000000 0.5000000 0.5454545  
## [1393] 0.7000000 0.7272727 0.5000000 0.7000000 0.6000000 0.9000000  
## [1399] 0.6000000 0.6000000 0.7000000 0.6000000 0.5454545 0.5454545  
## [1405] 0.5000000 0.9000000 0.7000000 0.7000000 0.6000000 0.7000000  
## [1411] 0.4000000 0.7000000 0.5454545 0.8000000 0.9000000 0.8000000  
## [1417] 0.9000000 0.8000000 0.9000000 0.5000000 0.9000000 0.9000000  
## [1423] 0.5000000 0.5000000 0.7272727 0.7000000 0.8000000 0.4000000  
## [1429] 0.8000000 0.7000000 0.7000000 0.6000000 0.6000000 0.9000000  
## [1435] 0.5000000 0.4000000 0.7000000 0.4000000 0.6000000 0.6000000  
## [1441] 0.4000000 0.9000000 0.6363636 0.6000000 0.9000000 0.7500000  
## [1447] 0.7000000 0.7000000 0.7272727 0.5000000 0.6363636 0.5000000  
## [1453] 0.8000000 0.7000000 0.5000000 0.7692308 0.7000000 1.0000000  
## [1459] 0.7000000 0.6000000 0.6000000 0.6000000 0.4545455 0.8181818  
## [1465] 0.6000000 0.5000000 0.7272727 0.8000000 0.6000000 0.5000000  
## Levels: high low medium

Evaluate the model performance.

CrossTable(x=winedata\_test\_labels, y=winedata\_test\_pred, prop.chisq=FALSE)

##   
##   
## Cell Contents  
## |-------------------------|  
## | N |  
## | N / Row Total |  
## | N / Col Total |  
## | N / Table Total |  
## |-------------------------|  
##   
##   
## Total Observations in Table: 1470   
##   
##   
## | winedata\_test\_pred   
## winedata\_test\_labels | high | low | medium | Row Total |   
## ---------------------|-----------|-----------|-----------|-----------|  
## high | 143 | 24 | 155 | 322 |   
## | 0.444 | 0.075 | 0.481 | 0.219 |   
## | 0.530 | 0.055 | 0.204 | |   
## | 0.097 | 0.016 | 0.105 | |   
## ---------------------|-----------|-----------|-----------|-----------|  
## low | 27 | 269 | 182 | 478 |   
## | 0.056 | 0.563 | 0.381 | 0.325 |   
## | 0.100 | 0.611 | 0.239 | |   
## | 0.018 | 0.183 | 0.124 | |   
## ---------------------|-----------|-----------|-----------|-----------|  
## medium | 100 | 147 | 423 | 670 |   
## | 0.149 | 0.219 | 0.631 | 0.456 |   
## | 0.370 | 0.334 | 0.557 | |   
## | 0.068 | 0.100 | 0.288 | |   
## ---------------------|-----------|-----------|-----------|-----------|  
## Column Total | 270 | 440 | 760 | 1470 |   
## | 0.184 | 0.299 | 0.517 | |   
## ---------------------|-----------|-----------|-----------|-----------|  
##   
##

k <- 1:10  
for(x in k){  
 winedata\_test\_pred <- knn(winedata\_train, winedata\_test,  
 winedata\_train\_labels, k = x)  
 accuracy <- mean(winedata\_test\_pred == winedata\_test\_labels)  
 cat('\n','When k =', x, "accuracy is ", accuracy)  
}

##   
## When k = 1 accuracy is 0.6632653  
## When k = 2 accuracy is 0.5884354  
## When k = 3 accuracy is 0.6  
## When k = 4 accuracy is 0.5884354  
## When k = 5 accuracy is 0.5809524  
## When k = 6 accuracy is 0.5857143  
## When k = 7 accuracy is 0.5707483  
## When k = 8 accuracy is 0.5666667  
## When k = 9 accuracy is 0.5714286  
## When k = 10 accuracy is 0.5761905

As we can see, the model has the highest accuracy of 65% when k=1.Then the next highest accuracy is 56% when k=3. When we repeat the process for k from 1 to 50 we see some decrease and then increase.

accuracy <- rep(0, 50)  
k <- 1:50  
for(x in k){  
 winedata\_test\_pred <- knn(winedata\_train, winedata\_test,  
 winedata\_train\_labels, k = x)  
 accuracy[x] <- mean(winedata\_test\_pred == winedata\_test\_labels)  
 }  
plot(k, accuracy, type = 'b')

