Chapter 2

Hendra Bunyamin

2023-09-09

A Short Simulation of OLS

```
library(tidyverse)

set.seed(1)
tb <- tibble(
  x = rnorm( 10000),
  u = rnorm( 10000),
  y = 5.5 * x + 12 *u)
```

Kita bentuk model regresi linier dengan me-regress y pada x, sbb:

```
reg_tb <- tb %>%
  lm( y ~ x, .) %>%
  print()
```

```
##
## Call:
## lm(formula = y ~ x, data = .)
##
## Coefficients:
## (Intercept)            x
##    -0.04991      5.55690
```

Kita tampilkan koefisien-koefisien dari `reg_tb`, sbb:

```
reg_tb$coefficients
```

```
## (Intercept)           x
## -0.04990882  5.55690164
```

Kita hitung hasil prediksi dari 2 model, yaitu:
1. model dengan y di-regress ke x, dan
2. model dengan $\beta_0 = 0.0732608$ dan $\beta_1 = 5.685033$.

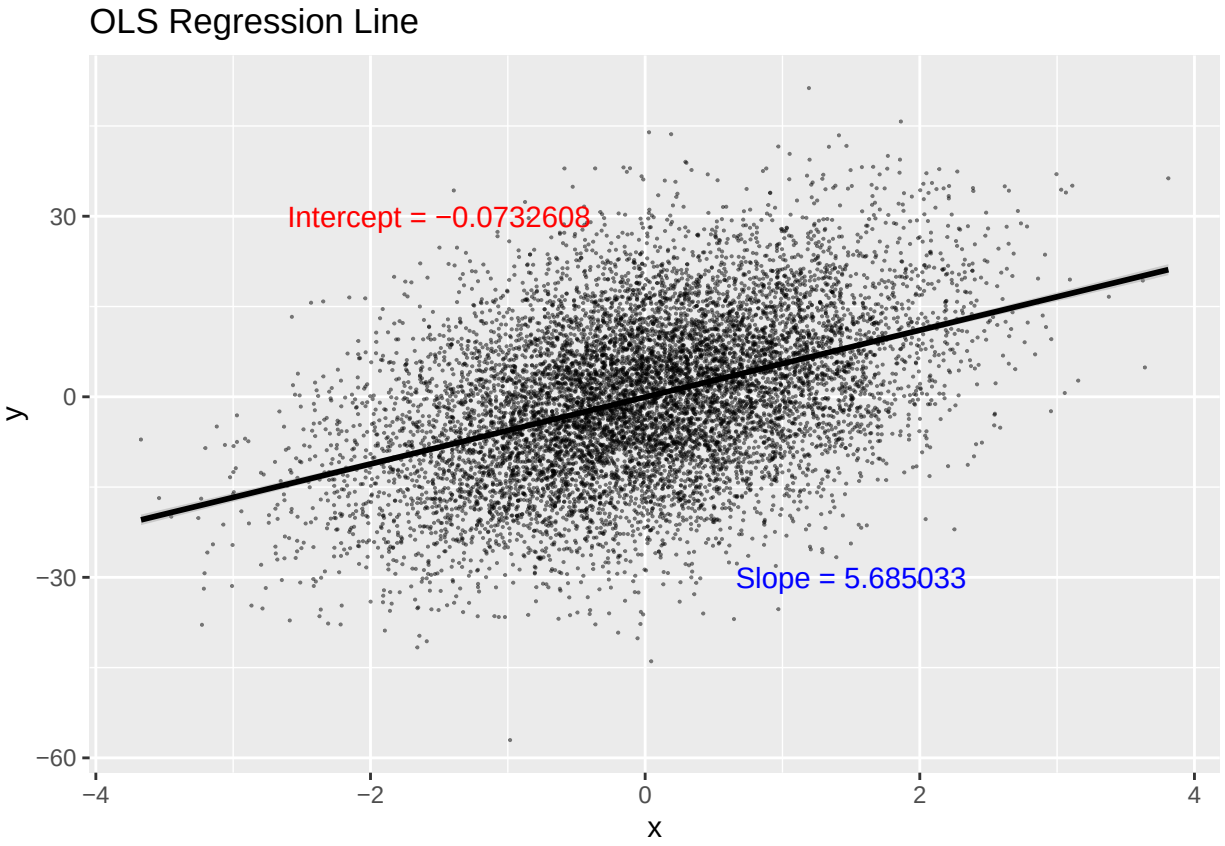
```
tb <- tb %>%
  mutate(
    yhat1 = predict(lm( y ~ x,.  )),
    yhat2 = 0.0732608 + 5.685033 * x,
    uhat1 = residuals( lm( y ~x, .  )  ),
    uhat2 = y - yhat2)
```

```
summary(tb[1:3])
```

```
##        x                   u                   y
##  Min.   :-3.671300   Min.   :-4.30278   Min.   :-57.04286
```

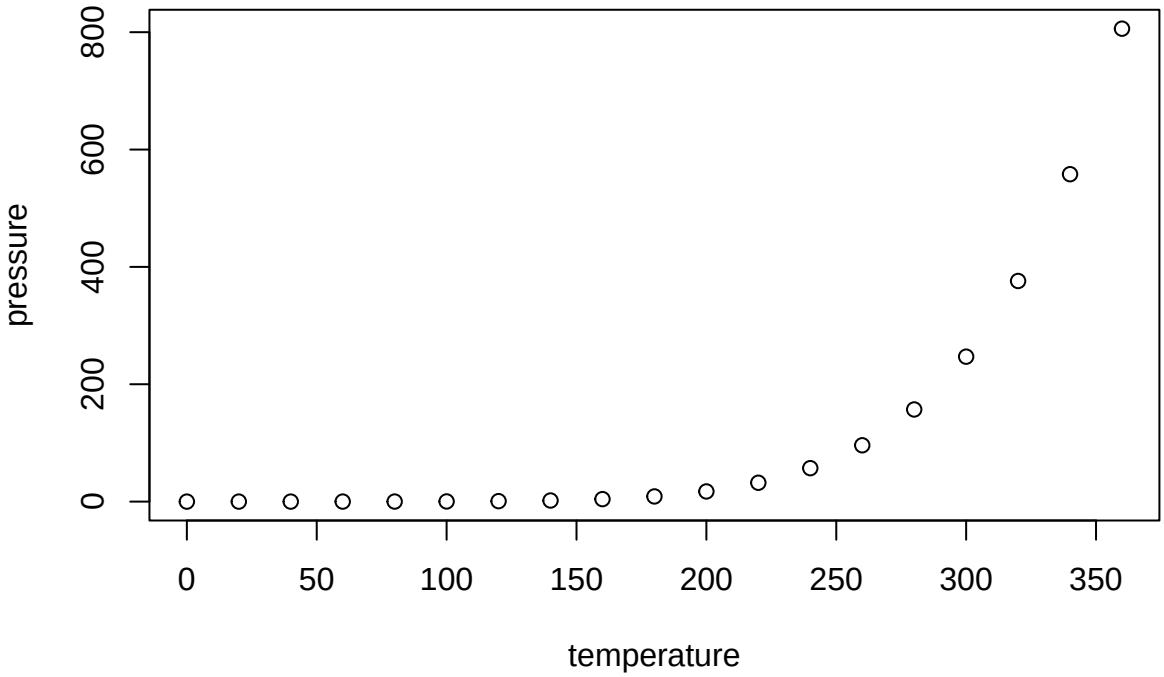
```
##  1st Qu.:-0.673394   1st Qu.:-0.68393   1st Qu.: -8.94045
##  Median :-0.015929   Median :-0.01906   Median : -0.14860
##  Mean   :-0.006537   Mean   :-0.00419   Mean   : -0.08623
##  3rd Qu.: 0.677660   3rd Qu.: 0.66140   3rd Qu.:  8.71725
##  Max.   : 3.810277   Max.   : 3.72796   Max.   : 51.29541
```

```
tb %>% lm(y ~ x, .) %>% ggplot(aes( x=x, y=y  )) +
  ggtitle("OLS Regression Line") +
  geom_point( size=0.05, color="black", alpha=0.5) +
  geom_smooth(method=lm, color="black") +
  annotate( "text", x = -1.5, y=30, color="red",
            label=paste( "Intercept =", -0.0732608  )) +
  annotate( "text", x = 1.5, y=-30, color="blue",
            label=paste( "Slope =", 5.685033  ))
```



Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.