R Assignment Template

Amelia

2022-05-10

#### Introduction

Here you will answer some fundamental questions like, but are not limited to: 1. What is the purpose of this assignment? 2. What data will you be using? Where is it from, how was it collected, when etc. 3. Are there any importatntn details about the data? 4. Why is this important?

#

#### Methods

1. What type of statistical test are you using?
2. Why does it make sense for this data?
3. What do you hope to gain out from this methodology?

# run tests

#### Results

1. What are the numeric results of the test?
2. Did everything go how you expected?
3. What do the results mean?
4. Is anything significant?
5. This is where you would but graphs

# report results   
##### Wine Quality In Class Assignment #####   
  
### Loading Data ###   
# Load datatable, corrplot, MASS, dplyr, and car libraries  
library(data.table)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:data.table':  
##   
## between, first, last

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

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

library('ggplot2')  
library('corrplot')

## corrplot 0.92 loaded

library('Rmisc')

## Loading required package: lattice

## Loading required package: plyr

## ------------------------------------------------------------------------------

## You have loaded plyr after dplyr - this is likely to cause problems.  
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:  
## library(plyr); library(dplyr)

## ------------------------------------------------------------------------------

##   
## Attaching package: 'plyr'

## The following objects are masked from 'package:dplyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

library('MASS')

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

library('car')

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

# Load data (pick either white or red wine)  
dt <- read.csv("C:\\Users\\jerem\\OneDrive\\Documents\\School\\\_REGIS\\2022-05\_Summer\\MSDS660\\Week3\\winequality-white.csv", sep = ";")  
  
# Change data to datatable or dataframe   
dt <- as.data.frame(dt)  
  
### Exploratory ###   
# Look at summary to see if there are any NAs or characters in the dataset.   
 # If there are NAs and/or characters remove them.  
head(dt)

## fixed.acidity volatile.acidity citric.acid residual.sugar chlorides  
## 1 7.0 0.27 0.36 20.7 0.045  
## 2 6.3 0.30 0.34 1.6 0.049  
## 3 8.1 0.28 0.40 6.9 0.050  
## 4 7.2 0.23 0.32 8.5 0.058  
## 5 7.2 0.23 0.32 8.5 0.058  
## 6 8.1 0.28 0.40 6.9 0.050  
## free.sulfur.dioxide total.sulfur.dioxide density pH sulphates alcohol  
## 1 45 170 1.0010 3.00 0.45 8.8  
## 2 14 132 0.9940 3.30 0.49 9.5  
## 3 30 97 0.9951 3.26 0.44 10.1  
## 4 47 186 0.9956 3.19 0.40 9.9  
## 5 47 186 0.9956 3.19 0.40 9.9  
## 6 30 97 0.9951 3.26 0.44 10.1  
## quality  
## 1 6  
## 2 6  
## 3 6  
## 4 6  
## 5 6  
## 6 6

nrow(dt)

## [1] 4898

ncol(dt)

## [1] 12

str(dt)

## '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(dt)

## 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 density   
## Min. :0.00900 Min. : 2.00 Min. : 9.0 Min. :0.9871   
## 1st Qu.:0.03600 1st Qu.: 23.00 1st Qu.:108.0 1st Qu.:0.9917   
## Median :0.04300 Median : 34.00 Median :134.0 Median :0.9937   
## Mean :0.04577 Mean : 35.31 Mean :138.4 Mean :0.9940   
## 3rd Qu.:0.05000 3rd Qu.: 46.00 3rd Qu.:167.0 3rd Qu.:0.9961   
## Max. :0.34600 Max. :289.00 Max. :440.0 Max. :1.0390   
## pH sulphates alcohol quality   
## Min. :2.720 Min. :0.2200 Min. : 8.00 Min. :3.000   
## 1st Qu.:3.090 1st Qu.:0.4100 1st Qu.: 9.50 1st Qu.:5.000   
## Median :3.180 Median :0.4700 Median :10.40 Median :6.000   
## Mean :3.188 Mean :0.4898 Mean :10.51 Mean :5.878   
## 3rd Qu.:3.280 3rd Qu.:0.5500 3rd Qu.:11.40 3rd Qu.:6.000   
## Max. :3.820 Max. :1.0800 Max. :14.20 Max. :9.000

which(is.na(dt$fixed.acidity))

## integer(0)

which(is.na(dt$volatile.acidity))

## integer(0)

which(is.na(dt$citric.acid))

## integer(0)

which(is.na(dt$residual.sugar))

## integer(0)

which(is.na(dt$chlorides))

## integer(0)

which(is.na(dt$free.sulfur.dioxide))

## integer(0)

which(is.na(dt$total.sulfur.dioxide))

## integer(0)

which(is.na(dt$density))

## integer(0)

which(is.na(dt$pH))

## integer(0)

which(is.na(dt$sulphates))

## integer(0)

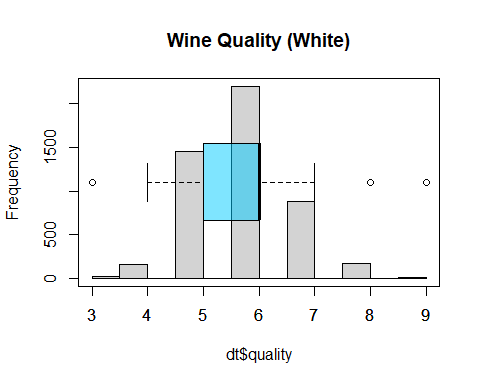
which(is.na(dt$alcohol))

## integer(0)

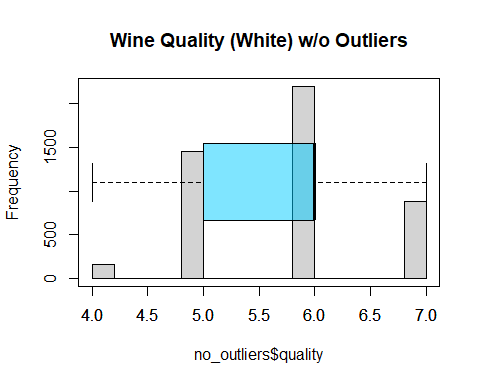
which(is.na(dt$quality))

## integer(0)

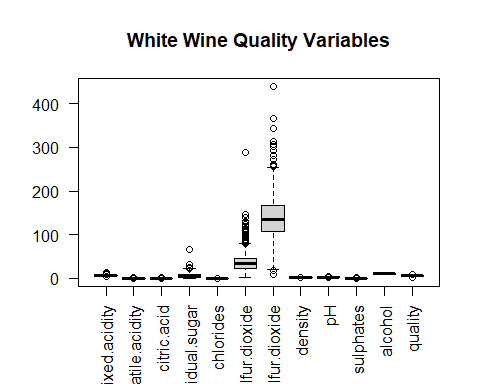
# there are NO null values!  
  
## Plot ##   
# We want to predict wine quality. Plot a histogram and box plot of Quality to look for outliers.  
 # If there are any outliers in quality comment on them and discuss how you will deal with outliers  
hist(dt$quality, main="Wine Quality (White)")  
  
par(new = TRUE)  
  
boxplot(dt$quality, horizontal = TRUE, col = rgb(0, 0.8, 1, alpha = 0.5))  
#title("Wine Quality (White)")  
box()



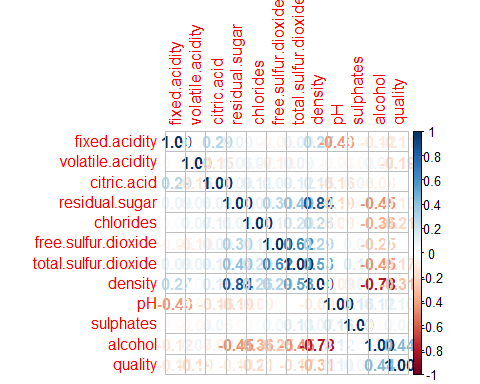
#I will now remove the outliers  
Q1 <- quantile(dt$quality, .25)  
Q3 <- quantile(dt$quality, .75)  
IQR <- IQR(dt$quality)  
#only keep rows in dataframe that have values within 1.5\*IQR of Q1 and Q3  
no\_outliers <- subset(dt, dt$quality> (Q1 - 1.5\*IQR) & dt$quality< (Q3 + 1.5\*IQR))  
  
hist(no\_outliers$quality, main="Wine Quality (White) w/o Outliers")  
par(new = TRUE)  
boxplot(no\_outliers$quality, horizontal = TRUE, col = rgb(0, 0.8, 1, alpha = 0.5))  
box()



par(mfrow = c(1,1))  
  
# Plot all the data to look for correlations  
boxplot(dt, las=2)  
title("White Wine Quality Variables")



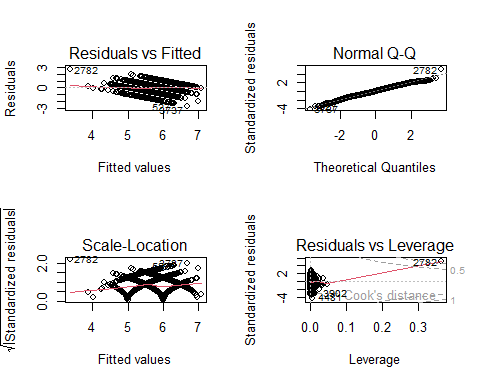
# Plot a correlation matrix  
corrplot(cor(dt), method = 'number')



#corrplot.mixed(cor(d, order = 'AOE')  
  
### Linear Model ###  
# Fit linear model to quality and all variables  
m1 <- lm(quality ~ fixed.acidity + volatile.acidity + citric.acid + residual.sugar + chlorides + free.sulfur.dioxide + total.sulfur.dioxide + density + pH + sulphates + alcohol, data = no\_outliers)  
  
  
## Check Assumptions ##   
# View the summary, VIF, and residual plots of the fit  
summary(m1)

##   
## Call:  
## lm(formula = quality ~ fixed.acidity + volatile.acidity + citric.acid +   
## residual.sugar + chlorides + free.sulfur.dioxide + total.sulfur.dioxide +   
## density + pH + sulphates + alcohol, data = no\_outliers)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.69239 -0.45116 -0.01175 0.45429 2.63200   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.240e+02 1.662e+01 7.462 1.01e-13 \*\*\*  
## fixed.acidity 5.477e-02 1.865e-02 2.936 0.00334 \*\*   
## volatile.acidity -1.797e+00 1.020e-01 -17.610 < 2e-16 \*\*\*  
## citric.acid -5.000e-03 8.492e-02 -0.059 0.95305   
## residual.sugar 6.644e-02 6.681e-03 9.943 < 2e-16 \*\*\*  
## chlorides -3.018e-01 4.854e-01 -0.622 0.53413   
## free.sulfur.dioxide 4.060e-03 7.790e-04 5.212 1.95e-07 \*\*\*  
## total.sulfur.dioxide -2.077e-04 3.373e-04 -0.616 0.53801   
## density -1.233e+02 1.686e+01 -7.311 3.10e-13 \*\*\*  
## pH 5.616e-01 9.403e-02 5.972 2.51e-09 \*\*\*  
## sulphates 6.285e-01 9.056e-02 6.940 4.46e-12 \*\*\*  
## alcohol 1.770e-01 2.145e-02 8.250 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6568 on 4686 degrees of freedom  
## Multiple R-squared: 0.282, Adjusted R-squared: 0.2803   
## F-statistic: 167.3 on 11 and 4686 DF, p-value: < 2.2e-16

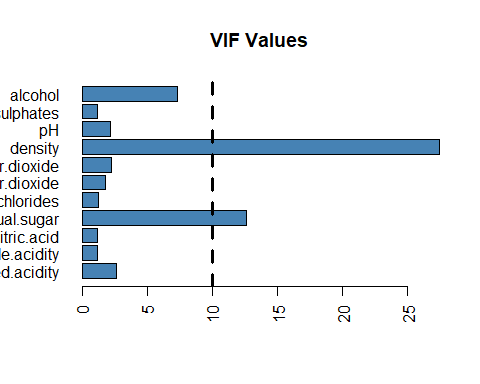
par(mfrow = c(2,2))  
plot(m1)



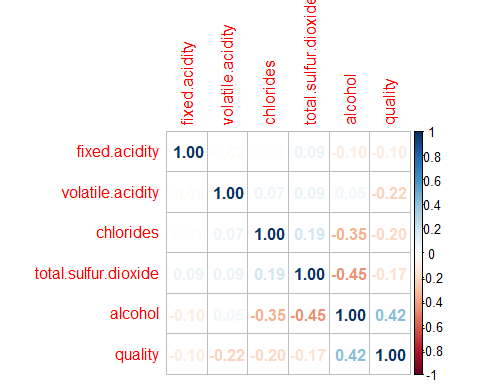
vif\_values <- vif(m1)  
vif\_values

## fixed.acidity volatile.acidity citric.acid   
## 2.653888 1.140338 1.174768   
## residual.sugar chlorides free.sulfur.dioxide   
## 12.635594 1.230366 1.788622   
## total.sulfur.dioxide density pH   
## 2.210407 27.449471 2.183314   
## sulphates alcohol   
## 1.135048 7.313533

par(mfrow = c(1,1))  
barplot(vif\_values, main= "VIF Values", horiz = TRUE, col = "steelblue", las=2)  
abline(v = 10, lwd = 3, lty = 2)



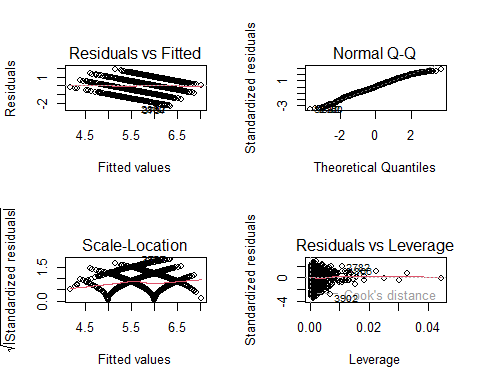
## Modifying Data ##   
# Are there any variables or observations you want to combine or remove based on VIF, p-value, and corrplot?  
 # If so adjust the variables and create a new model and view the summary, VIF, and residuals plots.  
  
#Using the VIF score, it would be smart to drop residual.sugar and density as they both have VIF values above 10.  
  
#Using corrplot, it would be wise to drop citric.acid, free.sulfur.dioxide, pH, residual.sugar and sulphates as they all have nearly 0 correlation with quality  
dt2 = no\_outliers[ , c('fixed.acidity', 'volatile.acidity', 'chlorides', 'total.sulfur.dioxide', 'alcohol', 'quality')]  
  
  
  
corrplot(cor(dt2), method = 'number')



m2 <- lm(quality ~ fixed.acidity + volatile.acidity + chlorides + total.sulfur.dioxide + alcohol, data = dt2)  
summary(m2)

##   
## Call:  
## lm(formula = quality ~ fixed.acidity + volatile.acidity + chlorides +   
## total.sulfur.dioxide + alcohol, data = dt2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.34726 -0.45663 -0.01449 0.47578 1.83954   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.6291152 0.1547646 23.449 < 2e-16 \*\*\*  
## fixed.acidity -0.0613876 0.0117810 -5.211 1.96e-07 \*\*\*  
## volatile.acidity -1.9547877 0.0990589 -19.734 < 2e-16 \*\*\*  
## chlorides -1.4463766 0.4804179 -3.011 0.00262 \*\*   
## total.sulfur.dioxide 0.0014030 0.0002629 5.337 9.89e-08 \*\*\*  
## alcohol 0.2879398 0.0096572 29.816 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6711 on 4692 degrees of freedom  
## Multiple R-squared: 0.2494, Adjusted R-squared: 0.2486   
## F-statistic: 311.8 on 5 and 4692 DF, p-value: < 2.2e-16

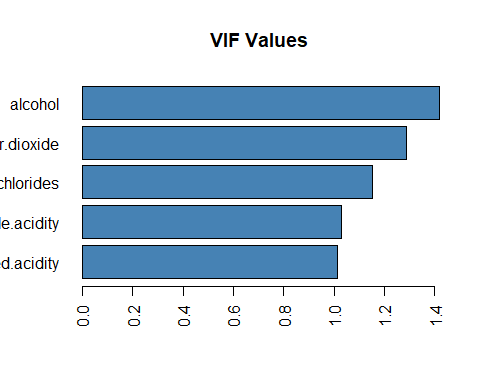
par(mfrow = c(2,2))  
plot(m2)



vif\_values2 <- vif(m2)  
vif\_values2

## fixed.acidity volatile.acidity chlorides   
## 1.013979 1.029832 1.154508   
## total.sulfur.dioxide alcohol   
## 1.286236 1.419496

par(mfrow = c(1,1))  
barplot(vif\_values2, main= "VIF Values", horiz = TRUE, col = "steelblue", las=2)  
abline(v = 10, lwd = 3, lty = 2)

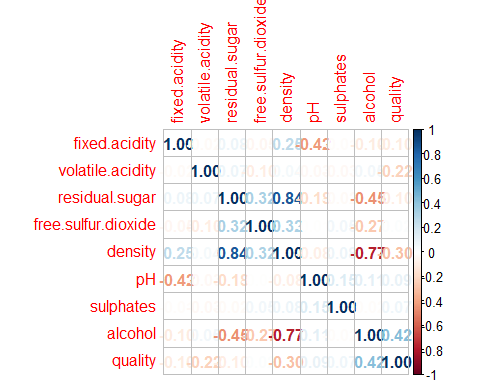


## AIC Plots ##   
# Use step AIC to see if there are any variables to remove and create a new model and view the summary, VIF, and residuals plots (if necessary).  
  
  
# View the AIC score of all the models  
stepAIC(m1, direction = "both")

## Start: AIC=-3937.83  
## quality ~ fixed.acidity + volatile.acidity + citric.acid + residual.sugar +   
## chlorides + free.sulfur.dioxide + total.sulfur.dioxide +   
## density + pH + sulphates + alcohol  
##   
## Df Sum of Sq RSS AIC  
## - citric.acid 1 0.001 2021.5 -3939.8  
## - total.sulfur.dioxide 1 0.164 2021.7 -3939.4  
## - chlorides 1 0.167 2021.7 -3939.4  
## <none> 2021.5 -3937.8  
## - fixed.acidity 1 3.719 2025.2 -3931.2  
## - free.sulfur.dioxide 1 11.717 2033.2 -3912.7  
## - pH 1 15.387 2036.9 -3904.2  
## - sulphates 1 20.777 2042.3 -3891.8  
## - density 1 23.060 2044.5 -3886.5  
## - alcohol 1 29.364 2050.8 -3872.1  
## - residual.sugar 1 42.652 2064.1 -3841.7  
## - volatile.acidity 1 133.782 2155.3 -3638.8  
##   
## Step: AIC=-3939.83  
## quality ~ fixed.acidity + volatile.acidity + residual.sugar +   
## chlorides + free.sulfur.dioxide + total.sulfur.dioxide +   
## density + pH + sulphates + alcohol  
##   
## Df Sum of Sq RSS AIC  
## - total.sulfur.dioxide 1 0.165 2021.7 -3941.4  
## - chlorides 1 0.172 2021.7 -3941.4  
## <none> 2021.5 -3939.8  
## + citric.acid 1 0.001 2021.5 -3937.8  
## - fixed.acidity 1 3.759 2025.2 -3933.1  
## - free.sulfur.dioxide 1 11.729 2033.2 -3914.6  
## - pH 1 15.540 2037.0 -3905.8  
## - sulphates 1 20.802 2042.3 -3893.7  
## - density 1 23.197 2044.7 -3888.2  
## - alcohol 1 29.583 2051.1 -3873.6  
## - residual.sugar 1 42.771 2064.3 -3843.5  
## - volatile.acidity 1 137.861 2159.3 -3631.9  
##   
## Step: AIC=-3941.44  
## quality ~ fixed.acidity + volatile.acidity + residual.sugar +   
## chlorides + free.sulfur.dioxide + density + pH + sulphates +   
## alcohol  
##   
## Df Sum of Sq RSS AIC  
## - chlorides 1 0.175 2021.8 -3943.0  
## <none> 2021.7 -3941.4  
## + total.sulfur.dioxide 1 0.165 2021.5 -3939.8  
## + citric.acid 1 0.003 2021.7 -3939.4  
## - fixed.acidity 1 3.808 2025.5 -3934.6  
## - free.sulfur.dioxide 1 15.447 2037.1 -3907.7  
## - pH 1 15.607 2037.3 -3907.3  
## - sulphates 1 20.638 2042.3 -3895.7  
## - density 1 24.908 2046.6 -3885.9  
## - alcohol 1 29.461 2051.1 -3875.5  
## - residual.sugar 1 44.284 2065.9 -3841.6  
## - volatile.acidity 1 145.241 2166.9 -3617.5  
##   
## Step: AIC=-3943.04  
## quality ~ fixed.acidity + volatile.acidity + residual.sugar +   
## free.sulfur.dioxide + density + pH + sulphates + alcohol  
##   
## Df Sum of Sq RSS AIC  
## <none> 2021.8 -3943.0  
## + chlorides 1 0.175 2021.7 -3941.4  
## + total.sulfur.dioxide 1 0.167 2021.7 -3941.4  
## + citric.acid 1 0.010 2021.8 -3941.1  
## - fixed.acidity 1 4.185 2026.0 -3935.3  
## - free.sulfur.dioxide 1 15.313 2037.2 -3909.6  
## - pH 1 16.676 2038.5 -3906.4  
## - sulphates 1 20.799 2042.6 -3897.0  
## - density 1 26.517 2048.3 -3883.8  
## - alcohol 1 29.458 2051.3 -3877.1  
## - residual.sugar 1 47.550 2069.4 -3835.8  
## - volatile.acidity 1 147.117 2168.9 -3615.1

##   
## Call:  
## lm(formula = quality ~ fixed.acidity + volatile.acidity + residual.sugar +   
## free.sulfur.dioxide + density + pH + sulphates + alcohol,   
## data = no\_outliers)  
##   
## Coefficients:  
## (Intercept) fixed.acidity volatile.acidity   
## 1.279e+02 5.687e-02 -1.813e+00   
## residual.sugar free.sulfur.dioxide density   
## 6.789e-02 3.754e-03 -1.273e+02   
## pH sulphates alcohol   
## 5.735e-01 6.258e-01 1.764e-01

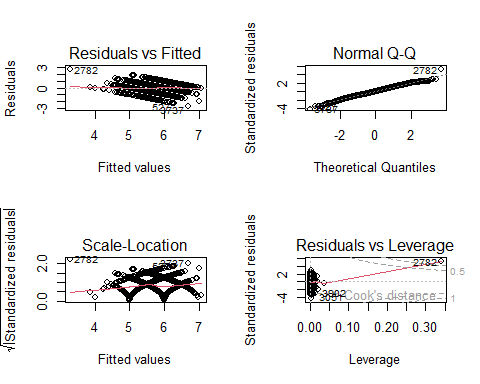
#using the stepAIC, i should remove citric.acid, total.sulfur.dioxide, chlorides  
dt3 = no\_outliers[ , c('fixed.acidity', 'volatile.acidity', 'residual.sugar', 'free.sulfur.dioxide', 'density', 'pH', 'sulphates', 'alcohol', 'quality')]  
corrplot(cor(dt3), method = 'number')



#using correlation, I should remove residual.sugar, free.sulfur.dioxide, pH, sulphates  
m3 <- lm(quality ~ fixed.acidity + volatile.acidity + residual.sugar + free.sulfur.dioxide + density + pH + sulphates + alcohol, data = dt3)  
summary(m3)

##   
## Call:  
## lm(formula = quality ~ fixed.acidity + volatile.acidity + residual.sugar +   
## free.sulfur.dioxide + density + pH + sulphates + alcohol,   
## data = dt3)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.69142 -0.45246 -0.00999 0.45162 2.70985   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.279e+02 1.601e+01 7.988 1.71e-15 \*\*\*  
## fixed.acidity 5.687e-02 1.825e-02 3.115 0.00185 \*\*   
## volatile.acidity -1.813e+00 9.814e-02 -18.471 < 2e-16 \*\*\*  
## residual.sugar 6.789e-02 6.465e-03 10.501 < 2e-16 \*\*\*  
## free.sulfur.dioxide 3.754e-03 6.299e-04 5.959 2.72e-09 \*\*\*  
## density -1.273e+02 1.623e+01 -7.842 5.45e-15 \*\*\*  
## pH 5.735e-01 9.221e-02 6.219 5.44e-10 \*\*\*  
## sulphates 6.258e-01 9.010e-02 6.945 4.30e-12 \*\*\*  
## alcohol 1.764e-01 2.134e-02 8.265 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6566 on 4689 degrees of freedom  
## Multiple R-squared: 0.2819, Adjusted R-squared: 0.2807   
## F-statistic: 230.1 on 8 and 4689 DF, p-value: < 2.2e-16

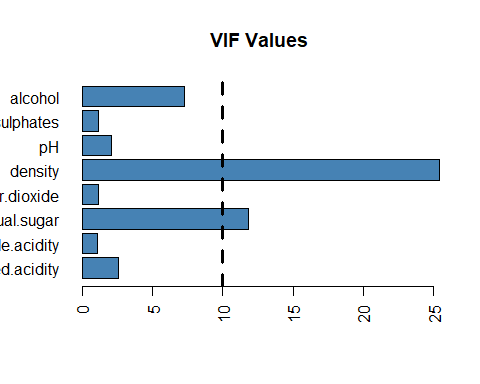
par(mfrow = c(2,2))  
plot(m3)



vif\_values3 <- vif(m3)  
vif\_values3

## fixed.acidity volatile.acidity residual.sugar free.sulfur.dioxide   
## 2.542763 1.055880 11.836008 1.169977   
## density pH sulphates alcohol   
## 25.436202 2.100631 1.123966 7.239737

par(mfrow = c(1,1))  
barplot(vif\_values3, main= "VIF Values", horiz = TRUE, col = "steelblue", las=2)  
abline(v = 10, lwd = 3, lty = 2)



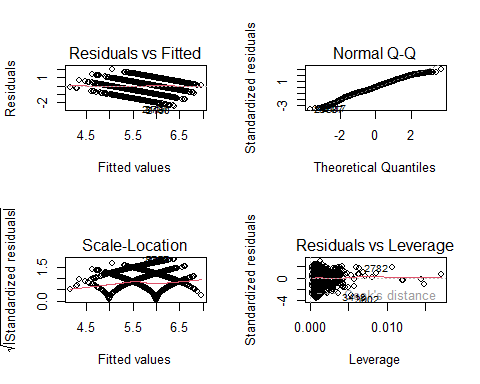
#using VIF, I should remove density and residual.sugar  
  
dt4 = dt3[ , c('fixed.acidity', 'volatile.acidity', 'alcohol', 'quality')]  
corrplot(cor(dt4), method = 'number')



m4 <- lm(quality ~ fixed.acidity + volatile.acidity + alcohol, data = dt4)  
summary(m4)

##   
## Call:  
## lm(formula = quality ~ fixed.acidity + volatile.acidity + alcohol,   
## data = dt4)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.38804 -0.45518 -0.00786 0.47447 1.93542   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.855104 0.126580 30.456 < 2e-16 \*\*\*  
## fixed.acidity -0.057577 0.011808 -4.876 1.12e-06 \*\*\*  
## volatile.acidity -1.913582 0.098114 -19.504 < 2e-16 \*\*\*  
## alcohol 0.274994 0.008189 33.581 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6736 on 4694 degrees of freedom  
## Multiple R-squared: 0.2435, Adjusted R-squared: 0.2431   
## F-statistic: 503.7 on 3 and 4694 DF, p-value: < 2.2e-16

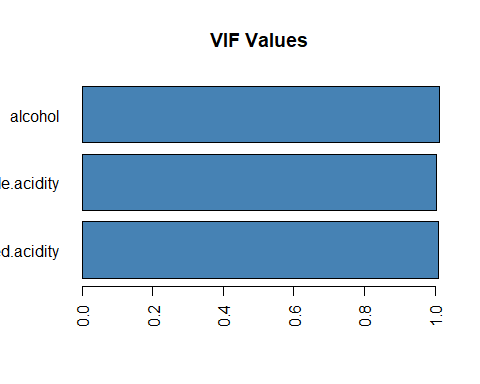
par(mfrow = c(2,2))  
plot(m4)



vif\_values4 <- vif(m4)  
vif\_values4

## fixed.acidity volatile.acidity alcohol   
## 1.011089 1.002867 1.013165

par(mfrow = c(1,1))  
barplot(vif\_values4, main= "VIF Values", horiz = TRUE, col = "steelblue", las=2)  
abline(v = 10, lwd = 3, lty = 2)



### Discussion   
# Post your R file to the discussion board and write a brief summary that includes:  
# Which wine data you used. An interpretation of the model and which variable is most positively   
# correlated to quality and which variable is most negatively correlated with quality.   
# Which variables (if any) did you remove or combine.   
# Why (if applicable) did you remove/combine the variables.  
# Also, post your final AIC and adjusted R^2 value.

#### Conclusion

1. What general conclusions can you make from the results?
2. What does this mean?
3. How is this significant?
4. How did the testing go?
5. What shoud be done different/going forward?