Test1.0

Juwon Lee

2023-01-13

First of all, we can make a dataset by the csv file.

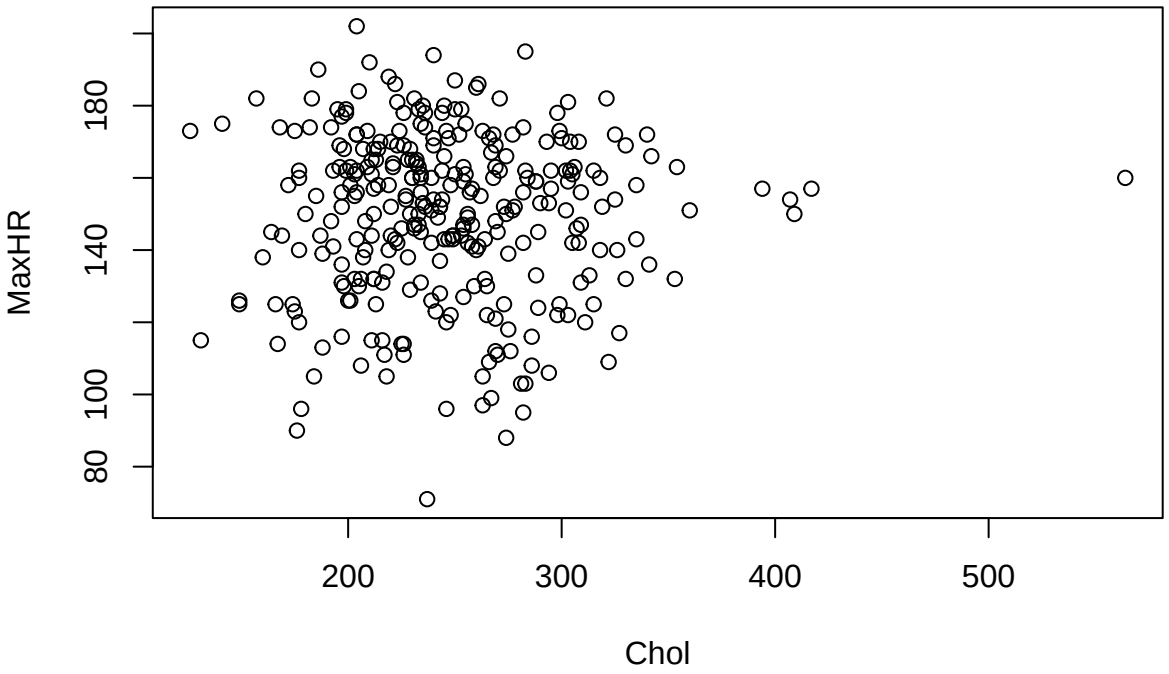
```
heart <- read.csv("/Users/user/Desktop/Yonsei/Junior/3-2/Introduction to Data Analysis and Regression/H

head(heart)
```

```
##   X Age Sex    ChestPain RestBP Chol Fbs RestECG MaxHR ExAng Oldpeak Slope Ca
## 1 1  63   1      typical    145  233   1       2   150     0     2.3     3  0
## 2 2  67   1 asymptomatic    160  286   0       2   108     1     1.5     2  3
## 3 3  67   1 asymptomatic    120  229   0       2   129     1     2.6     2  2
## 4 4  37   1   nonanginal    130  250   0       0   187     0     3.5     3  0
## 5 5  41   0   nontypical    130  204   0       2   172     0     1.4     1  0
## 6 6  56   1   nontypical    120  236   0       0   178     0     0.8     1  0
##         Thal AHD
## 1      fixed  No
## 2     normal Yes
## 3 reversable Yes
## 4     normal  No
## 5     normal  No
## 6     normal  No
```

```
heart1 <- heart[c(6,9)]
```

and we can make a plot,



and we can make Sxx, Syy, Sxy.

```
sum1 <- heart1[1] * heart1[2]

as.numeric(dim(heart1)[1])
```

```
## [1] 303
```

```
as.numeric(lapply(heart1[1], mean))
```

```
## [1] 246.6931
```

```
as.numeric(lapply(heart1[2], mean))
```

```
## [1] 149.6073
```

```
Sxy <- as.numeric(lapply(sum1, sum)) - as.numeric(dim(heart1)[1]) * as.numeric(lapply(heart1[1], mean))

Sxy
```

```
## [1] -1227.525
```

```
sum2 <- heart1[1]^2

Sxx <- as.numeric(lapply(sum2, sum)) - as.numeric(dim(heart1)[1]) * as.numeric(lapply(heart1[1], mean))

Sxx
```

```
## [1] 809616.5
```

```
beta0 <- as.numeric(lapply(heart1[2], mean)) - Sxy / Sxx * as.numeric(lapply(heart1[1], mean))

beta0
```

```
## [1] 149.9813
```

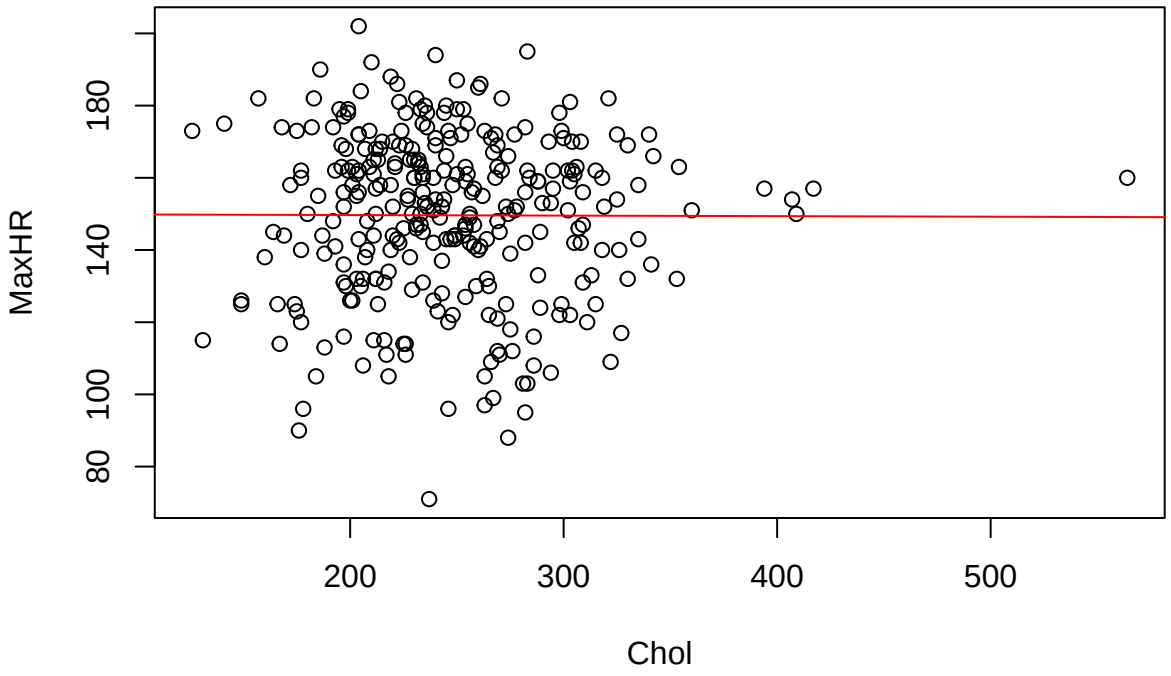
```
beta1 <- Sxy / Sxx

beta1
```

```
## [1] -0.001516181
```

Finally, we can make a simple linear regression line,

```
plot(heart1)
abline(beta0, beta1, col='red')
```



Moreover, we can get R-squared.

```
sum3 <- heart1[2]^2

Syy <- as.numeric(lapply(sum3, sum)) - as.numeric(dim(heart1)[2]) * as.numeric(lapply(heart1[2], mean))

Syy
```

```
## [1] 6895108
```

```
R <- beta1 * Sxy / Syy

R
```

```
## [1] 2.699231e-07
```

I want to check the outcome using lm, and I can get it.

```
MaxHR <- as.vector(heart1[2])
Chol <- as.vector(heart1[1])
```

```r
data_lm <- lm(MaxHR~Chol, data=heart1)
data_lm
```

```
##
## Call:
## lm(formula = MaxHR ~ Chol, data = heart1)
##
## Coefficients:
## (Intercept)         Chol
##  149.981292    -0.001516
```

```r
summary(data_lm)
```

```
##
## Call:
## lm(formula = MaxHR ~ Chol, data = heart1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -78.622 -16.079   3.375  16.412  52.328
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 149.981292   6.418400   23.37   <2e-16 ***
## Chol         -0.001516   0.025465   -0.06    0.953
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.91 on 301 degrees of freedom
## Multiple R-squared:  1.178e-05,  Adjusted R-squared:  -0.00331
## F-statistic: 0.003545 on 1 and 301 DF,  p-value: 0.9526
```

The R-squared in summary is 1.178e-05, which is also too small to assess.