Tugas Individu Minggu 7

Shabrina Shafwah Al-Rahmah G1401221083

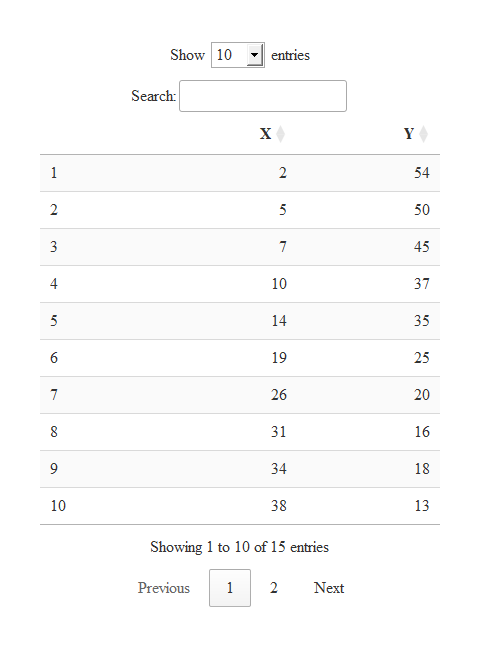
2024-03-05

# Data

X <- c(2, 5, 7, 10, 14, 19, 26, 31, 34, 38, 45, 52, 53, 60, 65)  
Y <- c(54, 50, 45, 37, 35, 25, 20, 16, 18, 13, 8, 11, 8, 4, 6)  
data <- data.frame(X,Y)  
library(DT)

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

datatable(data)



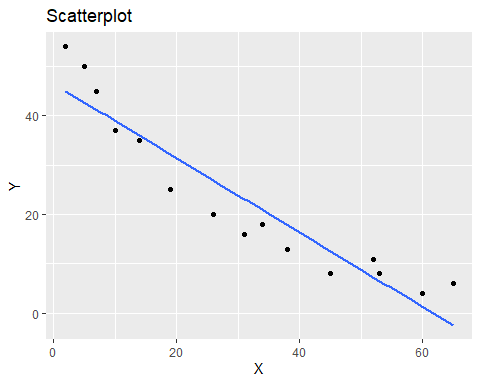
# Eksplorasi Data

model <- lm(Y ~ X, data = data)  
  
library(ggplot2)

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

ggplot(data, aes(x = X, y = Y)) +   
 geom\_point() +  
 geom\_smooth(method = "lm", se = FALSE) +  
 labs(title = "Scatterplot")

## `geom\_smooth()` using formula = 'y ~ x'



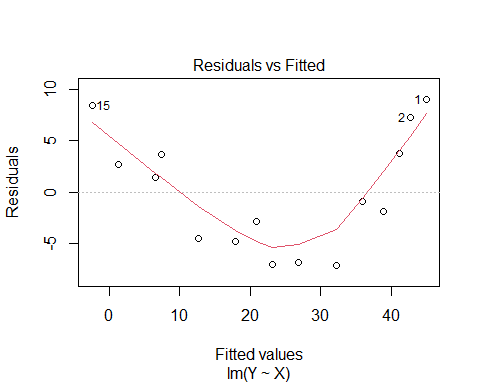
Berdasarkan *scatter plot* di atas, dapat dilihat bahwa peubah X dan peubah Y membentuk suatu pola linier dari kiri atas ke kanan bawah. Hubungan tersebut membentuk pola linier negatif.

# Pemeriksaan Asumsi

## Nilai harapan galat sama dengan nol ()

Plot sisaan vs

plot(model, 1)



Berdasarkan plot di atas, sisaan berada di sekitar nol sehingga nilai harapan sama dengan nol. Namun, sisaan membentuk pola kurva parabola sehingga ada kemungkinan ragam tidak homogen dan berarti model tidak pas. Perlu transformasi Y untuk membentuk model yang lebih baik.

Uji Formal

library(coin)

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

## Loading required package: survival

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

##   
## One Sample t-test  
##   
## data: model$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

Hasil bptest menunjukkan bahwa p-value > 0.05 sehingga tak tolak . Hal ini mengindikasikan bahwa nilai harapan galat sama dengan nol.

## Ragam galat homogen atau homoskedastisitas ()

Uji formal untuk mendeteksi homogenitas ragam sisaan dapat dilakukan dengan uji Breusch-Pagan menggunakan fungsi bptest yang memiliki hipotesis sebagai berikut.

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, data=data)

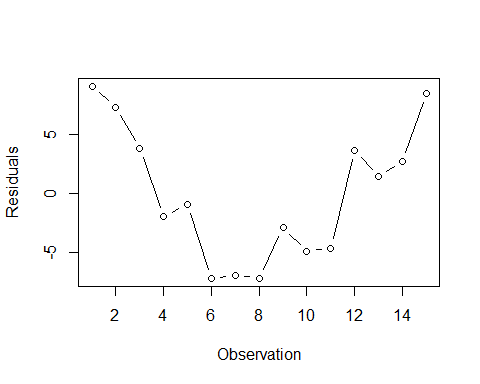
##   
## studentized Breusch-Pagan test  
##   
## data: model  
## BP = 0.52819, df = 1, p-value = 0.4674

Hasil bptest menunjukkan bahwa p-value > 0.05 sehingga tak tolak . Hal ini mengindikasikan bahwa ragam sisaan homogen.

## Galat saling bebas

Plot Sisaan vs Urutan

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



Uji Formal

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

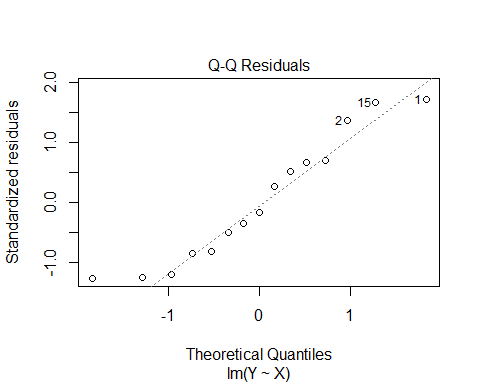
##   
## Runs Test  
##   
## data: model$residuals  
## statistic = -2.7817, runs = 3, n1 = 7, n2 = 7, n = 14, p-value =  
## 0.005407  
## alternative hypothesis: nonrandomness

Hasil runs test menunjukkan bahwa p-value < 0.05 sehingga tolak . Hal ini mengindikasikan bahwa ada autokorelasi atau sisaan tidak saling bebas pada model.

## Galat Menyebar Normal

Plot qq

plot(model,2)



Berdasarkan qq-plot tersebut, sisaan cenderung mendekati garis diagonal yang mewakili distribusi normal sehingga galat menyebar normal.

Uji Formal

shapiro.test(residuals(model))

##   
## Shapiro-Wilk normality test  
##   
## data: residuals(model)  
## W = 0.92457, p-value = 0.226

Berdasarkan hasil uji *Shapiro-Wilk* nilai p-value > 0 sehingga tak tolak . Hal ini menunjukkan bahwa galat menyebar normal.

# Dugaan Persamaan Regresi

summary(model)

##   
## 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

Didapatkan dugaan persamaan regresi sebagai berikut.

Hasil pendugaan parameter regresi menunjukkan bahwa rataan dugaan nilai y akan turun sebesar 0.752 jika berubah satu satuan. Adapun ketika x=0 (jika ada dalam selang pengamatan), maka dugaan rataan nilai y akan bernilai sebesar 46.460.

# Ukuran Kelayakan Model

summary\_model <- summary(model)  
  
(r\_squared <- summary\_model$r.squared)

## [1] 0.8855804

(adj\_r\_squared <- summary\_model$adj.r.squared)

## [1] 0.8767789

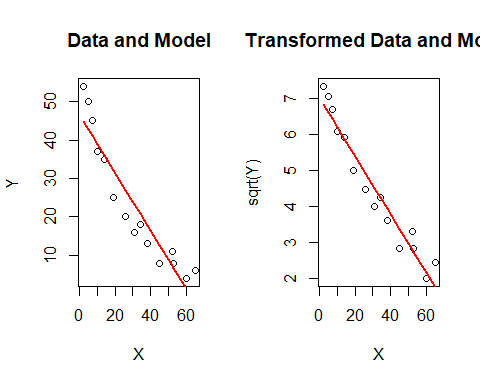
# Penanganan Masalah

Uji asumsi-asumsi yang telah dilakukan sebelumnya menunjukkan perlunya transformasi agar terbentuk model yang lebih baik, salah satunya karena terbentuknya pola parabola pada plot residual.

## Transformasi untuk Meluruskan: Pola Parabola

Transformasi ini dilakukan dengan Y diperkecil, yaitu .

data\_transformed <- data.frame(X, sqrt(data$Y))  
model2 <- lm(sqrt(Y) ~ X, data=data\_transformed)  
  
predicted\_values <- predict(model, data)  
predicted\_values2 <- predict(model2, data\_transformed)  
  
par(mfrow = c(1, 2))  
plot(X,Y,main="Data and Model", xlab="X", ylab="Y")  
lines(X, predicted\_values, col="red", lwd=2)  
  
plot(X, sqrt(data$Y), main="Transformed Data and Model", xlab="X", ylab="sqrt(Y)")  
lines(X, predicted\_values2, col="red", lwd=2)

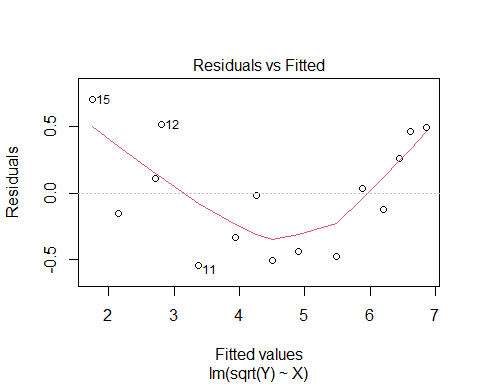


par(mfrow = c(1, 1))

# Uji Asumsi Setelah Transformasi

## Nilai harapan galat sama dengan nol

plot(model2, 1)



Uji Formal

library(coin)  
t.test(model2$residuals,mu=0,conf.level=0.95)

##   
## One Sample t-test  
##   
## data: model2$residuals  
## t = -1.0367e-16, df = 14, p-value = 1  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## -0.229541 0.229541  
## sample estimates:  
## mean of x   
## -1.1095e-17

Hasil uji t-test menunjukkan bahwa p-value > 0.05 sehingga tidak tolak H0. Hal ini mengindikasikan nilai harapan galat sama dengan 0.

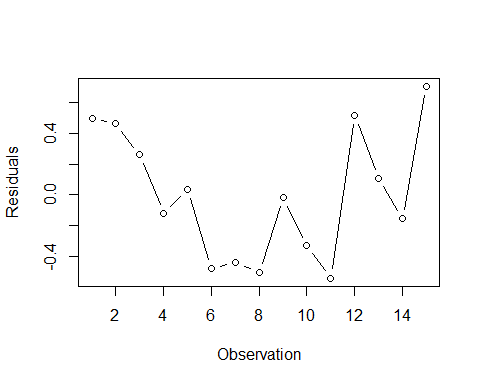
## Ragam galat homogen atau homoskedastisitas ()

library(lmtest)  
bptest(model2, data=data)

##   
## studentized Breusch-Pagan test  
##   
## data: model2  
## BP = 0.93605, df = 1, p-value = 0.3333

Hasil uji bptest menunjukkan bahwa p-value > 0.05 sehingga tidak tolak H0. Hal ini mengindikasikan ragam galat homogen.

plot(x = 1:dim(data)[1],  
 y = model2$residuals,  
 type = 'b',   
 ylab = "Residuals",  
 xlab = "Observation")



## Galat saling bebas

Uji Formal

library(randtests)  
runs.test(model2$residuals)

##   
## Runs Test  
##   
## data: model2$residuals  
## statistic = -0.55635, runs = 7, n1 = 7, n2 = 7, n = 14, p-value = 0.578  
## alternative hypothesis: nonrandomness

Hasil runs test menunjukkan bahwa p-value > 0.05 sehingga tidak tolak . Hal ini mengindikasikan bahwa tidak ada autokorelasi atau sisaan saling bebas.

## Galat Menyebar Normal

Uji Formal

shapiro.test(residuals(model2))

##   
## Shapiro-Wilk normality test  
##   
## data: residuals(model2)  
## W = 0.93128, p-value = 0.2852

Berdasarkan uji *shapiro-test* didapatkan bahwa p-value > 0.05 sehingga tidak tolak . Artinya, galat menyebar normal.

# Ukuran Kelayakan Model 2

summary\_model2 <- summary(model2)  
  
(r\_squared2 <- summary\_model2$r.squared)

## [1] 0.9439477

(adj\_r\_squared2 <- summary\_model2$adj.r.squared)

## [1] 0.939636

#Perbandingan Ukuran Kelayakan Model

(table\_data <- data.frame(  
 Model = c("Model 1", "Model 2"),  
 R\_squared = c(r\_squared, r\_squared2),  
 Adj\_R\_squared = c(adj\_r\_squared, adj\_r\_squared2)  
))

## Model R\_squared Adj\_R\_squared  
## 1 Model 1 0.8855804 0.8767789  
## 2 Model 2 0.9439477 0.9396360

Dapat terlihat bahwa baik maupun memiliki nilai yang lebih besar dan lebih mendekati nilai 1 pada model 2 hasil transformasi daripada model pratransformasi sehingga dapat dikatakan bahwa model 2 lebih layak dibandingkan model 1.

# Pendugaan Parameter Regresi

summary(model2)

##   
## Call:  
## lm(formula = sqrt(Y) ~ X, data = data\_transformed)  
##   
## 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

Dugaan persamaan regresi hasil transformasi sebagai berikut.

# Tranformasi Balik