Exercise_Regression_5.3

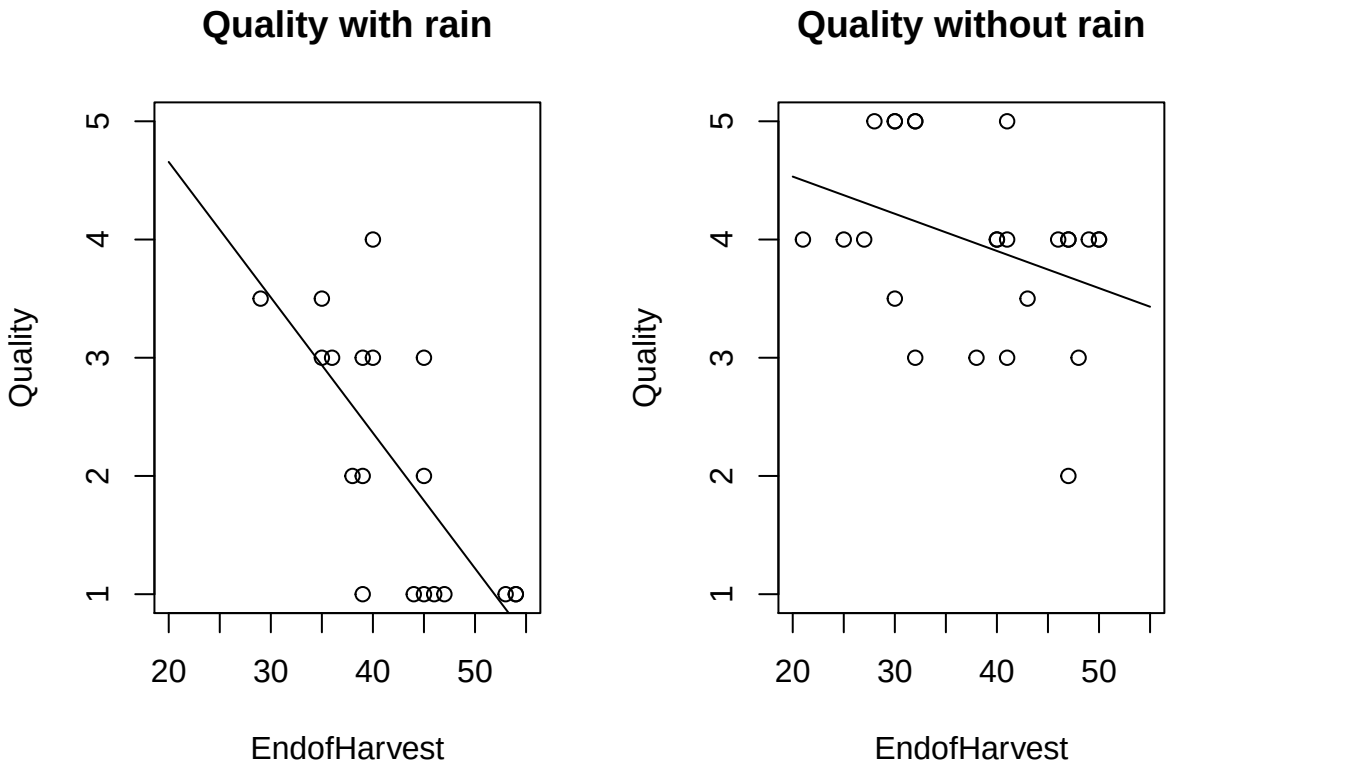
Emil H. Andersen, Nguyen Ho and Afsaneh MohammadNejad

May 20, 2016

The exercise uses information from the data set Latour.txt

Task a)

```
dataT <- read.table("Latour.txt",header=TRUE)
par(mfrow=c(1,2))
x <- seq(20,55,0.1)
yrain <- 5.16122 -0.03145*x + 1.78670 - 0.08314*x
yno <- 5.16122 -0.03145*x
plot(dataT[dataT$Rain==1,]$EndofHarvest,dataT[dataT$Rain==1,]$Quality,
     main='Quality with rain', xlab='EndofHarvest', ylab='Quality', xlim=c(20,55), ylim=c(1,5))
lines(x,yrain)
plot(dataT[dataT$Rain==0,]$EndofHarvest,dataT[dataT$Rain==0,]$Quality,
     main='Quality without rain', xlab='EndofHarvest', ylab='Quality', xlim=c(20,55), ylim=c(1,5))
lines(x,yno)
```



```
fit <- lm(dataT$Quality~dataT$EndofHarvest + dataT$Rain + dataT$EndofHarvest*dataT$Rain)
fit
```

```
##
## Call:
## lm(formula = dataT$Quality ~ dataT$EndofHarvest + dataT$Rain +
##     dataT$EndofHarvest * dataT$Rain)
##
## Coefficients:
##                    (Intercept)             dataT$EndofHarvest
##                        5.16122                        -0.03145
##                     dataT$Rain  dataT$EndofHarvest:dataT$Rain
##                        1.78670                        -0.08314
```

From this fit, we can see that $\beta_0 = 5.16122$, $\beta_1 = -0.03145$, $\beta_2 = 1.78670$, and $\beta_3 = -0.08314$, in the formula (5.10) given in the exercise. Based on the above 2 plots, as well as the information about the values for the β, the rate of change in quality clearly depends on whether there has been any unwanted rain during vintage.

Task b)

We invert the functions for yno and yrain, and check the values for x in the area $y = 4$ and $y = 3$, and see the difference. Since the functions yno and yrain are linear functions, this is going to give us an estimate relevant for question b).

$y_{rain} = 6.94792 - 0.11459x \Leftrightarrow x_{rain} = \frac{-y+6.94792}{0.11459}$

We get $days_{rain} = \frac{-3+6.94792}{0.11459} - \frac{-4+6.94792}{0.11459} = 8.726765$

$y_{norain} = 5.16122 - 0.03145x \Leftrightarrow x_{norain} = \frac{-y+5.16122}{0.03145}$

We get $days_{norain} = \frac{-3+5.16122}{0.03145} - \frac{-4+5.16122}{0.03145} = 31.7965$

So we have that you lose 1 entire quality point every 31.7965 days of harvest if there is no unwanted rain and every 8.726765 days of harvest if there IS unwanted rain.