

# wine quality analysis

John Henry Cruz

4/25/2020

## What Factors Contribute to A Top Tier Red Wine? An Analysis

John Henry Cruz

### Packages and Dataset

```
wine_dataset <- read.csv("/Users/johnhenrycruz/Documents/kid-codei/wine_quality/winequality-red.csv", sep = ";")
```

```
library(car)
```

```
## Loading required package: carData
```

### Regression Model and Model Fit (Adjusted R-squared)

```
# linear regression model with the predicting factors and the outcome variable  
mymodel_0 <- lm(quality~pH+residual.sugar+fixed.acidity*alcohol, data=wine_dataset)  
summary(mymodel_0)
```

```
##
## Call:
## lm(formula = quality ~ pH + residual.sugar + fixed.acidity +
##      alcohol, data = wine_dataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7470 -0.4027 -0.1116  0.5089  2.5206
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.360170   0.900184   2.622 0.008828 **
## pH             -0.558337   0.159099  -3.509 0.000462 ***
## residual.sugar -0.014286   0.012489  -1.144 0.252841
## fixed.acidity   0.143758   0.087747   1.638 0.101552
## alcohol         0.461324   0.068569   6.728 2.39e-11 ***
## fixed.acidity:alcohol -0.009741  0.008161  -1.194 0.232832
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6971 on 1593 degrees of freedom
## Multiple R-squared:  0.2571, Adjusted R-squared:  0.2548
## F-statistic: 110.3 on 5 and 1593 DF,  p-value: < 2.2e-16
```

## Correlation Analysis and Test

```
# correlation between each predicting variable and the wine quality
cor(wine_dataset$pH,wine_dataset$quality)
```

```
## [1] -0.05773139
```

```
cor(wine_dataset$alcohol,wine_dataset$quality)
```

```
## [1] 0.4761663
```

```
cor(wine_dataset$residual.sugar,wine_dataset$quality)
```

```
## [1] 0.01373164
```

```
cor(wine_dataset$fixed.acidity,wine_dataset$quality)
```

```
## [1] 0.1240516
```

```
# correlation tests between each predicting variable and the wine quality
cor.test(wine_dataset$pH,wine_dataset$quality)
```

```
##  
## Pearson's product-moment correlation  
##  
## data: wine_dataset$pH and wine_dataset$quality  
## t = -2.3109, df = 1597, p-value = 0.02096  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.106451268 -0.008734972  
## sample estimates:  
## cor  
## -0.05773139
```

```
cor.test(wine_dataset$alcohol,wine_dataset$quality)
```

```
##  
## Pearson's product-moment correlation  
##  
## data: wine_dataset$alcohol and wine_dataset$quality  
## t = 21.639, df = 1597, p-value < 2.2e-16  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.4373540 0.5132081  
## sample estimates:  
## cor  
## 0.4761663
```

```
cor.test(wine_dataset$residual.sugar,wine_dataset$quality)
```

```
##  
## Pearson's product-moment correlation  
##  
## data: wine_dataset$residual.sugar and wine_dataset$quality  
## t = 0.5488, df = 1597, p-value = 0.5832  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.03531327 0.06271056  
## sample estimates:  
## cor  
## 0.01373164
```

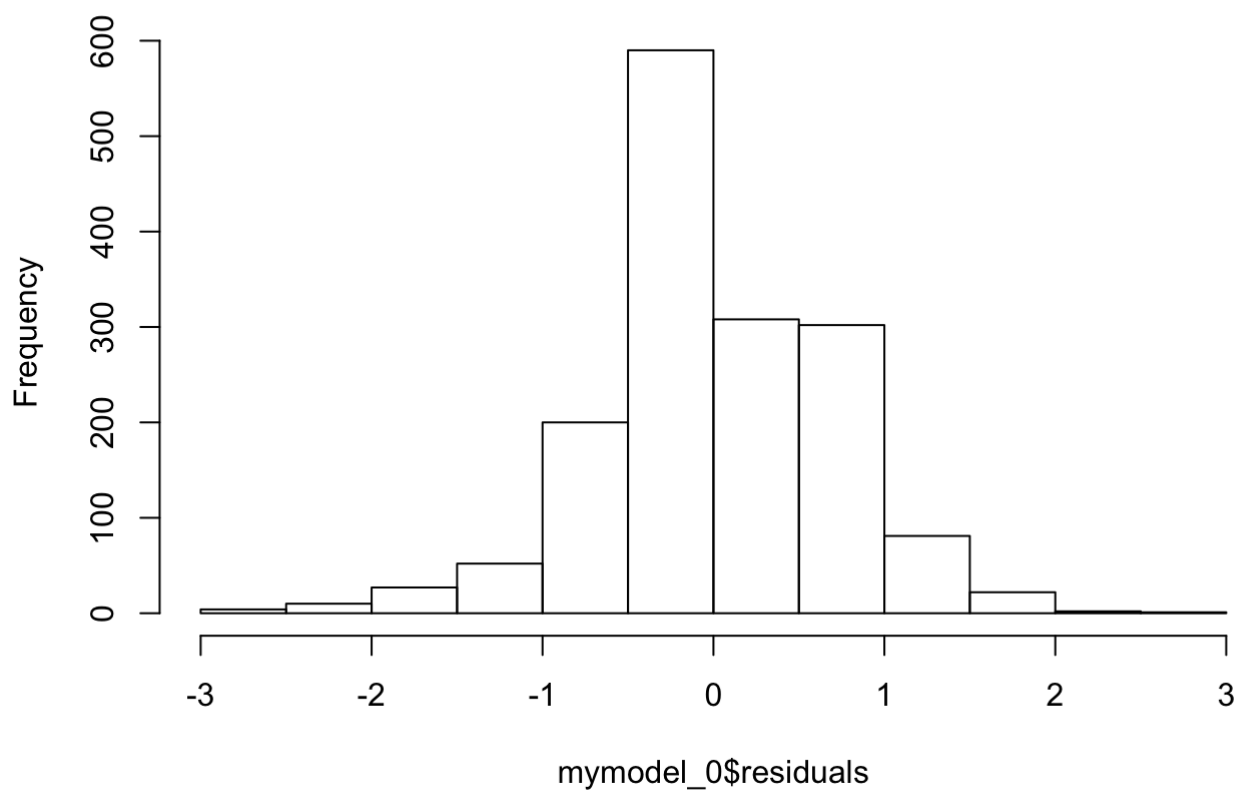
```
cor.test(wine_dataset$fixed.acidity,wine_dataset$quality)
```

```
##  
## Pearson's product-moment correlation  
##  
## data: wine_dataset$fixed.acidity and wine_dataset$quality  
## t = 4.996, df = 1597, p-value = 6.496e-07  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.07548957 0.17202667  
## sample estimates:  
## cor  
## 0.1240516
```

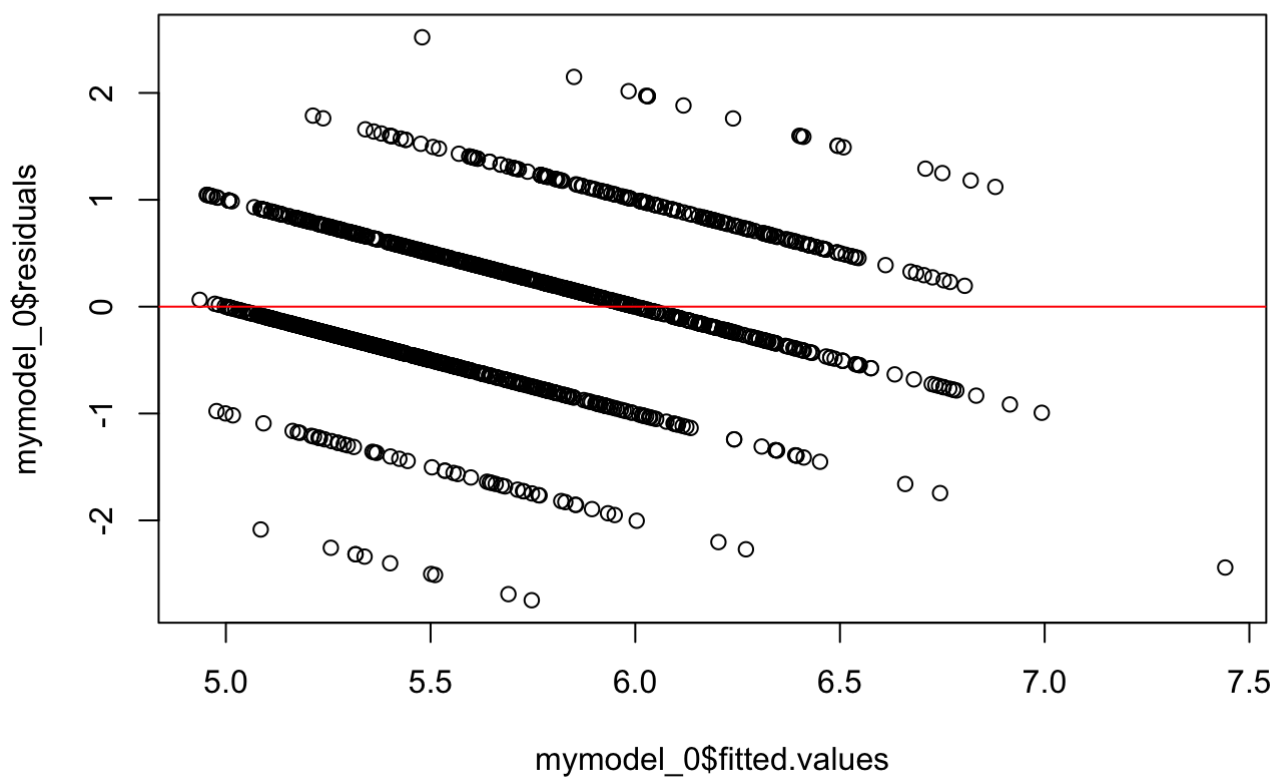
## Assumptions

```
#checking for normality  
hist(mymodel_0$residuals)
```

**Histogram of mymodel\_0\$residuals**



```
#checking for equal variance  
plot(mymodel_0$fitted.values, mymodel_0$residuals)  
abline(h=0, col='red')
```



```
#check for possible colinearity
vif(mymodel_0)
```

```
##           pH          residual.sugar      fixed.acidity
##      1.983798         1.019477         76.745904
##      alcohol fixed.acidity:alcohol
##      17.556829         89.125755
```