R Notebook

Dataset z https://rpubs.com/joelrudinas03/WQD

Obsahuje mereni ruznych parametru vin a jejich hodnoceni (kvalitu).

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

Oznacime vina s kvalitou <6 jako spatna, s kvalitou >= 6 jako dobra.

```
winequality$bad = factor(winequality$quality<6, levels=c(TRUE, FALSE), labels=c('bad','good'))
```

Summary pro spatna vina

```
summary(winequality[winequality$bad == 'bad',])
```

```
##  fixed.acidity    volatile.acidity  citric.acid     residual.sugar  
##  Min.   : 4.200   Min.   :0.1000   Min.   :0.0000   Min.   : 0.600  
##  1st Qu.: 6.400   1st Qu.:0.2400   1st Qu.:0.2400   1st Qu.: 1.700  
##  Median : 6.800   Median :0.2900   Median :0.3200   Median : 6.625  
##  Mean   : 6.962   Mean   :0.3103   Mean   :0.3343   Mean   : 7.054  
##  3rd Qu.: 7.500   3rd Qu.:0.3500   3rd Qu.:0.4100   3rd Qu.:11.025  
##  Max.   :11.800   Max.   :1.1000   Max.   :1.0000   Max.   :23.500  
##    chlorides       free.sulfur.dioxide total.sulfur.dioxide    density      
##  Min.   :0.00900   Min.   :  2.00      Min.   :  9.0        Min.   :0.9872  
##  1st Qu.:0.04000   1st Qu.: 20.00      1st Qu.:117.0        1st Qu.:0.9932  
##  Median :0.04700   Median : 34.00      Median :149.0        Median :0.9951  
##  Mean   :0.05144   Mean   : 35.34      Mean   :148.6        Mean   :0.9952  
##  3rd Qu.:0.05300   3rd Qu.: 49.00      3rd Qu.:182.0        3rd Qu.:0.9971  
##  Max.   :0.34600   Max.   :289.00      Max.   :440.0        Max.   :1.0024  
##        pH         sulphates         alcohol         quality        bad      
##  Min.   :2.79   Min.   :0.2500   Min.   : 8.00   Min.   :3.000   bad :1640  
##  1st Qu.:3.08   1st Qu.:0.4100   1st Qu.: 9.20   1st Qu.:5.000   good:   0  
##  Median :3.16   Median :0.4700   Median : 9.60   Median :5.000              
##  Mean   :3.17   Mean   :0.4815   Mean   : 9.85   Mean   :4.876              
##  3rd Qu.:3.24   3rd Qu.:0.5300   3rd Qu.:10.40   3rd Qu.:5.000              
##  Max.   :3.79   Max.   :0.8800   Max.   :13.60   Max.   :5.000              
```

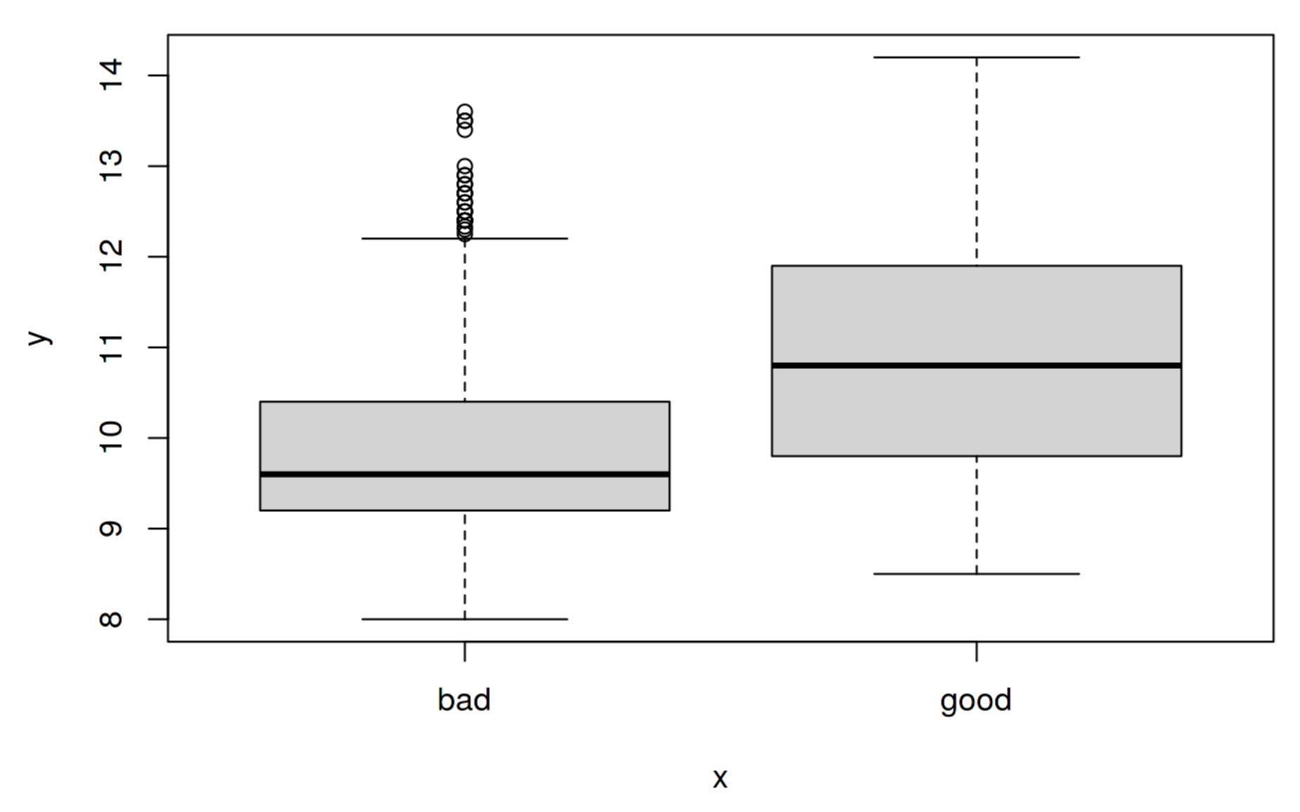
Summary pro dobra vina

```
summary(winequality[winequality$bad == 'good',])
```

```
##  fixed.acidity    volatile.acidity  citric.acid     residual.sugar  
##  Min.   : 3.800   Min.   :0.0800   Min.   :0.0000   Min.   : 0.700  
##  1st Qu.: 6.300   1st Qu.:0.2000   1st Qu.:0.2700   1st Qu.: 1.700  
##  Median : 6.800   Median :0.2500   Median :0.3200   Median : 4.750  
##  Mean   : 6.801   Mean   :0.2621   Mean   :0.3341   Mean   : 6.058  
##  3rd Qu.: 7.300   3rd Qu.:0.3100   3rd Qu.:0.3800   3rd Qu.: 9.100  
##  Max.   :14.200   Max.   :0.9650   Max.   :1.6600   Max.   :65.800  
##    chlorides       free.sulfur.dioxide total.sulfur.dioxide    density      
##  Min.   :0.01200   Min.   :  3.00      Min.   : 18.0        Min.   :0.9871  
##  1st Qu.:0.03400   1st Qu.: 25.00      1st Qu.:105.0        1st Qu.:0.9912  
##  Median :0.04000   Median : 34.00      Median :128.0        Median :0.9930  
##  Mean   :0.04292   Mean   : 35.29      Mean   :133.2        Mean   :0.9935  
##  3rd Qu.:0.04800   3rd Qu.: 45.00      3rd Qu.:158.0        3rd Qu.:0.9954  
##  Max.   :0.25500   Max.   :112.00      Max.   :294.0        Max.   :1.0390  
##        pH          sulphates        alcohol         quality        bad      
##  Min.   :2.720   Min.   :0.220   Min.   : 8.50   Min.   :6.000   bad :   0  
##  1st Qu.:3.090   1st Qu.:0.410   1st Qu.: 9.80   1st Qu.:6.000   good:3258  
##  Median :3.190   Median :0.480   Median :10.80   Median :6.000              
##  Mean   :3.197   Mean   :0.494   Mean   :10.85   Mean   :6.382              
##  3rd Qu.:3.300   3rd Qu.:0.560   3rd Qu.:11.90   3rd Qu.:7.000              
##  Max.   :3.820   Max.   :1.080   Max.   :14.20   Max.   :9.000              
```

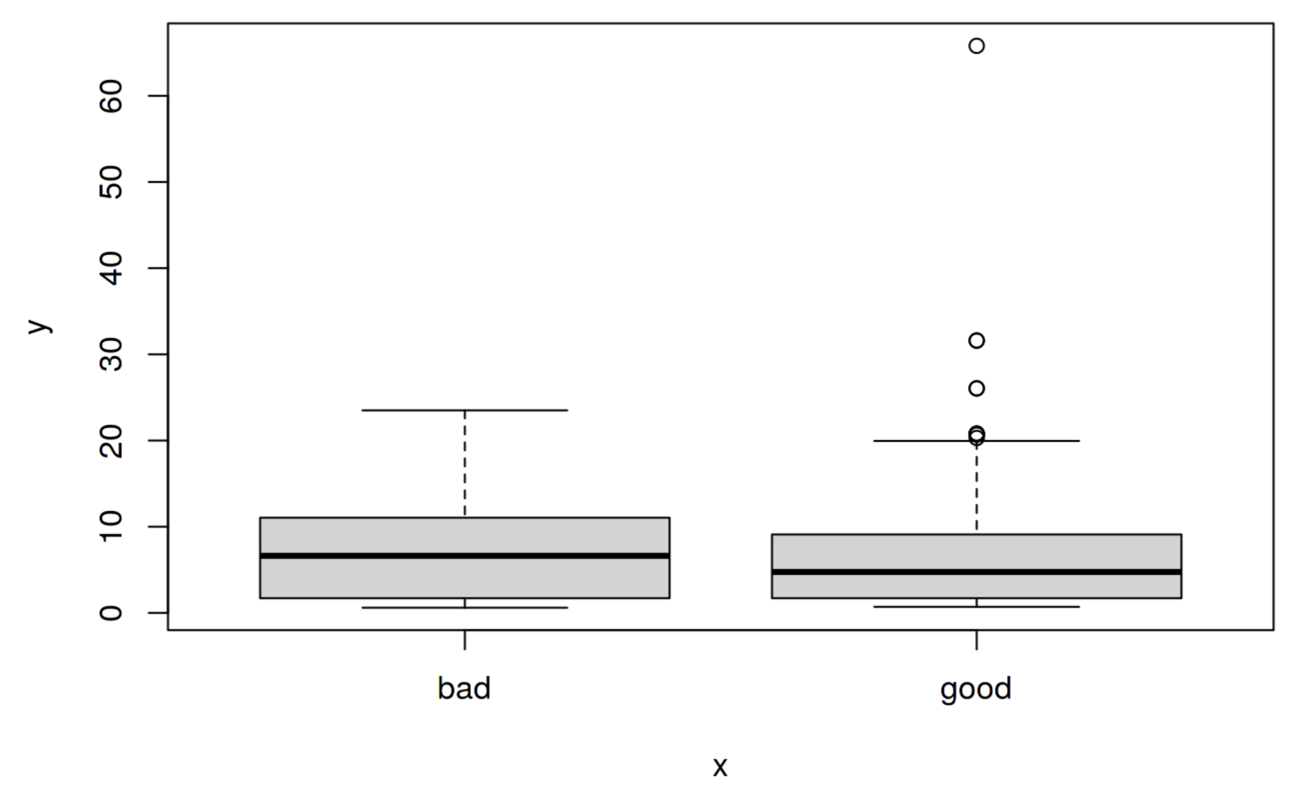
Alcohol

```
plot(winequality$bad, winequality$alcohol)
```



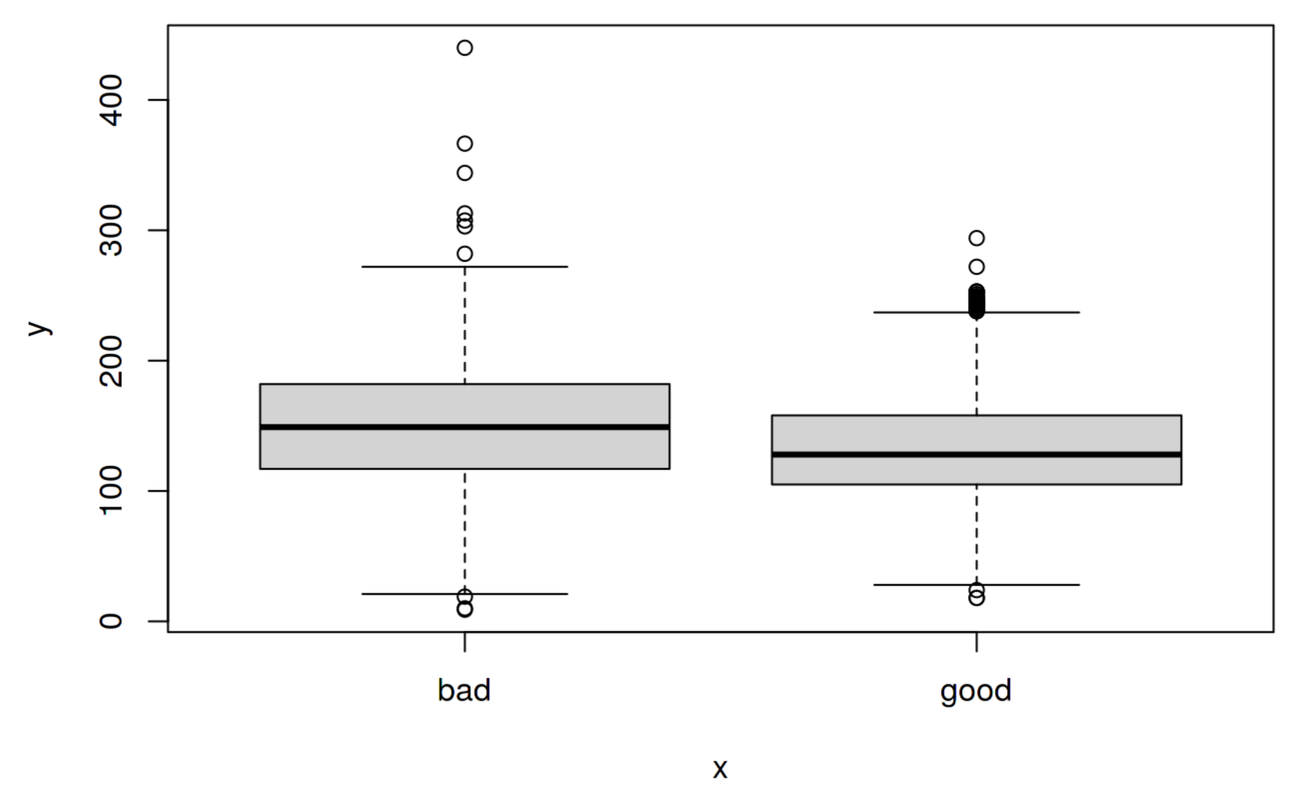
Residual sugar

```
plot(winequality$bad, winequality$residual.sugar)
```



Total sulfur dioxide

```
plot(winequality$bad, winequality$total.sulfur.dioxide)
```



Logisticka regrese

```
glm(winequality$bad ~ winequality$alcohol, family = binomial)
```

```
## 
## Call:  glm(formula = winequality$bad ~ winequality$alcohol, family = binomial)
## 
## Coefficients:
##         (Intercept)  winequality$alcohol  
##              -7.823                0.826  
## 
## Degrees of Freedom: 4897 Total (i.e. Null);  4896 Residual
## Null Deviance:       6245 
## Residual Deviance: 5434  AIC: 5438
```

```
glm(winequality$bad~winequality$alcohol + winequality$residual.sugar + winequality$total.sulfur.dioxide, family =
binomial)
```

```
## 
## Call:  glm(formula = winequality$bad ~ winequality$alcohol + winequality$residual.sugar + 
##     winequality$total.sulfur.dioxide, family = binomial)
## 
## Coefficients:
##                      (Intercept)               winequality$alcohol  
##                        -9.487625                          0.964290  
##       winequality$residual.sugar  winequality$total.sulfur.dioxide  
##                         0.062596                         -0.001147  
## 
## Degrees of Freedom: 4897 Total (i.e. Null);  4894 Residual
## Null Deviance:       6245 
## Residual Deviance: 5364  AIC: 5372
```

```
glm(winequality$bad ~ winequality$fixed.acidity + winequality$volatile.acidity + winequality$citric.acid + winequ
ality$residual.sugar + winequality$chlorides + winequality$free.sulfur.dioxide + winequality$total.sulfur.dioxide
+ winequality$total.sulfur.dioxide + winequality$density + winequality$pH + winequality$sulphates + winequality$a
lcohol, family = binomial)
```

```
## 
## Call:  glm(formula = winequality$bad ~ winequality$fixed.acidity + winequality$volatile.acidity + 
##     winequality$citric.acid + winequality$residual.sugar + winequality$chlorides + 
##     winequality$free.sulfur.dioxide + winequality$total.sulfur.dioxide + 
##     winequality$total.sulfur.dioxide + winequality$density + 
##     winequality$pH + winequality$sulphates + winequality$alcohol, 
##     family = binomial)
## 
## Coefficients:
##                      (Intercept)         winequality$fixed.acidity  
##                        2.582e+02                         3.648e-02  
##     winequality$volatile.acidity           winequality$citric.acid  
##                       -6.459e+00                         1.158e-01  
##       winequality$residual.sugar             winequality$chlorides  
##                        1.701e-01                         8.852e-01  
##  winequality$free.sulfur.dioxide  winequality$total.sulfur.dioxide  
##                        9.601e-03                        -1.333e-03  
##              winequality$density                    winequality$pH  
##                       -2.709e+02                         1.090e+00  
##            winequality$sulphates               winequality$alcohol  
##                        1.797e+00                         7.429e-01  
## 
## Degrees of Freedom: 4897 Total (i.e. Null);  4886 Residual
## Null Deviance:       6245 
## Residual Deviance: 4933  AIC: 4957
```