

# TUGAS ANREG KULIAH TRANSFORMASI REGRESI

Ria Yunita(G11401221115)

2024-03-05

```
library(readxl)

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

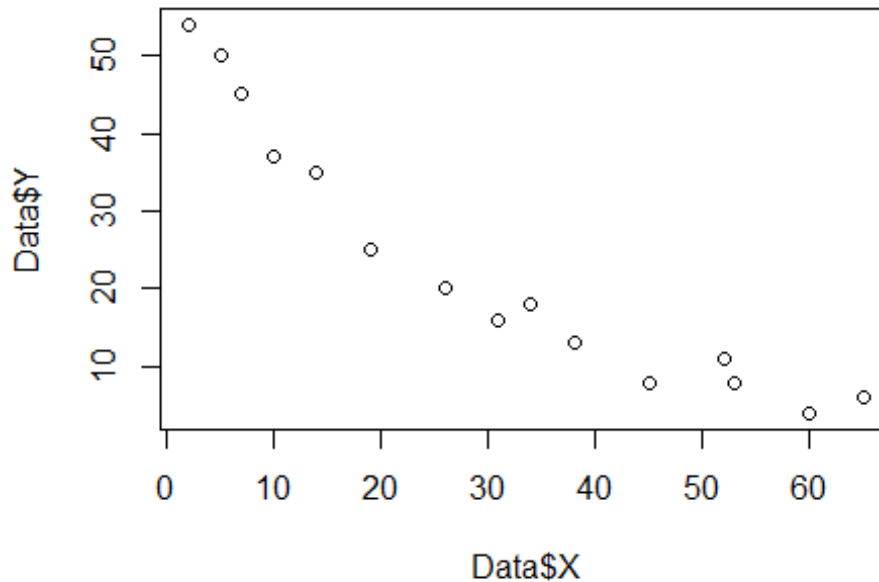
Data <- read_xlsx("D:/Semester 4/Analisis Regresi/Data_Tugas.xlsx")
Data

## # 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.reg= lm(formula = Y ~ X, data = Data)
summary(model.reg)

##
## 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
plot(x = Data$X,y = Data$Y)
```



## Uji Formal Kondisi Gauss-Markov

### 1 Nilai harapan sisaan sama dengan nol

H0: Nilai harapan sisaan sama dengan 0 H1: Nilai harapan sisaan tidak sama dengan 0

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

##
## One Sample t-test
##
## data:  model.reg$residuals
## t = -4.9493e-16, df = 14, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -3.143811  3.143811
## sample estimates:
##      mean of x
## -7.254614e-16
```

Karena p-value > 0,05 maka terima  $H_0$  yang artinya nilai harapan sisaan sama dengan 0.

## 2 Ragam sisaan homogen

$H_0$  : Homogen  $H_1$  : Tidak Homogen

```
cek.homogen = lm(formula = abs(model.reg$residuals) ~ X, # y: abs residual
  data = Data)
summary(cek.homogen)

##
## Call:
## lm(formula = abs(model.reg$residuals) ~ X, data = Data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.2525 -1.7525  0.0235  2.0168  4.2681
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.45041     1.27241   4.284  0.00089 ***
## X            -0.01948     0.03456  -0.564  0.58266
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.714 on 13 degrees of freedom
## Multiple R-squared:  0.02385,    Adjusted R-squared:  -0.05124
## F-statistic: 0.3176 on 1 and 13 DF,  p-value: 0.5827

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

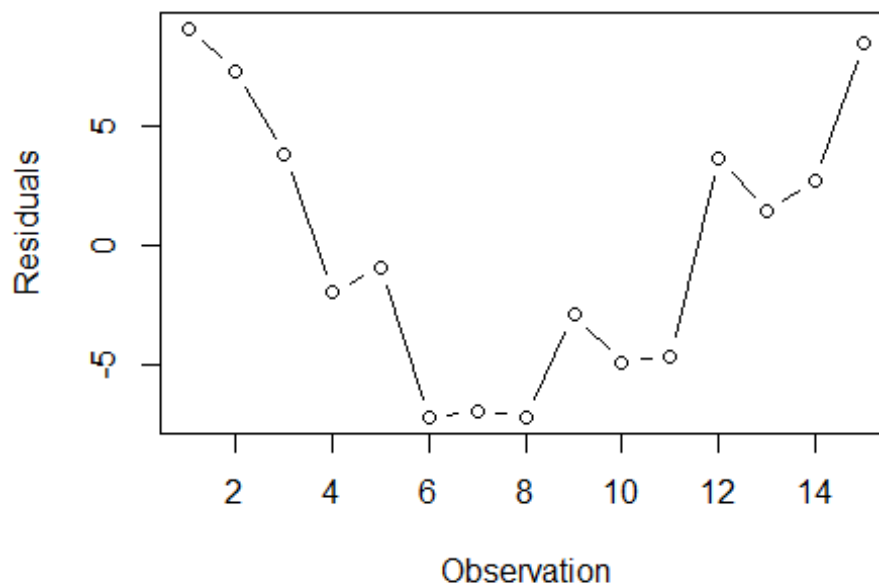
bptest(model.reg)

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

Berdasarkan hasil dua uji kehomogenan sisaan diatas, p-value > 0,05 artinya Sisaannya homogen.

### 3 Sisaan saling bebas

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



H0 :Saling bebas

H1 :Tidak saling bebas

```
library(randtests)  
runs.test(model.reg$residuals)
```

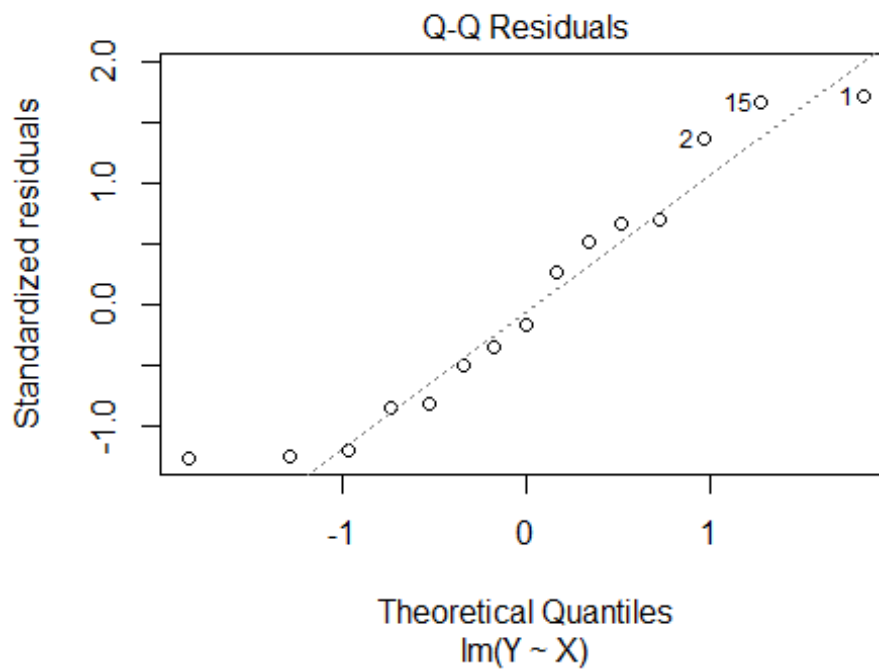
```
##  
##  Runs Test  
##  
## data:  model.reg$residuals  
## statistic = -2.7817, runs = 3, n1 = 7, n2 = 7, n = 14, p-value =  
## 0.005407  
## alternative hypothesis: nonrandomness
```

Berdasarkan uji randtest diatas, p-value < 0,05 (tolak h0) yang artinya Sisaan tidak saling bebas.

#### 4 Uji Normalitas Sisaan

H0: Sisaan Normal H1: Sisaan Tidak Normal

```
plot(model.reg,2)
```



```
ks.test(model.reg$residuals, "pnorm", mean=mean(model.reg$residuals),
sd=sd(model.reg$residuals))

##
##  Exact one-sample Kolmogorov-Smirnov test
##
## data:  model.reg$residuals
## D = 0.12432, p-value = 0.9521
## alternative hypothesis: two-sided

tseries::jarque.bera.test(model.reg$residuals)

## Registered S3 method overwritten by 'quantmod':
##   method      from
##   as.zoo.data.frame zoo

##
##  Jarque Bera Test
##
## data:  model.reg$residuals
## X-squared = 1.1142, df = 2, p-value = 0.5729
```

```
nortest::ad.test(model.reg$residuals)

##
## Anderson-Darling normality test
##
## data: model.reg$residuals
## A = 0.35232, p-value = 0.4178

nortest::lillie.test(model.reg$residuals)

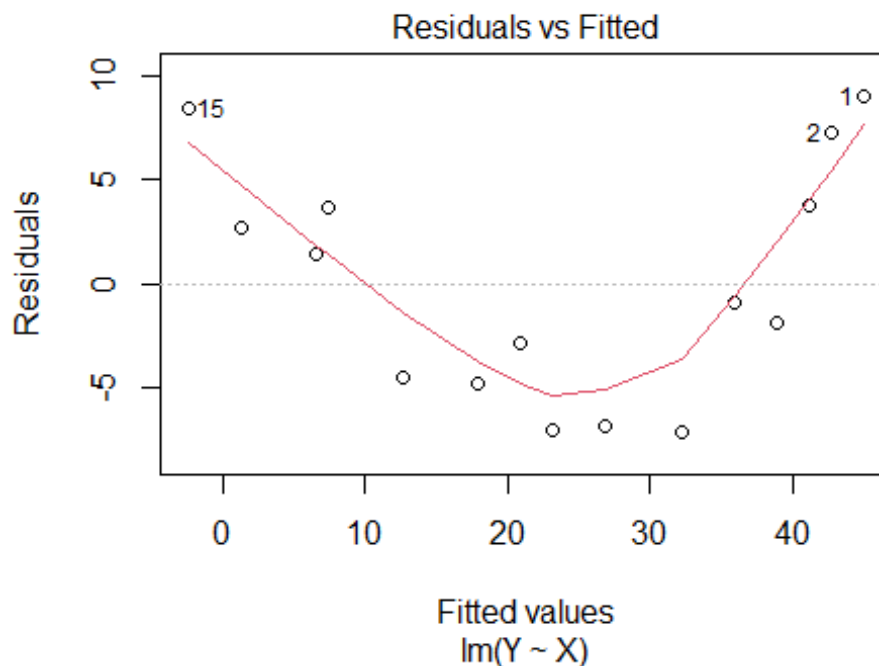
##
## Lilliefors (Kolmogorov-Smirnov) normality test
##
## data: model.reg$residuals
## D = 0.12432, p-value = 0.7701
```

Berdasarkan hasil uji2 normalitas diatas,  $p\text{-value} > 0,05$  yang artinya sisaan menyebar Normal.

#### *Kesimpulan Uji formal kondisi gaus markov*

Berdasarkan hasil uji2 diatas, semua kondisi gaus markov terpenuhi kecuali pada kondisi sisaannya yang tidak saling bebas.

```
plot(model.reg,1)
```



Berdasarkan grafik plot sisaan vs yduga tersebut membentuk kurva parabola yang berarti modelnya tidak pas sehingga diperlukan transformasi terhadap Y.

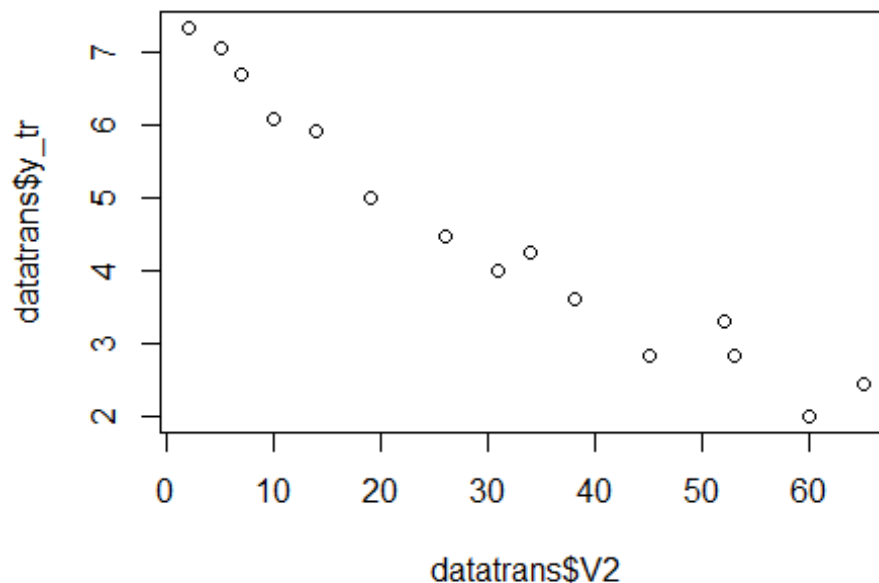
## Penanganan Masalah

*Transformasi terhadap Y dengan memperkecil y*

```
y_tr <- sqrt(Data$Y)
datatrans <- data.frame(cbind(y_tr, Data$X))
head(datatrans)
```

```
##      y_tr V2
## 1 7.348469  2
## 2 7.071068  5
## 3 6.708204  7
## 4 6.082763 10
## 5 5.916080 14
## 6 5.000000 19
```

```
plot(datatrans$V2, datatrans$y_tr)
```



## Diagnostik Sisaan Hasil Transformasi

Selanjutnya dilakukan pemodelan dan diagnostik sisaan kembali pada data baru dengan peubah respons yang sudah ditransformasi.

```
model_tr <- lm(y_tr ~ Data$X, datatrans)
summary(model_tr)
```

```
##
## Call:
```

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

#### *Autokorelasi: Durbin-Watson*

```
library(car)

## Warning: package 'car' was built under R version 4.3.3
## Loading required package: carData

(uji_autokol3 <- durbinWatsonTest(model_tr,
                                   alternative="two.sided"))

## lag Autocorrelation D-W Statistic p-value
## 1      0.2362986      1.220617    0.032
## Alternative hypothesis: rho != 0

ifelse(uji_autokol3$p < 0.05, "Ada Autokorelasi", "Tidak Ada Autokorelasi")

## [1] "Ada Autokorelasi"
```

#### *Homogenitas: Breusch-Pagan*

```
library(lmtest)
(homogen_model3 <- bptest(model_tr))

##
## studentized Breusch-Pagan test
##
## data: model_tr
## BP = 0.93605, df = 1, p-value = 0.3333

ifelse(homogen_model3$p.value < 0.05, "Ragam Tidak Homogen", "Ragam Homogen")

## BP
## "Ragam Homogen"
```



#### Normalitas: Kolmogorov-Smirnov

```
library(nortest)
sisaan_model_tr <- resid(model_tr)
(norm_model_tr<- lillie.test(sisaan_model_tr))

##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  sisaan_model_tr
## D = 0.13356, p-value = 0.6692

ifelse(norm_model_tr$p.value < 0.05, "Sisaan Tidak Menyebar Normal", "Sisaan
Menyebar Normal")

## [1] "Sisaan Menyebar Normal"
```

#### Sisaan saling bebas

```
library(randtests)
runs.test(as.numeric(model_tr$residuals))

##
##  Runs Test
##
## data:  as.numeric(model_tr$residuals)
## statistic = -0.55635, runs = 7, n1 = 7, n2 = 7, n = 14, p-value = 0.578
## alternative hypothesis: nonrandomness
```

Berdasarkan uji randtest diatas, p-value > 0,05 (tolak  $H_0$ ) yang artinya Sisaan saling bebas.

#### Kesimpulan transformasi

Dengan demikian, hasil tranformasi dengan memperkecil y dapat mengatasi permasalahan Sisaan yang tidak saling bebas sebelumnya. Sehingga setelah transformasi dengan memperkecil y semua kondisi asumsi Gauss-Markov telah terpenuhi.

```
model_tr <- lm(y_tr~Data$X, datatrans)
summary(model_tr)

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

### *Interpretasi hasil model Regresi Transformasi*

Berdasarkan hasil transformasi dengan memperkecil y, dapat dilihat bahwa adjusted R-Square nya sebesar 0,9396 atau sebesar 93,96% keragaman yang terdapat pada variabel y dapat dijelaskan oleh variabel penjelas x. Kemudian didapatkan juga Persamaan regresi hasil transformasinya dapat dituliskan sebagai berikut :

$$Y^*duga = 7.015455 - 0.81339X + \varepsilon$$

### **Transformasi Balik**

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

$$\sqrt{Y}duga = 7.015455 - 0.081045X + \varepsilon$$

$$Yduga = (7.015455 - 0.081045X)^2 + \varepsilon$$

### **Kesimpulan Interpretasi**

Interpretasi dari persamaan model transformasi balik diatas menggambarkan hubungan kuadratik. Kemudian ketika peubah respon X mengalami kenaikan maka Y akan semakin kecil karena koefisien pada peubah X yang bernilai negatif.