

Tugas Kuliah Anreg Pertemuan 6

Deden Ahmad Rabani

2024-03-06

Baca Data

```
data.anreg <- read.csv("C:/Users/acer/OneDrive - apps.ipb.ac.id/Semester  
4/Anreg/Kuliah 6.csv", sep = ";")
```

```
data.anreg
```

```
##      X  Y  
## 1    2 54  
## 2    5 50  
## 3    7 45  
## 4   10 37  
## 5   14 35  
## 6   19 25  
## 7   26 20  
## 8   31 16  
## 9   34 18  
## 10  38 13  
## 11  45  8  
## 12  52 11  
## 13  53  8  
## 14  60  4  
## 15  65  6
```

```
Y<-data.anreg$Y
```

```
X<-data.anreg$X
```

```
n <- nrow(data.anreg)
```

```
n
```

```
## [1] 15
```

Summary Model Regresi

```
model <- lm(Y~X, data.anreg)
```

```
summary(model)
```

```
##
```

```
## Call:
```

```
## lm(formula = Y ~ X, data = data.anreg)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -7.1628 -4.7313 -0.9253  3.7386  9.0446
```

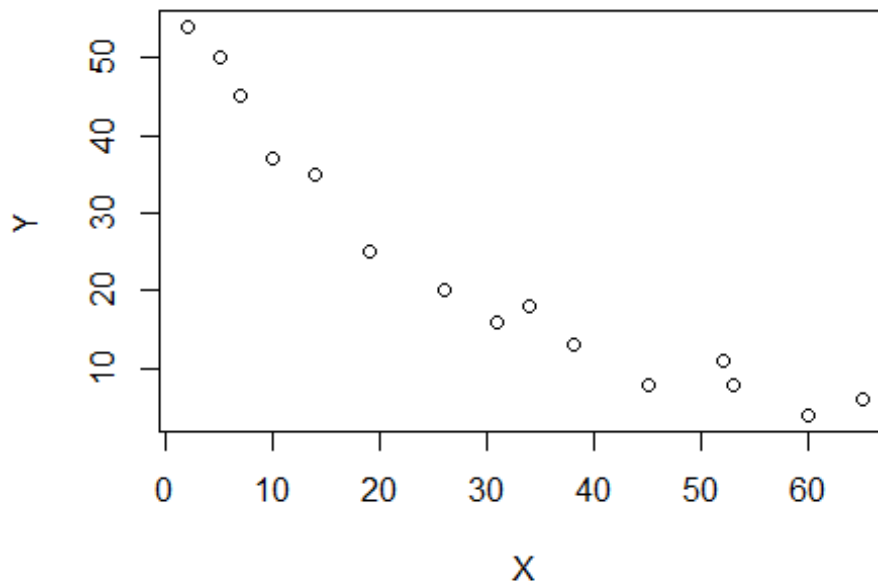
```
##
```

```
## Coefficients:
```

```
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 46.46041    2.76218   16.82 3.33e-10 ***
## X           -0.75251    0.07502  -10.03 1.74e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.891 on 13 degrees of freedom
## Multiple R-squared:  0.8856, Adjusted R-squared:  0.8768
## F-statistic: 100.6 on 1 and 13 DF,  p-value: 1.736e-07
```

Scatter Plot

```
y.bar <- mean(Y)
plot(X,Y)
```



Terlihat dari Scatter plot diatas, hubungan antara X dan Y tidak linear, melainkan membentuk pola ekponensial

Uji Formal Normalitas: Kolmogorov-Smirnov

Hipotesis:

H0: sisaan menyebar Normal

H1: sisaan tidak menyebar Normal

```
library(nortest)
sisaan_model <- resid(model)
(norm_model <- lillie.test(sisaan_model))

##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  sisaan_model
## D = 0.12432, p-value = 0.7701
```

Karena p-value > 0,05 maka Tak Tolak H0. Artinya dalam taraf nyata 5%, tidak cukup bukti untuk menyatakan bahwa sisaan tidak menyebar Normal.

Uji Homogenitas: Breusch-Pagan

Hipotesis:

H0: Ragam Homogen

H1: Ragam tidak Homogen

```
library(lmtest)

## Warning: package 'lmtest' was built under R version 4.3.2
## Loading required package: zoo
## Warning: package 'zoo' was built under R version 4.3.2
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric
(homogen_model <- bptest(model))

##
##  studentized Breusch-Pagan test
##
## data:  model
## BP = 0.52819, df = 1, p-value = 0.4674
```

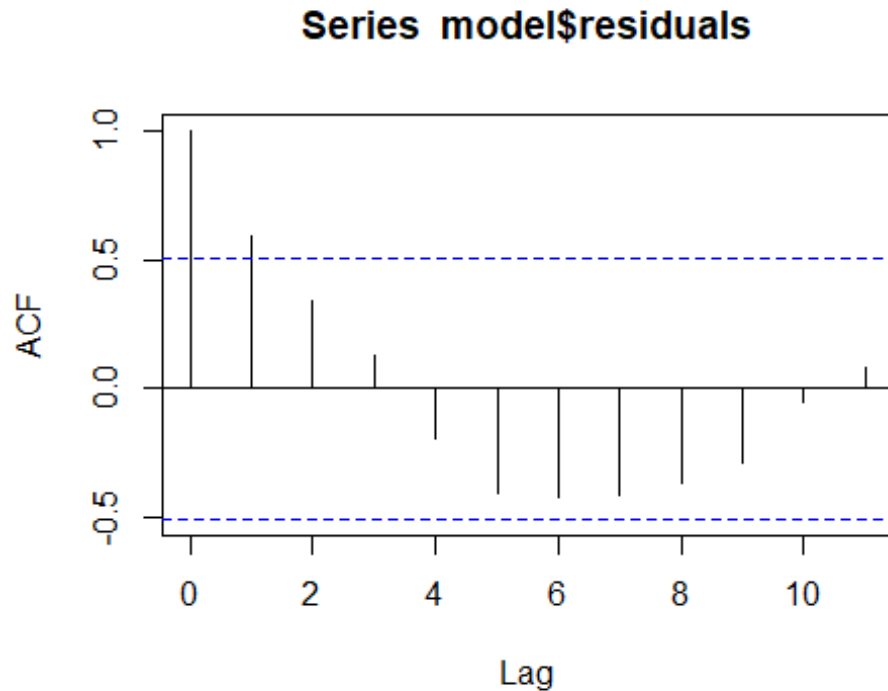
Karena p-value > 0,05 bermakna tak tolak H0. Artinya pada taraf nyata 5%, tidak cukup bukti untuk menyatakan bahwa ragam sisaan tidak homogen.

Autokorelasi

```
dwtest(model)

##
##  Durbin-Watson test
##
```

```
## data: model
## DW = 0.48462, p-value = 1.333e-05
## alternative hypothesis: true autocorrelation is greater than 0
acf(model$residuals)
```



Grafik tersebut menunjukkan bahwa autokorelasi pada lag 1 mencapai 0.5, sedangkan pada lag 2 mencapai 0.4. Kedua nilai tersebut melampaui batas kepercayaan 95%, menunjukkan signifikansi autokorelasi pada lag 1 dan 2. Hal ini berarti bahwa asumsi Gauss-Markov tidak terpenuhi, terutama dalam hal non-autokorelasi.

Percobaan 1: Transformasi Weighted Least Square

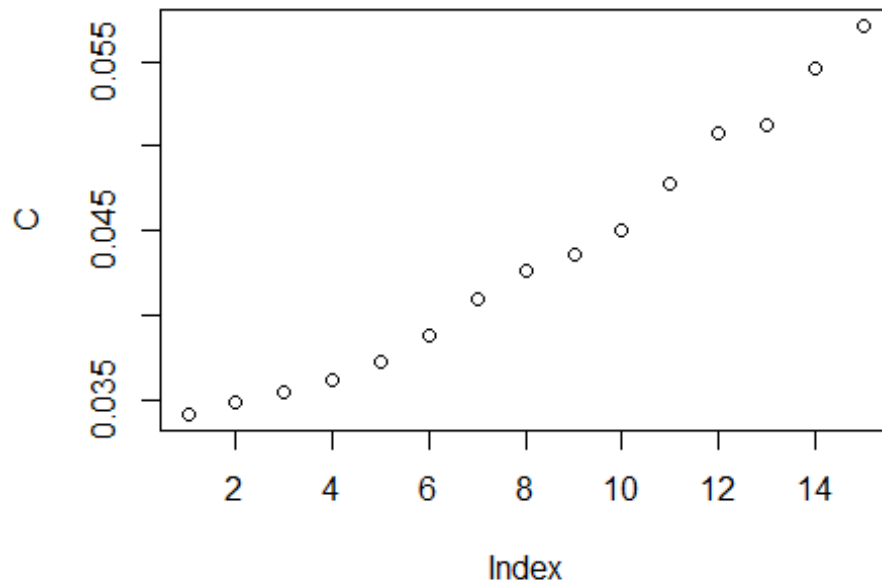
```
A <- abs(model$residuals)
B <- model$fitted.values
fit <- lm(A ~ B, data.anreg)
C <- 1 / fit$fitted.values^2
C
```

	1	2	3	4	5	6
##	0.03414849	0.03489798	0.03541143	0.03620311	0.03730067	0.03874425
7	0.04091034					
	8	9	10	11	12	13
##	0.04257072	0.04361593	0.04507050	0.04779711	0.05077885	0.05122749
14	0.05454132					

```
##          15  
## 0.05710924
```

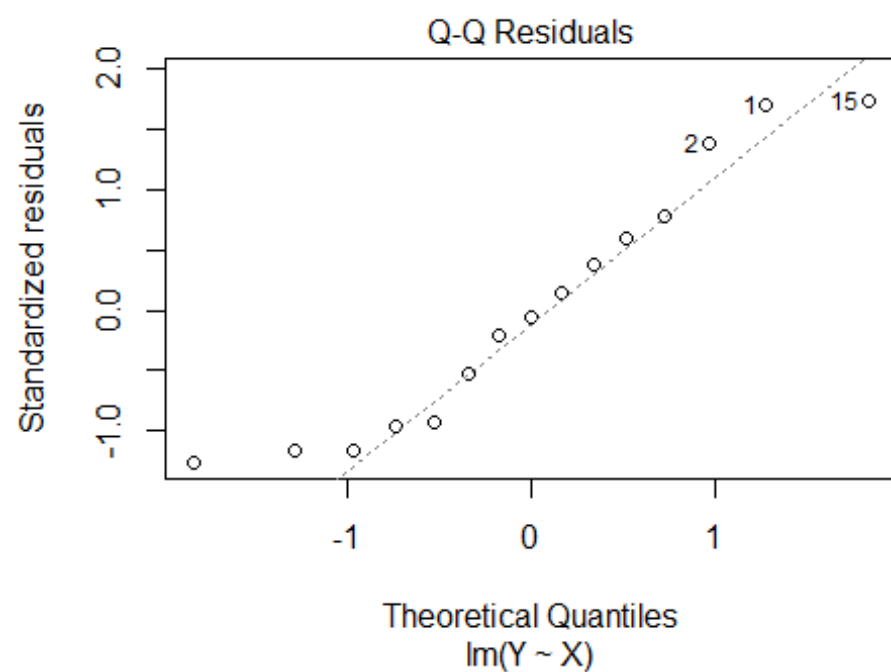
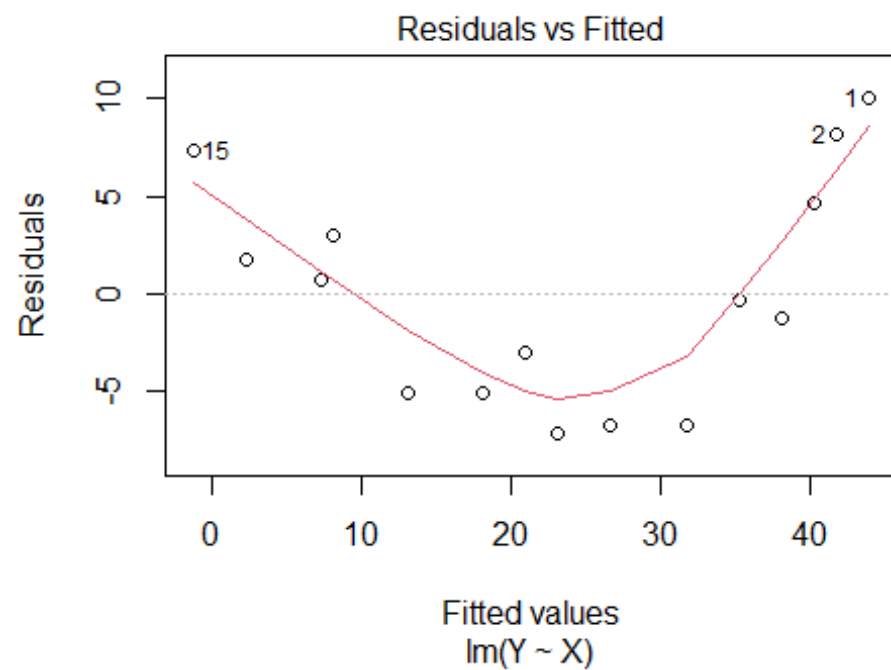
Scatter Plot

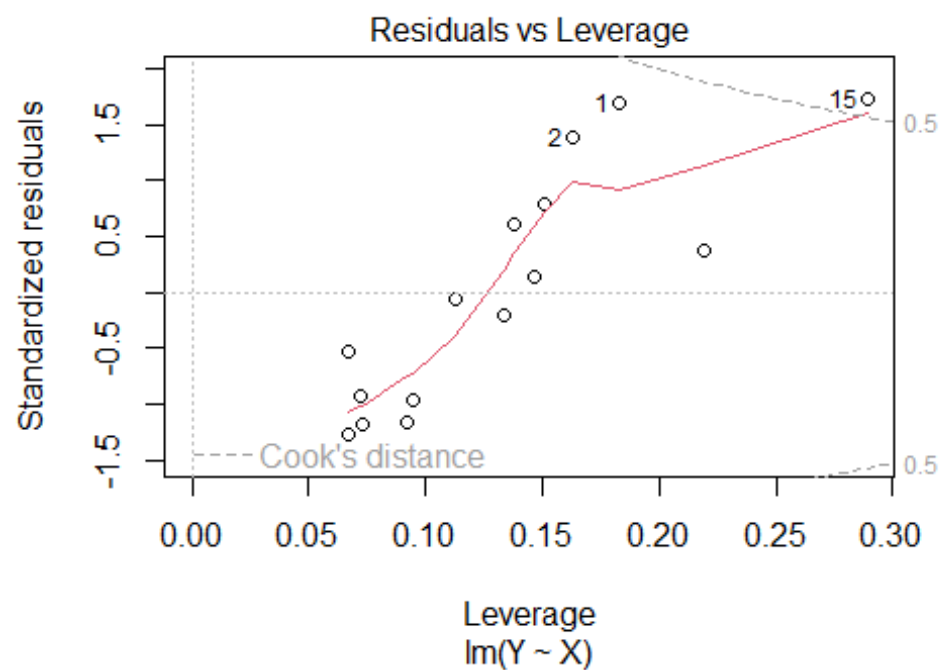
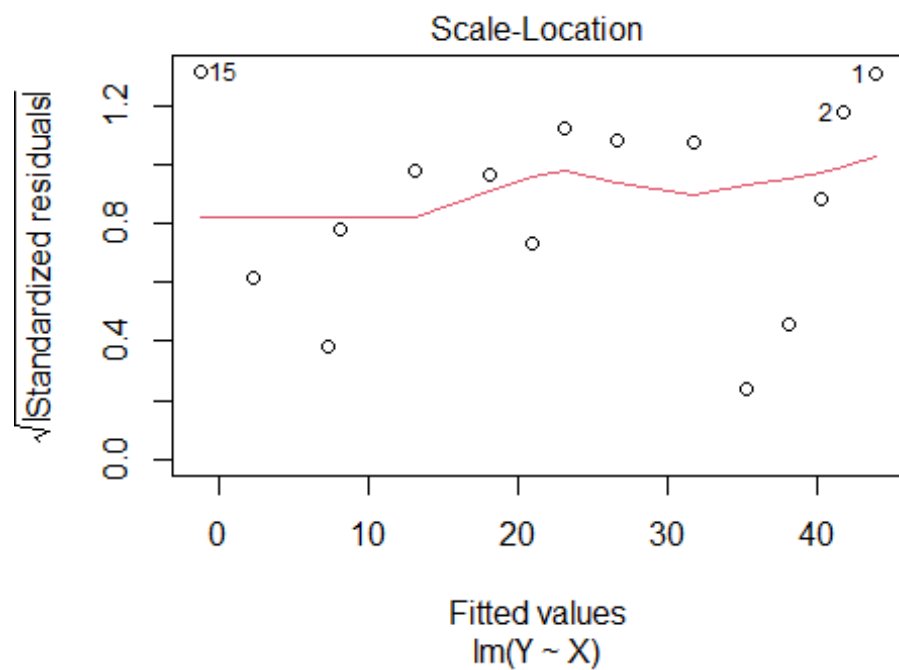
```
plot(C)
```



Model

```
model2 <- lm(Y~X, data=data.anreg, weights = C)  
plot(model2)
```





Summary

```
summary(model2)
```

```
##
## Call:
## lm(formula = Y ~ X, data = data.anreg, weights = C)
##
## Weighted Residuals:
##      Min        1Q    Median        3Q        Max
## -1.46776 -1.09054 -0.06587  0.77203  1.85309
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 45.41058    2.90674  15.623 8.35e-10 ***
## X           -0.71925    0.07313   -9.835 2.18e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.204 on 13 degrees of freedom
## Multiple R-squared:  0.8815, Adjusted R-squared:  0.8724
## F-statistic: 96.73 on 1 and 13 DF,  p-value: 2.182e-07
```

WLS belum efektif karena belum memenuhi asumsi Gauss-Markov

Percobaan 2: Transformasi Akar pada X, Y, atau X dan Y

```
library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.3.2
## Warning: package 'ggplot2' was built under R version 4.3.2
## Warning: package 'tidyr' was built under R version 4.3.2
## Warning: package 'readr' was built under R version 4.3.2
## Warning: package 'dplyr' was built under R version 4.3.2
## Warning: package 'stringr' was built under R version 4.3.2
## Warning: package 'forcats' was built under R version 4.3.2
## Warning: package 'lubridate' was built under R version 4.3.2

## — Attaching core tidyverse packages — tidyverse
2.0.0 —
## ✓ dplyr      1.1.4      ✓ readr      2.1.4
## ✓ forcats   1.0.0      ✓ stringr    1.5.1
## ✓ ggplot2    3.4.4      ✓ tibble     3.2.1
## ✓ lubridate 1.9.3      ✓ tidyr      1.3.0
## ✓ purrr     1.0.2
## — Conflicts —
tidyverse_conflicts() —
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()    masks stats::lag()
```



```

## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
## conflicts to become errors

library(ggribes)
library(GGally)

## Warning: package 'GGally' was built under R version 4.3.2

## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg      ggplot2

library(plotly)

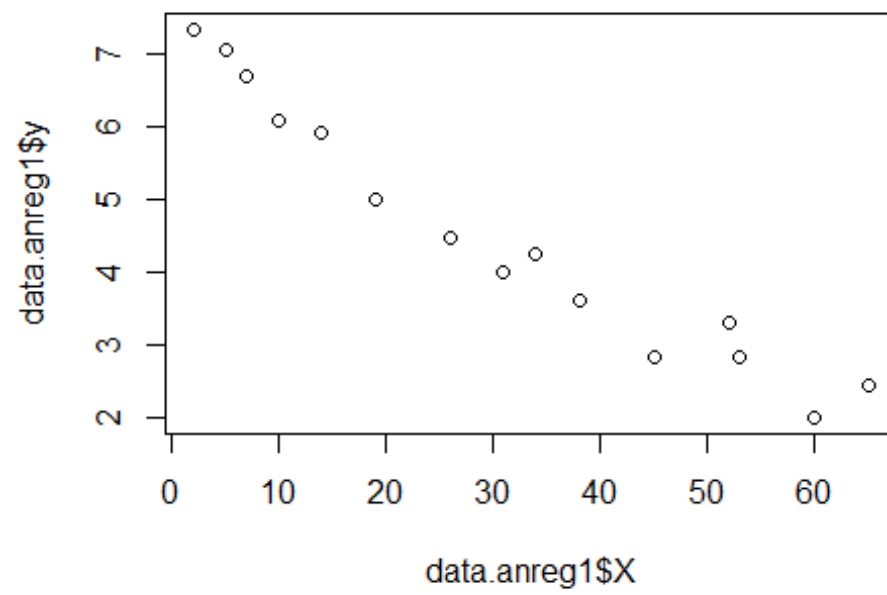
## Warning: package 'plotly' was built under R version 4.3.2

##
## Attaching package: 'plotly'
##
## The following object is masked from 'package:ggplot2':
##
##   last_plot
##
## The following object is masked from 'package:stats':
##
##   filter
##
## The following object is masked from 'package:graphics':
##
##   layout

library(dplyr)
library(lmtest)
library(stats)

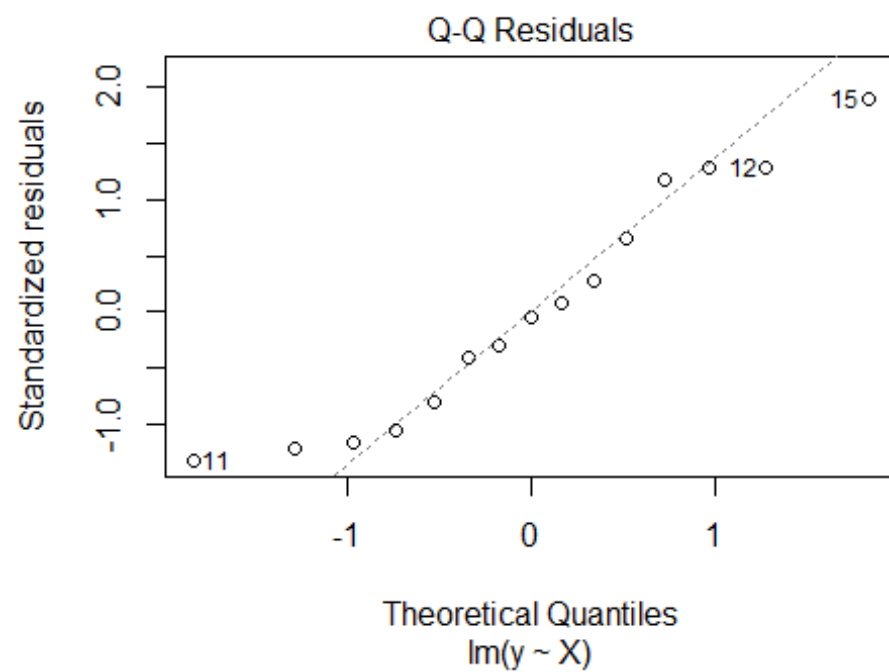
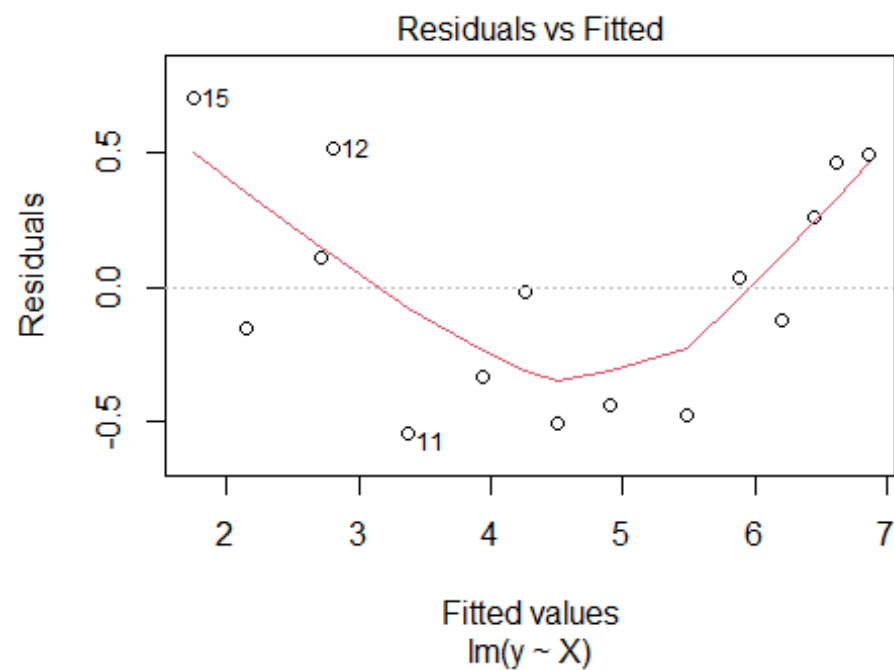
data.anreg1 <- data.anreg %>%
  mutate(y = sqrt(Y)) %>%
  mutate(x = sqrt(X))
model3 <- lm(y ~ X, data = data.anreg1)
plot(x = data.anreg1$X, y = data.anreg1$y)

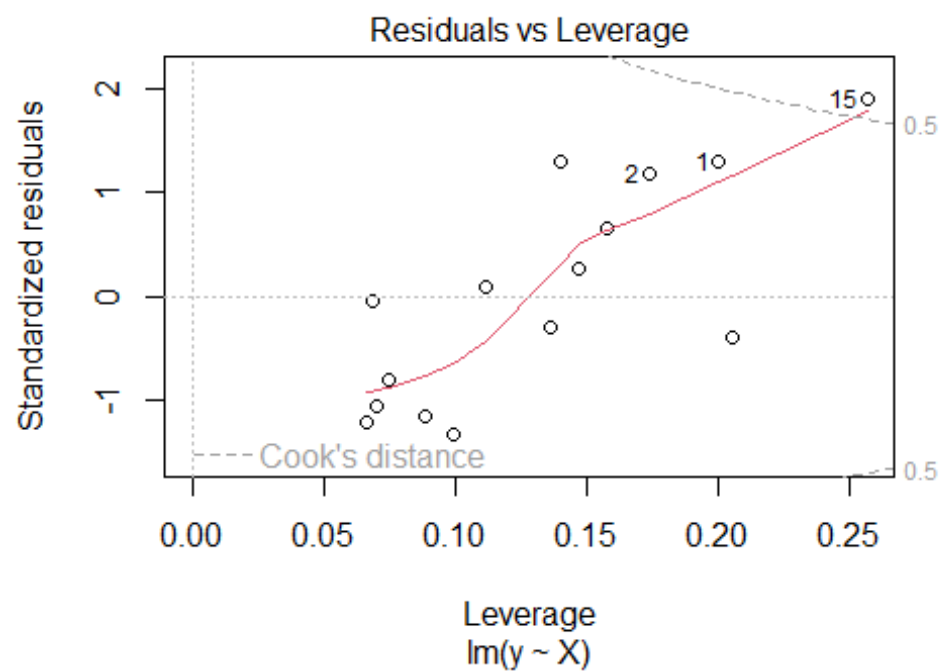
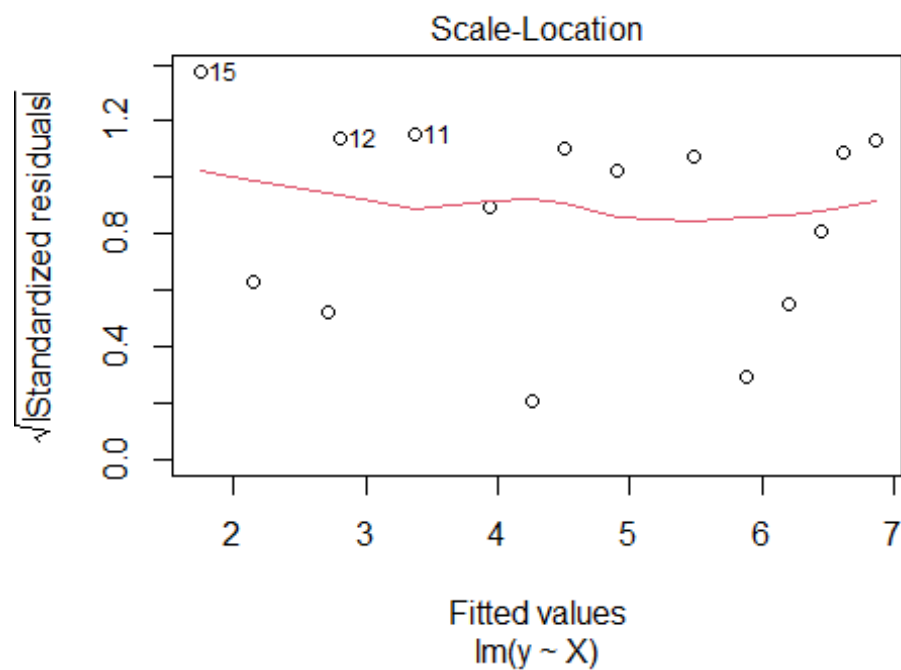
```



Plot

```
plot(model3)
```





Summary

```
summary(model13)
```

```
##
## Call:
## lm(formula = y ~ X, data = data.anreg1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53998 -0.38316 -0.01727  0.36045  0.70199
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.015455   0.201677   34.79 3.24e-14 ***
## X           -0.081045   0.005477  -14.80 1.63e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4301 on 13 degrees of freedom
## Multiple R-squared:  0.9439, Adjusted R-squared:  0.9396
## F-statistic: 218.9 on 1 and 13 DF, p-value: 1.634e-09
```

Uji Autokorelasi

```
dwtest(model3)
```

```
##
## Durbin-Watson test
##
## data: model3
## DW = 1.2206, p-value = 0.02493
## alternative hypothesis: true autocorrelation is greater than 0
```

Kesimpulan: Karena p-Value lebih besar dari 0.05, tidak ada cukup bukti untuk menolak hipotesis nol, artinya tidak adanya autokorelasi. Dari transformasi yang dilakukan, terlihat bahwa menggunakan akar Y sebagai transformasi membuat model regresi lebih efisien.

Model regresi setelah transformasi tersebut dinyatakan sebagai:

$$Y^* = 8.71245 - 0.81339X_1 + \epsilon$$

$$Y^* = \sqrt{Y}$$

$$X^* = \sqrt{X}$$

Dengan melakukan transformasi balik, kita mendapatkan:

$$Y = \left(8.7124535 - 0.8133888X^{\frac{1}{2}}\right)^2$$

Interpretasi dari model menunjukkan bahwa Y berkorelasi negatif dengan akar kuadrat dari X, menunjukkan hubungan yang bersifat kuadratik. Semakin besar nilai akar kuadrat dari X, semakin kecil nilai rata-rata Y, dengan tingkat penurunan yang semakin besar. Puncak kurva menunjukkan nilai rata-rata maksimum Y untuk nilai tertentu dari X. Nilai konstanta 8.71245 mewakili Y ketika X sama dengan 0, sementara koefisien regresi -

0.81339 menunjukkan hubungan negatif antara Y dan akar kuadrat dari X. Koefisien yang dikuadratkan menunjukkan hubungan kuadratik antara Y dan X, menandakan bahwa perubahan Y tidak proporsional dengan perubahan X, melainkan berubah dengan tingkat peningkatan yang semakin tinggi.