

## Analisis Regresi: Tugas Kuliah 7

### Packages

```
library(randtests)
library(nortest)
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

library(readxl)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(plotly)

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

## Loading required package: ggplot2

##
## 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(lmtest)  
library(car)  
  
## Loading required package: carData  
  
##  
## Attaching package: 'car'  
  
## The following object is masked from 'package:dplyr':  
##  
##      recode
```

## Data

```
dataku <- read_excel("C:/Users/Asus/Downloads/Anreg Individu.xlsx")  
dataku
```

```
## # A tibble: 15 × 2  
##       x     y  
##   <dbl> <dbl>  
## 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
```

## Model

```
model <- lm(y~., dataku)  
model  
  
##  
## Call:  
## lm(formula = y ~ ., data = dataku)  
##  
## Coefficients:  
## (Intercept)                x  
##    46.4604         -0.7525
```

```
summary(model)
```

```
##  
## Call:  
## lm(formula = y ~ ., data = dataku)  
##  
## 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
```

Model:

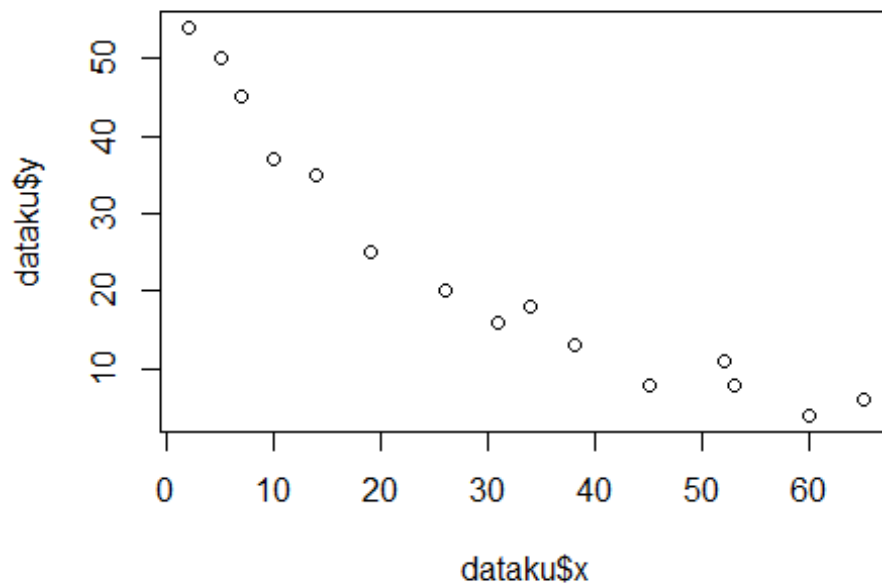
$$\hat{Y} = 46.4604 - 0.7525X + e$$

Hasil tersebut belum pasti menjadi model terbaik karena belum diuji secara menyeluruh. Perlu dilakukan pengeksplorasian dan pengujian asumsi Gaus Markov dan normalitas, untuk menciptakan model terbaik

## Exploration of Condition

### Exploration of Plot X and Y

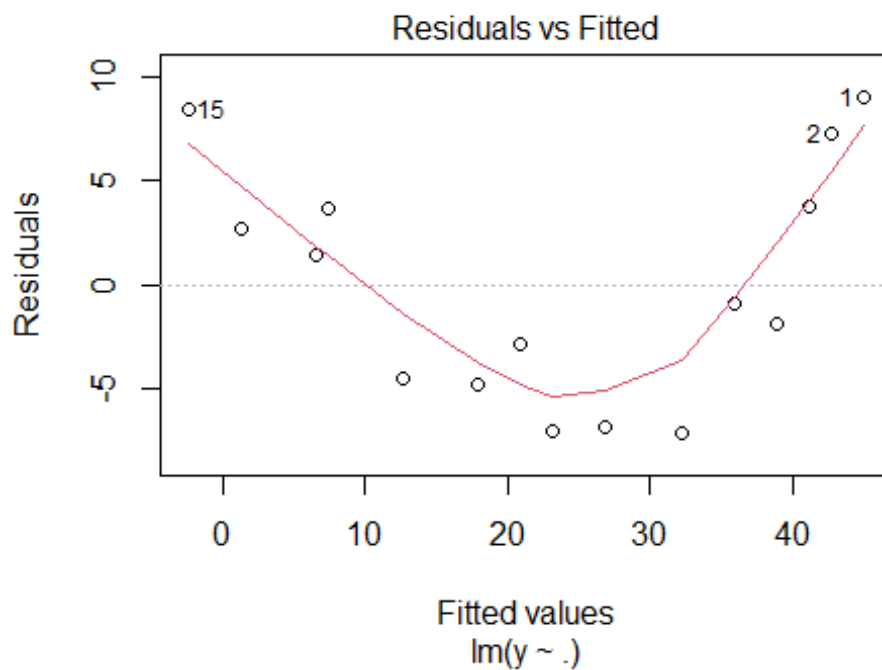
```
plot(x = dataku$x, y = dataku$y)
```



Dari hasil eksplorasi di atas, dapat dilihat bahwa X dan Y tidak berhubungan linier.

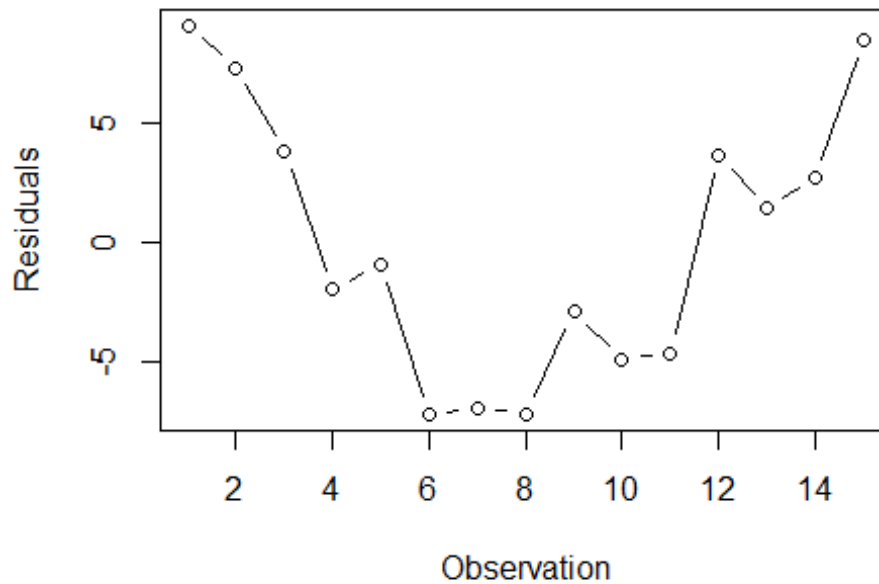
#### Exploration of Residual Plot and Y

`plot(model, 1)`



### Exploration of Residual Plot and Sequences

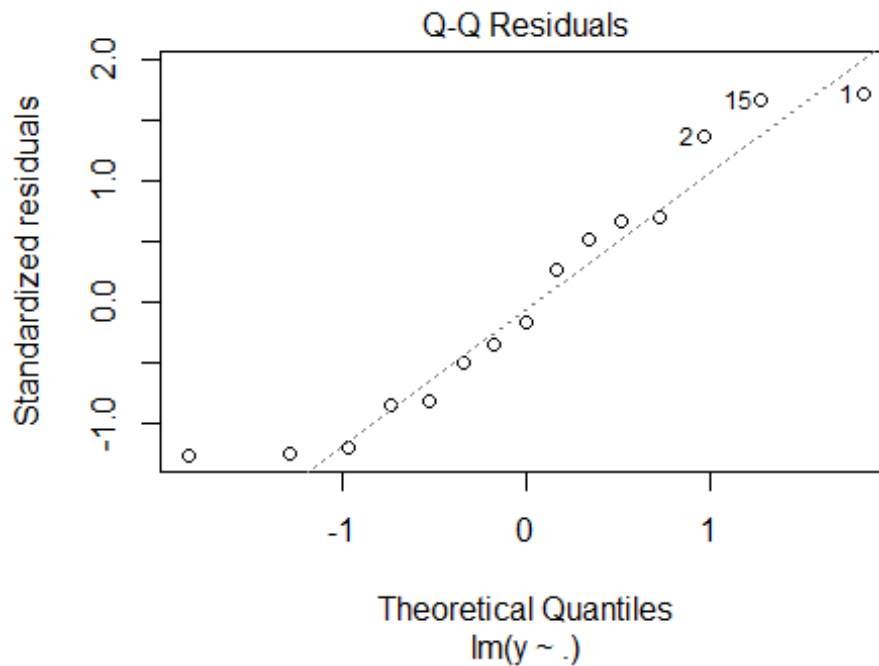
```
plot(x = 1:dim(dataku)[1],  
     y = model$residuals,  
     type = 'b',  
     ylab = "Residuals",  
     xlab = "Observation")
```



Dari Plot di atas dapat dilihat bahwa tebaran membentuk sebuah pola, maka model tidak pas dan sisaan tidak saling bebas

### Exploration of Residuals Normality and QQ-Plot

```
plot(model, 2)
```



```
shapiro.test(model$residual)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  model$residual
## W = 0.92457, p-value = 0.226
```

## Uji Formal

### Uji Autokorelasi

```
dwtest(model)
```

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

### Uji Homoskedastisitas

```
bptest(model)
```

```
##
##  studentized Breusch-Pagan test
##
```

```
## data: model
## BP = 0.52819, df = 1, p-value = 0.4674

ncvTest(model)

## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 0.1962841, Df = 1, p = 0.65774
```

## Transformasi

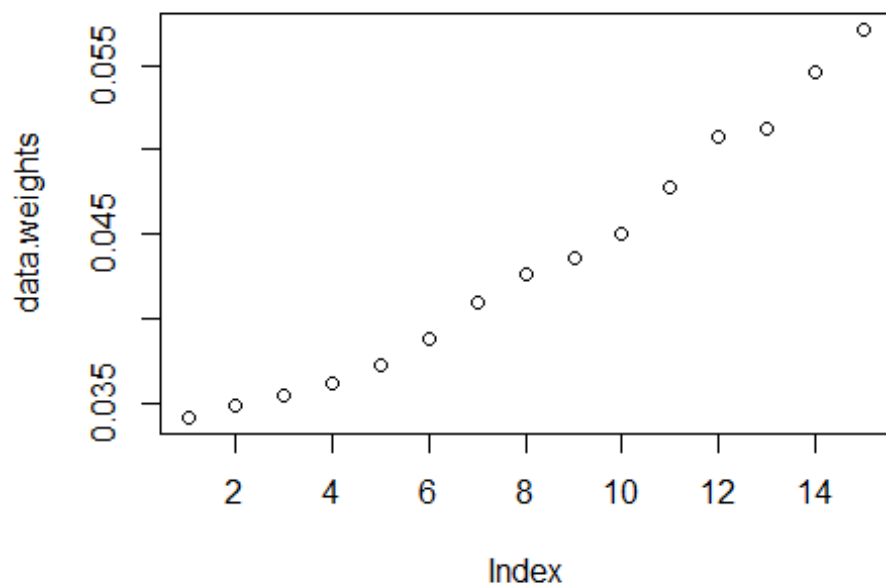
### WLS (Weighted Least Squares)

Step WLS hanya untuk perbandingan, karena ragam yang tersisa sudah seragam, pembobotan tidak diperlukan, cukup dengan mentransformasi data langsung.

```
resid_abs <- abs(model$residuals)
fitted_val <- model$fitted.values
fit <- lm(resid_abs ~ fitted_val, dataku)
data.weights <- 1 / fit$fitted.values^2
data.weights

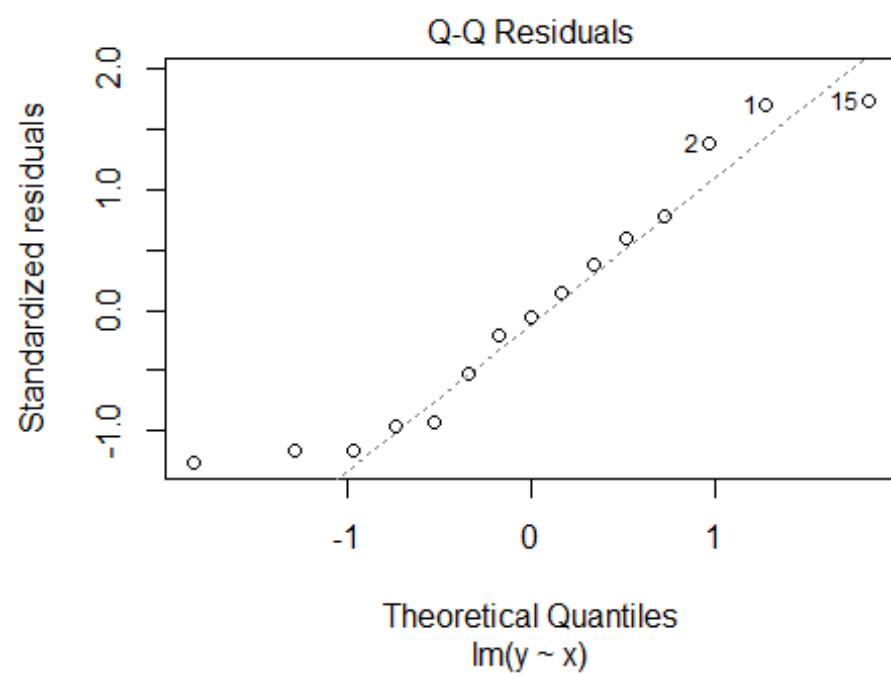
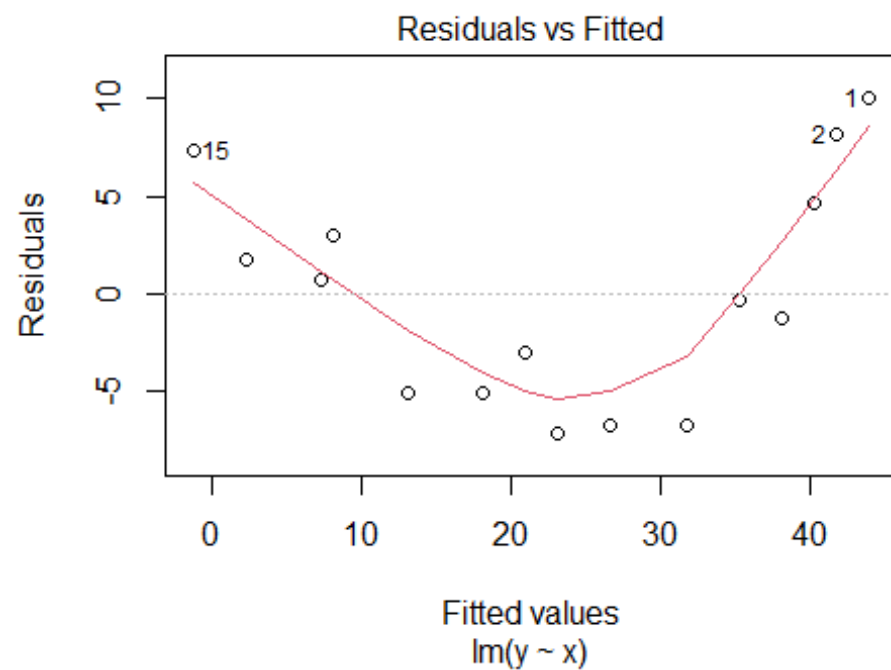
##           1           2           3           4           5           6
7
## 0.03414849 0.03489798 0.03541143 0.03620311 0.03730067 0.03874425
0.04091034
##           8           9          10          11          12          13
14
## 0.04257072 0.04361593 0.04507050 0.04779711 0.05077885 0.05122749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

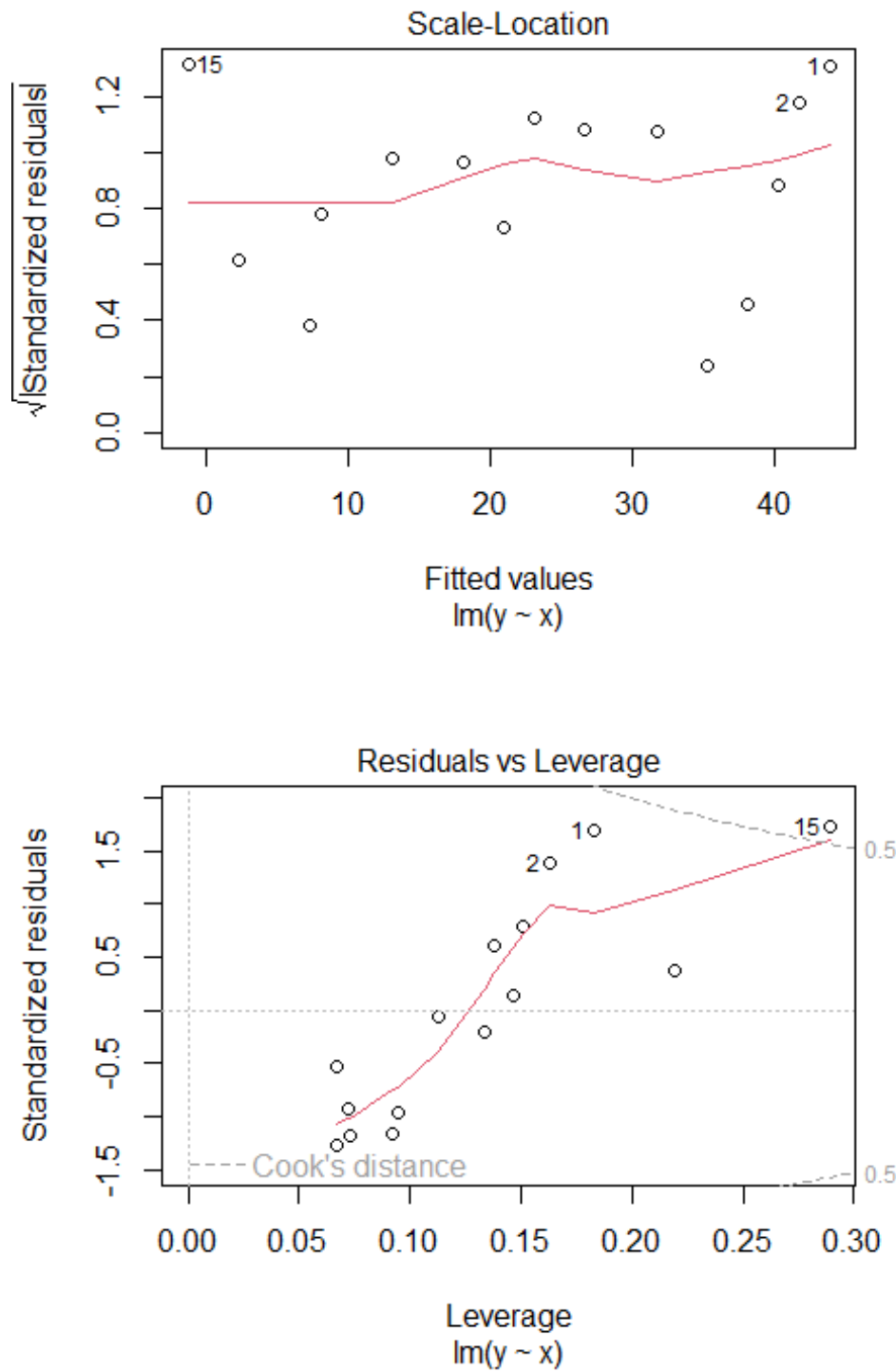
plot(data.weights)
```



```
model.weighted <- lm(y~x, data = dataku, weights = data.weights)
plot(model.weighted)
```







Dari perhitungan di atas, didapatkan model sebagai berikut:

```
model.lmw <- lm(y~x,
data = dataku,
```

```

weights = data.weights)
summary(model.lmw)

##
## Call:
## lm(formula = y ~ x, data = dataku, weights = data.weights)
##
## 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

```

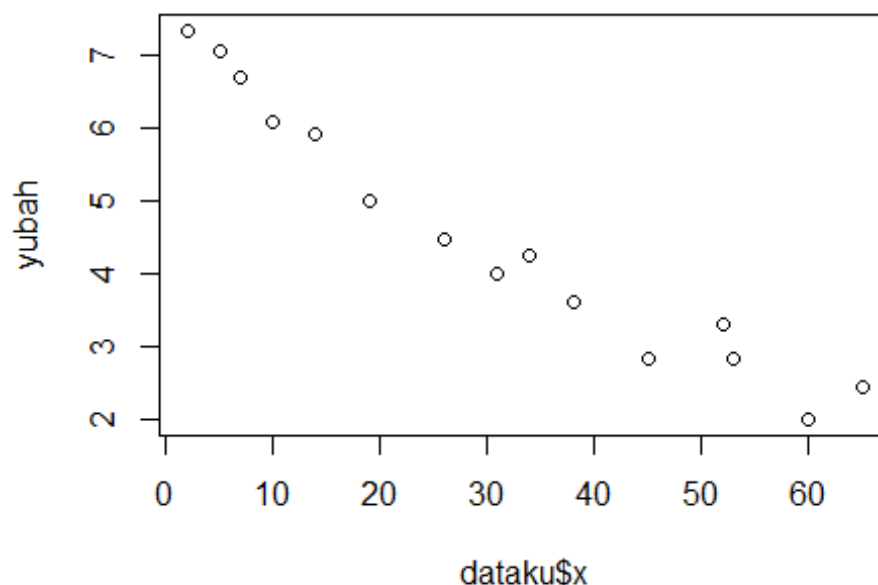
### Transformasi Data

```

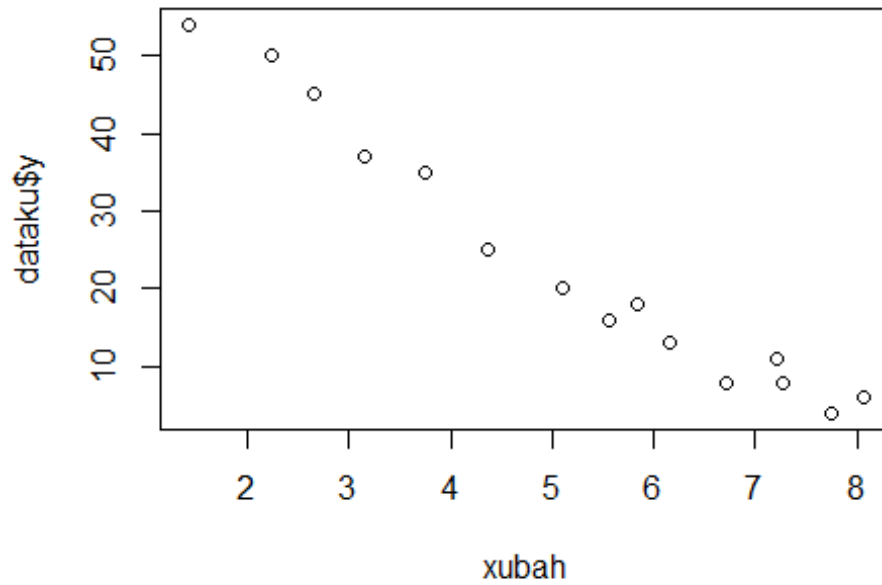
yubah = sqrt(dataku$y)
xubah = sqrt(dataku$x)

plot(x = dataku$x, y = yubah)

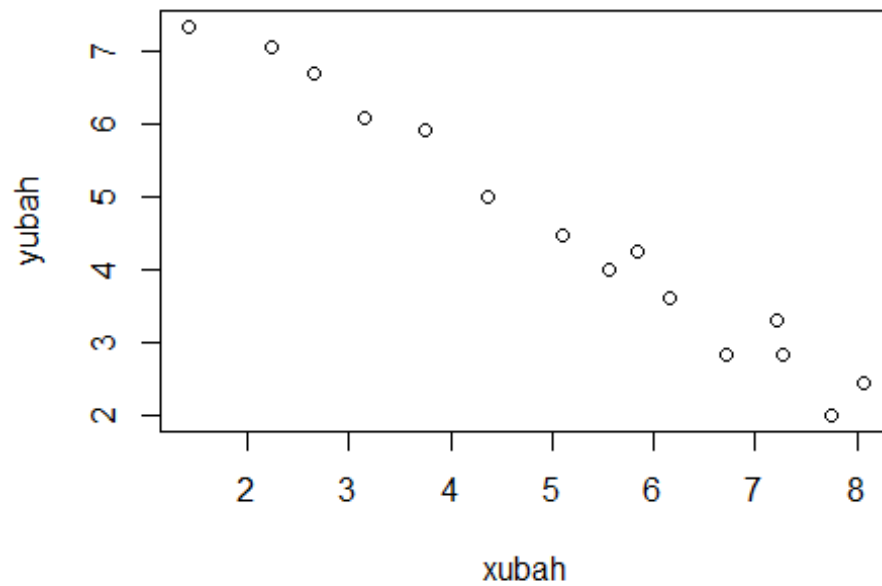
```



```
plot(x = xubah, y = dataku$y)
```



```
plot(x = xubah, y = yubah)
```



```
data.sqrt <- data.frame(xubah, yubah)
```

Karena hubungan X dan Y cenderung membentuk sebuah parabola dan nilai  $B1 < 0$ , maka data dapat ditransformasi dengan mengecilkan nilai X dan/atau Y dengan membentuknya menjadi pangkat setengah atau akar dari data asli. Terdapat perbedaan antara hasil plot hubungan Xubah dengan Y, X dengan Yubah, dan Xubah dengan Yubah sehingga perlu ditelusuri lebih lanjut untuk memperoleh model terbaik melalui pemeriksaan asumsi pada data dengan sisaan paling bebas

## Model and Assumption Checking

### Checking xubah with y

```
modellagi = lm(formula = dataku$y ~ xubah)
summary(modellagi)
```

```
##
## Call:
## lm(formula = dataku$y ~ xubah)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.4518 -2.8559  0.7657  2.0035  5.2422
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   63.2250     2.2712   27.84 5.67e-13 ***
## xubah         -7.7481     0.4097  -18.91 7.68e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.262 on 13 degrees of freedom
## Multiple R-squared:  0.9649, Adjusted R-squared:  0.9622
## F-statistic: 357.7 on 1 and 13 DF, p-value: 7.684e-11
```

Model:

$$\hat{Y} = 63.2250 - 7.7481X + e$$

```
dwtest(modellagi)
```

```
##
## Durbin-Watson test
##
## data:  modellagi
## DW = 1.1236, p-value = 0.01422
## alternative hypothesis: true autocorrelation is greater than 0
```

Interpretasi: dikarenakan nilai  $p\text{-value} = 0.01422 < \alpha = 0.05$ , maka tolak  $H_0$ . Hal ini menandakan bahwa sisaan tidak saling bebas, yang berarti asumsi tidak terpenuhi pada tingkat signifikansi 5%, sehingga model tersebut bukanlah model terbaik.

### Checking x with yubah

```
modellagi2 = lm(formula = yubah ~ dataku$x)
summary(modellagi2)
```

```
##
## Call:
## lm(formula = yubah ~ dataku$x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53998 -0.38316 -0.01727  0.36045  0.70199
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.015455   0.201677   34.79 3.24e-14 ***
## dataku$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
```

Model:

$$\hat{Y} = 7.015455 - 0.081045X + e$$

```
dwtest(modellagi2)
```

```
##
## Durbin-Watson test
##
## data:  modellagi2
## DW = 1.2206, p-value = 0.02493
## alternative hypothesis: true autocorrelation is greater than 0
```

Interpretasi: dikarenakan nilai p-value = 0.02493 dari uji Durbin-Watson (DW) < alpha = 0.05, maka tolak H0. Hal ini menunjukkan bahwa sisaan tidak saling bebas, sehingga asumsi tidak terpenuhi pada tingkat signifikansi 5%, sehingga model tersebut bukanlah model terbaik.

### Checking xubah with yubah

```
modellagi3 = lm(formula = yubah ~ xubah)
summary(modellagi3)
```

```
##
## Call:
## lm(formula = yubah ~ xubah)
##
## 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 ***
## xubah       -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

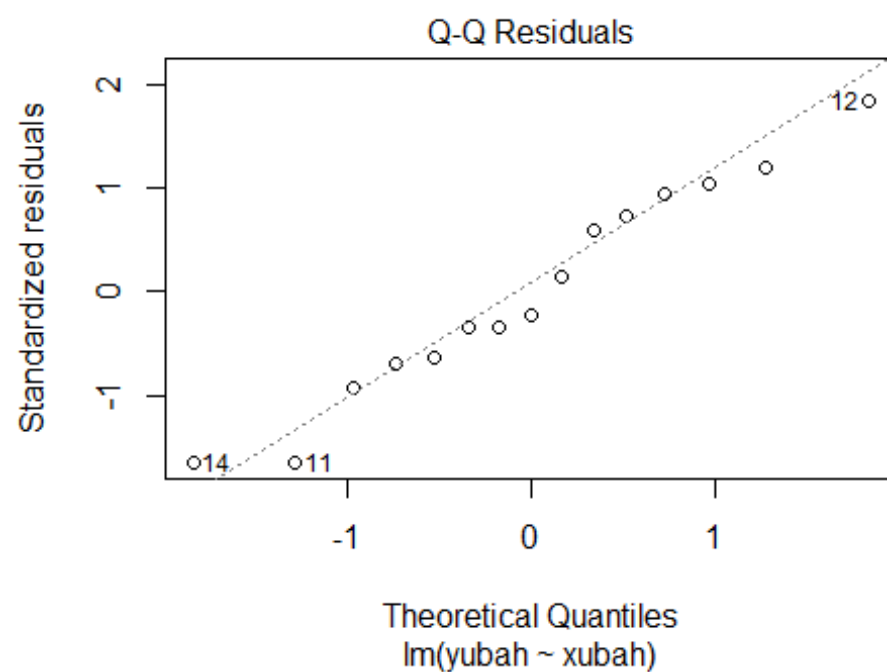
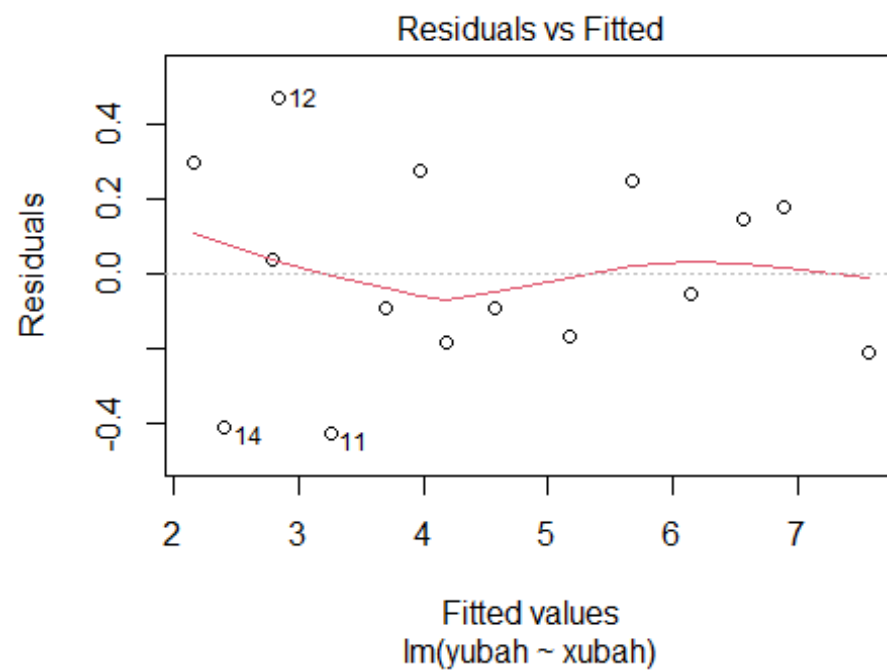
dwtest(modellagi3)

##
## Durbin-Watson test
##
## data:  modellagi3
## DW = 2.6803, p-value = 0.8629
## alternative hypothesis: true autocorrelation is greater than 0
```

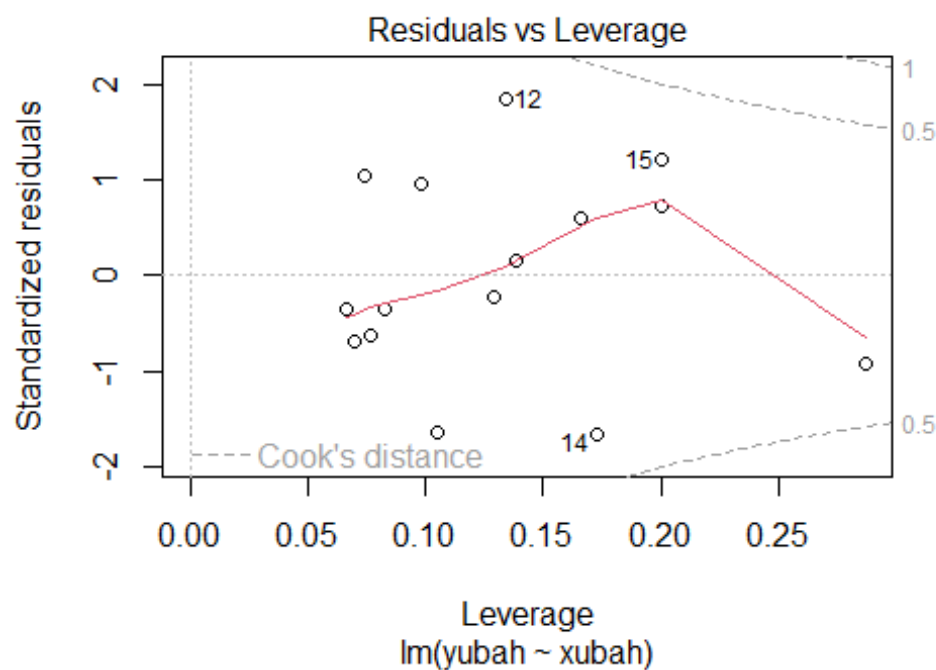
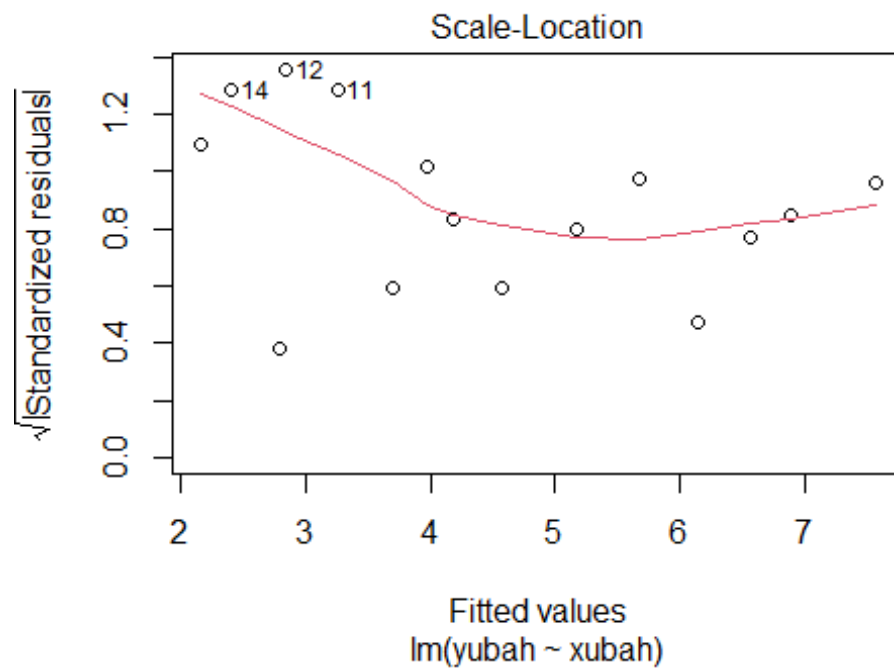
Interpretasi: dikarenakan nilai p-value = 0.8629 dari uji Durbin-Watson (DW) > alpha = 0.05, maka tak tolak H0. Ini menandakan bahwa sisaan adalah saling bebas. Meskipun uji autokorelasi menunjukkan bahwa sisaan saling bebas, perlu untuk memeriksa kembali dengan menggunakan uji asumsi lainnya untuk memastikan bahwa model yang dipilih adalah yang terbaik.

Karena dari 3 hasil perhitungan di atas didapat salah satu model yang tak tolak H0 yaitu pada modellagi3, maka model tersebut divisualisasikan pada plot.

```
plot(modellagi3)
```







T-test

```
t.test(modellagi3$residuals, mu = 0, conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data: modellagi3$residuals
## t = 2.0334e-16, df = 14, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -0.1463783 0.1463783
## sample estimates:
## mean of x
## 1.387779e-17
```

### DW-Test

```
ncvTest(modellagi3)
```

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 2.160411, Df = 1, p = 0.14161
```

### Colmogrov Test

```
residual.modellagi3 <- resid(modellagi3)
(norm.modellagi3 <- lillie.test(residual.modellagi3))

##
## Lilliefors (Kolmogorov-Smirnov) normality test
##
## data: residual.modellagi3
## D = 0.11948, p-value = 0.817
```

## Conclusion and Transformation

Model terbaik adalah ketika kedua variabel X dan Y ditransformasikan menjadi bentuk akar atau pangkat 1/2, dan memenuhi semua asumsi dalam analisis regresi linier sederhana. Oleh karena itu, model untuk data ini adalah:

$$\hat{Y}^{\frac{1}{2}} = 8.71245 - 0.81339X^{\frac{1}{2}}$$

Lalu terdapat transformasi balik pada model yang telah dibentuk untuk mengembalikan model untuk menjelaskan variabel respons sebelum dilakukan transformasi .

$$\hat{Y} = (8.71245 - 0.81339X^{\frac{1}{2}} + e)^2$$

Dari model tersebut dapat terlihat bahwa adanya hubungan terbalik antara Y dan  $X^{\frac{1}{2}}$  sebagai hubungan kuadratik. Semakin besar nilai  $X^{\frac{1}{2}}$ , semakin kecil nilai rata-rata dugaan Y. Ketika nilai  $X^{\frac{1}{2}}$  sama dengan nol dan berada dalam rentang amatan, nilai rata-rata dugaan Y adalah 8.71245. Setiap peningkatan satu satuan  $X^{\frac{1}{2}}$  akan mengurangi nilai rata-rata dugaan Y sebesar 0.81339