

# EKONOMETRİ

Bu çalışmada bağımlı değişkeni maaş , bağımsız değişkenleri eğitim ve deneyim olan veri seti kullanılmıştır. Aşağıdaki maddeler uygulanmıştır:

1.  $MAAŞ = \beta_0 + \beta_1 EGITIM + \beta_2 DENEYIM$  tahmin ediniz
2. Artıkların X'e karşı grafiğini çiziniz
3. Artıkların histogramını çiziniz
4. Artıklara Normalik Testi
5. Park Testi
6. Goldfeld-Quandt Testi (c=3)
7. Glejser Testi
8. Spearman Sıra Korelasyon Testi
9. Breusch-Pagan-Godfrey Testi
10. White Testi
11. Kalınlıların bir önceki kalıntıya göre grafiğini çiziniz.
12. Ağırlıklandırılmış EKK uygulayınız.
13.  $MAAŞ = \beta_1 EGITIM + \beta_2 DENEYIM$  tahmin ediniz. 1.'de elde ettiklerinizle karşılaştırınız.
14. MAAŞ ve DENEYIM arasında Log-Log Model, Inverse ve Quadtrac Modeli (sabit terimli ve sabit terimsiz) uygulayıp karşılaştırınız.

Kullanılan Paketler:

```
install.packages("tidyverse")#verilere ulaşmanıza yardımcı olan dizi işlev sağlar.
```

Installing package into '/cloud/lib/x86\_64-pc-linux-gnu-library/4.3'  
(as 'lib' is unspecified)

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2    3.5.1      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr      1.0.2
```

```
-- Conflicts ----- tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
```

```
x dplyr::lag()     masks stats::lag()
```

```
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
install.packages("ggplot2")#Grafik oluşturmak için kullanılır.
```

Installing package into '/cloud/lib/x86\_64-pc-linux-gnu-library/4.3'  
(as 'lib' is unspecified)

```
library(ggplot2)
```

```
install.packages("olsrr")#Regresyon modelleri, testler içerir.
```

Installing package into '/cloud/lib/x86\_64-pc-linux-gnu-library/4.3'  
(as 'lib' is unspecified)

```
library(olsrr)
```

Attaching package: 'olsrr'

The following object is masked from 'package:datasets':

```
rivers
```

```
install.packages("skedastic")#Heteroskedasticity için kullanılır.
```

Installing package into '/cloud/lib/x86\_64-pc-linux-gnu-library/4.3'  
(as 'lib' is unspecified)

```
library(skedastic)  
install.packages("lmtest")#Doğrusal regresyon modellerinde testler için.
```

Installing package into '/cloud/lib/x86\_64-pc-linux-gnu-library/4.3'  
(as 'lib' is unspecified)

```
library(lmtest)
```

Loading required package: zoo

Attaching package: 'zoo'

The following objects are masked from 'package:base':

as.Date, as.Date.numeric

```
install.packages("stats")#istatistik paketi.
```

Installing package into '/cloud/lib/x86\_64-pc-linux-gnu-library/4.3'  
(as 'lib' is unspecified)

Warning: package 'stats' is a base package, and should not be updated

```
library(stats)
```

Veri Girişi

```
maas <- c(9746, 10299, 10752, 10561, 10580, 10666,  
          10704, 10737, 10762, 10784, 10784, 10829, 10844, 10846,  
          10858, 10858, 10858, 10958, 10878, 10878)  
egitim <- c(14, 17, 18, 18, 19, 19, 19, 20, 19,  
            19, 21, 21, 22, 19, 22, 23, 24, 25, 24, 18)  
deneyim <- c(17, 19, 15, 29, 26, 17, 24, 18, 24,
```

```

19, 18, 12, 15, 24, 22, 24, 24, 26, 13, 15)

data <- (data.frame(MAAS = maas, EGITIM = egitim, DENEYIM = deneyim))

data <- na.omit(data)

```

```
str(data)
```

```

'data.frame':  20 obs. of  3 variables:
 $ MAAS   : num  9746 10299 10752 10561 10580 ...
 $ EGITIM : num   14  17  18  18  19  19  19  20  19  19 ...
 $ DENEYIM: num   17  19  15  29  26  17  24  18  24  19 ...

```

Bu veri seti 20 satır ve 3 sütundan oluşur. Bu üç değişkenin de nümerik veri tipindedir.

### 1.Model Tahmini:

MAAŞ= Beta0 + Beta1EGITIM+ Beta2DENEYIM modelini tahmin edelim.

```

model <- lm(MAAS ~ EGITIM + DENEYIM, data = data)
model

```

Call:

```
lm(formula = MAAS ~ EGITIM + DENEYIM, data = data)
```

Coefficients:

(Intercept)	EGITIM	DENEYIM
9222.953	76.016	-1.893

Modelin özeti:

```
summary(model)
```

Call:

```
lm(formula = MAAS ~ EGITIM + DENEYIM, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-508.98	-79.93	14.52	96.70	315.17

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	9222.953	352.373	26.174	3.55e-15	***
EGITIM	76.016	15.895	4.782	0.000173	***
DENEYIM	-1.893	8.803	-0.215	0.832261	

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 186.6 on 17 degrees of freedom

Multiple R-squared: 0.5742, Adjusted R-squared: 0.5241

F-statistic: 11.46 on 2 and 17 DF, p-value: 0.0007056

H0 : Beta = 0 (Model anlamlı değildir.)

H1 : Beta  $\neq$  0 (Model anlamlıdır.)

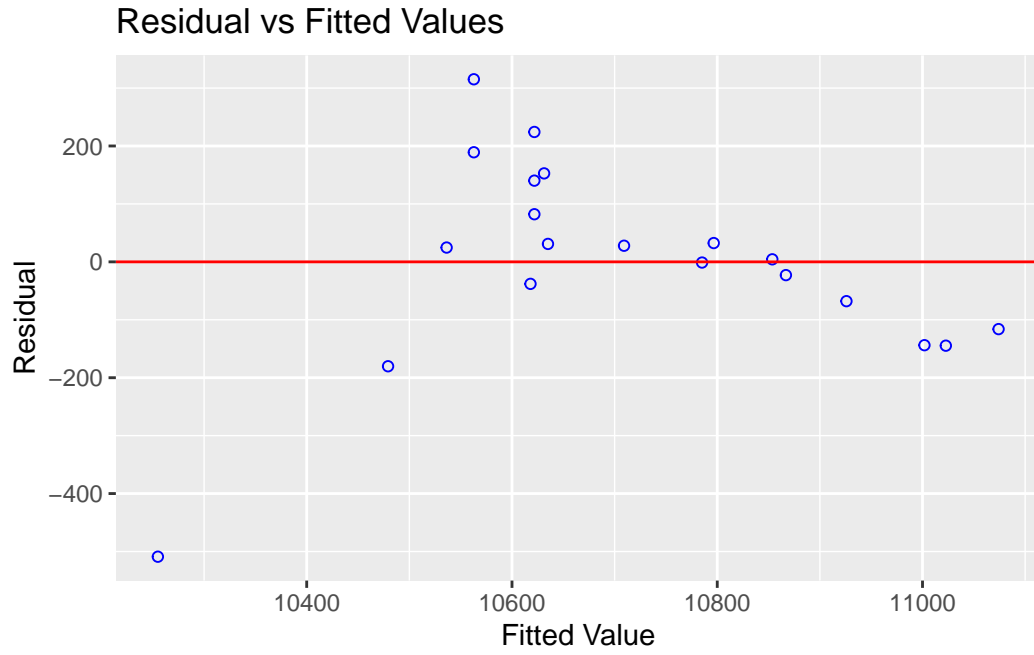
p – değeri = 0.0007056 yani 0.05'ten çok küçük bir sayı olduğundan H0 hipotezimiz reddedilir, model anlamlıdır.

R<sup>2</sup> değeri bağımlı değişkenin yüze kaç oranında bağımsız değişkenler tarafından açıklandığını belirtir. Modeldeki bağımlı değişken olan maaş %57.432 oranında bağımsız değişkenler tarafından açıklanmaktadır.

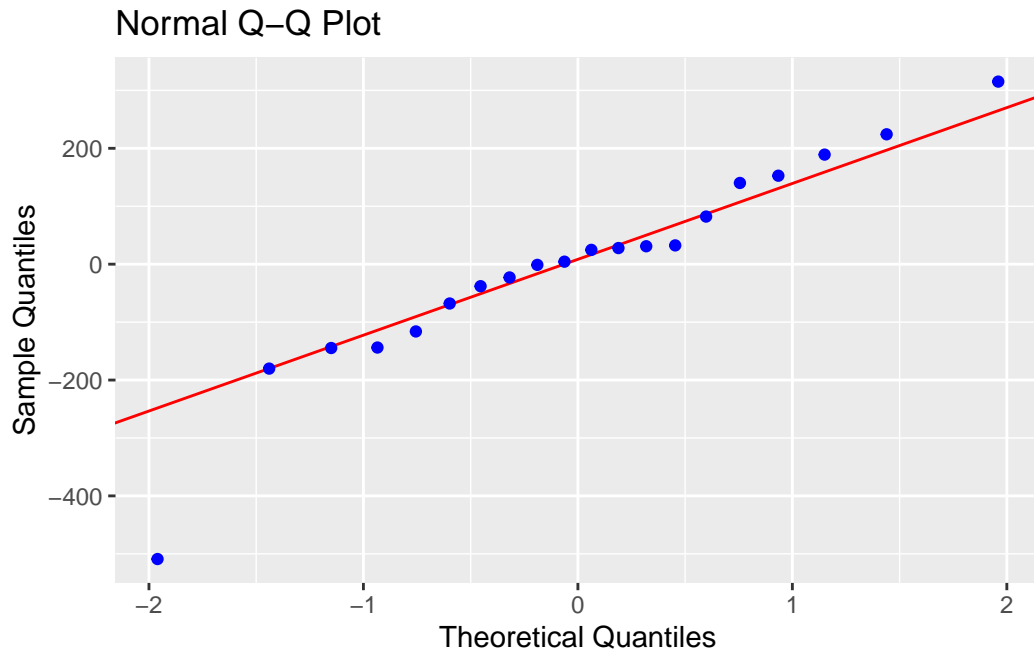
## 2.Artıkların X'e Karşı Grafikleri

Artık grafikler genellikle bir regresyon analizindeki artıkların normal olarak dağılıp dağılmadığını ve heteroskedasticity gösterip göstermediklerini değerlendirmek için kullanılır.

```
ols_plot_resid_fit(model)
```



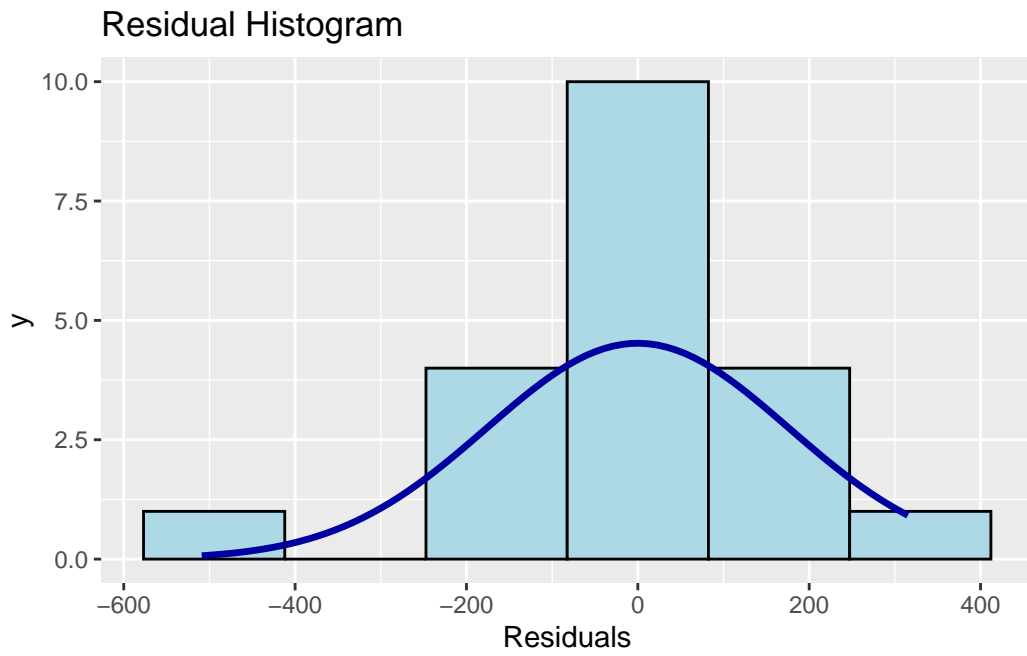
```
ols_plot_resid_qq(model)
```



Artıkların grafiğine bakarak değişen varyans olabilir diyebiliriz.

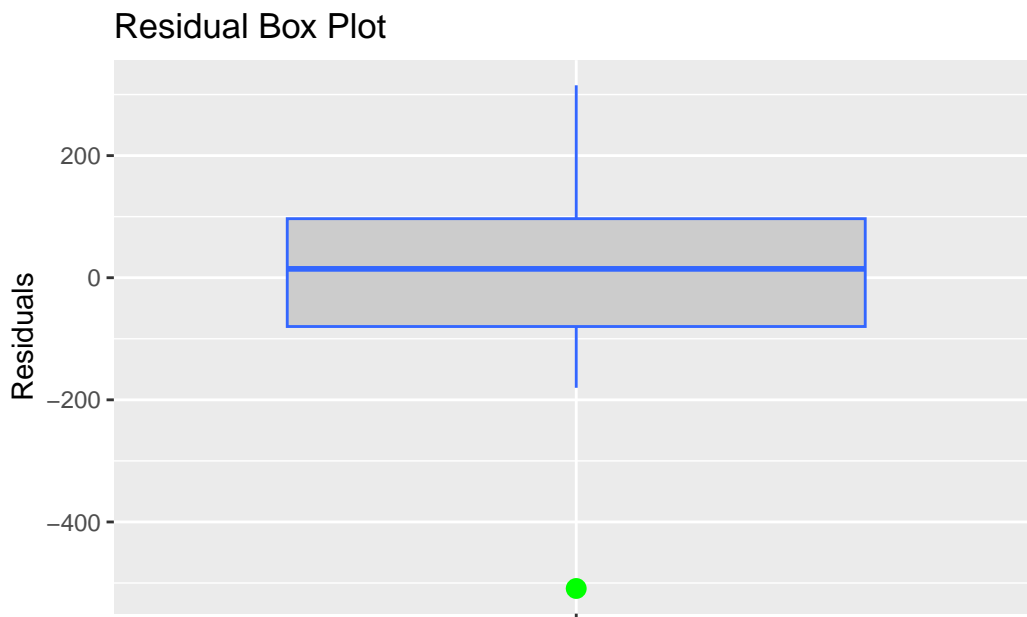
### 3. Artıkların Histogramı

```
ols_plot_resid_hist(model)
```



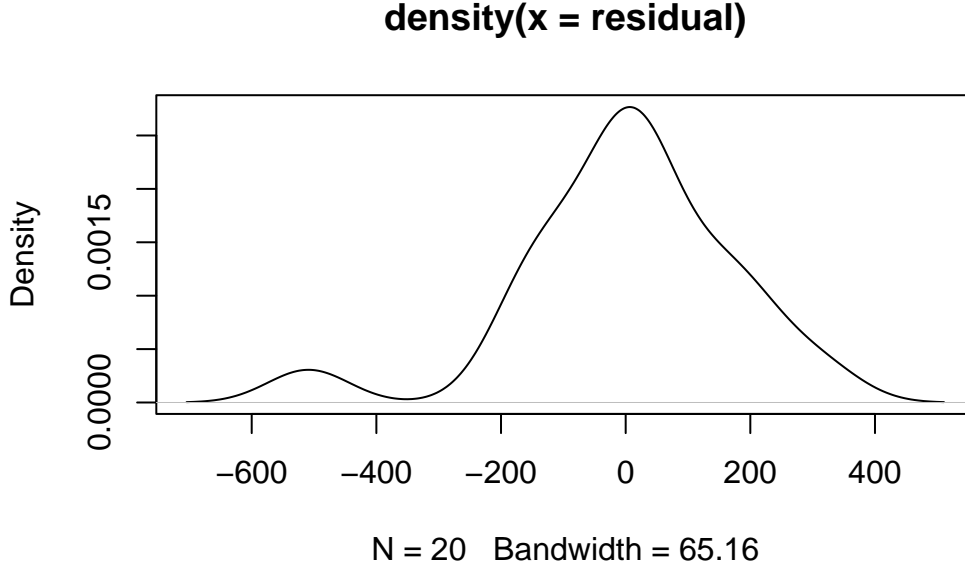
Artıkların boxplotını çizelim:

```
ols_plot_resid_box(model)
```



Artıkların dağılım eğrisini çizelim:

```
residual <- (model$residuals)
plot(density(residual))
```



Yukarıdaki grafiklerde görüldüğü üzere artıklar nokta sıfır noktasında yoğunlaşmaktadır. Bu grafiklerden aldığımız bilgiye dayanarak verilerin normal dağıldığını söyleyebiliriz.

#### 4. Artıklara Normalik Testi

```
ols_test_normality(model)
```

Test	Statistic	pvalue
Shapiro-Wilk	0.9333	0.1785
Kolmogorov-Smirnov	0.1271	0.8636
Cramer-von Mises	1.7167	0.0000
Anderson-Darling	0.4367	0.2678

H0: Artıklar Normal dağılır. (p-value > 0.05)

H1: Artıklar Normal dağılmaz. (p-value < 0.05)

Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling test istatistiklerinin p-value değerleri 0.05 değerinden büyük olduğundan H0 dağıldığı hipotezini kabul ederiz.



## 5. Park Testi

Park testinin özel bir fonksiyonu olmadığından modelini oluşturmamız gerekir.

```
res <- residuals(model)
res2 <- res^2
lnres2 <- log(res2)
y_tahmin <- model$fitted.values
lny_tahmin <- log(y_tahmin)
yardımcı_model <- lm(lnres2 ~ lny_tahmin)
summary(yardımcı_model)
```

Call:

```
lm(formula = lnres2 ~ lny_tahmin)
```

Residuals:

Min	1Q	Median	3Q	Max
-7.592	-1.479	1.081	2.102	3.047

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	454.74	317.65	1.432	0.169
lny_tahmin	-48.11	34.23	-1.405	0.177

Residual standard error: 2.856 on 18 degrees of freedom

Multiple R-squared: 0.09889, Adjusted R-squared: 0.04882

F-statistic: 1.975 on 1 and 18 DF, p-value: 0.1769

p değeri > 0.05 olduğu için “H0:Sabit varyans varsayımı geçerlidir” hipotezi reddedilemez.

## 6. Goldfeld-Quandt testi (c=3) Testi

Goldfeld-Quandt testi (c=3) testi için gqtest() fonksiyonu kullanılır.

```
gqtest(model, order.by = ~ EGITIM + DENEYIM, data = data, fraction = 3)
```

Goldfeld-Quandt test

data: model

GQ = 0.070711, df1 = 6, df2 = 5, p-value = 0.9971

alternative hypothesis: variance increases from segment 1 to 2

H0: Sabit Varyans vardır. (p değeri > 0.05)

H1: Değişen varyans vardır. (p değeri < 0.05) p değeri 0.05'ten büyük olduğundan H0 varsayımını kabul eder sabit varyanslıdır diyebiliriz.

## 7. Glejser Testi

Glejser Testi için `glejser()` fonksiyonu kullanılır.

```
glejser(model)
```

```
# A tibble: 1 x 4
  statistic p.value parameter alternative
    <dbl>    <dbl>    <dbl> <chr>
1      8.40  0.0150         2 greater
```

P-değeri 0.05'ten küçük olduğundan heteroscedasticity'nin mevcut olduğunu söyleyebiliriz.

## 8. Spearman Sıra Korelasyon Testi

```
cor.test(model$fitted.values, model$residuals, method = "spearman")
```

```
Warning in cor.test.default(model$fitted.values, model$residuals, method =
"spearman"): Cannot compute exact p-value with ties
```

Spearman's rank correlation rho

```
data: model$fitted.values and model$residuals
S = 1680.7, p-value = 0.2614
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
-0.263654
```

Tahmin edilen X'ler ile artıklar arasında anlamlı bir ilişki yoktur.

## 9. Breusch-Pagan-Godfrey Testi

```
bptest(model)
```

```
studentized Breusch-Pagan test
```

```
data: model
```

```
BP = 7.609, df = 2, p-value = 0.02227
```

p değeri 0.05'ten küçük olduğu için modelde değişen varyansın mevcut olduğunu söyleyebiliriz.

## 10. White Testi

```
white(model)
```

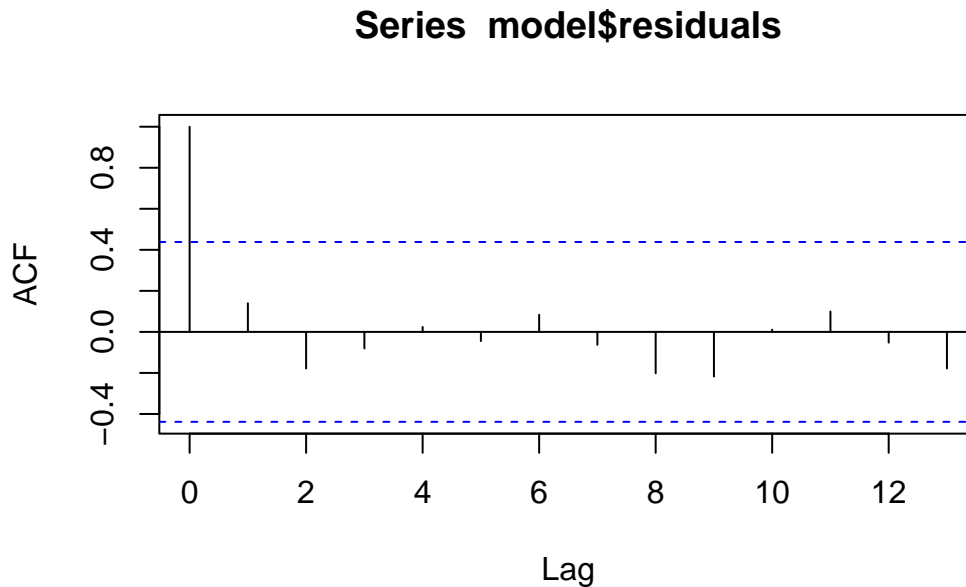
```
# A tibble: 1 x 5
```

	statistic	p.value	parameter	method	alternative
	<dbl>	<dbl>	<dbl>	<chr>	<chr>
1	17.4	0.00158	4	White's Test	greater

p değeri 0.05'ten küçük olduğu için modelde değişen varyansın mevcut olduğunu söyleyebiliriz.

## 11. Kalıntıların Bir Önceki Kalıntıya Göre Grafiği

```
acf(model$residuals, type = "correlation")
```



## 12. Ağırlıklandırılmış EKK

```
wt <- 1 / lm(abs(model$residuals) ~ model$fitted.values)$fitted.values^2
wls_model <- lm( MAAS ~ . , data = data, weights=wt)
summary(wls_model)
```

Call:

```
lm(formula = MAAS ~ ., data = data, weights = wt)
```

Weighted Residuals:

Min	1Q	Median	3Q	Max
-2.6080	-0.4126	0.2011	0.6430	1.5220

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	9781.709	161.687	60.498	< 2e-16 ***
EGITIM	44.989	7.351	6.120	1.13e-05 ***
DENEYIM	1.930	1.993	0.968	0.347

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.066 on 17 degrees of freedom

Multiple R-squared: 0.7833, Adjusted R-squared: 0.7578

F-statistic: 30.73 on 2 and 17 DF, p-value: 2.26e-06

H0 : Beta = 0 (Model anlamlı değildir.) H1 : Beta = 0 (Model anlamlıdır.) Basit doğrusal regresyon modelinden çok daha küçük bir p değeri görmekteyiz. H0 hipotezi reddedilir. Model anlamlıdır. Ağırlıklı en küçük kareler modeli daha anlamlı bir modeldir sonucuna varabiliriz.

13.  $MAAS = \text{Beta1}EGITIM + \text{Beta2}DENEYIM$  tahmin ediniz. 1'de elde ettiğinizle karşılaştırmamız.

```
lm.model <- lm(MAAS ~ EGITIM + DENEYIM - 1, data = data)
summary(lm.model)
```

Call:

```
lm(formula = MAAS ~ EGITIM + DENEYIM - 1, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-2328.2	-267.0	234.9	776.0	2075.2

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
EGITIM	433.21	50.89	8.512	1e-07 ***
DENEYIM	94.47	49.94	1.891	0.0748 .

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1165 on 18 degrees of freedom

Multiple R-squared: 0.9894, Adjusted R-squared: 0.9882

F-statistic: 836.2 on 2 and 18 DF, p-value: < 2.2e-16

Katsayılar incelendiğinde Beta0 ın çıkarıldığı modelin istatistiksel olarak daha anlamlı sonuçlar ürettiğini söyleyebiliriz.

**14. MAAS ve DENEYIM arasında log-log model, inverse ve quadtrac model i (sabit terimli ve sabit terimsiz) uygulayıp karşılaştırınız.**

```
#(Sabit Terimli)
log_model1 <- lm(log(MAAS) ~ log(DENEYIM), data = data)
#(Sabit Terimsiz)
log_model2 <- lm(log(MAAS) ~ log(DENEYIM) -1 , data =data)
```

```
summary(log_model1)
```

Call:

```
lm(formula = log(MAAS) ~ log(DENEYIM), data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.093707	-0.001243	0.007367	0.013796	0.022837

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	9.273864	0.072546	127.834	<2e-16 ***
log(DENEYIM)	0.001573	0.024353	0.065	0.949

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02681 on 18 degrees of freedom  
Multiple R-squared: 0.0002316, Adjusted R-squared: -0.05531  
F-statistic: 0.00417 on 1 and 18 DF, p-value: 0.9492

```
summary(log_model2)
```

Call:

```
lm(formula = log(MAAS) ~ log(DENEYIM) - 1, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.1874	-0.5753	0.1230	0.5794	1.5766

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
log(DENEYIM)	3.10408	0.05904	52.58	<2e-16 ***
---				

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7866 on 19 degrees of freedom  
Multiple R-squared: 0.9932, Adjusted R-squared: 0.9928  
F-statistic: 2764 on 1 and 19 DF, p-value: < 2.2e-16

Sabit terimli modelide hatalar sabit terim çıkartıldığı zaman artmıştır. Sabit terimli modelimde R kare sabit terim çıkartıldığı zaman R kare artmıştır. Bu da sabit terimi olmayan modelin daha ilişkili yani varyansı daha iyi açıkladığını göstermektedir.