

Fen Fakültesi Aktüerya Bilimleri Bölümü

İST359 Regresyon Çözümlemesi

Regresyon Süreci Ödevi

Prof. Dr. Duru KARASOY

Hakkı KONDAK – 2220381067

17/12/2024

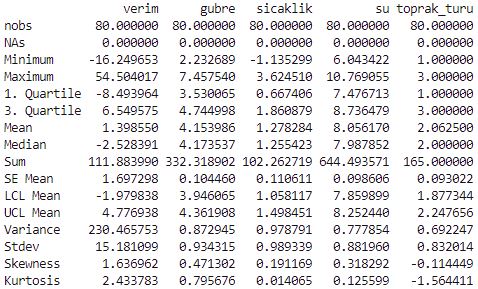
# İndeks

1. [Senaryo](#_Senaryo)
2. [Tanımlayıcı İstatistikler](#_Tanımlayıcı_İstatistikler)
3. [Varsayım Kontrolleri](#_Varsayım_Kontrolleri)
   1. [Normallik Varsayımı](#_Normallik_Varsayımı)
   2. [Doğrusallık Varsayımı](#_Doğrusallık_Varsayımı)
4. [Doğrusal Modelin Kuruluşu](#_Doğrusallık_Modelinin_Kuruluşu)
5. [Artıkların İncelenmesi](#_Artıkların_İncelenmesi)
   1. [Hat Uzaklığı](#_Hat_Uzaklığı)
   2. Student Türü Artıklar
   3. Cook Uzaklığı
   4. Standart Türü Artıklar
6. Yeni Veri Setinde Tanımlayıcı İstatistikler
7. Yeni Veri Setinde Varsayım Kontrolleri
   1. Normallik Varsayımı
   2. Doğrusallık Varsayımı

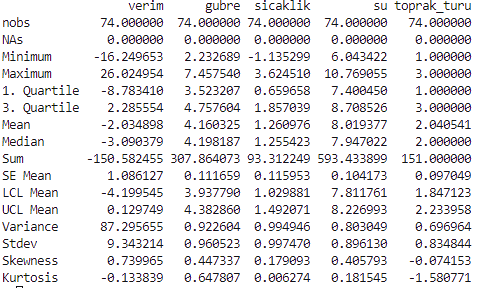
# Senaryo

# Tanımlayıcı İstatistikler

İlk datamızın tamamlayıcı istatistikleri



Yeni datamızın tamamlayıcı istatistikleri

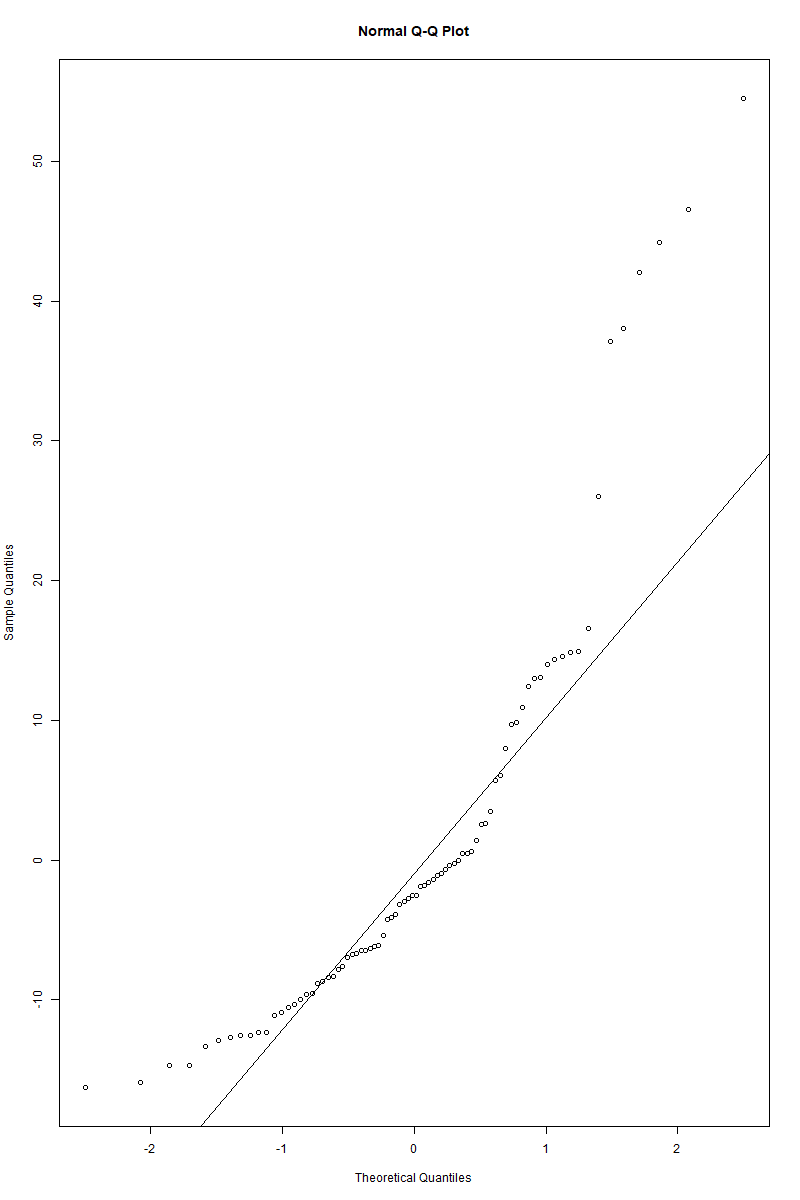


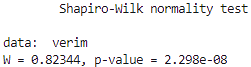
Bağımlı ve bağımsız değişkenlerimize ait en büyük, en küçük, ortalama, ortanca, varyans, standart sapma, basıklık ve çarpıklık değerleri gibi tanımlayıcı istatistikler tabloda görülmektedir.

# Varsayım Kontrolleri

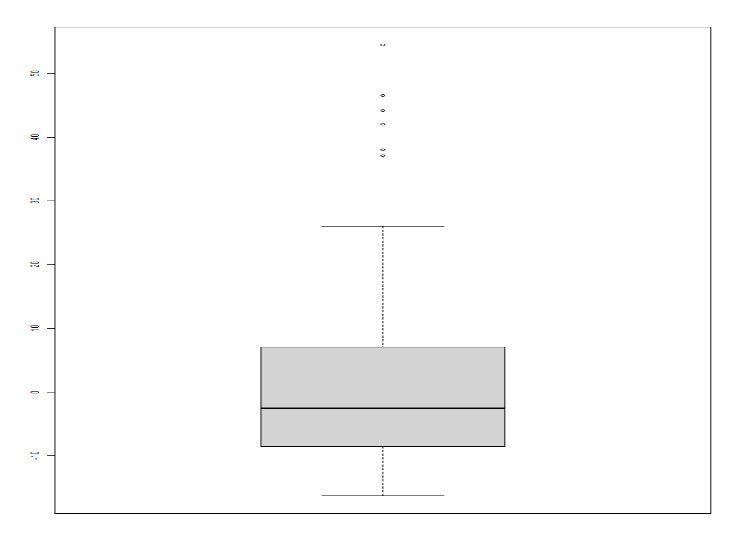
## Normallik Varsayımı

Shapiro-wilk testimize bakıyoruz

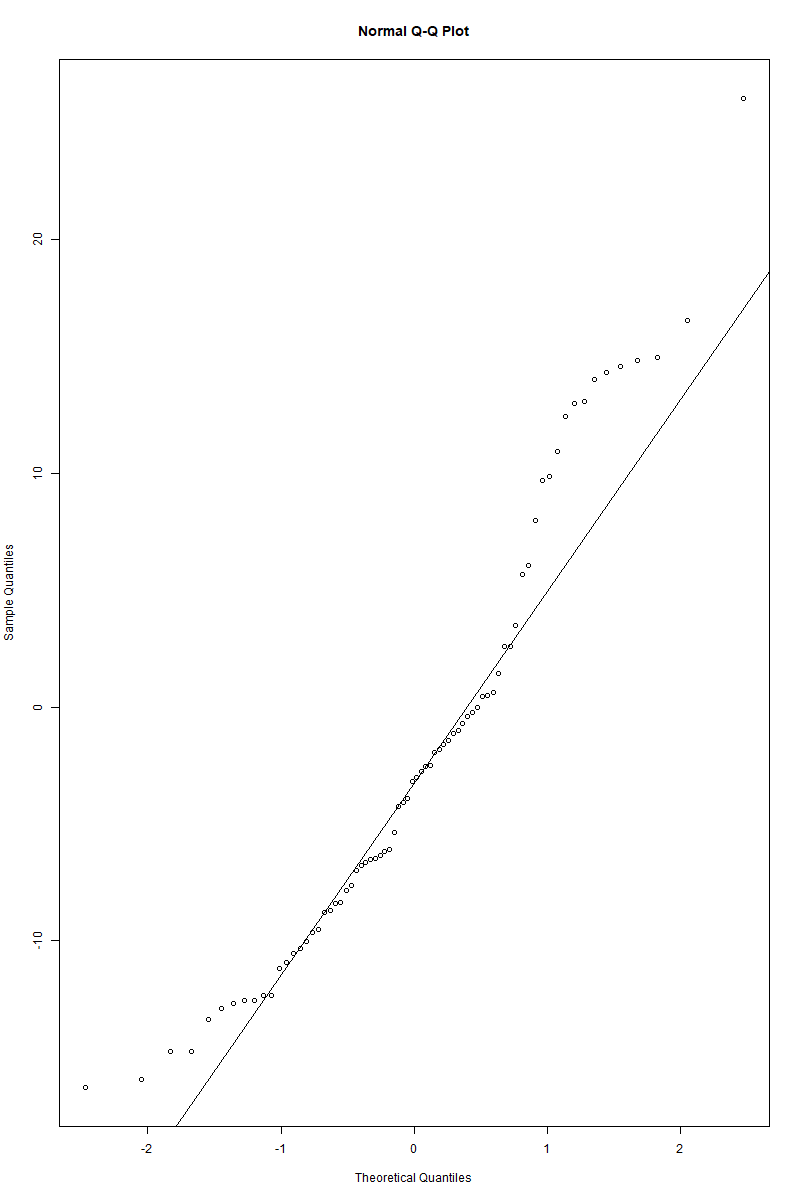




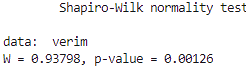
**Normallik sağlamazdığını görüyoruz ve Boxplot grafiğinden artık değerlerini bulup çıkarıyoruz**

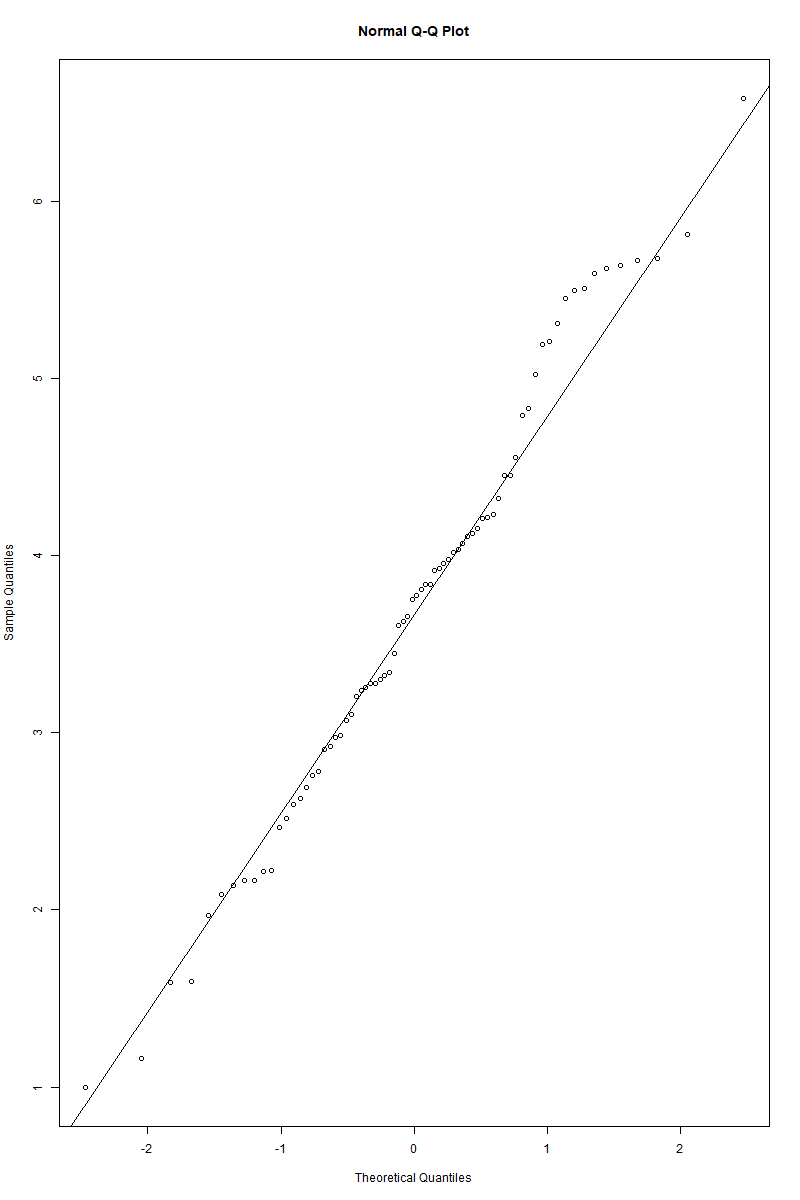
Görüldüğü üzere 17, 33, 46, ,64, 69 ve 72 artık değerlerimizdir. Bunları silip tekrardan datamızı oluşturacaz.

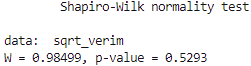


Grafikte normal dağılıyor gibi gözükse de shapiro normallik testi yapınca p-value 0.05’den küçük olduğu için H0 reddedilir. Modelimiz normal dağılıma uygun değildir.



**Hala normalik sağlamıyor o yüzden karekök dönüşümü uyguluyoruz**.





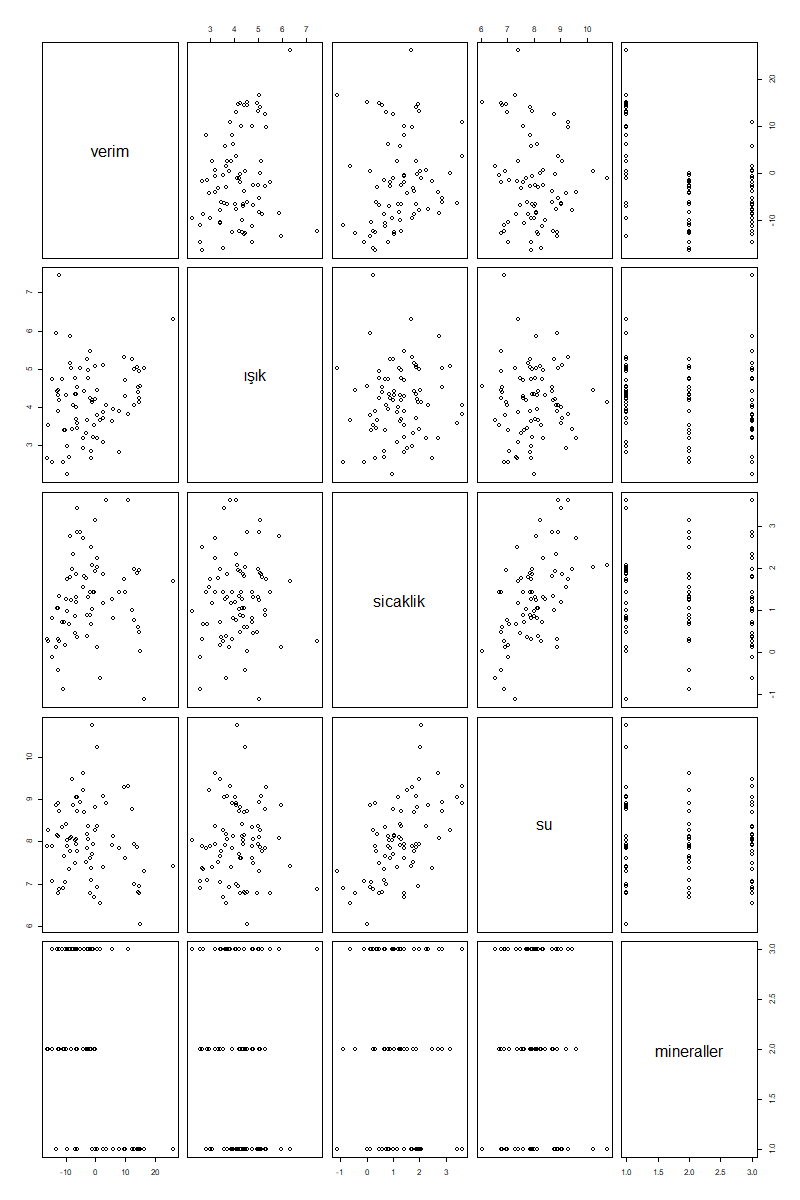
Hipotez testimizi kuruyoruz.

H0: Verilerimiz ile normal dağılım arasında fark yoktur.

H1: Verilerimiz ile normal dağılım arasında fark vardır.

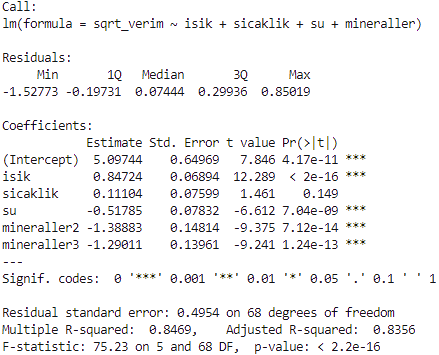
p değerimiz anlamlılık düzeyimiz olan %5’ten büyüktür. Bu durumda H0 hipotezimizi reddedemeyiz. Bağımlı değişkenimizin %95 güven düzeyinde anlamlı olduğunu söyleyebiliriz.

## Doğrusallık Varsayımı



Grafiklerde görüldüğü üzere bağımlı değişkenimiz ile bağımsız değişkenlerimiz arasında doğrusallık vardır. Doğrusal model kurulabilir. Ancak bağımsız değişkenlerimiz arasında da ilişki durumu söz konusudur.

# Doğrusallık Modelinin Kuruluşu



Model anlamlılığı hipotezimizi kuruyoruz.

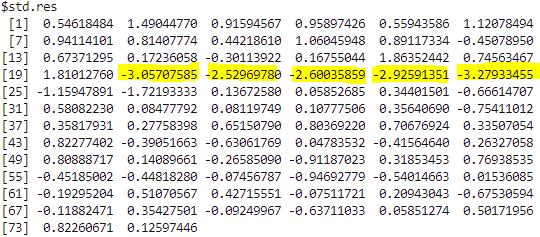
H0: β1=β2=β3=β4=β5=0 (Model anlamlı değildir.)

H1: β1,β2,β3,β4,β5’ten en az birisi sıfırdan farklıdır. (Model anlamlıdır.)

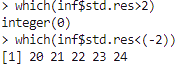
Modelimizdeki p değeri anlamlılık düzeyimiz olan %5’ten küçüktür. H0 hipotezimizi reddederiz. Modelimiz %95 güven düzeyinde anlamlıdır. Toprak\_turu

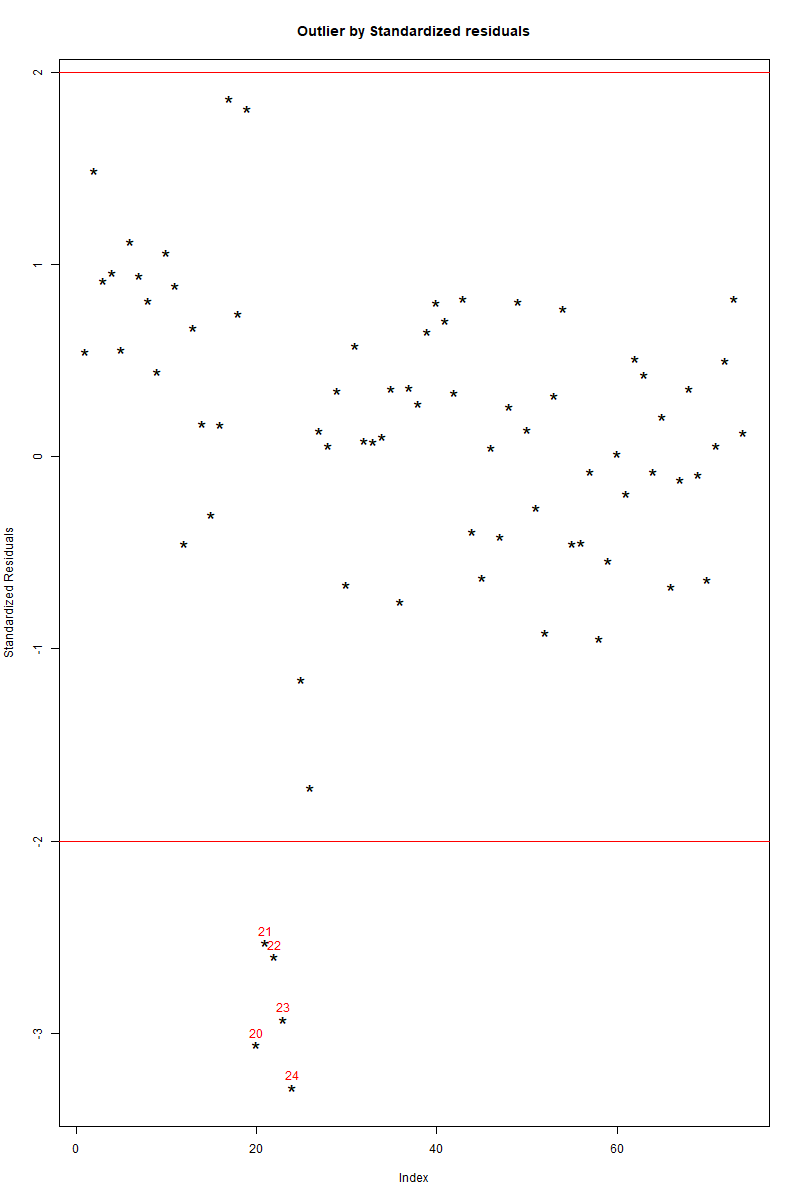
# Artıkların İncelenmesi

## Standart Türü Artıklar



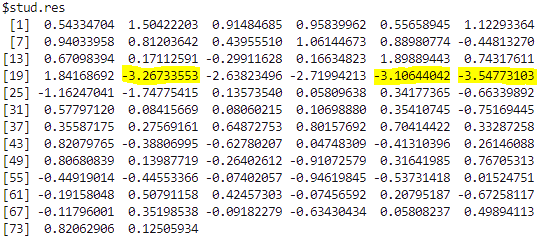
Standart Türü artıklara göre değerlerimiz (-2,2) aralığında olmalıdır. Aralık dışındaki değerler Standart Türü artıklar tablomuza göre artık değerlerdir.



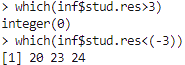


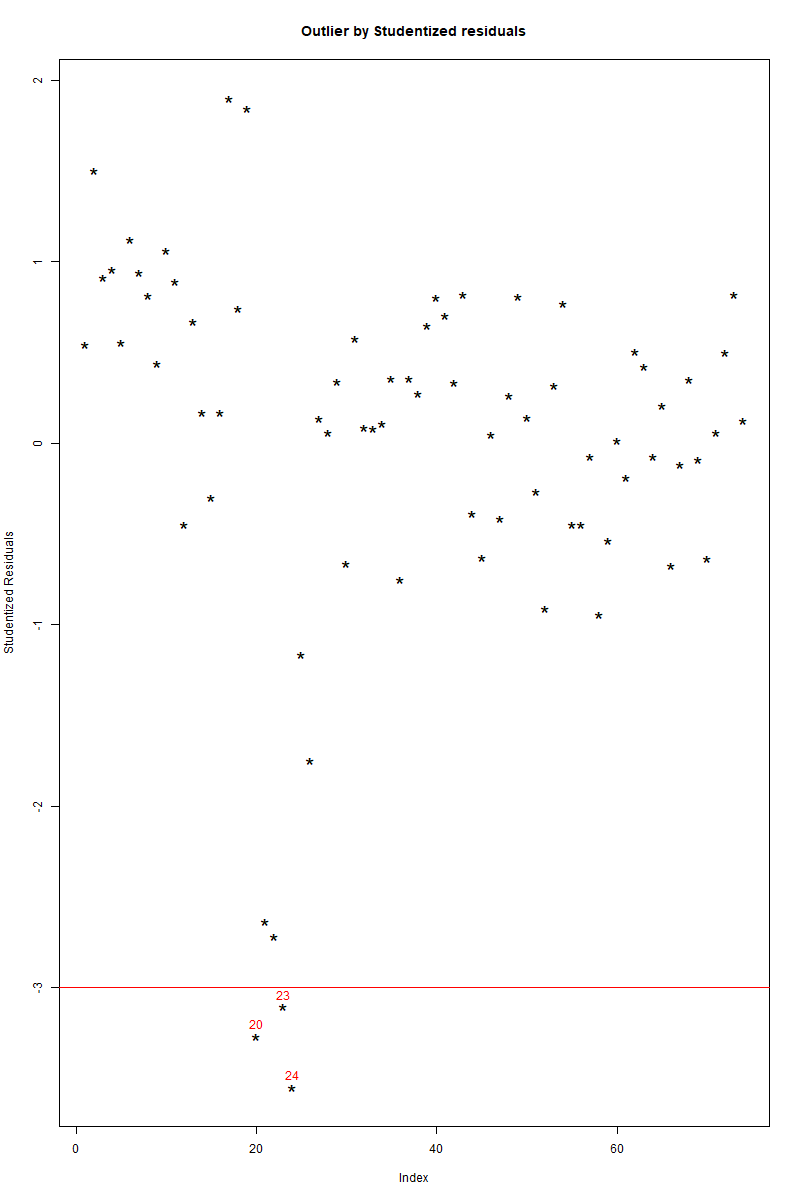
Grafikte ve kod üzerinde de görülüğü üzere 20, 21, 22, 23 ve 24. gözlemlerimiz Standart Türü artık incelmesine göre artık değerlerdir.

## Student Türü Artıklar



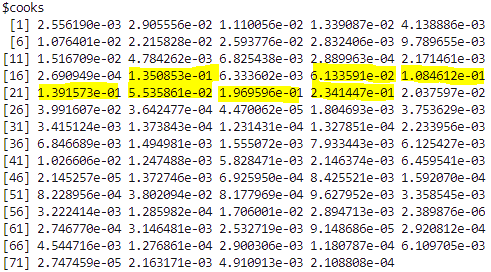
Student Türü artık incelemesine göre (-3,3) aralığı dışında kalan değerler aykırı değerlerdir.



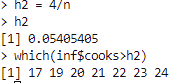


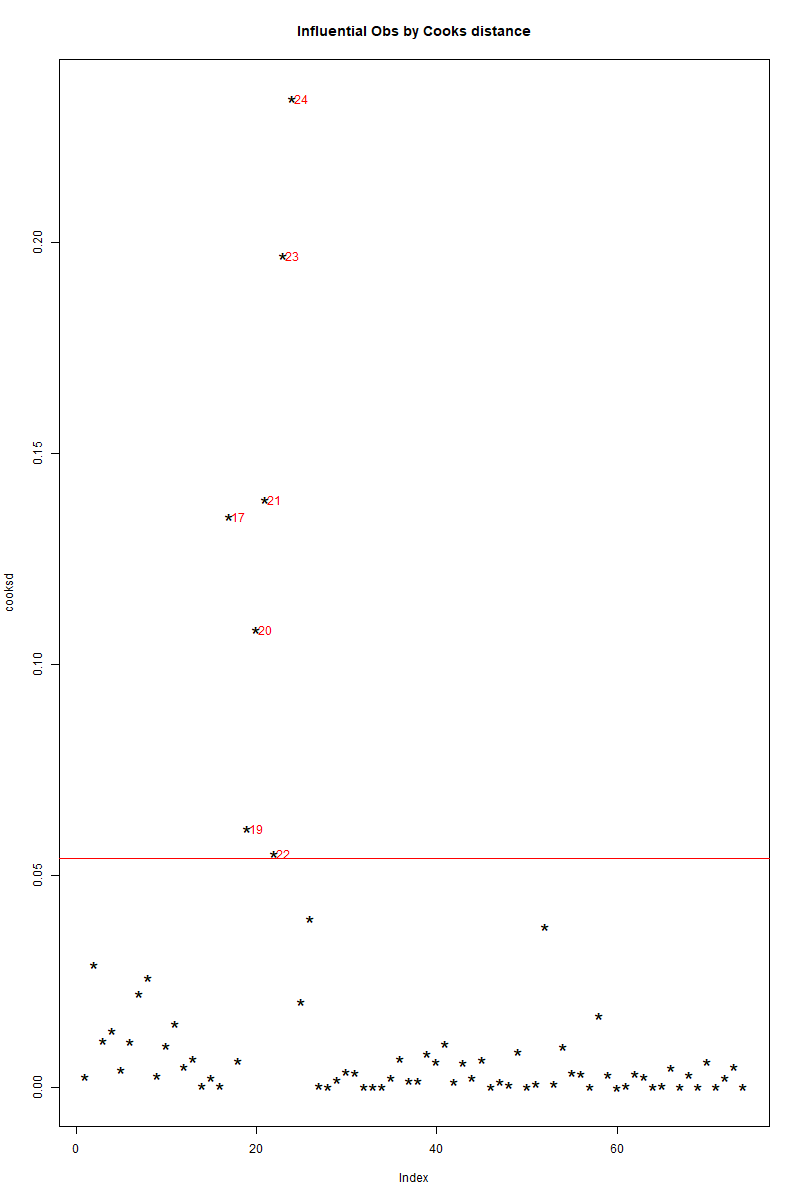
Grafikte ve kod üzerinde de görüldüğü üzere 20, 23 ve 24. Gözlemlerimiz Student Türü artık incelemesine göre artık değerlerdir.

## Cook Uzaklığı



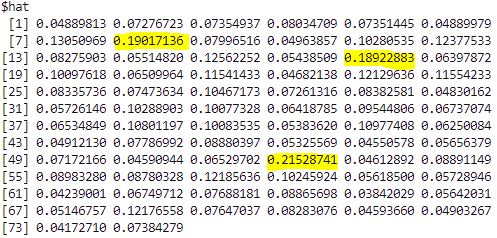
Cook Uzaklığı artık incelemesine göre n>50 olduğu için değerinden büyük gözlemlerimiz etkin değerdir.



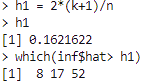


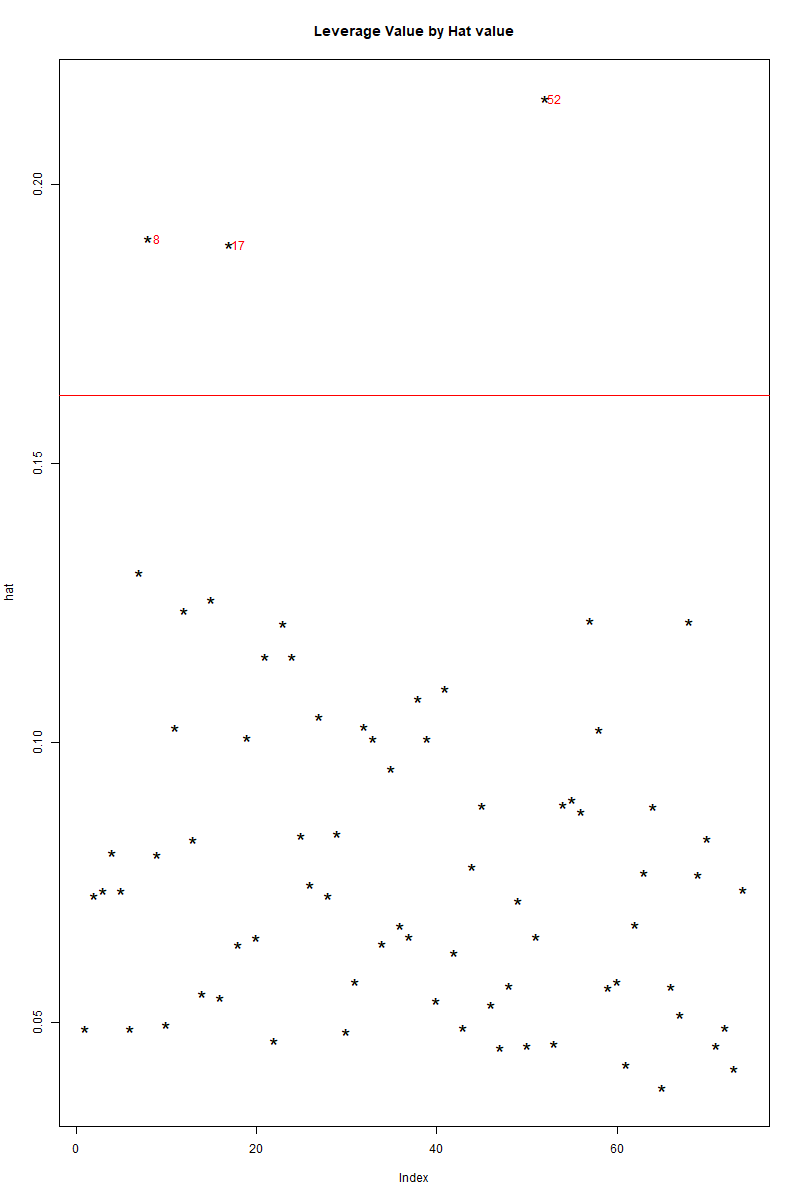
Kod üzerinde ve grafikte de görüldüğü üzere Cook Uzaklığı artık incelemesine göre 17, 19, 20, 21, 22, 23 ve 24. gözlemler artık değerdir.

## Hat Uzaklığı



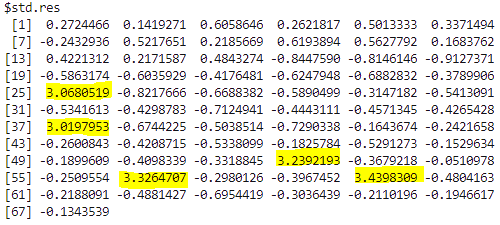
Hat Uzaklığı artık incelemesine göre değerinden büyük olan gözlemlerimiz uç değerlerdir.



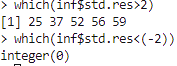


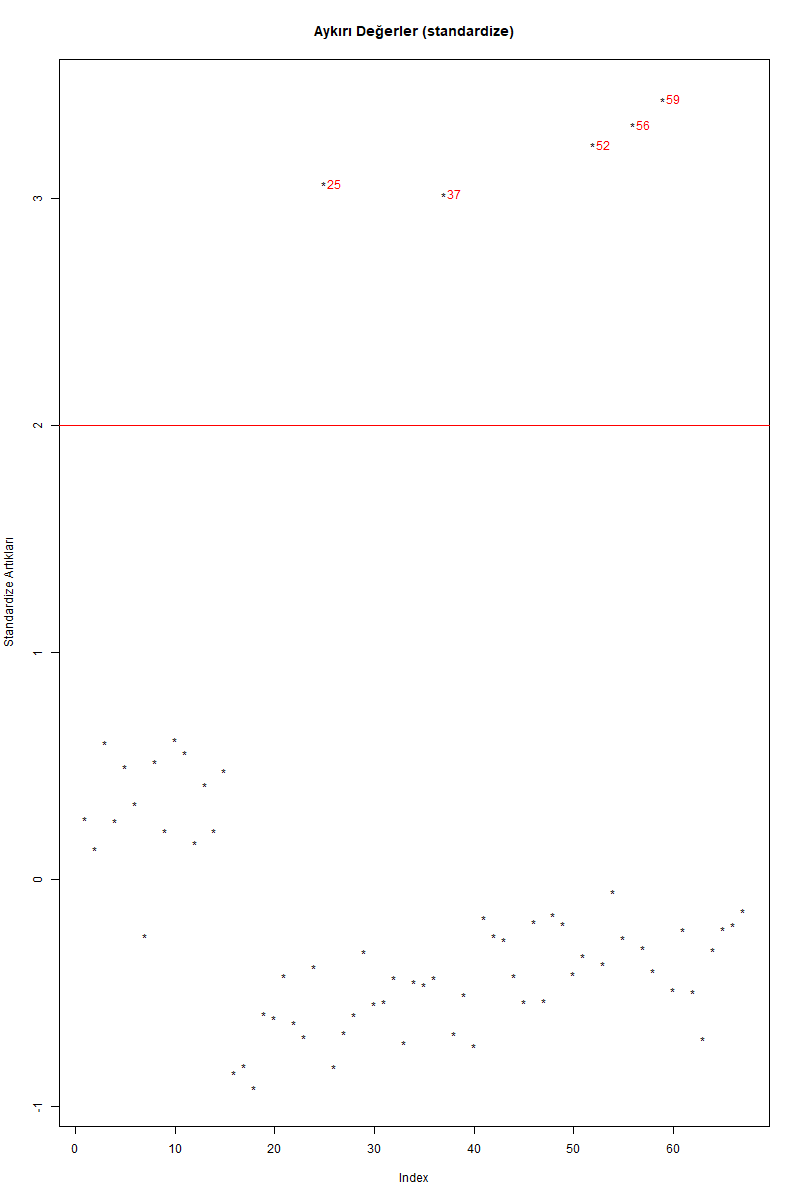
Kod üzerinde ve grafikte de görüldüğü üzere Hat Uzaklığı artık incelmesine göre 8, 17 ve 52. gözlem uç değerdir.

## Standart Türü Artıklar



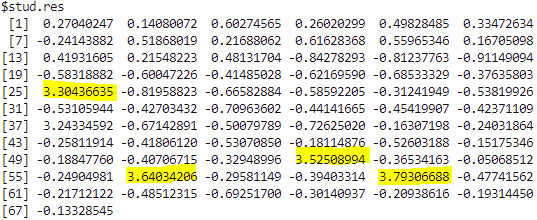
Standart artık incelemesine göre (-2,2) aralığı dışında kalan gözlemlerimiz artık değerdir.



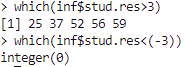


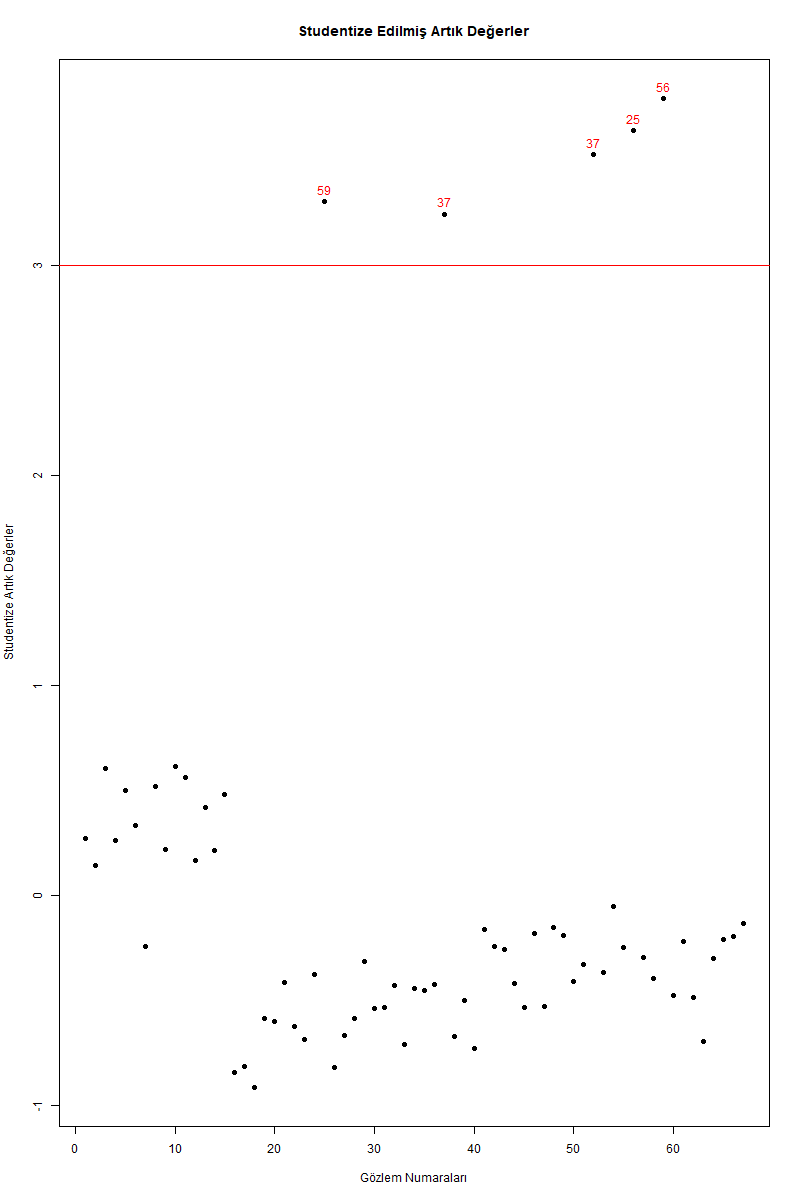
R kodu üzerinde ve grafikte de görüldüğü üzere 25, 37, 52, 56 ve 93. gözlemler Standart Türü artık incelemesine göre artık değerlerdir.

## Student Türü Artık İncelemesi



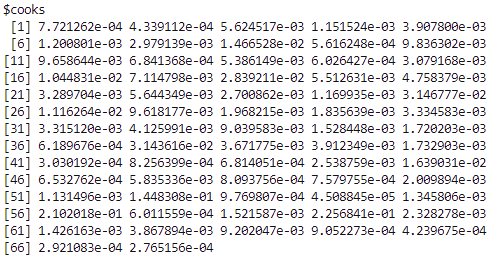
Student Türü artık incelemesine göre (-3,3) aralığı dışında kalan değerlerimiz artıktır.



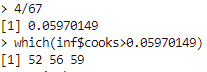


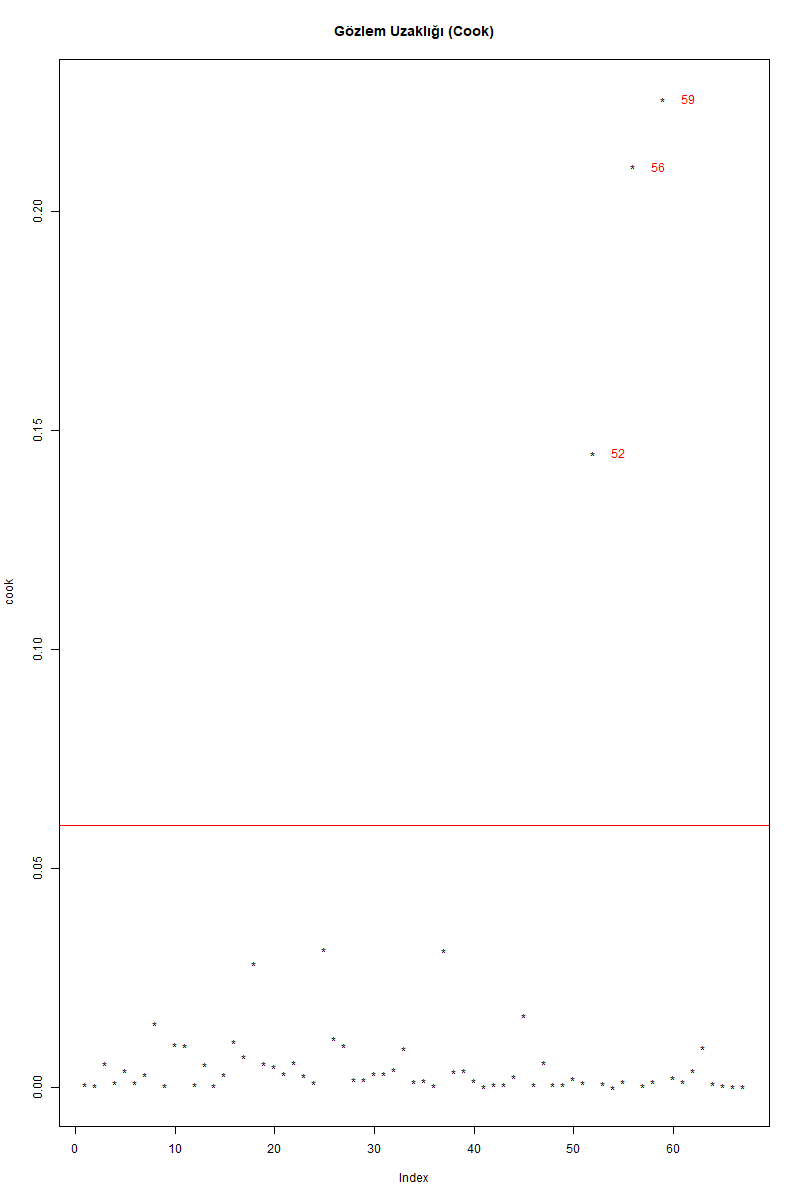
R kodu üzerinde ve grafikte de görüldüğü üzere 25, 37, 52, 56 ve 93. gözlemler Student Türü artık incelemesine göre artık değerlerdir.

## Cook Uzaklığı



Cook Uzaklığı artık incelemesine göre n>50 olduğu için değerinden büyük gözlemlerimiz etkin değerdir.





R kodu ve grafik üzerinde de görüldüğü üzere 52, 56 ve 59. gözlemlerimiz Cook Uzaklığı artık incelemesine göre artık değerlerdir.

Yeni verilerimiz üzerinde yaptığımız artık incelemesi sonucunda 19, 28, 38, 45, 51, 54, 62, 64, 84 ve 93. gözlemlerimiz atık değerlerimizdir.