Tugas Individu Analisis Regresi_P6

Delita Nur Hasanah G1401221018

2024-03-06

INPUT DATA

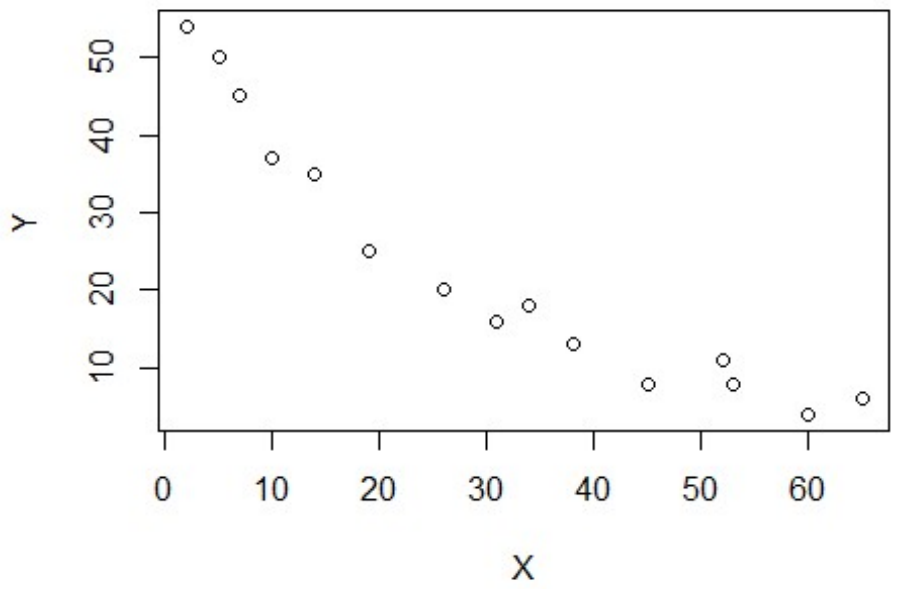
```
data <- read.csv("C:/Users/Delita Nur Hasanah/Downloads/Tugas
Individu.csv",sep = ";")
data
```

```
##     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$Y
X <- data$X
n <- nrow(data)
```

EKSPLORASI DATA

```
ybar <- mean(Y)
plot(X,Y)
```



Terlihat bahwa hubungan antara X dan Y tidak linear dan membentuk pola ekponensial

ANALISIS REGRESI SEDERHANA

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

##
## Call:
## lm(formula = Y ~ X, data = data)
##
## 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
```

UJI FORMAL

1) NORMALITAS: KOLMOGOROV-SMIRNOV

$$H_0: N$$

(sisaan menyebar normal)

$$H_1: N$$

(sisaan tidak menyebar normal)

```
library(nortest)
sisa <- resid(anreg)
(norm <- lillie.test(sisa))

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

Dapat dilihat bahwa P-Value > 0.05, yang berarti Tak Tolak H0. Hal ini menyatakan tidak cukup bukti bahwa sisaan tidak menyebar normal dalam taraf nyata 5%.

2) HOMOGENITAS: BREUSCH-PAGAN

$$H_0: var[\varepsilon] = \sigma^2 I$$

(ragam homogen)

$$H_1: var[\varepsilon] \neq \sigma^2 I$$

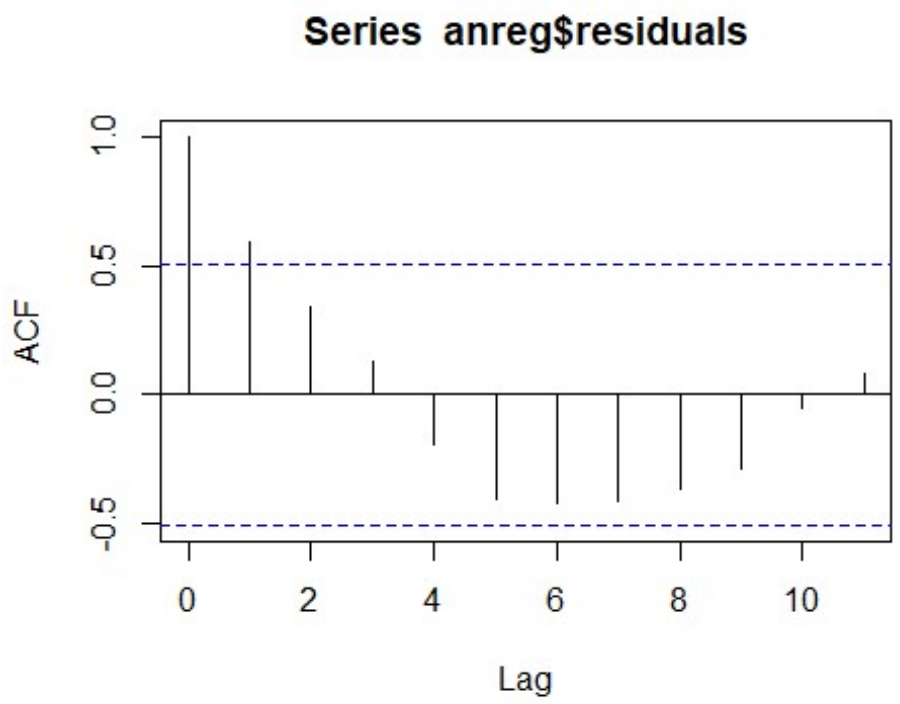
(ragam tidak homogen)

```
library(lmtest)

## Warning: package 'lmtest' was built under R version 4.2.3

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 4.2.3

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

(homogen <- bptest(anreg))

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

Dapat dilihat bahwa P-Value > 0.05, yang berarti Tak Tolak H0. Hal ini menyatakan tidak cukup bukti bahwa sisaan tidak homogen dalam taraf nyata 5%.

AUTOKORELASI

```
dwtest(anreg)

##
##  Durbin-Watson test
##
## data:  anreg
## DW = 0.48462, p-value = 1.333e-05
## alternative hypothesis: true autocorrelation is greater than 0

acf(anreg$residuals)
```

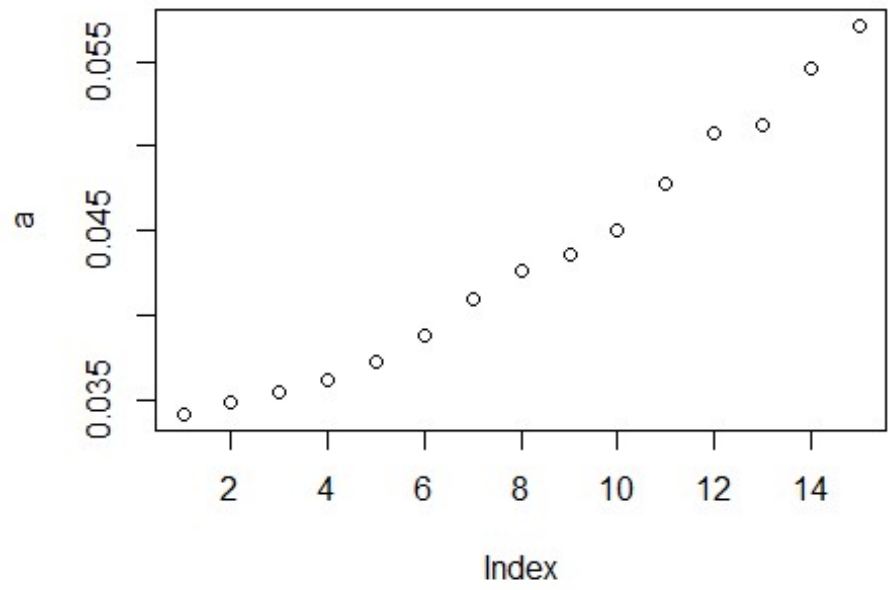


Grafik tersebut menunjukan bahwa autokorelasi pada lag 1adalah 0.5 dan pada lag 2 adalah 0.4. Kedua lag tersebut melebihi batas kepercayaan 95% yang menandakan bahwa autokorelasi pada lag 1 dan 2 signifikan. Hal ini tidak memenuhi asumsi Gauss-Markov.

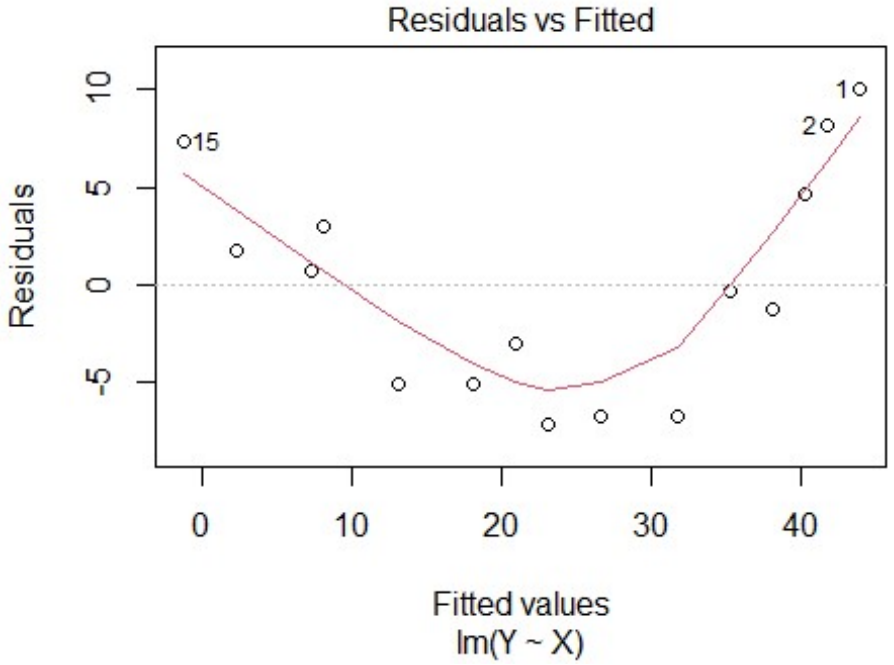
PENANGANAN KONDISI TAK STANDAR 1) TRANSFORMASI WEIGHTED LEAST SQUARE

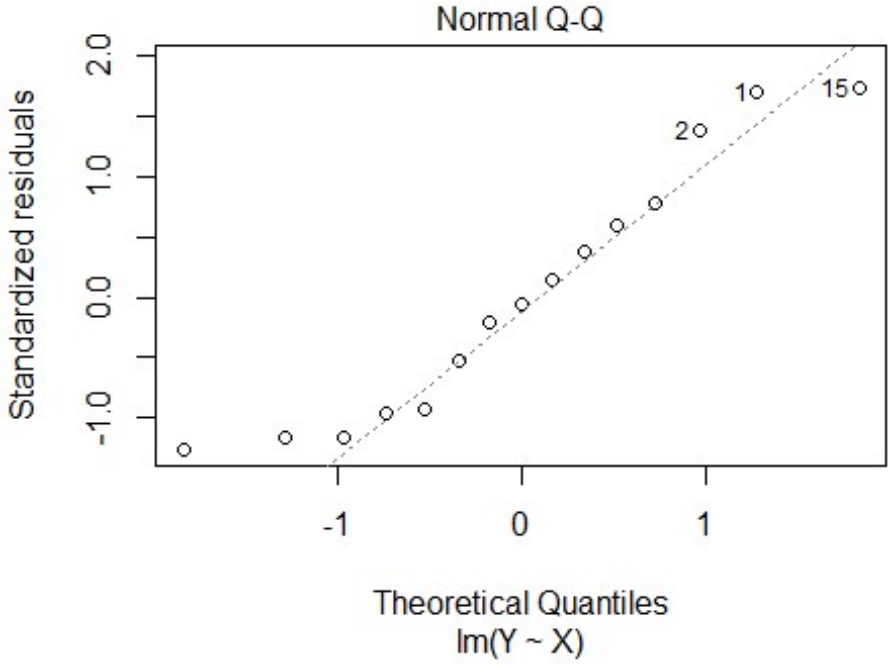
```
sisa2 <- abs(anreg$residuals)
value <- anreg$fitted.values
wls <- lm(sisa2~value, data)
a <- 1/wls$fitted.values^2
a
```

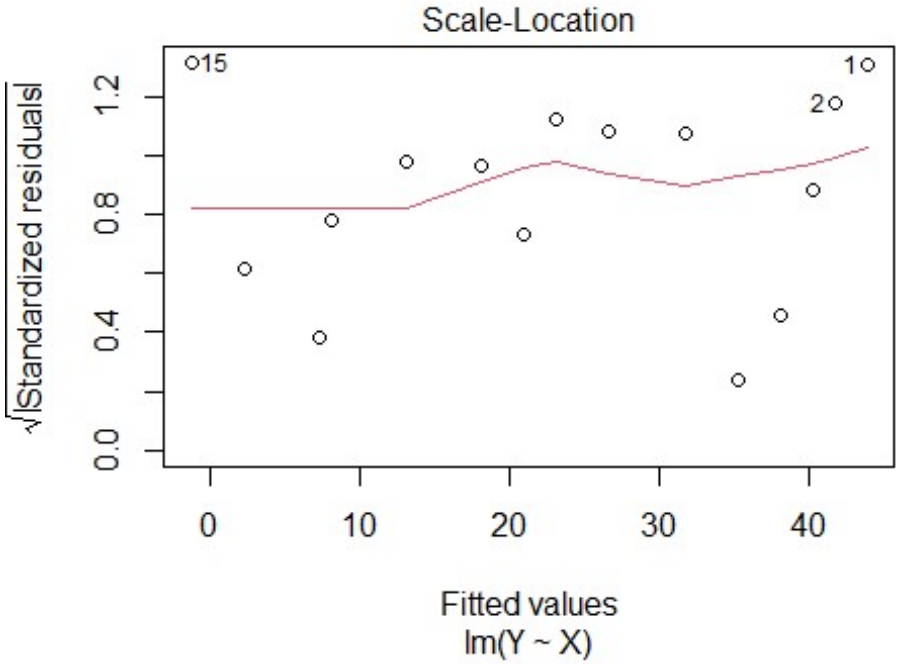
```
##           1           2           3           4           5           6
7
## 0.03414849 0.03489798 0.03541143 0.03620311 0.03730067 0.03874425
0.04091034
##           8           9          10          11          12          13
14
## 0.04257072 0.04361593 0.04507050 0.04779711 0.05077885 0.05122749
0.05454132
##          15
## 0.05710924

plot(a)
```

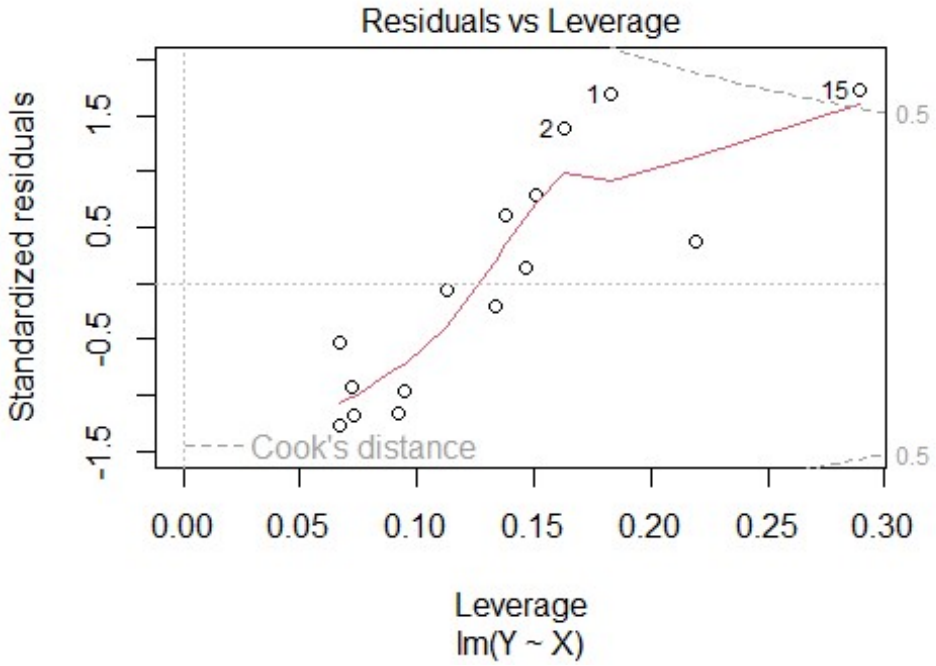



```
anreg2 <- lm(Y~X, data = data, weights=a)
plot(anreg2)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Y ~ X)
15
1
2
10
5
0
-5
0
10
20
30
40


Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(Y ~ X)
15
1
2
2.0
1.0
0.0
-1.0
-1
0
1






```
summary(anreg2)

##
## Call:
```

```
## lm(formula = Y ~ X, data = data, weights = a)
##
## 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
```

Berdasarkan analisis WLS diatas, dapat diketahui bahwa WLS belum efektif karena asumsi Gauss-Markov belum terpenuhi.

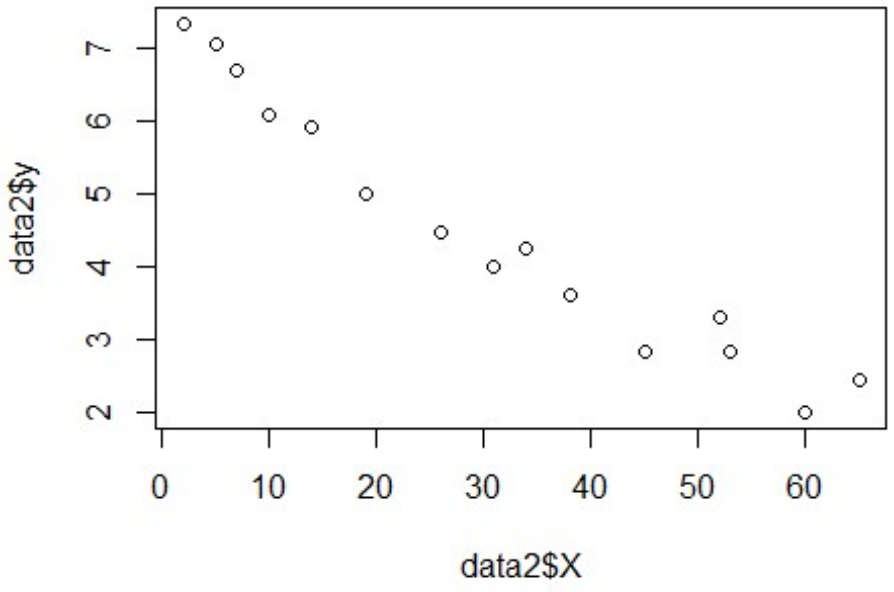
2) TRANSFOMASI AKAR

```
library(lmtest)
library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.2.3

## Warning: package 'ggplot2' was built under R version 4.2.3

## Warning: package 'tibble' was built under R version 4.2.3

## Warning: package 'tidyr' was built under R version 4.2.3

## Warning: package 'readr' was built under R version 4.2.3

## Warning: package 'purrr' was built under R version 4.2.3

## Warning: package 'dplyr' was built under R version 4.2.3

## Warning: package 'stringr' was built under R version 4.2.3

## Warning: package 'forcats' was built under R version 4.2.3

## Warning: package 'lubridate' was built under R version 4.2.3

## ── Attaching core tidyverse packages ──────────────────────── tidyverse
2.0.0 ──
## ✓ dplyr     1.1.3     ✓ readr     2.1.5
## ✓ 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 ────────────────────────────────────────
```

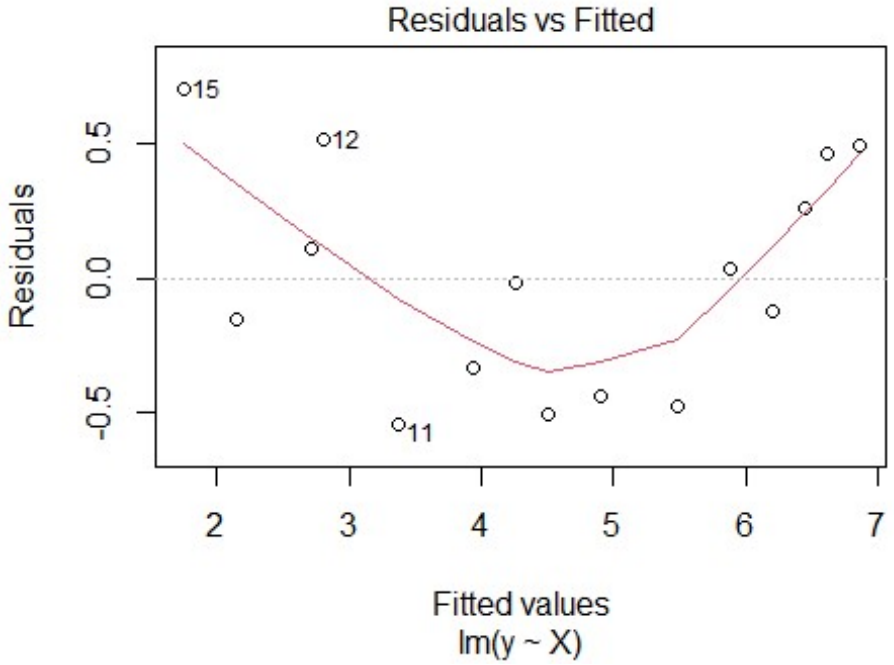
```
tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors

library(stats)
library(ggridges)

## Warning: package 'ggridges' was built under R version 4.2.3

library(dplyr)
library(GGally)

## Warning: package 'GGally' was built under R version 4.2.3

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

library(plotly)

## Warning: package 'plotly' was built under R version 4.2.3

##
## 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

data2 <- data %>% mutate(y = sqrt(Y)) %>% mutate(x= sqrt(X))
anreg3 <- lm(y~X, data=data2)
summary(anreg3)

##
## Call:
## lm(formula = y ~ X, data = data2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53998 -0.38316 -0.01727  0.36045  0.70199
##
## Coefficients:
```

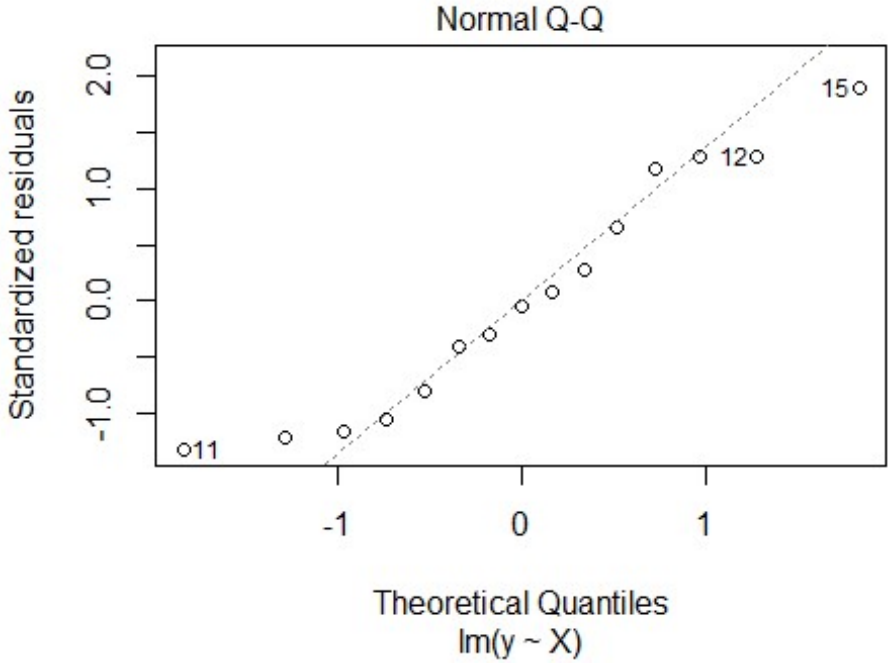
```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.015455   0.201677   34.79 3.24e-14 ***
## X           -0.081045   0.005477  -14.80 1.63e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4301 on 13 degrees of freedom
## Multiple R-squared:  0.9439, Adjusted R-squared:  0.9396
## F-statistic: 218.9 on 1 and 13 DF,  p-value: 1.634e-09
```

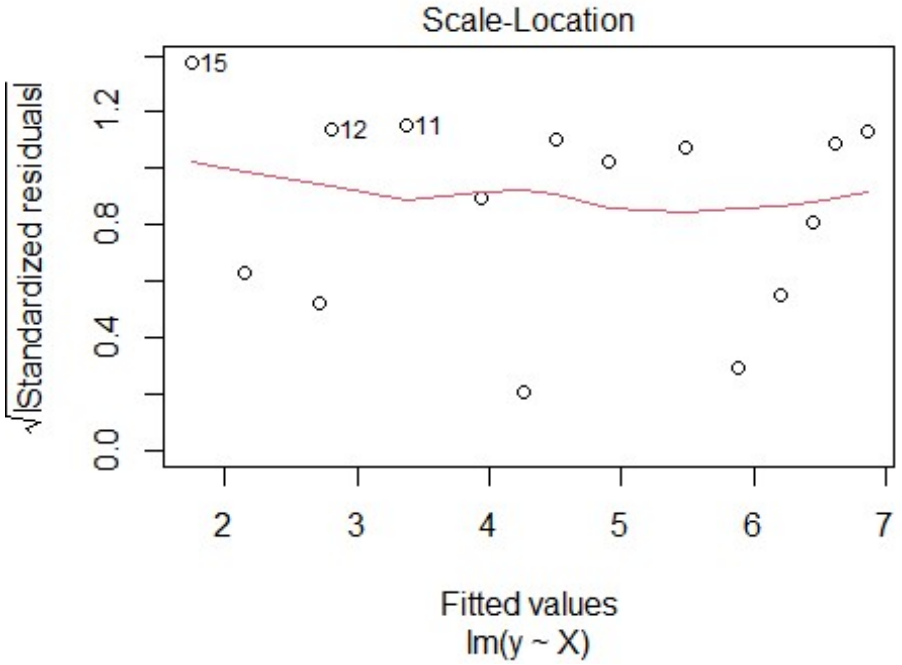
```
plot(x=data2$X, y=data2$y)
```

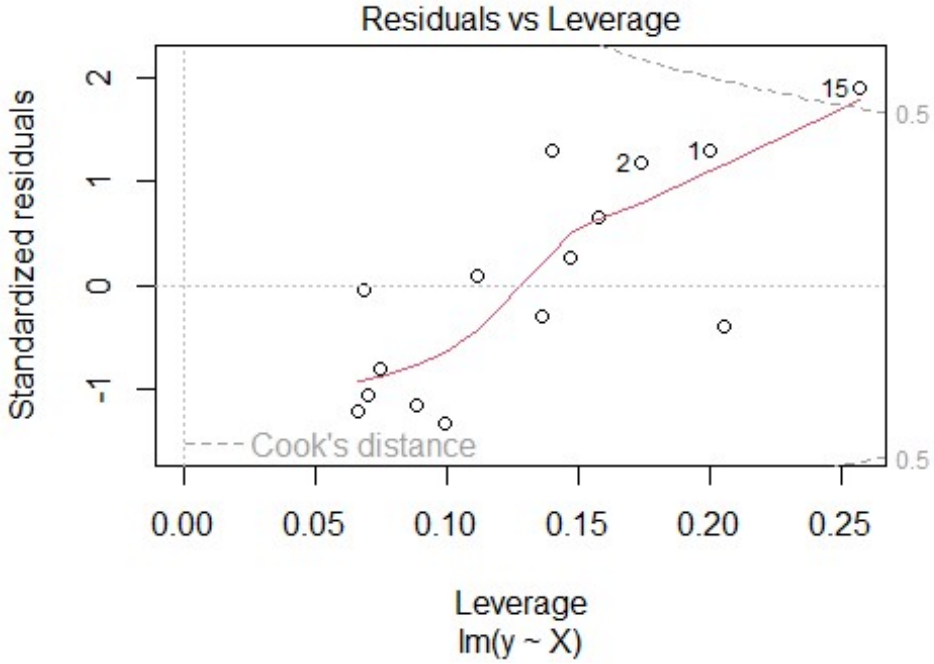



```
plot(anreg3)
```


Residuals vs Fitted
Residuals
Fitted values
lm(y ~ X)
15
12
11
0.5
0.0
-0.5
2
3
4
5
6
7


Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(y ~ X)
15
12
11
2.0
1.0
0.0
-1.0
-1
0
1



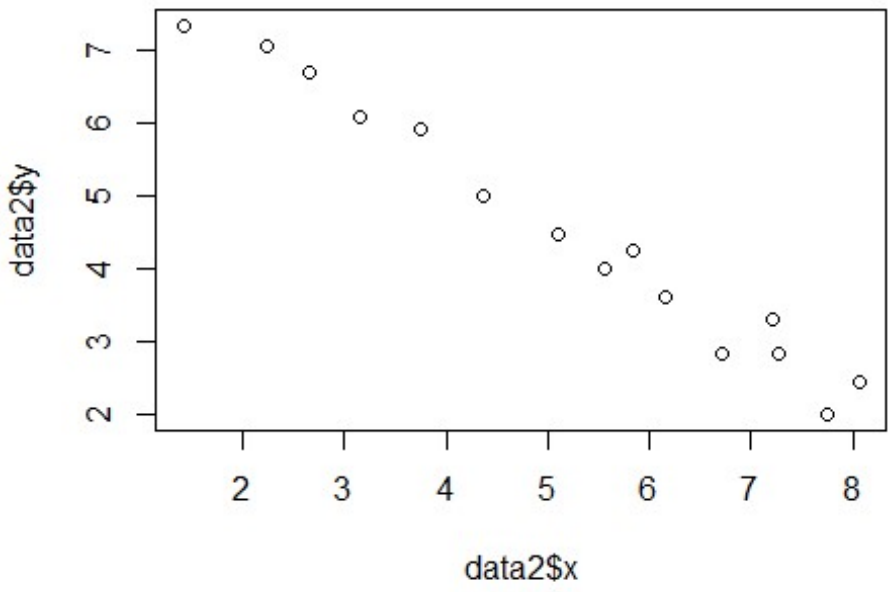


UJI AUTOKORELASI MODEL REGRESI TRANSFORMASI

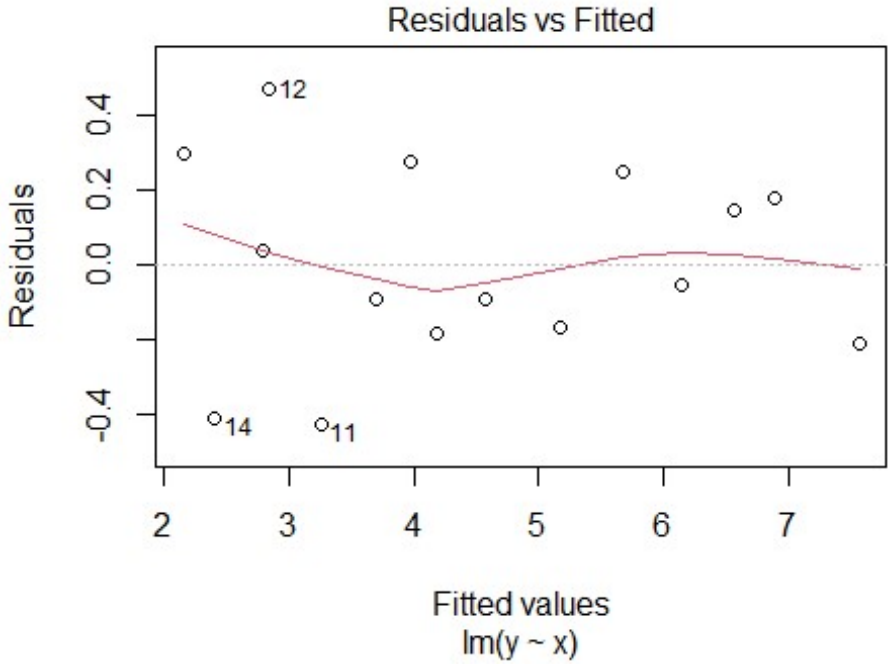
```
dwtest(anreg3)
```

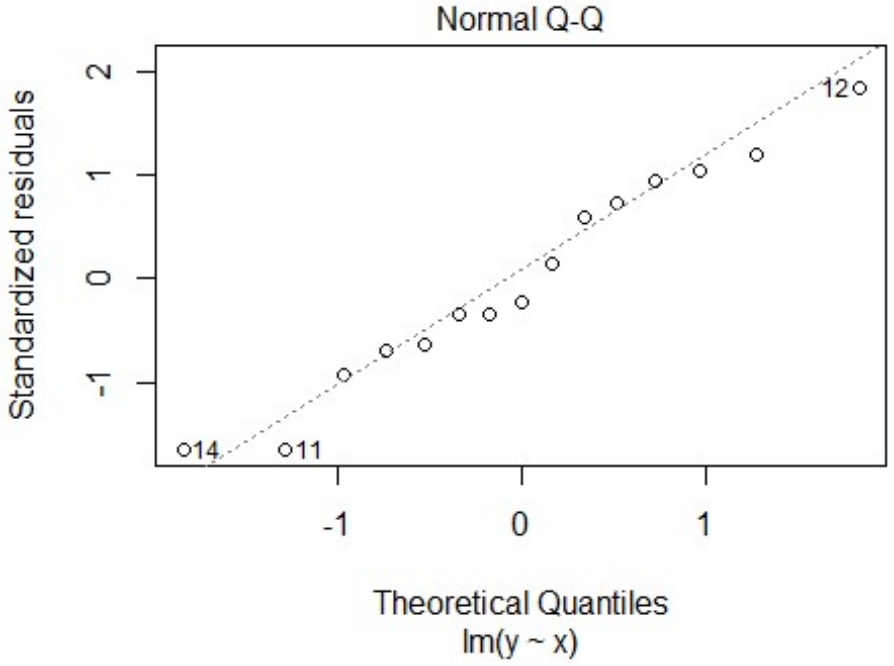
```
##
##  Durbin-Watson test
##
## data:  anreg3
## DW = 1.2206, p-value = 0.02493
## alternative hypothesis: true autocorrelation is greater than 0
```

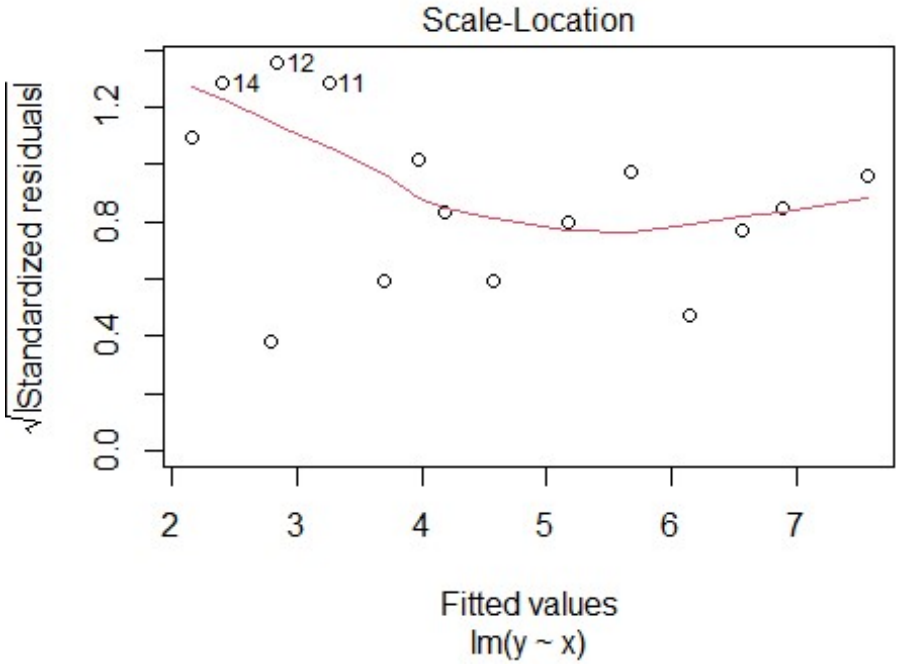
```
anreg3 <- lm(y~x, data=data2)
plot(x=data2$x, y=data2$y)
```

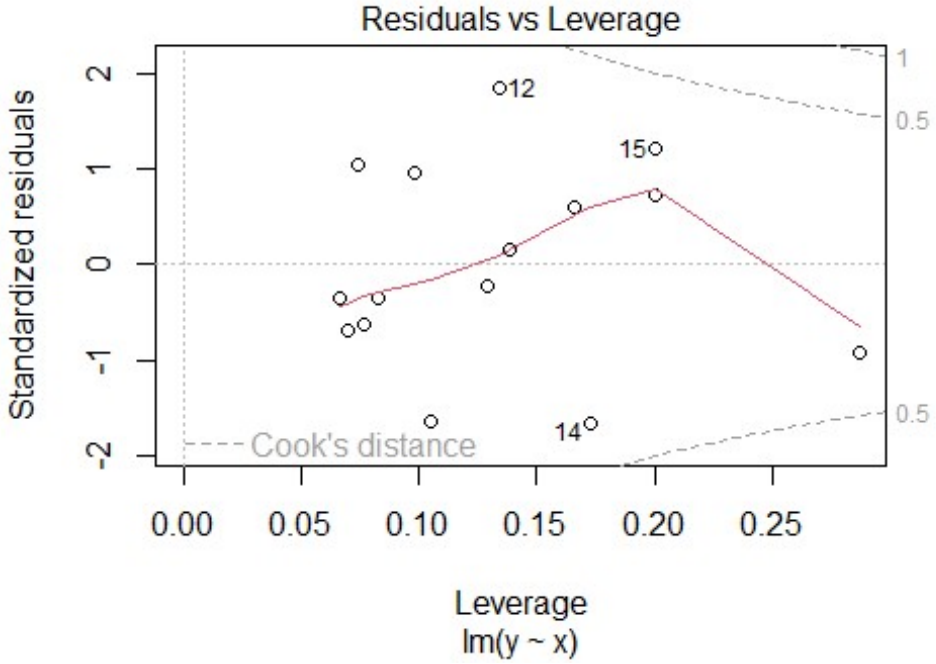



```
plot(anreg3)
```


Residuals vs Fitted
Residuals
Fitted values
lm(y ~ x)
12
14
11
-0.4
0.0
0.2
0.4
2
3
4
5
6
7


Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(y ~ x)
12
14
11
-1
0
1
2






```
summary(anreg3)

##
## Call:
```

```
## lm(formula = y ~ x, data = data2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42765 -0.17534 -0.05753  0.21223  0.46960
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.71245    0.19101   45.61 9.83e-16 ***
## x           -0.81339    0.03445  -23.61 4.64e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2743 on 13 degrees of freedom
## Multiple R-squared:  0.9772, Adjusted R-squared:  0.9755
## F-statistic: 557.3 on 1 and 13 DF,  p-value: 4.643e-12
```

KESIMPULAN

Berdasarkan transformasi akar y, dapat dihasilkan P-Value > 0.05. Hal ini menyatakan bahwa tidak cukup bukti untuk menolak H0 yang menyatakan tidak ada autokorelasi. Transformasi akar Y membuat persamaan regresi menjadi lebih efektif. Model regresi setelah transformasi dapat dinyatakan seperti berikut

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

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

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

$$Y = (8.71245 - 0.81339X^2)2 + e$$

INTERPRETASI

Dapat dilihat bahwa Y berkorelasi terbalik dengan akar kuadrat dari X. Hubungan yang dihasilkan adalah kuadratik. Nilai akar kuadrat dari X yang semakin besar akan membuat rata-rata nilai Y semakin kecil diikuti dengan tingkat penurunan yang semakin meningkat. Ketika nilai X = 0, niai Y akan berada pada nilai 8.71245. Nilai negatif pada -0.81339X^2 menunjukkan adanya hubungan terbalik antara Y dan akar kuadrat X. Semakin besar akar kuadrat X, akan membuat nilai Y semakin kecil. Perubahan Y yang semakin tinggi tidak akan proporsional dengan perubahan X.