

Tuman Kuliah 7 Anreg

Salsabila Fayiza_G1401221036

2024-03-05

```
library(readxl)

## Warning: package 'readxl' was built under R version 4.3.3

dataxy <- read_excel(path = "D:/IPB/SEMESTER 4/Analisis Regresi/pertemuan 7/d
ata.xlsx")
dataxy

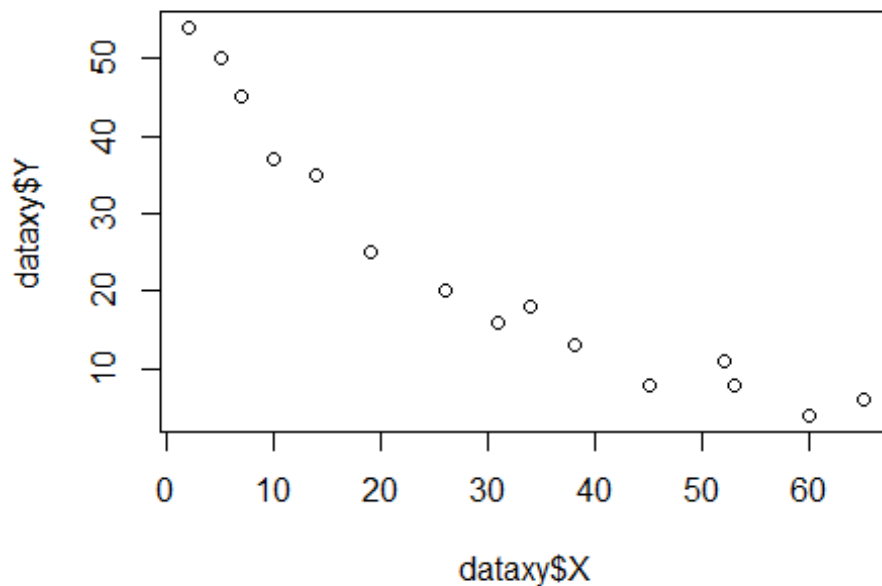
## # A tibble: 15 × 3
##       no      X      Y
##   <dbl> <dbl> <dbl>
## 1     1     2    54
## 2     2     5    50
## 3     3     7    45
## 4     4    10    37
## 5     5    14    35
## 6     6    19    25
## 7     7    26    20
## 8     8    31    16
## 9     9    34    18
## 10    10    38    13
## 11    11    45     8
## 12    12    52    11
## 13    13    53     8
## 14    14    60     4
## 15    15    65     6
```

Eksplorasi Data

```
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.3.2

plot(x = dataxy$X, y = dataxy$Y)
```



Scatter plot di atas menunjukkan X dan Y tidak berhubungan linear. Data menyebar secara eksponensial, tanpa ada data pencilan.

```
e <- 2.718281828
x <- dataxy$X
y <- dataxy$Y
alfa <- 56.665
beta <- -0.038
y_duga <- round(alfa*e^(beta*x), 3)
galat <- round(y - y_duga, 3)

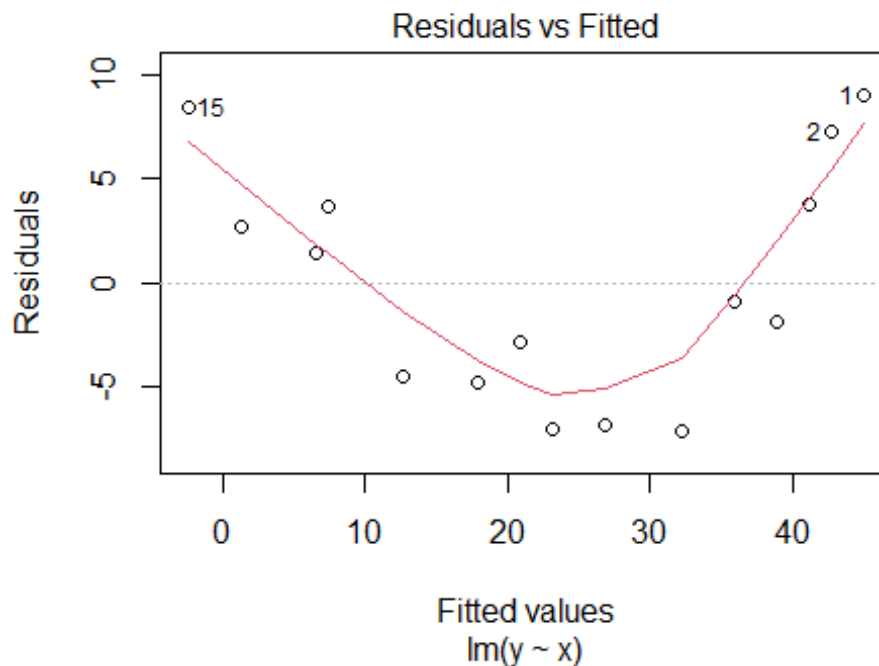
tabelnilai <- data.frame(x, y, y_duga, galat)
tabelnilai
```

##	x	y	y_duga	galat
## 1	2	54	52.518	1.482
## 2	5	50	46.860	3.140
## 3	7	45	43.430	1.570
## 4	10	37	38.751	-1.751
## 5	14	35	33.287	1.713
## 6	19	25	27.527	-2.527
## 7	26	20	21.098	-1.098
## 8	31	16	17.447	-1.447
## 9	34	18	15.567	2.433
## 10	38	13	13.372	-0.372
## 11	45	8	10.249	-2.249

```
## 12 52 11 7.855 3.145
## 13 53 8 7.562 0.438
## 14 60 4 5.796 -1.796
## 15 65 6 4.793 1.207
```

Uji Homoskedastisitas

```
model_lm <- lm(y~x)
plot(model_lm, which = 1)
```



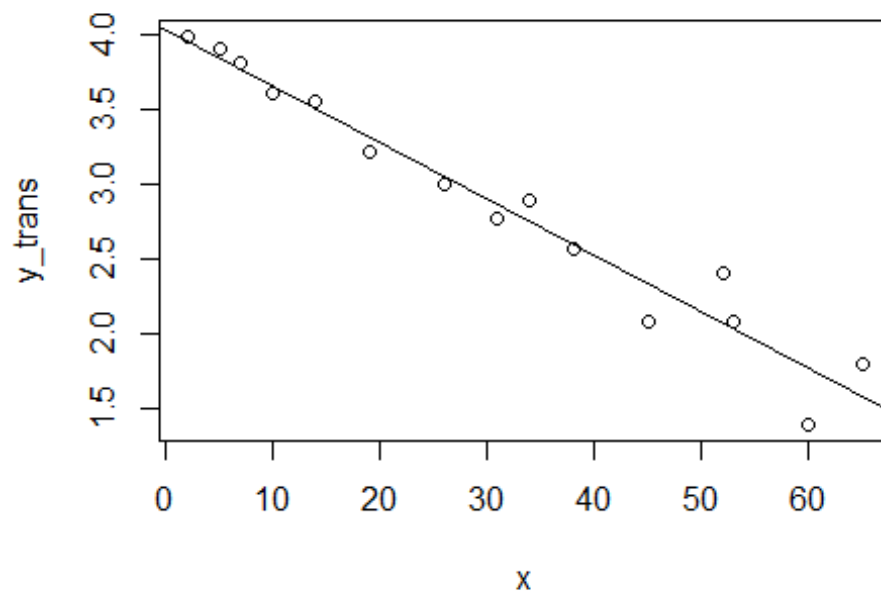
Dari Uji

Homoskedastisitas ini, dapat dilihat bahwa varians residual bernilai konstan. Varian residual ini cenderung meningkat seiring dengan nilai prediksi. Sehingga dapat disimpulkan adanya Homoskedastisitas.

Transformasi

Transformasi Akar pada x, y, atau X dan y

```
y_trans <- log(y)
plot(x, y_trans)
abline(lm(y_trans ~ x))
```



Setelah dilakukan transformasi terlihat data bergerak linear.

transformasi pada parameter

```
b0 <- log(alfa)
b1 <- beta
round(b0,3)

## [1] 4.037

round(b1,3)

## [1] -0.038

yduga_trans <- b0 + b1*x
galat_trans <- y_trans - yduga_trans

tabelnilai_trans <- data.frame(x, y_trans, yduga_trans, galat_trans)
tabelnilai_trans
```

	x	y_trans	yduga_trans	galat_trans
## 1	2	3.988984	3.961157	0.02782731
## 2	5	3.912023	3.847157	0.06486627
## 3	7	3.806662	3.771157	0.03550575
## 4	10	3.610918	3.657157	-0.04623882
## 5	14	3.555348	3.505157	0.05019133
## 6	19	3.218876	3.315157	-0.09628091

```
## 7 26 2.995732 3.049157 -0.05342446
## 8 31 2.772589 2.859157 -0.08656801
## 9 34 2.890372 2.745157 0.14521502
## 10 38 2.564949 2.593157 -0.02820738
## 11 45 2.079442 2.327157 -0.24771519
## 12 52 2.397895 2.061157 0.33673854
## 13 53 2.079442 2.023157 0.05628481
## 14 60 1.386294 1.757157 -0.37086238
## 15 65 1.791759 1.567157 0.22460273
```

Setelah perhitungan diatas, diperoleh persamaan regresi linear sederhana dari eksponensial setelah ditransformasi adalah : **$y \text{ duga} = 4.037 - 0.038X$**

```
model <- lm(y_trans ~ x)
summary(model)

##
## Call:
## lm(formula = y_trans ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.37241 -0.07073  0.02777  0.05982  0.33539
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.037159   0.084103   48.00 5.08e-16 ***
## x           -0.037974   0.002284  -16.62 3.86e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1794 on 13 degrees of freedom
## Multiple R-squared:  0.9551, Adjusted R-squared:  0.9516
## F-statistic: 276.4 on 1 and 13 DF, p-value: 3.858e-10
```

Dapat dilihat bahwa benar nilai intersep dan slope pada anova sama dengan nilai beta0 dan beta1 yang ditransformasi secara manual. Persamaan regresi ini cenderung merupakan model regresi terbaik jika dilihat melalui nilai R-Squared mendekati 1. Setelah melakukan transformasi seperti diatas, dapat disimpulkan jika transformasi akar Y membuat persamaan regresi menjadi lebih efektif. Model regresi setelah transformasi :

$$Y^* = 4.037159 - 0.0379745X$$

$$Y^* = \text{akar } Y$$

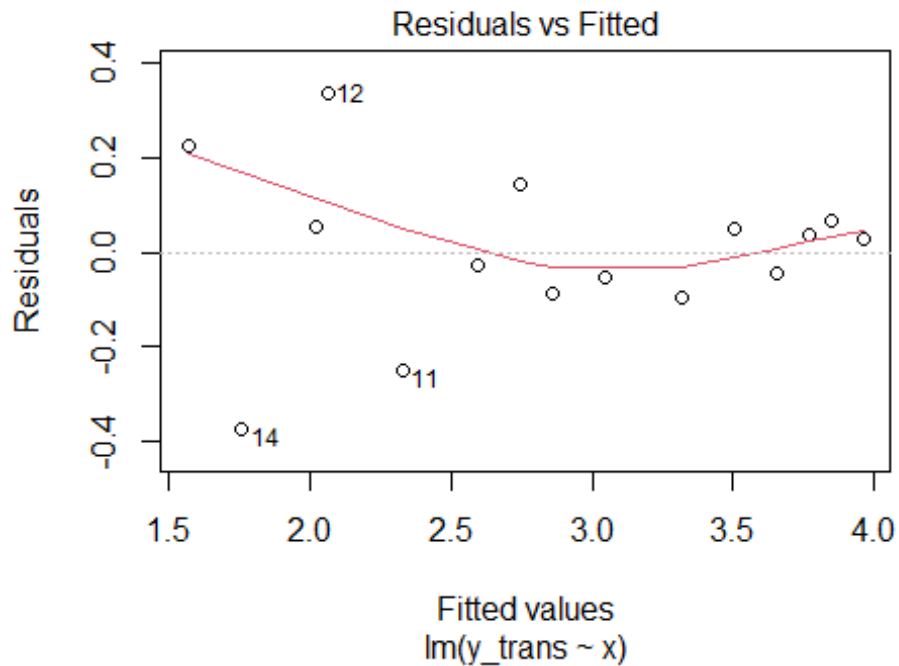
$$\text{jika dilakukan transformasi balik menjadi: } Y = (4.037159 - 0.037974X)^2$$

Interpretasi Persamaan regresi menunjukkan hubungan kuadrat negatif antara Y dan X. Ketika nilai X bertambah, nilai Y cenderung berkurang dengan kecepatan yang semakin cepat. Koefisien regresi dan perubahan Y untuk perubahan X membantu mengukur pengaruh X pada Y. Konstanta 4.037159 mewakili nilai Y ketika X sama dengan 0. Koefisien

-0.0379745 menunjukkan pengaruh perubahan X terhadap Y. Semakin besar nilai absolut koefisien, semakin besar pengaruh X terhadap Y.

Uji Asumsi

```
# plot persebaran nilai ragam  
plot(model, 1)
```



Dapat dilihat bahwa data memiliki lebar keragaman yang berbeda (ragam cenderung berbentuk corong) dan masih berada di sekitar 0. Ragam bersifat Heterodixcity

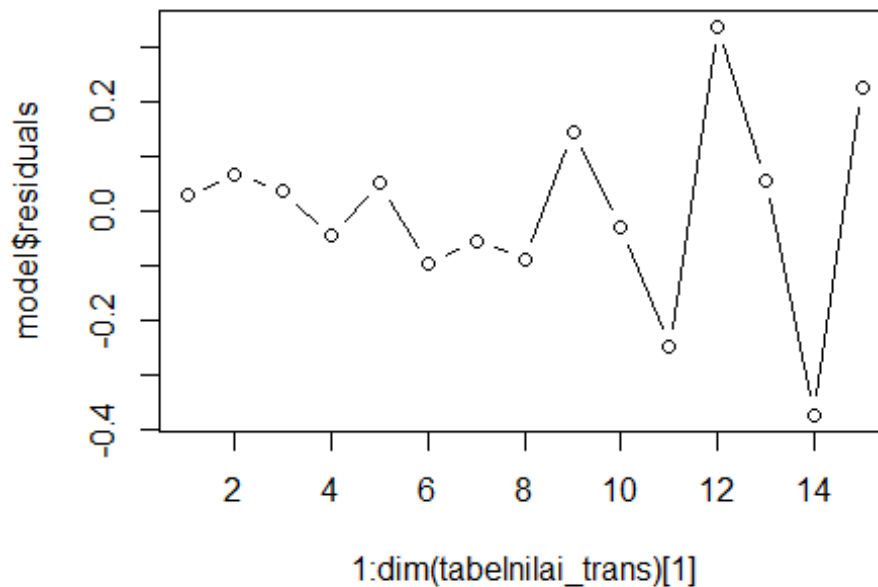
H0 : Ragam Homogen H1 : Ragam tidak homogen

```
library(lmtest)  
## Warning: package 'lmtest' was built under R version 4.3.3  
## 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 <- bptest(model)  
homogen
```

```
##
## studentized Breusch-Pagan test
##
## data: model
## BP = 6.9535, df = 1, p-value = 0.008365
```

Nilai P-value pada uji kehomogonan itu menyatakan ada perbedaan yang cukup signifikan dalam galat yakni bernilai 0.008365. Galat tidak homogen dan cenderung berbentuk corong. Dibutuhkan analisis kuadrat tekecil terboboti.

```
plot(x = 1:dim(tabelnilai_trans)[1],
     y = model$residuals,
     type = 'b')
```



Ragam membentuk suatu pola dari kecil ke besar : H_0 : Galat acak saling bebas H_1 : Galat tak acak dan tak saling bebas

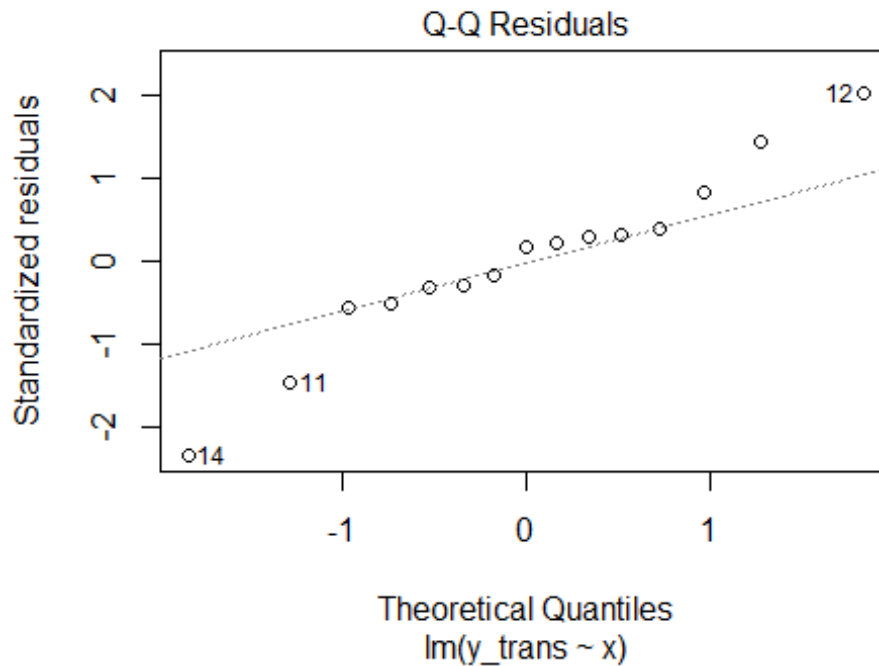
uji runs.test

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

##
## Runs Test
##
## data: model$residuals
## statistic = 0.55635, runs = 9, n1 = 7, n2 = 7, n = 14, p-value = 0.578
## alternative hypothesis: nonrandomness
```

Nilai p-value > 0.05, maka terima H_0 dengan galat acak saling bebas

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



Plot QQ galat

terlihat sebaran yang cenderung simetris dengan ujung-ujung menipis. H_0 = data menyebar normal H_1 = Data tidak menyebar normal

shapiro-wilk test

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

```
##  
## Shapiro-Wilk normality test  
##  
## data: model$residuals  
## W = 0.96487, p-value = 0.7762
```

p-value > 0.05. Maka terima H_0 yaitu data menyebar normal. Dari 3 asumsi Gauss-Markov, 1 asumsi tidak terpenuhi yaitu asumsi galat menyebar homogen.

WLS

```
bobot <- 1/((model$fitted.values)^2)  
  
wls_model <- lm(y_trans ~ x, data = tabelnilai_trans, weights = bobot)  
summary(wls_model)  
  
##  
## Call:
```



```
## lm(formula = y_trans ~ x, data = tabelnilai_trans, weights = bobot)
##
## Weighted Residuals:
##      Min        1Q      Median        3Q        Max
## -0.22264 -0.02085  0.01294  0.02128  0.15618
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.012204   0.143682   27.92 5.44e-13 ***
## x           -0.037239   0.003058  -12.18 1.75e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09149 on 13 degrees of freedom
## Multiple R-squared:  0.9194, Adjusted R-squared:  0.9132
## F-statistic: 148.3 on 1 and 13 DF,  p-value: 1.751e-08

homogen2 <- bptest(wls_model)
homogen2

##
## studentized Breusch-Pagan test
##
## data:  wls_model
## BP = 2.3997, df = 1, p-value = 0.1214
```

Setelah dilakukan pembobotan, galat bernilai homogen. Hal ini dapat dilihat melalui standar error sisaan yang berkurang dari 0.1794 menjadi 0.09149. Akan tetapi nilai R-Squared juga berkurang dari 0.9551 menjadi 0.9194 yang menandakan bahwa regresi terbobot ini tak lebih mewakilkan nilai y terhadap x dibandingkan regresi awal sebelum dilakukan pembobotan.

Persamaan regresi : $y \text{ duga} = 4.012 - 0.037X$.

```
yduga <- 4.012 - 0.037*x
tabelnilairegresi <- data.frame(y_trans, yduga)
tabelnilairegresi

##      y_trans yduga
## 1  3.988984 3.938
## 2  3.912023 3.827
## 3  3.806662 3.753
## 4  3.610918 3.642
## 5  3.555348 3.494
## 6  3.218876 3.309
## 7  2.995732 3.050
## 8  2.772589 2.865
## 9  2.890372 2.754
## 10 2.564949 2.606
## 11 2.079442 2.347
## 12 2.397895 2.088
```

```
## 13 2.079442 2.051
## 14 1.386294 1.792
## 15 1.791759 1.607
```