

# Tek Değişkenli Zaman Serisi Analizi

Emircan ATALAY

## Contents

Veri ve Metodoloji . . . . .	1
Kütüphanelerin Yüklenmesi ve Verinin Aktarılması . . . . .	2
Zaman Serisi, ACF ve PACF Grafikleri . . . . .	2
Farklı Alınmış Seri . . . . .	5
Toplamsal Winters Yöntemi . . . . .	7
Box-Ljung Test . . . . .	9
Hata serisinin ACF grafiği . . . . .	9
Artıkların Zaman Serisi Grafiği ve Dağılım Grafiği . . . . .	10
Tahmin Değerleri . . . . .	11
Çarpımsal Winters Yöntemi . . . . .	13
Box-Ljung Test . . . . .	14
Hata serisinin ACF grafiği . . . . .	14
Artıkların Zaman Serisi Grafiği ve Dağılım Grafiği . . . . .	15
Tahmin Değerleri . . . . .	16
Mevsimsel ARIMA Modelleri (SARIMA) . . . . .	18
Anlamlı Bulunan Modeller . . . . .	24
ACF,PACF, Artıkların Dağılım Grafiği ve Box-Ljung Testi . . . . .	28
Tahmin Değerleri . . . . .	31
Kullanılan Modellerin Tahmin Değerleri ile Gerçek Değerlerin Karşılaştırılması . . . . .	32
Ortalama Mutlak Hata (MAE) . . . . .	34
Kök Ortalama Kare Hata (RMSE) . . . . .	34
Ortalama Yüzde Hata (MAPE) . . . . .	35
Sonuç . . . . .	36

## Veri ve Metodoloji

Bu çalışmada kullanılan veri, Türkiye İstatistik Kurumu (TÜİK) tarafından yayımlanan, 2015 (Haziran)-2023 (Aralık) yılları arasındaki tavuk yumurtası üretim miktarını (bin adet) kapsamaktadır. Veriler aylık olarak düzenlenmiş olup, toplamda 103 gözlemden oluşmaktadır. Veriler, trend, mevsimsellik ve rastgele bileşenlerden oluşan tipik bir zaman serisi yapısına sahiptir. Analiz sürecinde, öncelikle zaman serisinin genel görünümünü anlamak amacıyla bir zaman serisi grafiği oluşturulmuştur. Daha sonra, serinin otokorelasyon ve kısmi otokorelasyon fonksiyonları (ACF ve PACF) incelenerek serinin durağanlık durumu ve otoregresif yapısı analiz edilmiştir. Serinin durağan olmaması durumunda fark alma yöntemi uygulanarak durağanlık sağlanmıştır. Ayrıştırma modelleri, Regresyon modelleri(üstel regresyon, kübik regresyon, karesel regresyon, lojistik regresyon modelleri), Üstel düzleştirme yöntemleri(Winters yöntemleri), otoregresif modeller (AR, MA, ARMA, ARIMA) kullanılarak serinin yapısına en uygun model belirlenmiştir. Model seçimi sırasında Akaike Bilgi Kriteri (AIC) ve Bayes Bilgi Kriteri (BIC) , Ortalama Mutlak Hata (MAE), Ortalama Yüzde Hata (MAPE) gibi değerlendirme kriterleri göz önünde bulundurulmuştur.

## Kütüphanelerin Yüklenmesi ve Verinin Aktarılması

```
library(fpp)
library(forecast)
library(haven)
library(stats)
library(ggplot2)
```

```
veri = read_sav("veri.sav")
head(veri)
```

```
## # A tibble: 6 x 1
##   degerler
##   <dbl>
## 1  1198324
## 2  1286892
## 3  1294923
## 4  1340401
## 5  1417880
## 6  1408095
```

```
summary(veri)
```

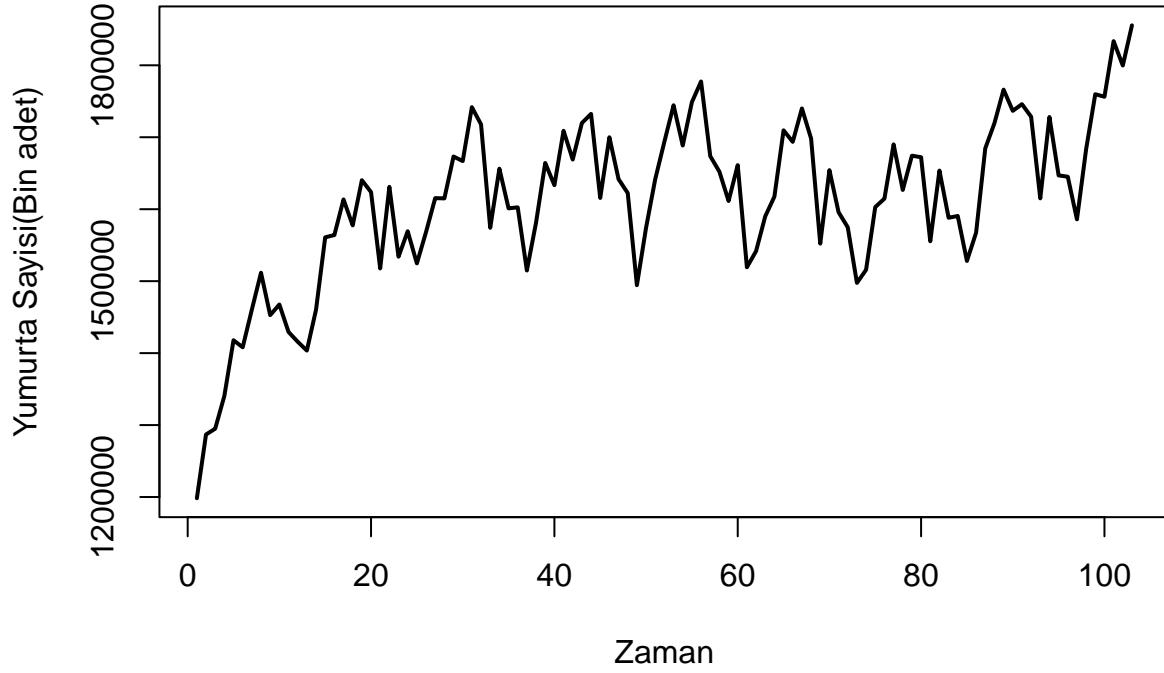
```
##   degerler
## Min.   :1198324
## 1st Qu.:1558174
## Median :1617495
## Mean   :1610287
## 3rd Qu.:1689308
## Max.   :1855505
```

## Zaman Serisi, ACF ve PACF Grafikleri

```
veri_ts = ts(veri)
```

```
ts.plot(veri_ts, gpars=list(xlab="Zaman", ylab="Yumurta Sayısı(Bin adet)", main= "Zaman Serisi Grafiği", l
```

## Zaman Serisi Grafiği

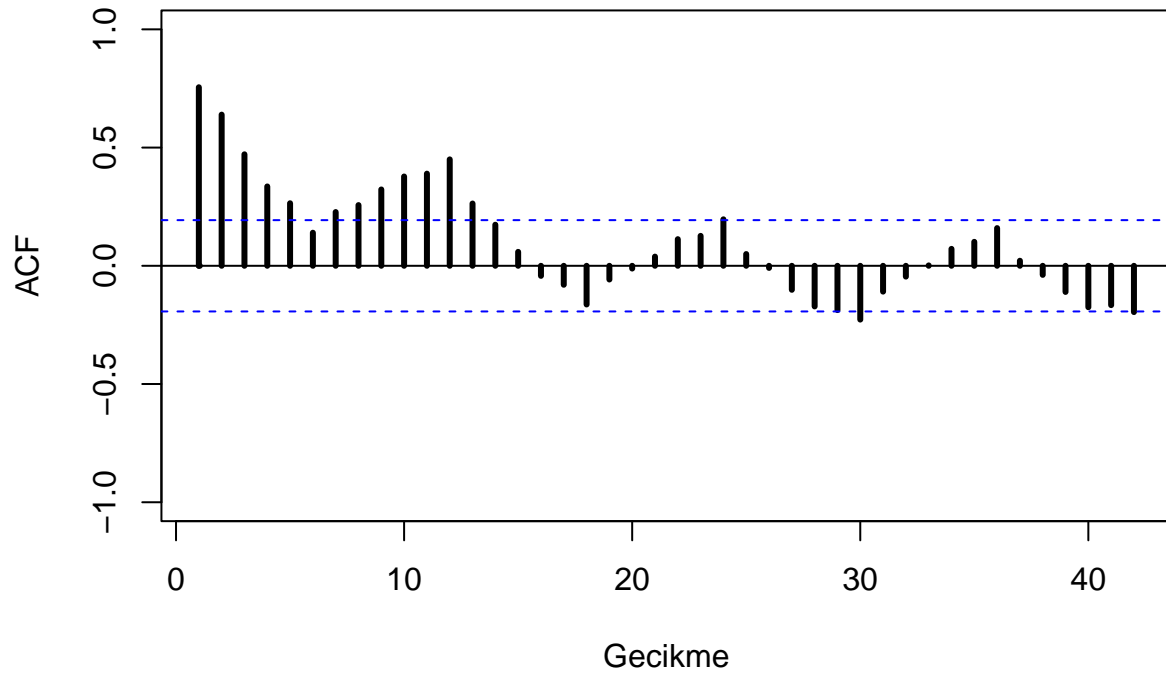


Zaman serisi grafiği incelendiğinde, genel olarak bir artış eğiliminin olduğu gözlemlenmektedir. Ayrıca, belirli aralıklarla artış ve azalışların meydana gelmesi, serinin hem mevsimsel dalgalanmalara hem de uzun vadeli bir trende sahip olabileceğini düşündürmektedir.

Serinin geçmiş değerleri ile olan ilişkilerini daha net bir şekilde gözlemleyebilmek ve mevsimsel ile trend bileşenleri hakkında daha güçlü çıkarımlar yapabilmek adına ACF (Autocorrelation Function) ve PACF (Partial Autocorrelation Function) grafikleri incelenecektir.

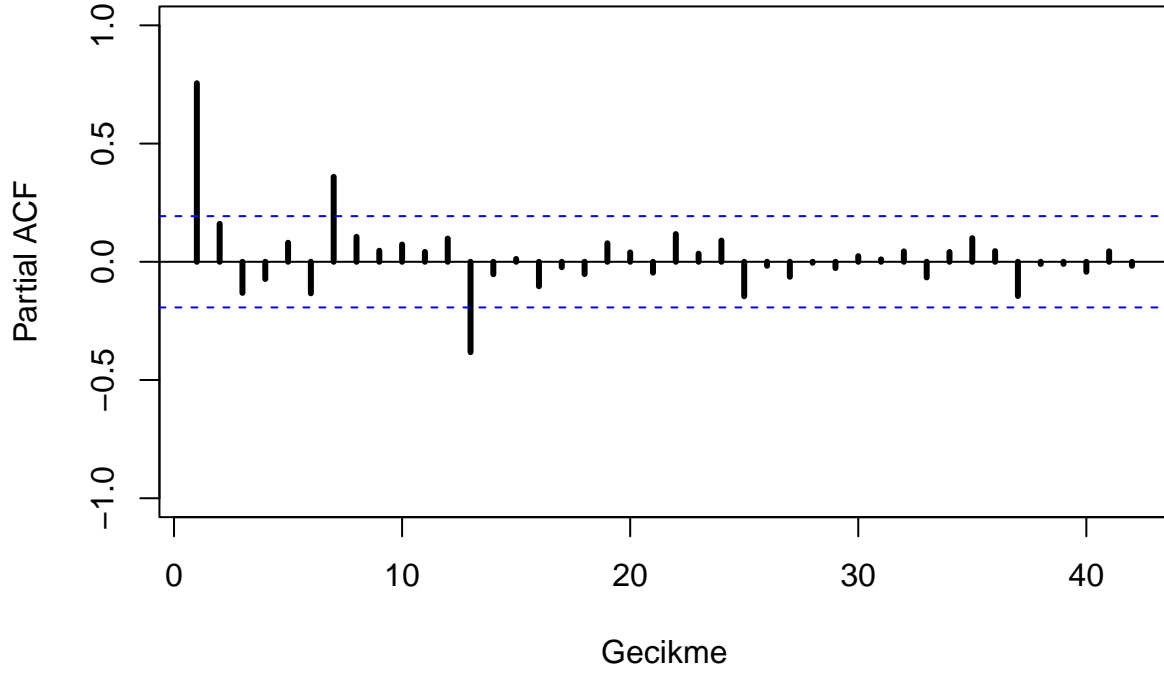
```
Acf(veri_ts,lag.max = 42, ylim=c(-1,1), lwd=3, main = "ACF Grafiği", xlab="Gecikme")
```

## ACF Grafiği



```
Pacf(veri_ts, lag.max = 42, ylim=c(-1,1), lwd=3, main = "PACF Grafiği", xlab="Gecikme")
```

## PACF Grafiği



ACF grafiği incelendiğinde, ilk dört gecikme değerinin sınır dışında kaldığı gözlemlenmektedir. Bu durum, zaman serisinde bir trendin mevcut olduğuna dair güçlü bir gösterge sunmaktadır.

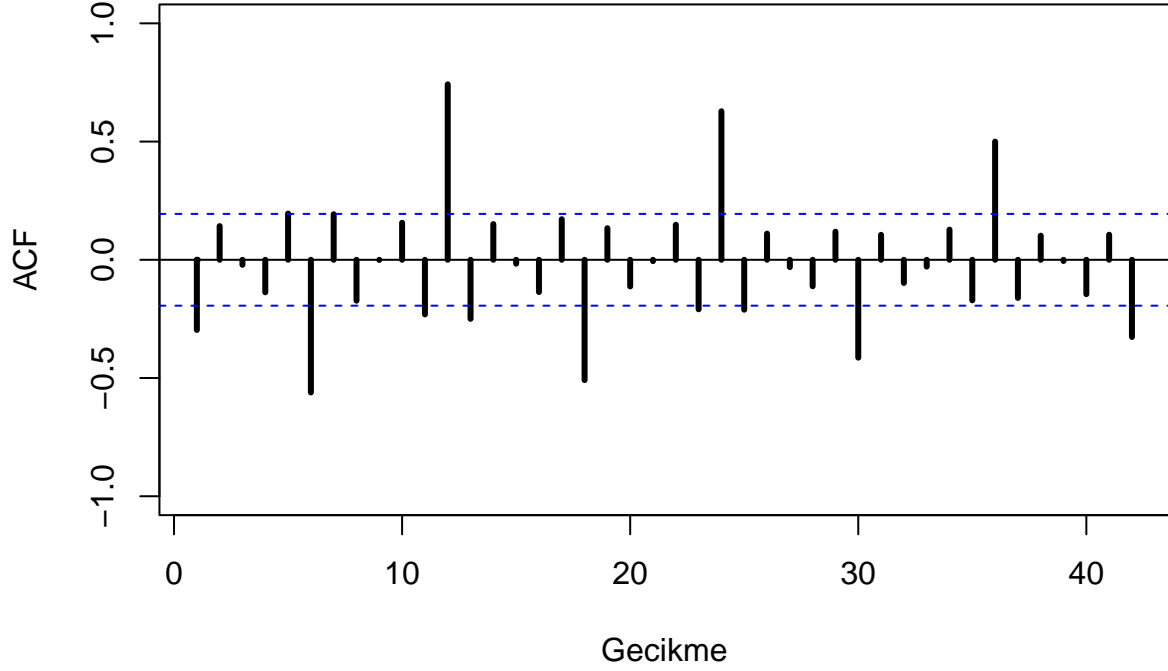
Trend mevcut olduğu için seriyi durağan hale getirmek adına fark alma yöntemi kullanılacaktır ve fark alınmış serinin ACF grafiği tekrar incelenecektir.

### Farkı Alınmış Seri

```
veri_trend = diff(veri_ts)
```

```
Acf(veri_trend, lag.max = 42, ylim=c(-1,1), lwd=3, main = "ACF Grafiği", xlab="Gecikme")
```

## ACF Grafiği



Birinci derece farkı alınmış serinin ACF grafiği yukarıda verilmiştir. Grafikte, ilk dört gecikme değerinin sınırlar içerisinde kaldığı gözlemlenmektedir, bu da serinin trendden arındırıldığını göstermektedir. Ancak, belirli aralıklarla bazı gecikme değerlerinin sınırların dışına çıktığı da dikkat çekmektedir. Bu durum, serinin mevsimsel bileşenler içerdiğini ortaya koymaktadır.

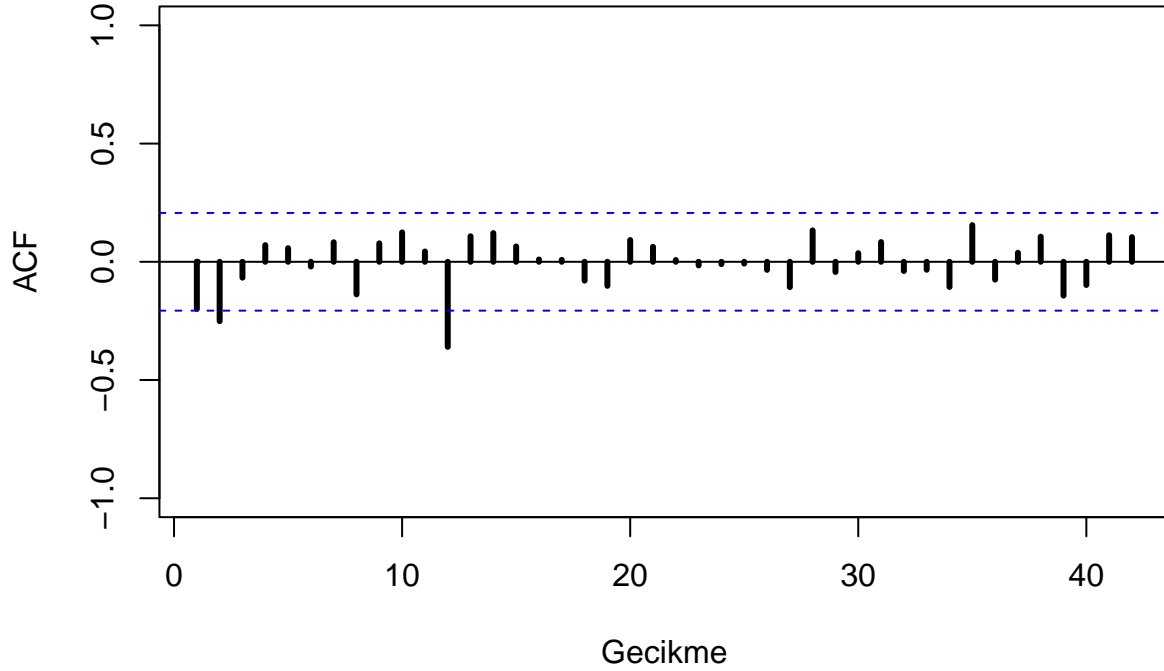
Belirli aralıklarla sınırların dışına çıkan gecikme değerleri göz önünde bulundurularak, periyot belirlenip mevsimsel fark alınarak seri mevsimsellikten arındırılacak ve durağan hale getirilecektir.

Periyot belirlenirken, sırasıyla sınır dışına çıkan en büyük gecikme değeri ile ikinci en büyük gecikme değeri arasındaki gecikmelerin sayısı göz önünde bulundurulmuş ve periyot 12 olarak belirlenmiştir.

```
veri_mevsimsel = diff(veri_trend, lag = 12)
```

```
Acf(veri_mevsimsel, lag.max = 42, ylim=c(-1,1), lwd=3, main = "ACF Grafiği", xlab="Gecikme")
```

## ACF Grafiği



Farklı alınmış serinin ACF grafiği yukarıda verilmiştir. Grafiğe bakıldığında, serinin trend ve mevsimsellikten arındırıldığı ve böylece durağan hale geldiği gözlemlenmektedir.

Serinin durağan hale getirilmesinin ardından, uygun bir model seçebilmek için çeşitli yöntemler uygulanacaktır. Bu süreçte farklı modelleme yöntemleri, serinin yapısına en iyi uyum sağlayacak modelin belirlenmesi amacıyla değerlendirilecektir. Analiz sonuçları, model seçiminin doğruluğunu destekleyen kriterler ve görsellerle detaylı şekilde sunulacaktır.

## Toplamsal Winters Yöntemi

```
veri_ts2 = ts(veri, start = c(2015,6), frequency = 12)
```

```
Winters1 = ets(veri_ts2, model = "AAA")
```

```
summary(Winters1)
```

```
## ETS(A,Ad,A)
##
## Call:
## ets(y = veri_ts2, model = "AAA")
##
## Smoothing parameters:
##   alpha = 0.7359
##   beta  = 1e-04
##   gamma = 1e-04
##   phi   = 0.9672
##
```

```
## Initial states:
## l = 1378314.9342
## b = 13232.1117
## s = -38687.45 -38690.9 27470.19 -41285.48 76430.04 83399.3
##      36746.94 74711.9 13537.23 2058.913 -72053.65 -123637.1
##
## sigma: 30116.37
##
##      AIC      AICc      BIC
## 2619.240 2627.382 2666.665
##
## Training set error measures:
##      ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 182.5407 27519.04 20413.43 -0.02101148 1.297803 0.3224222
##      ACF1
## Training set 0.04697417
```

Kullanılan veri kümesine gerekli düzeltmeler yapıldıktan sonra Toplamsal Winters yöntemi uygulandığında testin özet sonuçları yukarıdaki gibidir.

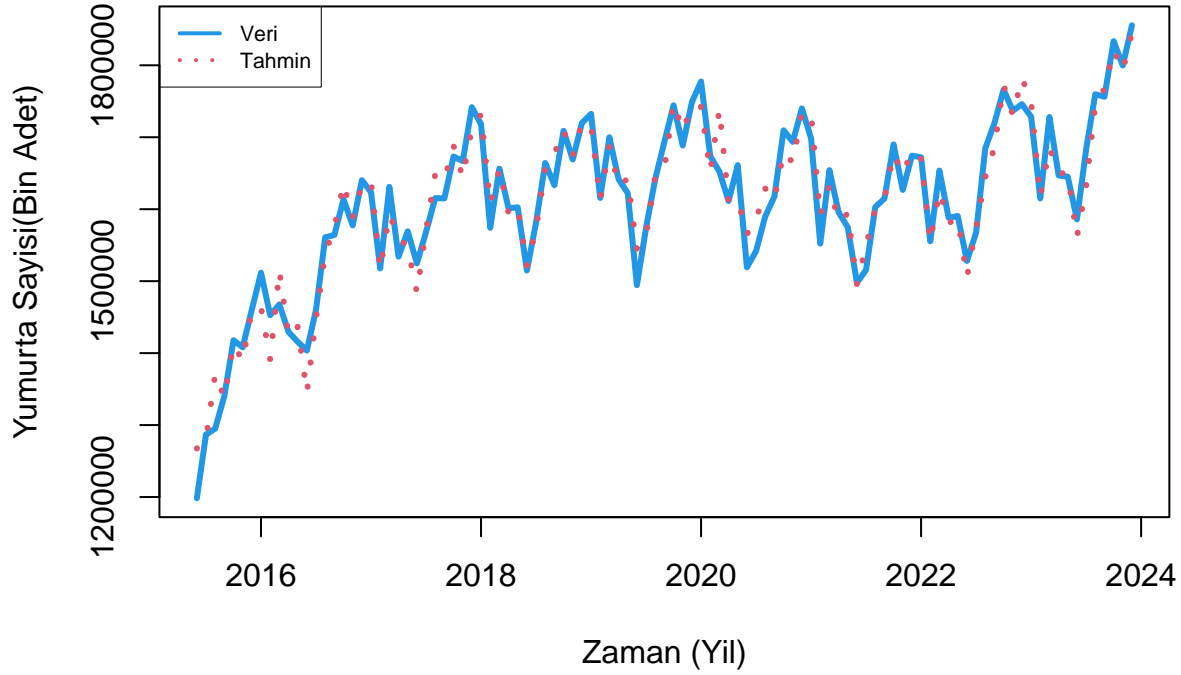
Modelin uygunluğunu kontrol etmek için ilk olarak orijinal seri ile tahmin serisinin uyumu incelenecektir.

```
tahmint = Winters1[["fitted"]]
```

```
plot(window(veri_ts2), main = "Zaman Serisi Uyum Grafiği",
      xlab="Zaman (Yıl)", ylab="Yumurta Sayısı(Bin Adet)",lty=1, col=4, lwd=3,ylim=c(min(veri_ts2), max(
      cex.main = 1)
lines( window(tahmint) ,lty=3,col=2,lwd=3)
legend("topleft",c(expression(paste(Veri)),
                    expression(paste(Tahmin))),
      lwd=c(2,2),lty=c(1,3), cex=0.7, col=c(4,2),box.lwd = 0.5)
```



## Zaman Serisi Uyum Grafiği



Grafik incelendiğinde, orijinal seri ile tahmin serisinin yeteri kadar uyum içerisinde olduğu söylenebilir.

Artıkların, model varsayımlarına uygunluğunu incelerken genellikle kullanılan üç temel grafik bulunmaktadır. ACF Grafiği, Artıkların Zaman Grafiği ve Artıkların Dağılım Grafiği. Bu grafikler incelenecek ve Box-Ljung testi uygulanacaktır.

```
hata = Winters1[["residuals"]]
```

### Box-Ljung Test

$H_0$  : Gecikmeler arasında ilişki yoktur.

$H_1$  : Gecikmeler arasında ilişki vardır.

```
Box.test(hata, lag = 42, type = "Ljung")
```

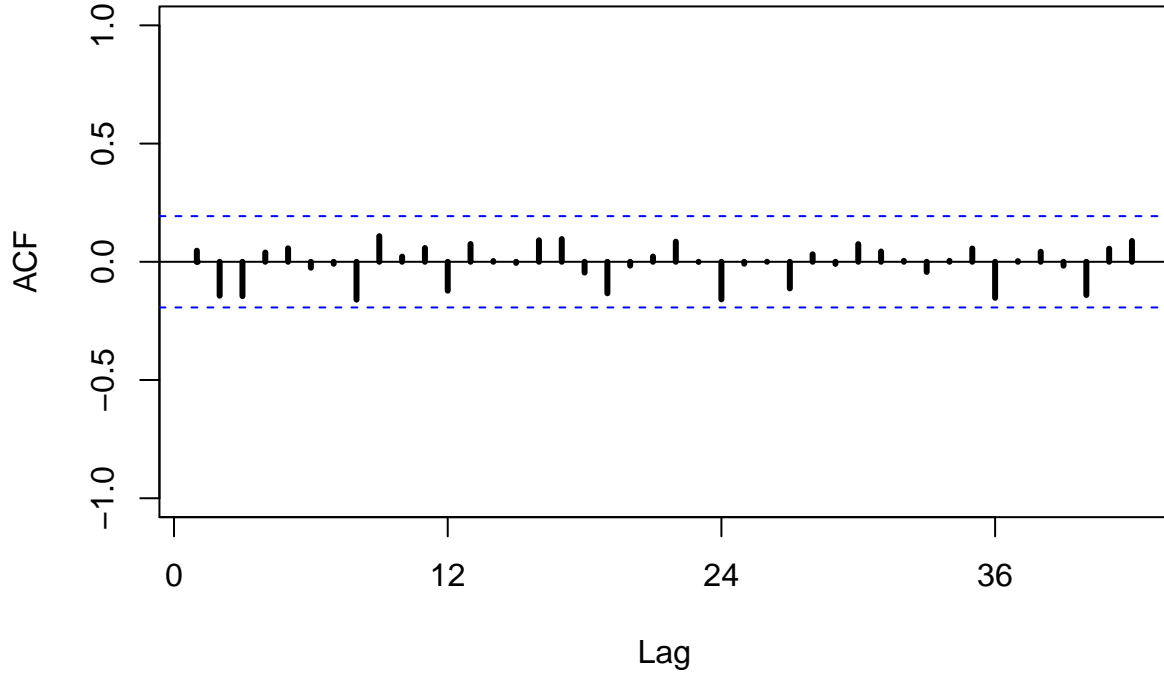
```
##  
## Box-Ljung test  
##  
## data: hata  
## X-squared = 35.025, df = 42, p-value = 0.7685
```

Test sonucunda, p-value > 0.05 olduğundan dolayı  $H_0$  **REDDEDİLEMEZ**. Yani gecikmeler arasında ilişki olmadığı söylenebilir.

### Hata serisinin ACF grafiği

```
Acf(hata, main="Hata Serisi", lag.max = 42, ylim=c(-1,1), lwd=3)
```

## Hata Serisi

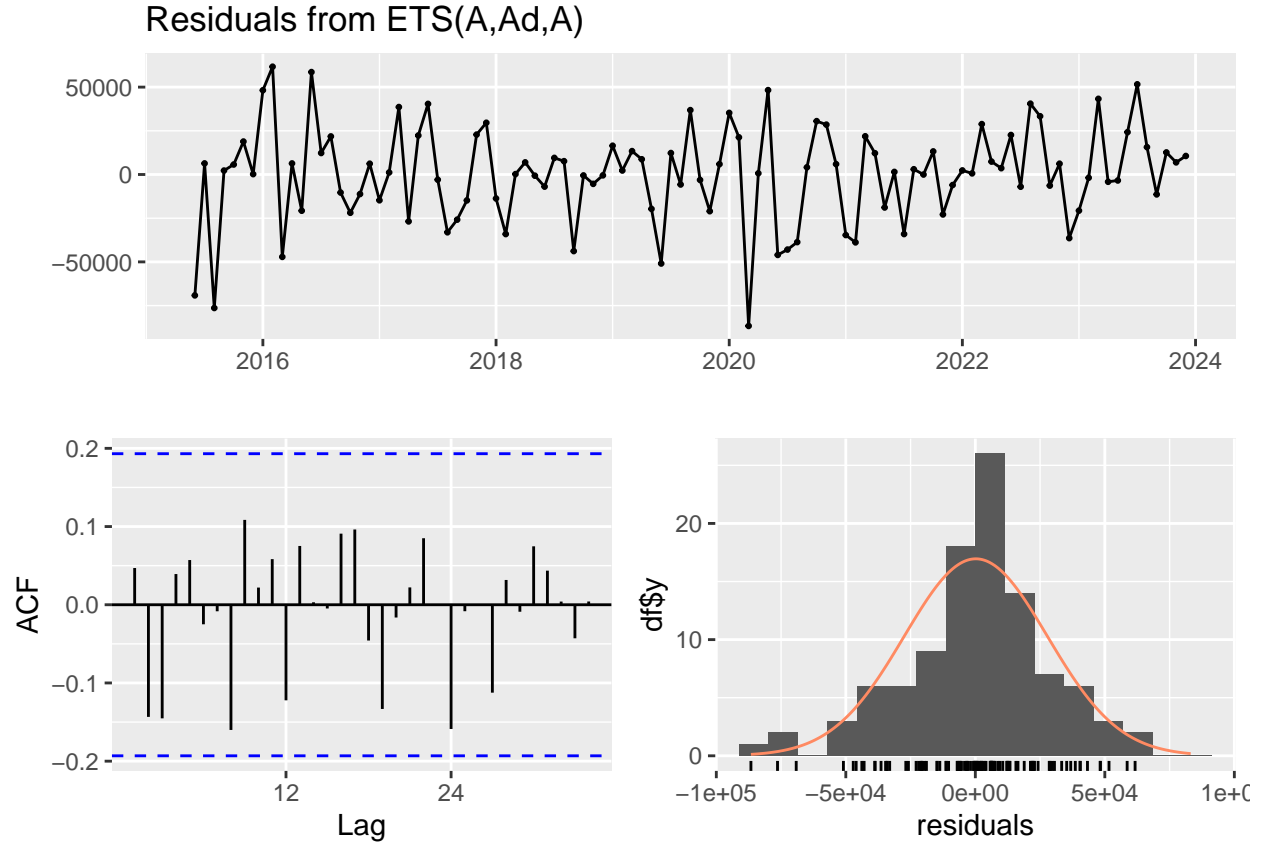


Grafikte görüldüğü üzere, gecikmeler sınırların içerisinde olduğu için hata serisi, akgürültü serisidir.

Box-Ljung testi sonucunda da gecikmeler arasında ilişki olmadığını söylemiştik. İncelediğimiz ACF grafiğinde de hata serisinin akgürültü serisi olduğu anlaşılmıştır.

### Artıkların Zaman Serisi Grafiği ve Dağılım Grafiği

```
checkresiduals(Winters1, lag = 42)
```



```
##
##  Ljung-Box test
##
## data:  Residuals from ETS(A,Ad,A)
## Q* = 35.025, df = 42, p-value = 0.7685
##
## Model df: 0.   Total lags used: 42
```

Residuals from ETS grafiği incelendiğinde, artıkların belirgin bir trend, mevsimsellik veya yapı göstermediği bu sebeple rastgele bir şekilde dağılmış olduğu söylenebilir. Bu istenilen bir durumdur.

Artıkların dağılım grafiğinde, artıkların dağılımının normal dağılıma uyumlu olması beklenir. Yukarıdaki grafik incelendiğinde, artıkların dağılımının ,neredeyse, normal dağılıma uyumlu olduğu söylenebilir.

Tüm bu gerekçeler doğrultusunda, modelin uygun bir yapıya sahip olduğu söylenebilir. Bu uygunluk temelinde, model kullanılarak sonraki 12 ay için tahminler gerçekleştirilecektir.

### Tahmin Değerleri

```
ongoru = forecast(Winters1,h=12)

ongoru[["mean"]]
```

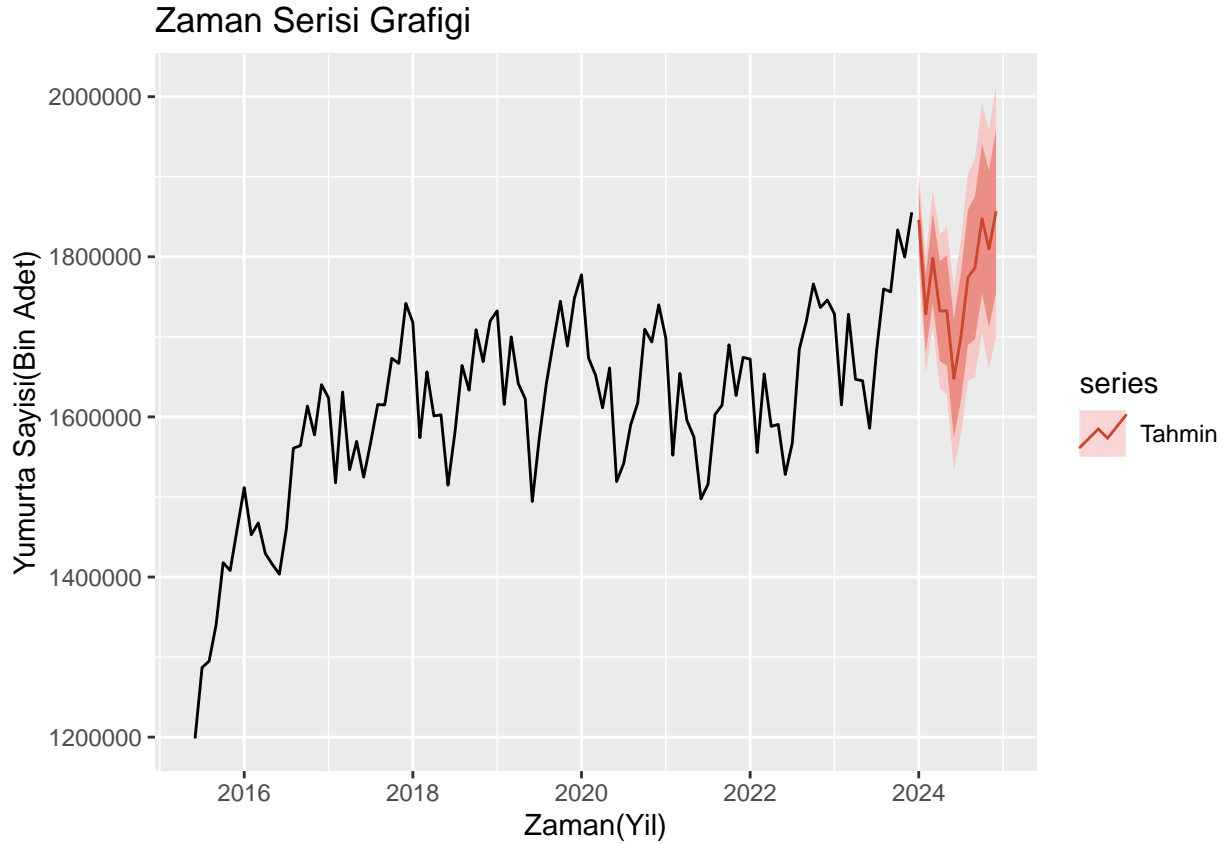
```
##          Jan      Feb      Mar      Apr      May      Jun      Jul      Aug      Sep
## 2024 1846160 1728852 1798002 1732222 1732595 1647998 1699930 1774369 1786176
##          Oct      Nov      Dec
## 2024 1847666 1810003 1856947
```

ongoru

##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## Jan 2024	1846160	1807565	1884756	1787133	1905187
## Feb 2024	1728852	1680929	1776775	1655560	1802144
## Mar 2024	1798002	1742290	1853715	1712797	1883207
## Apr 2024	1732222	1669682	1794763	1636575	1827870
## May 2024	1732595	1663901	1801289	1627536	1837654
## Jun 2024	1647998	1573657	1722340	1534302	1761694
## Jul 2024	1699930	1620339	1779521	1578207	1821654
## Aug 2024	1774369	1689854	1858885	1645114	1903624
## Sep 2024	1786176	1697006	1875345	1649803	1922548
## Oct 2024	1847666	1754073	1941259	1704527	1990804
## Nov 2024	1810003	1712186	1907821	1660404	1959603
## Dec 2024	1856947	1755079	1958815	1701154	2012741

Model kullanılarak yapılan tahminlerin %80 güven düzeyindeki alt sınırı (Lo80), üst sınırı (Hi80), %95 güven düzeyindeki alt sınırı (Lo95), üst sınırı (Hi95) ve nokta tahmin değerleri yukarıda sunulmuştur.

```
autoplot(veri_ts2) +  
  autolayer(ongoru, series = "Tahmin", PI=TRUE) +  
  ggtitle("Zaman Serisi Grafiği") +  
  xlab("Zaman(Yıl)") +  
  ylab("Yumurta Sayısı(Bin Adet)")
```



## Çarpımsal Winters Yöntemi

```
Winters2 = ets(abs(veri_ts2), model = "MAM")
```

```
summary(Winters2)
```

```
## ETS(M,Ad,M)
##
## Call:
## ets(y = abs(veri_ts2), model = "MAM")
##
## Smoothing parameters:
##   alpha = 0.692
##   beta  = 2e-04
##   gamma = 1e-04
##   phi   = 0.9778
##
## Initial states:
##   l = 1362314.9923
##   b = 13415.5565
##   s = 0.9771 0.9761 1.0169 0.9743 1.0466 1.0506
##         1.0223 1.0457 1.0075 1.0003 0.9587 0.9239
##
## sigma: 0.0189
##
##      AIC      AICc      BIC
## 2621.452 2629.595 2668.878
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -1764.553 26933.01 20425.78 -0.1387963 1.297466 0.3226173
##              ACF1
## Training set 0.07260807
```

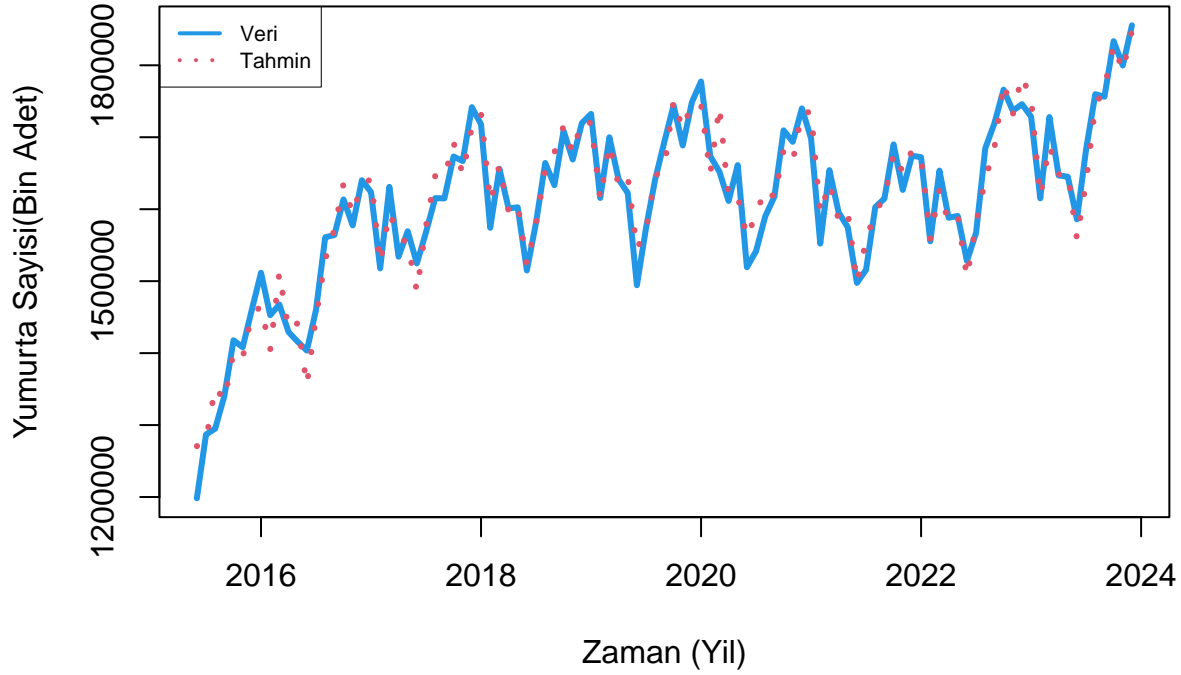
Kullanılan veri kümesine gerekli düzeltmeler yapıldıktan sonra **Çarpımsal Winters** yöntemi uygulandığında testin özet sonuçları yukarıdaki gibidir.

Modelin uygunluğunu kontrol etmek için ilk olarak orijinal seri ile tahmin serisinin uyumu incelenecektir.

```
tahmin1 = Winters2[["fitted"]]
```

```
plot(window(veri_ts2), main = "Zaman Serisi Uyum Grafiği",
      xlab="Zaman (Yıl)", ylab="Yumurta Sayısı(Bin Adet)",lty=1, col=4, lwd=3,ylim=c(min(veri_ts2), max(veri_ts2)),
      cex.main = 1)
lines( window(tahmin1) ,lty=3,col=2,lwd=3)
legend("topleft",c(expression(paste(Veri)),
                    expression(paste(Tahmin))),
      lwd=c(2,2),lty=c(1,3), cex=0.7, col=c(4,2),box.lwd = 0.5)
```

## Zaman Serisi Uyum Grafiği



Grafik incelendiğinde, orijinal seri ile tahmin serisinin yeteri kadar uyum içerisinde olduğu söylenebilir.

Artıkların, model varsayımlarına uygunluğunu incelerken genellikle kullanılan üç temel grafik bulunmaktadır. ACF Grafiği, Artıkların Zaman Grafiği ve Artıkların Dağılım Grafiği. Bu grafikler incelenecek ve Box-Ljung testi uygulanacaktır.

```
hata1<- Winters2[["residuals"]]
```

### Box-Ljung Test

$H_0$  : Gecikmeler arasında ilişki yoktur.

$H_1$  : Gecikmeler arasında ilişki vardır.

```
Box.test(hata1, lag = 42, type = "Ljung")
```

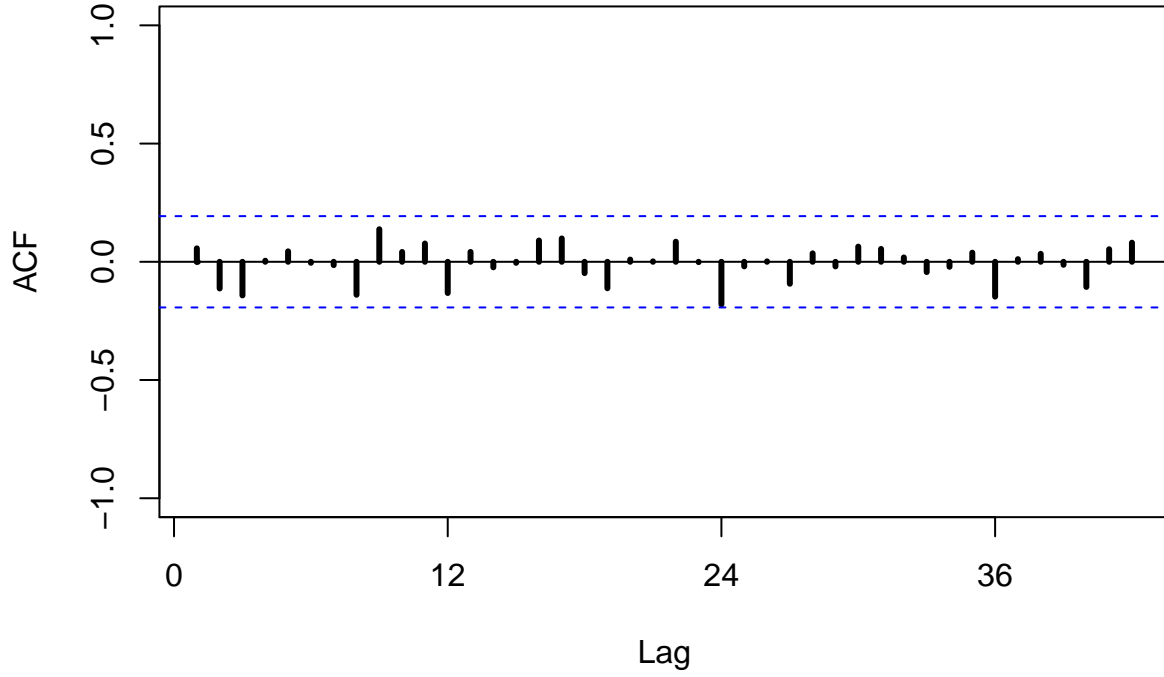
```
##  
## Box-Ljung test  
##  
## data: hata1  
## X-squared = 31.927, df = 42, p-value = 0.8702
```

Test sonucunda, p-value > 0.05 olduğundan dolayı  $H_0$  **REDDEDİLEMEZ**. Yani gecikmeler arasında ilişki olmadığı söylenebilir.

### Hata serisinin ACF grafiği

```
Acf(hata1,main="Hata Serisi", lag.max = 42, ylim=c(-1,1), lwd=3)
```

## Hata Serisi

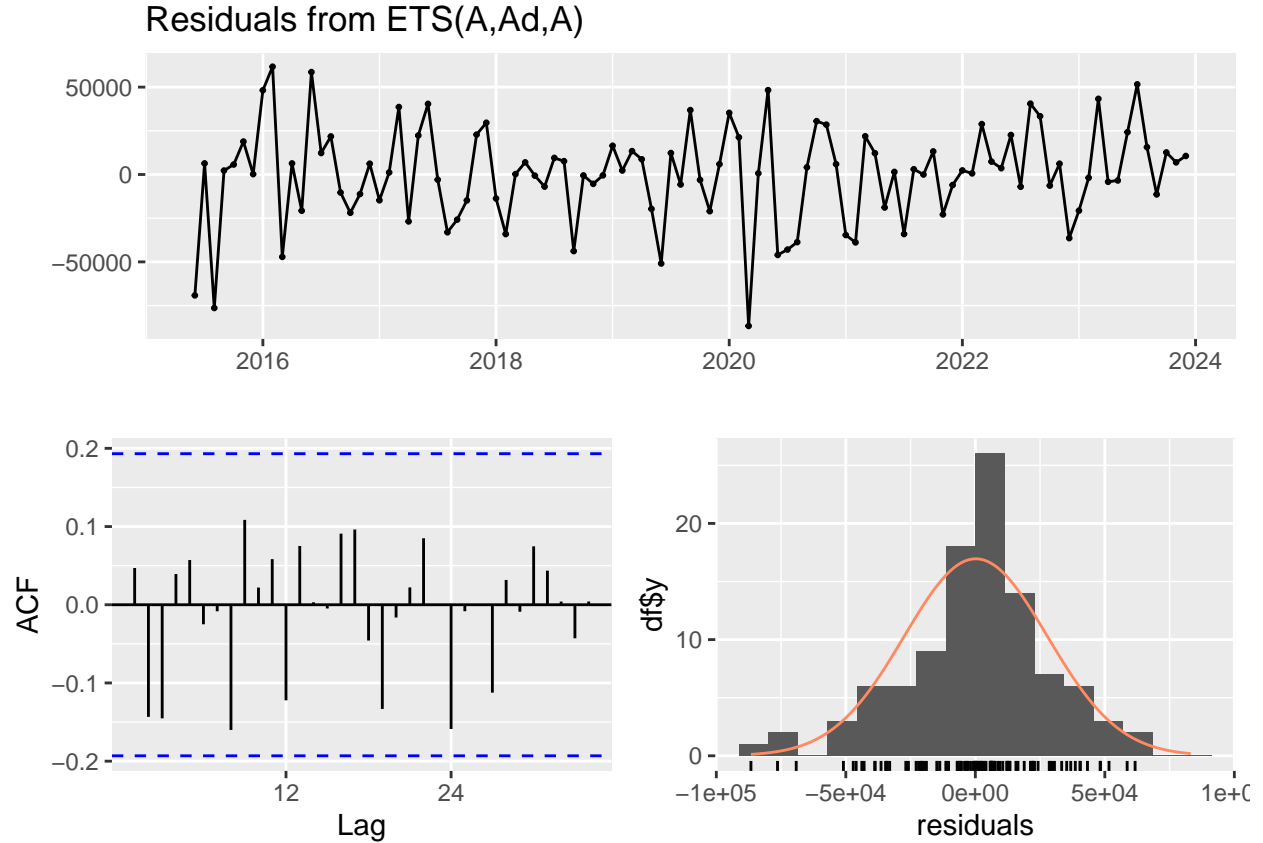


Grafikte görüldüğü üzere, gecikmeler sınırların içerisinde olduğu için hata serisi, akgürültü serisidir.

Box-Ljung testi sonucunda da gecikmeler arasında ilişki olmadığını söylemiştik. İncelediğimiz ACF grafiğinde de hata serisinin akgürültü serisi olduğu anlaşılmıştır.

### Artıkların Zaman Serisi Grafiği ve Dağılım Grafiği

```
checkresiduals(Winters1, lag = 42)
```



```
##
##  Ljung-Box test
##
## data:  Residuals from ETS(A,Ad,A)
## Q* = 35.025, df = 42, p-value = 0.7685
##
## Model df: 0.   Total lags used: 42
```

Residuals from ETS grafiđi incelendiđinde, artıkların belirgin bir trend, mevsimsellik veya yapı göstermediđi bu sebeple rastgele bir řekilde dađılmış olduđu sđylenebilir. Bu istenilen bir durumdur.

Artıkların dađılım grafiđinde, artıkların dađılımının normal dađılıma uyumlu olması beklenir. Yukarıdaki grafik incelendiđinde, artıkların dađılımının ,neredeyse, normal dađılıma uyumlu olduđu sđylenebilir.

Tüm bu gerekçeler dođrultusunda, modelin uygun bir yapıya sahip olduđu sđylenebilir. Bu uygunluk temelinde, model kullanılarak sonraki 12 ay için tahminler gerçekteřtirilecektir.

### Tahmin Deđerleri

```
ongoru <- forecast(Winters2,h=12)
ongoru[["mean"]]
```

```
##          Jan      Feb      Mar      Apr      May      Jun      Jul      Aug      Sep
## 2024 1847630 1721407 1797963 1726962 1730021 1636771 1699620 1774461 1788347
##          Oct      Nov      Dec
## 2024 1857313 1816958 1868307
```

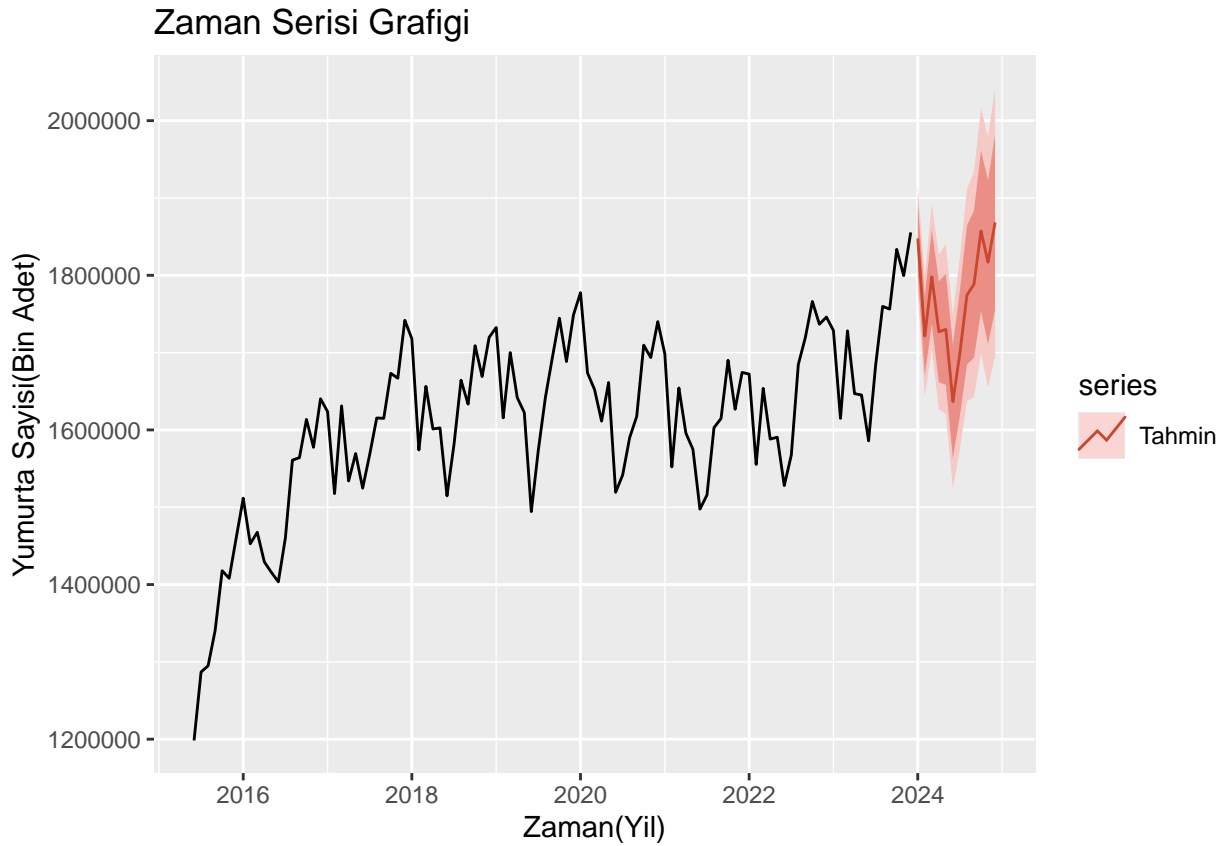


ongoru

##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## Jan 2024	1847630	1802800	1892460	1779068	1916191
## Feb 2024	1721407	1670620	1772194	1643735	1799079
## Mar 2024	1797963	1736939	1858987	1704635	1891291
## Apr 2024	1726962	1661581	1792343	1626971	1826954
## May 2024	1730021	1658388	1801655	1620467	1839575
## Jun 2024	1636771	1563655	1709887	1524949	1748592
## Jul 2024	1699620	1618529	1780710	1575602	1823637
## Aug 2024	1774461	1684731	1864192	1637230	1911692
## Sep 2024	1788347	1693081	1883613	1642650	1934043
## Oct 2024	1857313	1753598	1961027	1698695	2015930
## Nov 2024	1816958	1711033	1922883	1654959	1978957
## Dec 2024	1868307	1754985	1981629	1694996	2041618

Model kullanılarak yapılan tahminlerin %80 güven düzeyindeki alt sınırı (Lo80), üst sınırı (Hi80), %95 güven düzeyindeki alt sınırı (Lo95), üst sınırı (Hi95) ve nokta tahmin değerleri yukarıda sunulmuştur.

```
autoplot(veri_ts2) +  
  autolayer(ongoru, series = "Tahmin", PI=TRUE) +  
  ggtitle("Zaman Serisi Grafiği") +  
  xlab("Zaman(Yıl)") +  
  ylab("Yumurta Sayısı(Bin Adet)")
```



## Mevsimsel ARIMA Modelleri (SARIMA)

SARIMA modeli, ARIMA modelinin bir genişletmesi olup, mevsimsel bileşenleri de içerecek şekilde modelin parametrelerini düzenler. SARIMA, mevsimsel özellikleri AR, MA, ve differencing (farklılaştırma) terimleri ile doğrudan modele dahil eder.

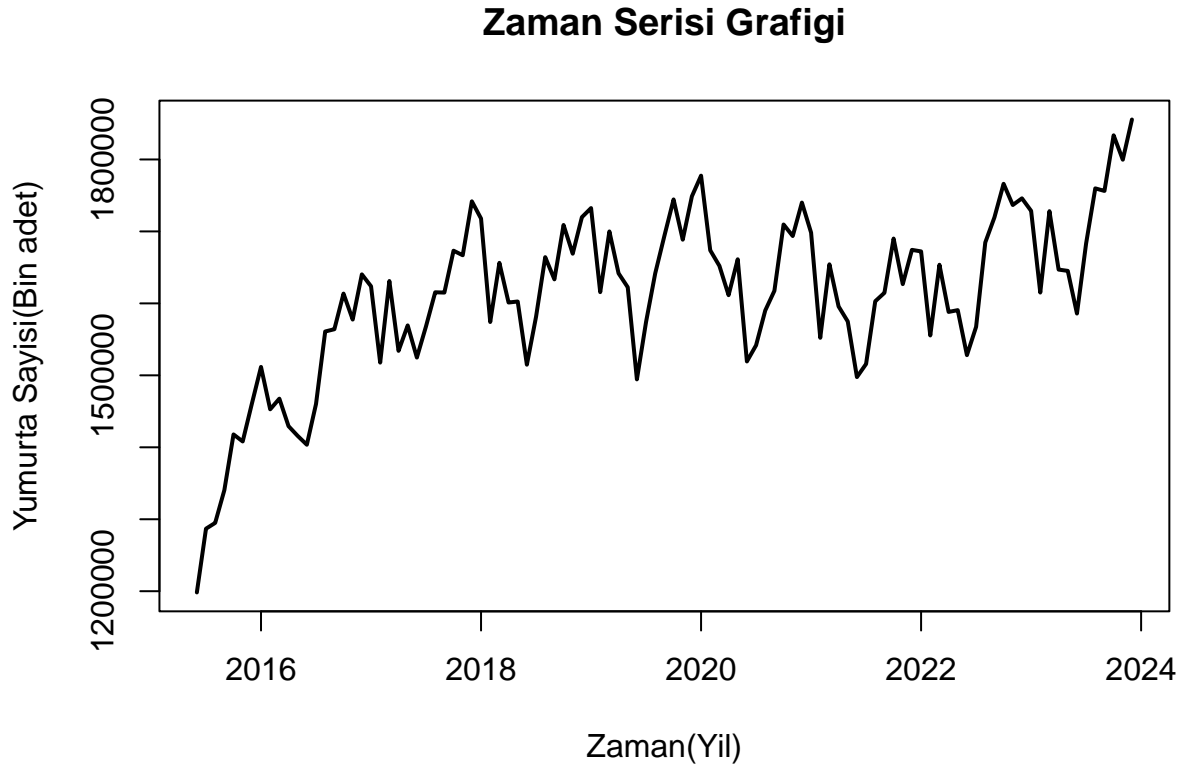
SARIMA modelinin genel formu şu şekildedir:

$$\text{SARIMA}(p,d,q)(P,D,Q)$$

İlk adım olarak, veri setinin zaman serisi grafiği incelenmiştir. Zaman serisi grafiği üzerinde görsel bir inceleme yapıldığında, mevsimsel dalgalanmalar ve trend gibi yapılar gözlemlenmiştir. Bu yapıların varlığı, serinin durağan olmadığına işaret etmektedir. Zaman serisinin durağan hale getirilmesi için, trendin ve mevsimselliğin ortadan kaldırılması gerekmektedir. Zaman serisini durağanlaştırmak amacıyla, hem trend farkı hem de mevsimsel fark alınmıştır. Bu işlem, zaman serisinin durağan hale gelmesini sağlamak için gereklidir. Trend farkı almak, serinin zamanla artan veya azalan yapısını ortadan kaldırırken; mevsimsel fark almak, belirli bir periyot boyunca tekrar eden döngüsel yapıları düzeltir.

```
verisar = ts(veri, start = c(2015,6), frequency =12)
```

```
ts.plot(verisar,main="Zaman Serisi Grafiği", xlab="Zaman(Yıl)", ylab="Yumurta Sayısı(Bin adet)", lwd=2)
```



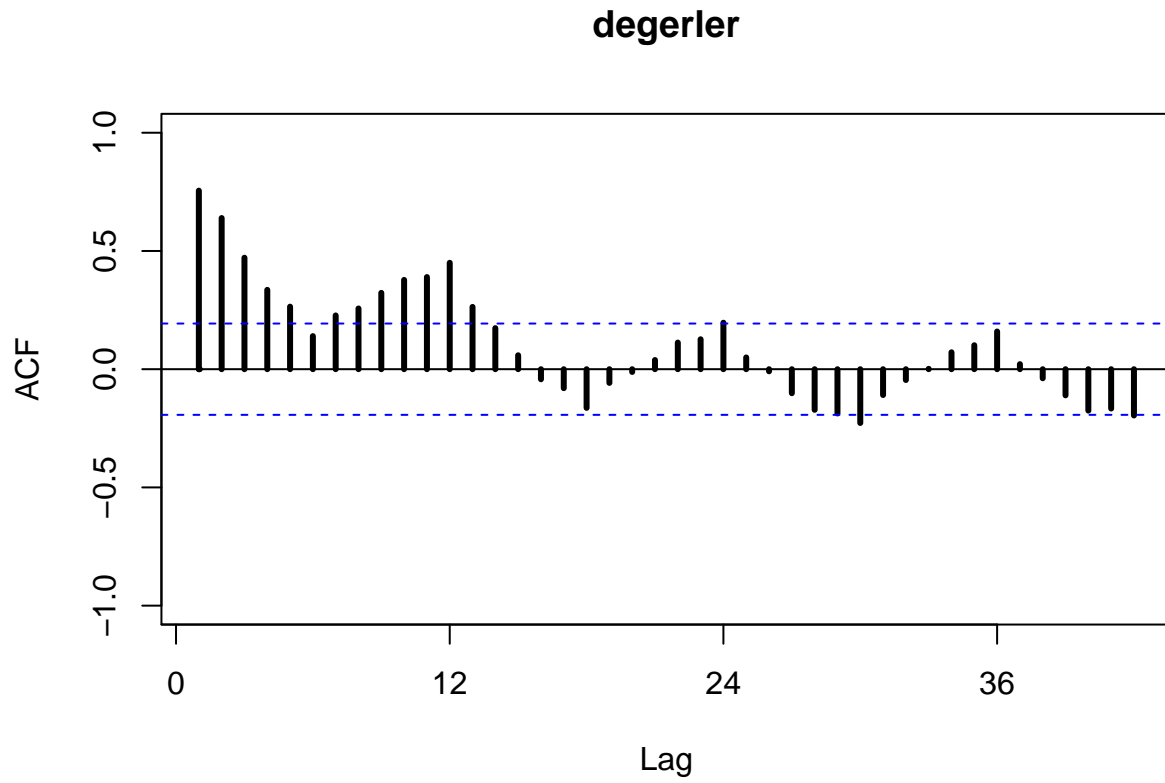
Serinin genel yönelimindeki (trend) değişiklikleri ortadan kaldırmak için birinci dereceden fark alınmıştır. Bu işlem, trendi etkili bir şekilde gidererek veriyi daha düzgün hale getirmiştir.

Serideki mevsimsel bileşenleri ortadan kaldırmak amacıyla birinci dereceden mevsimsel fark alınmıştır. Bu, mevsimsel dalgalanmaları ortadan kaldırarak seriyi daha düzgün ve durağan hale getirmiştir.

Seriyi durağan hale getirmek için birinci dereceden farkların alınması yeterli olmuştur. Bu durumda ilk parametre değerlerimiz  $d=1$ ,  $D=1$  olarak belirlenmiştir. ACF ve PACF grafiklerinin incelenmesi, ARIMA

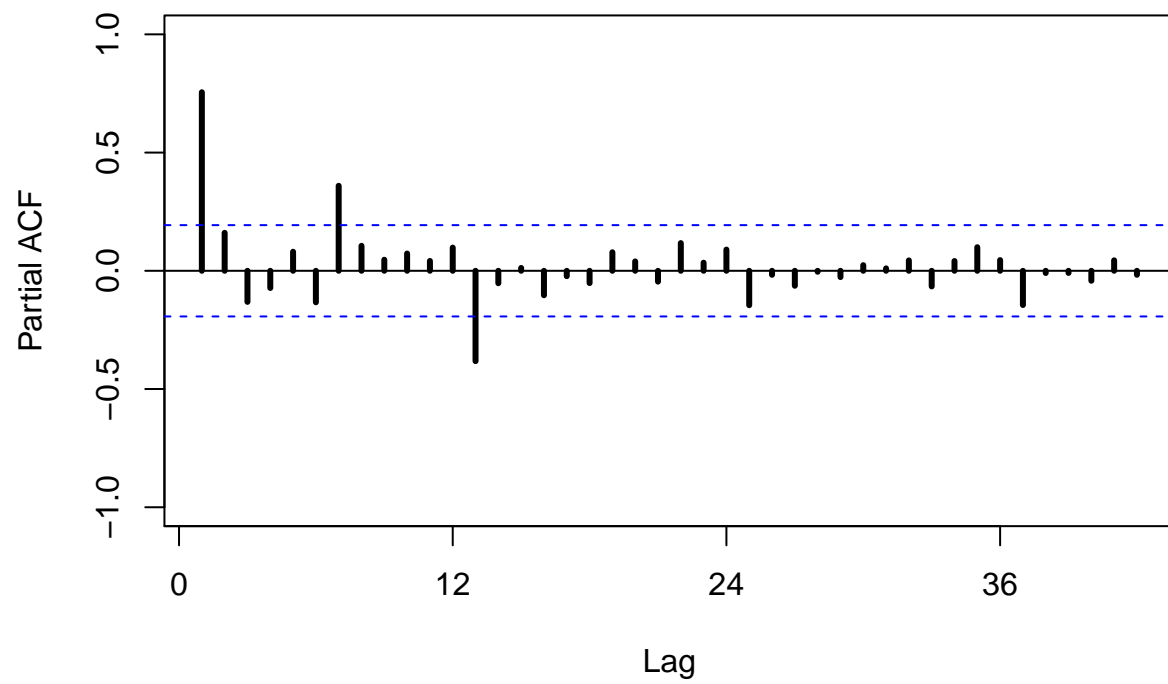
modelinin parametrelerinin seçilmesinde önemli bir adımdır.

```
Acf(verisar,lag.max = 42, ylim=c(-1,1), lwd=3)
```



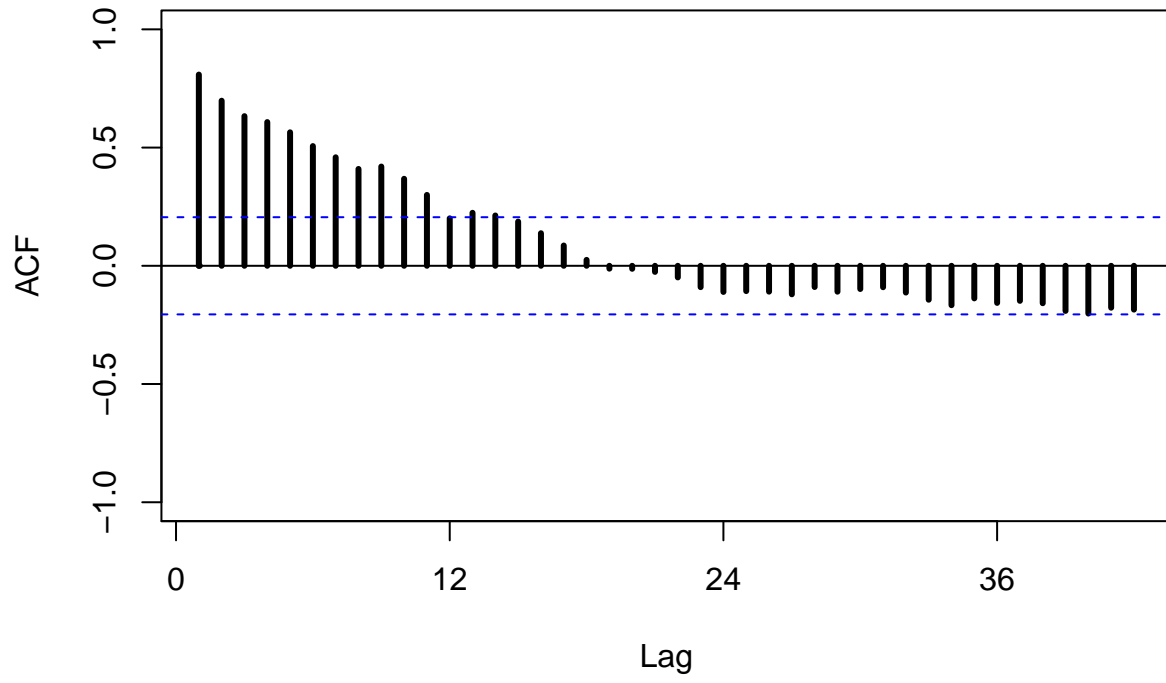
```
Pacf(verisar,lag.max = 42, ylim=c(-1,1), lwd=3)
```

## Series verisar



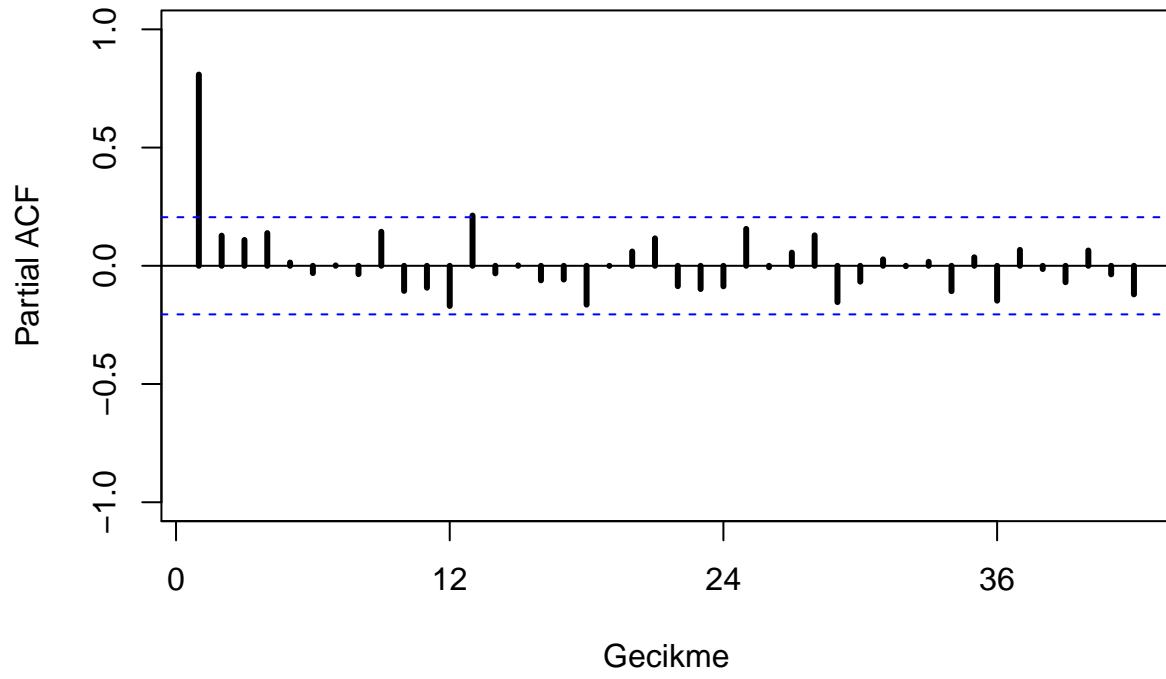
```
Acf(diff(verisar,12),lag.max = 42, ylim=c(-1,1), lwd=3)
```

## degerler

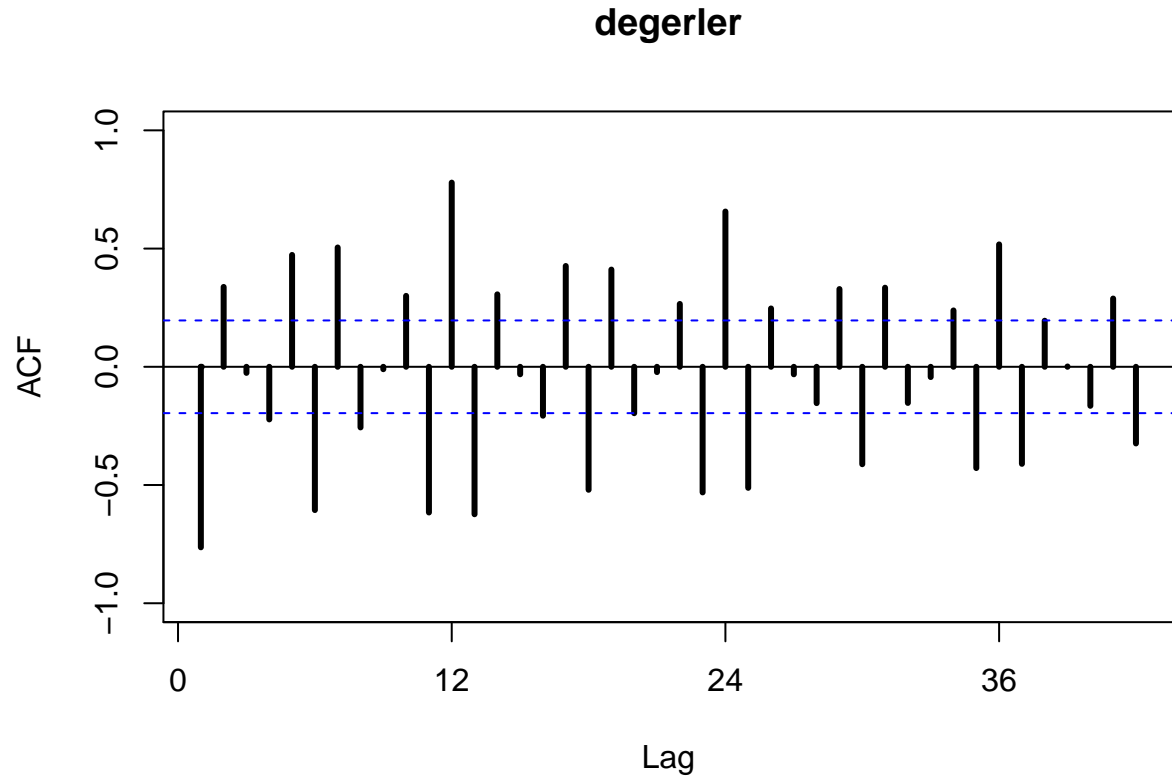


```
Pacf(diff(verisar,12),lag.max = 42, ylim=c(-1,1), lwd=3, main = "PACF Grafiği", xlab = "Gecikme")
```

## PACF Grafigi

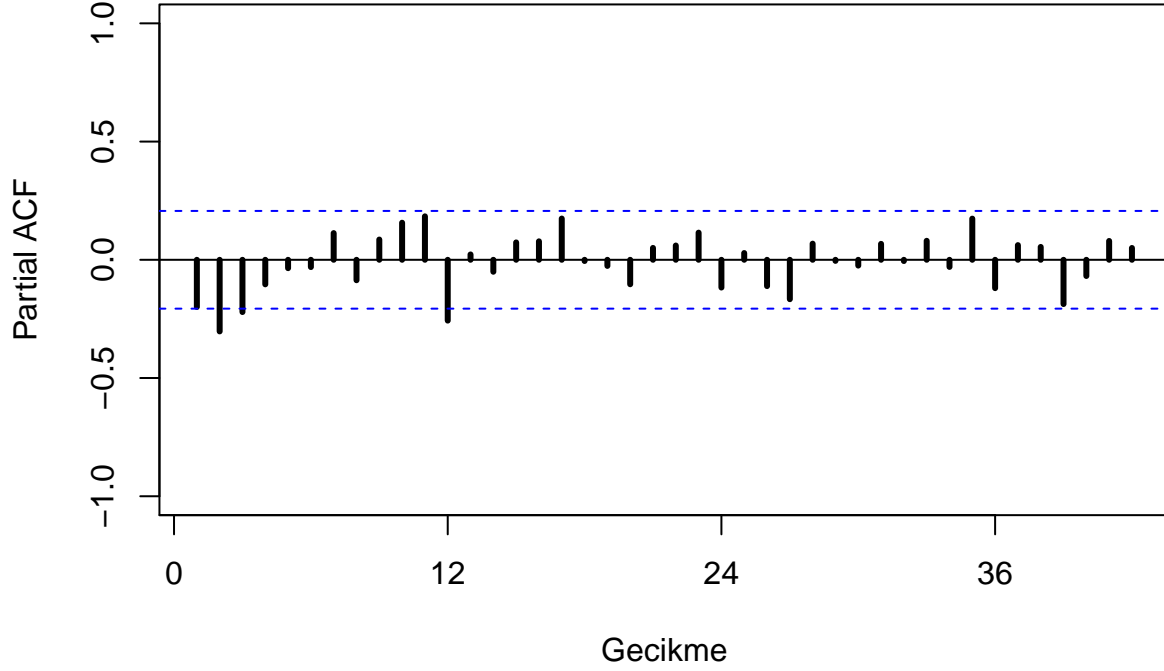


```
Acf(diff(diff(verisar,,2,12)),lag.max = 42, ylim=c(-1,1), lwd=3)
```



```
Pacf(diff(diff(verisar,12)),lag.max = 42, ylim=c(-1,1), lwd=3, main = "PACF Grafiği", xlab = "Gecikme")
```

## PACF Grafigi



ACF ve PACF grafiklerinden elde edilen bulgulara dayanarak, q parametresinin en fazla “1”, p parametresinin ise en fazla “2” değerini alabileceğine karar verilmiştir.

Yukarıda belirtilen parametre kombinasyonlarıyla toplamda on farklı ARIMA modeli test edilmiştir. Bu modeller arasından beş tanesi istatistiksel olarak anlamlı bulunmuş ve test sonuçları aşağıda sunulmuştur.

### Anlamlı Bulunan Modeller

*BIC = 2154*

```
deneme_arima2 <- Arima(verisar, order = c(0,1,1), seasonal= c(1,1,0), include.constant=TRUE)
coeftest(deneme_arima2)
```

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ma1  -0.37055    0.12887 -2.8754 0.0040348 **
## sar1  -0.39869    0.10598 -3.7618 0.0001687 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(deneme_arima2)
```

```
## Series: verisar
## ARIMA(0,1,1)(1,1,0)[12]
##
## Coefficients:
```



```

##          ma1      sar1
##        -0.3706  -0.3987
## s.e.    0.1289   0.1060
##
## sigma^2 = 1.253e+09: log likelihood = -1070.48
## AIC=2146.97  AICc=2147.25  BIC=2154.47
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -2365.122 32715.03 24416.37 -0.1677534 1.502921 0.3856471
##              ACF1
## Training set 0.0778036

BIC = 2162

deneme_arima3 <- Arima(verisar, order = c(0,1,1), seasonal= c(0,1,0), include.constant=TRUE)
coeftest(deneme_arima3)

##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ma1 -0.49689      0.11505 -4.3189 1.568e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(deneme_arima3)

## Series: verisar
## ARIMA(0,1,1)(0,1,0)[12]
##
## Coefficients:
##          ma1
##        -0.4969
## s.e.    0.1151
##
## sigma^2 = 1.448e+09: log likelihood = -1076.55
## AIC=2157.11  AICc=2157.25  BIC=2162.11
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -2269.303 35376.99 27056.42 -0.1607267 1.664266 0.4273456
##              ACF1
## Training set 0.1208607

BIC=2155.86

deneme_arima4 <- Arima(verisar, order = c(2,1,0), seasonal= c(1,1,0), include.constant=TRUE)
coeftest(deneme_arima4)

##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1 -0.22761      0.10276 -2.2149 0.0267647 *
## ar2 -0.26780      0.10369 -2.5827 0.0098026 **
## sar1 -0.40700      0.10640 -3.8251 0.0001307 ***

```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(deneme_arima4)

## Series: verisar
## ARIMA(2,1,0)(1,1,0)[12]
##
## Coefficients:
##          ar1          ar2          sar1
##      -0.2276  -0.2678  -0.4070
## s.e.    0.1028   0.1037   0.1064
##
## sigma^2 = 1.222e+09:  log likelihood = -1068.93
## AIC=2145.86   AICc=2146.33   BIC=2155.86
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -2285.442 32132.41 23499.35 -0.1624452 1.44884 0.3711631
##              ACF1
## Training set -0.04688018

BIC=2164.08

deneme_arima9 <- Arima(verisar, order = c(2,1,0), seasonal= c(0,1,0), include.constant=TRUE)
coeftest(deneme_arima9)

##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1 -0.25828     0.10076 -2.5634  0.01036 *
## ar2 -0.31430     0.10283 -3.0564  0.00224 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(deneme_arima9)

## Series: verisar
## ARIMA(2,1,0)(0,1,0)[12]
##
## Coefficients:
##          ar1          ar2
##      -0.2583  -0.3143
## s.e.    0.1008   0.1028
##
## sigma^2 = 1.425e+09:  log likelihood = -1075.29
## AIC=2156.58   AICc=2156.86   BIC=2164.08
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -1726.544 34889.92 26171.99 -0.1262105 1.607969 0.4133764
##              ACF1
## Training set -0.07448556

BIC=2156.49
```

```
deneme_arima10<- Arima(verisar, order = c(2,1,0), seasonal= c(2,1,0), include.constant=TRUE)
coeftest(deneme_arima10)
```

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1  -0.23494    0.10316 -2.2775  0.02276 *
## ar2  -0.25686    0.10456 -2.4565  0.01403 *
## sar1 -0.50938    0.11830 -4.3057 1.664e-05 ***
## sar2 -0.23366    0.11524 -2.0276  0.04260 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(deneme_arima10)
```

```
## Series: verisar
## ARIMA(2,1,0)(2,1,0)[12]
##
## Coefficients:
##          ar1      ar2      sar1      sar2
##      -0.2349  -0.2569  -0.5094  -0.2337
## s.e.    0.1032   0.1046   0.1183   0.1152
##
## sigma^2 = 1.166e+09: log likelihood = -1067
## AIC=2144   AICc=2144.71   BIC=2156.49
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -3129.212 31204.5 22725.45 -0.2147529 1.403074 0.3589397
##              ACF1
## Training set -0.04478997
```

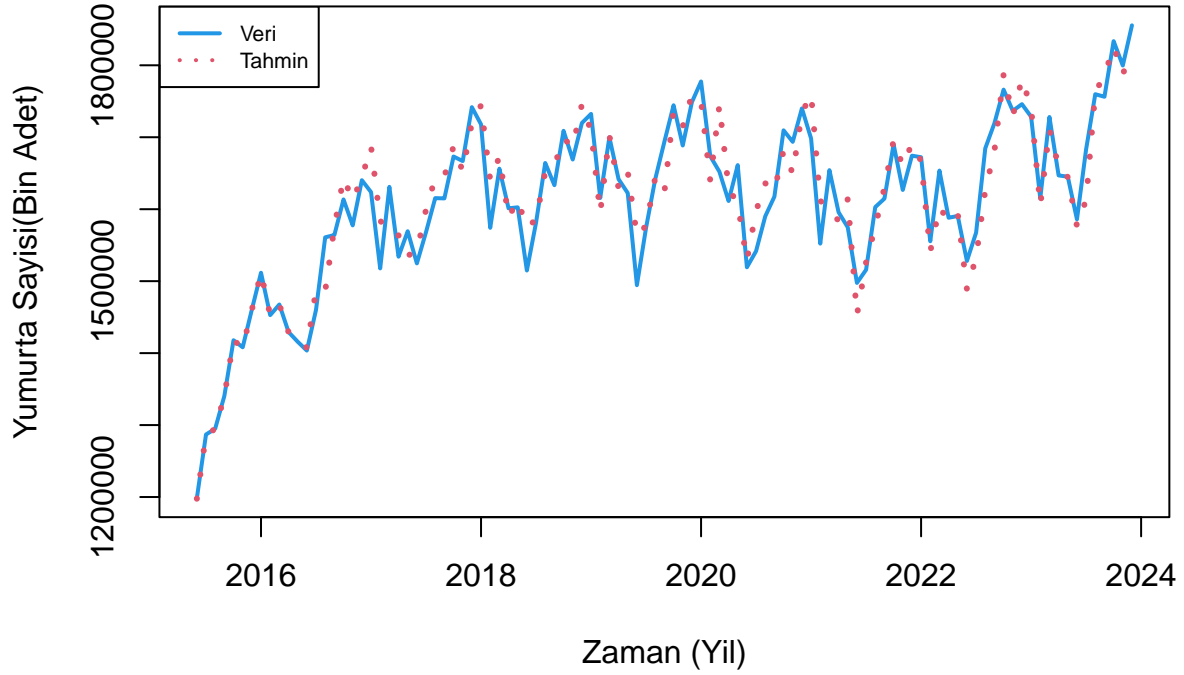
Yukarıda, istatistiksel olarak anlamlı bulunan beş modelin istatistikleri verilmiştir. En iyi sonuç veren modeli belirlemek için Bayes Bilgi Kriterinden (BIC) yararlanılmıştır. Beş modelin BIC değerlerine bakıldığında, en küçük değere sahip model “BIC = 2154,47” değeri ile ARIMA(0,1,1)(1,1,0)[12] parametrelerine sahip modeldir.

En iyi model belirlendiğine göre şimdi modelin geçerliliği kontrol edilecektir. Bu amaçla, tahmin serisi ile orijinal serinin zaman grafiği, hata serisinin ACF grafiği ve Box-Ljung test istatistiği aşağıda sunulmuştur.

```
tahmin = deneme_arima2[["fitted"]]
hata = deneme_arima2[["residuals"]]
```

```
plot(window(verisar), main = "Uyum Grafiği",
      xlab="Zaman (Yıl)", ylab="Yumurta Sayısı(Bin Adet)",lty=1, col=4, lwd=2)
lines(window(tahmin) ,lty=3,col=2,lwd=3)
legend("topleft",c(expression(paste("Veri")),
                    expression(paste("Tahmin"))),
      lwd=c(2,2),lty=c(1,3), cex=0.7, col=c(4,2))
```

## Uyum Grafiği



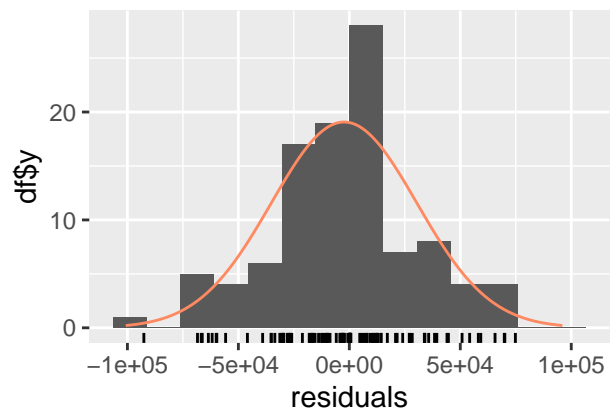
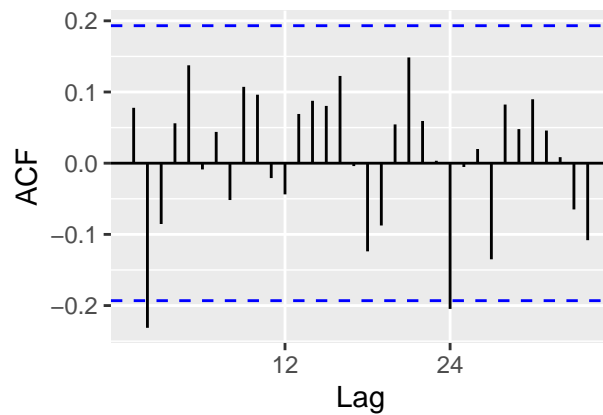
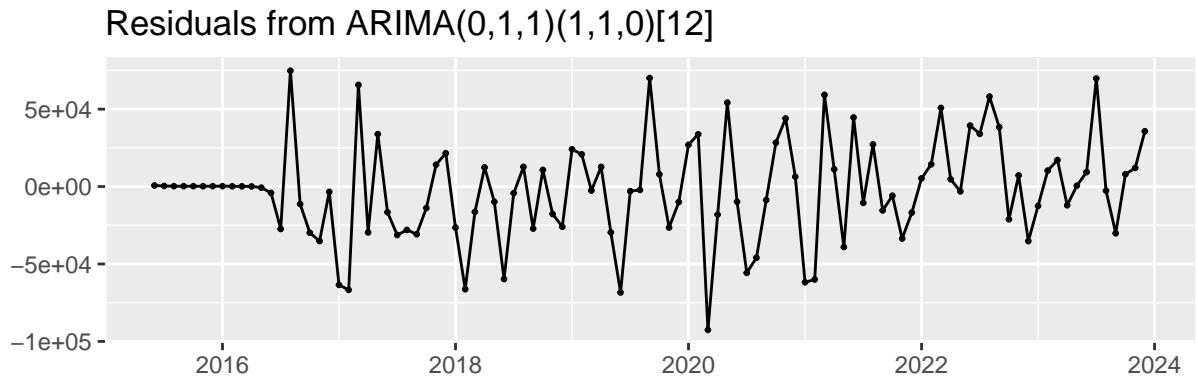
Grafik incelendiğinde, orijinal seri ile tahmin serisinin yeteri kadar uyum içerisinde olduğu söylenebilir. Fakat daha kesin kararlar verebilmek için hata serisinin ACF grafiği ve BoxLjung testi incelenecektir.

### ACF,PACF, Artıkların Dağılım Grafiği ve Box-Ljung Testi

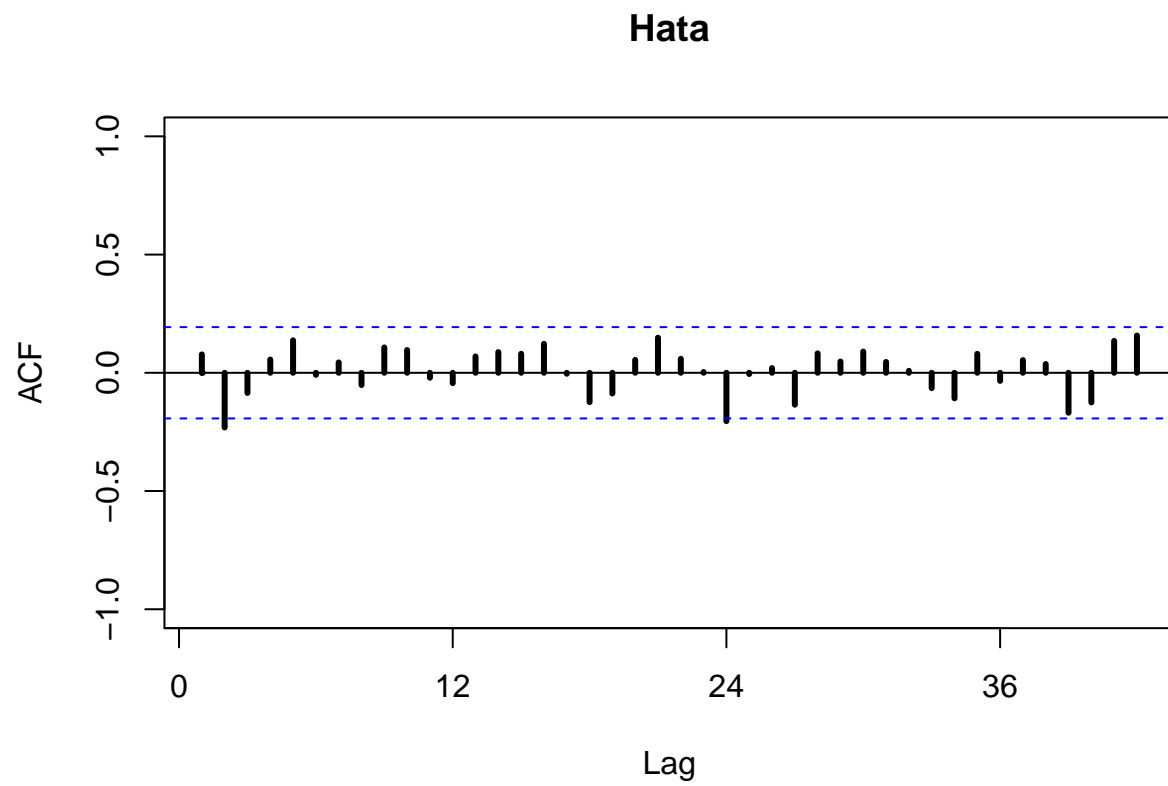
```
Box.test(hata, lag = 42, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: hata  
## X-squared = 54.408, df = 42, p-value = 0.09496
```

```
checkresiduals(deneme_arima2)
```

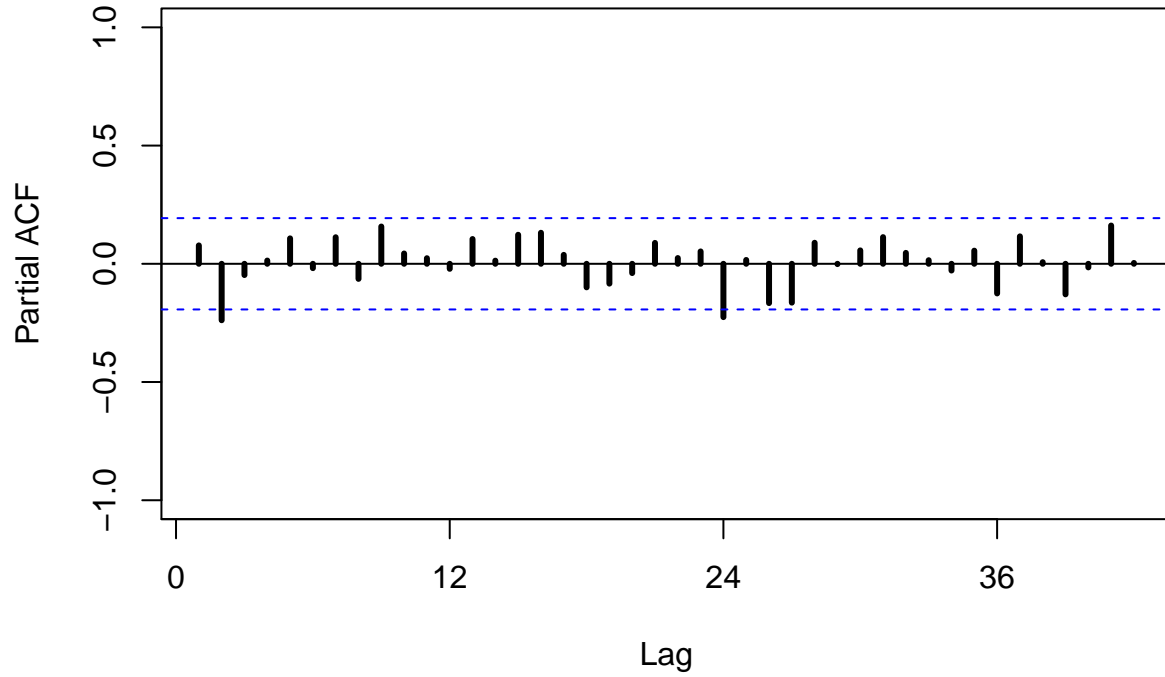


```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,1,1)(1,1,0)[12]
## Q* = 23.2, df = 19, p-value = 0.2287
##
## Model df: 2.   Total lags used: 21
Acf(hata,main="Hata", lag.max = 42, ylim=c(-1,1), lwd=3)
```



```
Pacf(hata,main="Hata",lag.max = 42, ylim=c(-1,1), lwd=3)
```

## Hata



Grafikler ve Box-Ljung test istatistiği incelendiğinde, artıkların zaman içerisinde rastgele dağıldığı ve gecikmeler arasında herhangi bir ilişki bulunmadığı sonucuna ulaşılmıştır. Bu nedenle, modelin incelenen veri seti için uygun olduğu ifade edilebilir.

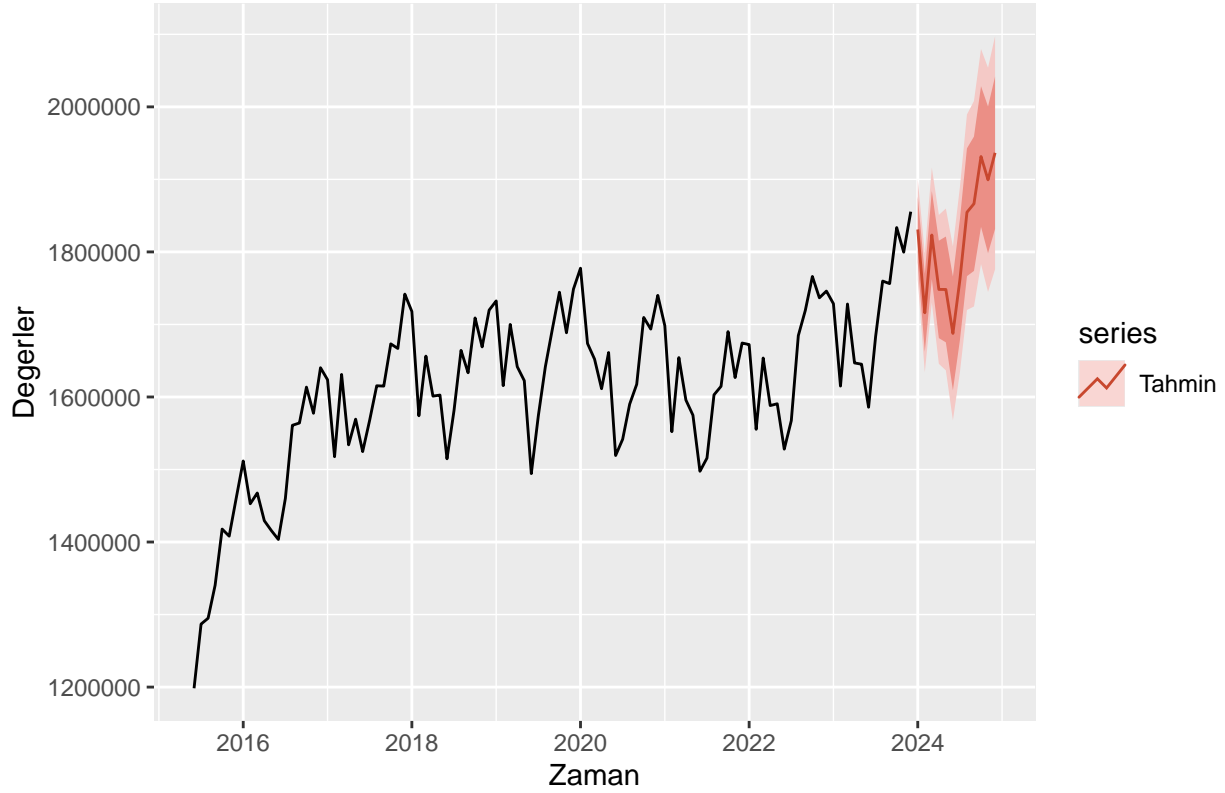
## Tahmin Değerleri

```
ongoru = forecast(deneme_arima2 , h=12)
ongoru["mean"]
```

```
## $mean
##      Jan      Feb      Mar      Apr      May      Jun      Jul      Aug      Sep
## 2024 1830934 1716096 1823293 1748329 1748248 1687653 1761861 1854680 1866432
##      Oct      Nov      Dec
## 2024 1931520 1899396 1936669
```

```
autoplot(verisar) +
  autolayer(ongoru, series="Tahmin", PI=TRUE) +
  ggtitle("Zaman Serisi ve SARIMA Tahminler") +
  xlab("Zaman") +
  ylab("Degerler")
```

## Zaman Serisi ve SARIMA Tahminler



### Kullanılan Modellerin Tahmin Değerleri ile Gerçek Değerlerin Karşılaştırılması

Modelin geçerliliğini anlayabilmek için gerçek değerler ile tahmin değerlerinin çizgi grafiği, tahmin hatalarının dağılımı, ortalama mutlak hata(MAE), kök ortalama hata(RMSE) ve ortalama yüzde hata(MAPE) değerleri incelenmiştir. Elde edilen bulgular aşağıda sunulmuştur.

```
original_values = ts(c(1838422, 1733464
, 1790081
, 1707799
, 1764629
, 1694374
, 1706493
, 1808368
, 1828030))

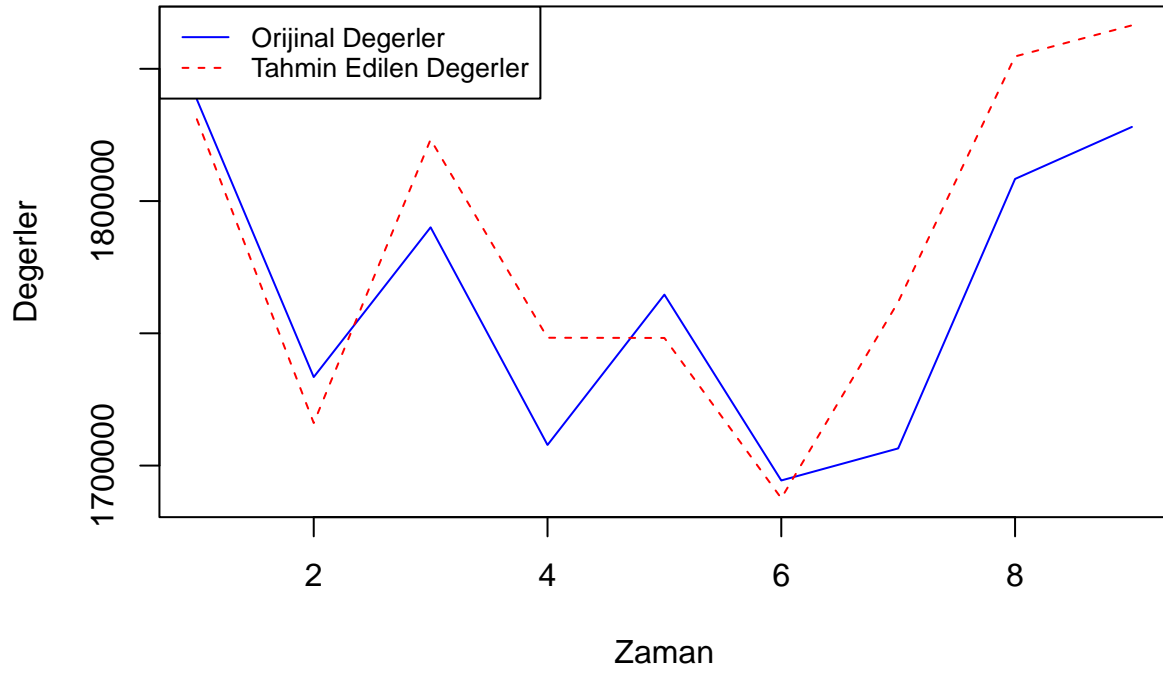
forecasted_values = ts(c(1830934 ,1716096 ,1823293 ,1748329, 1748248, 1687653, 1761861 ,1854680, 1866430))

forecasted_valueswinters = ts(c(1846160, 1728852, 1798002, 1732222, 1732595, 1647998, 1699930, 1774369,

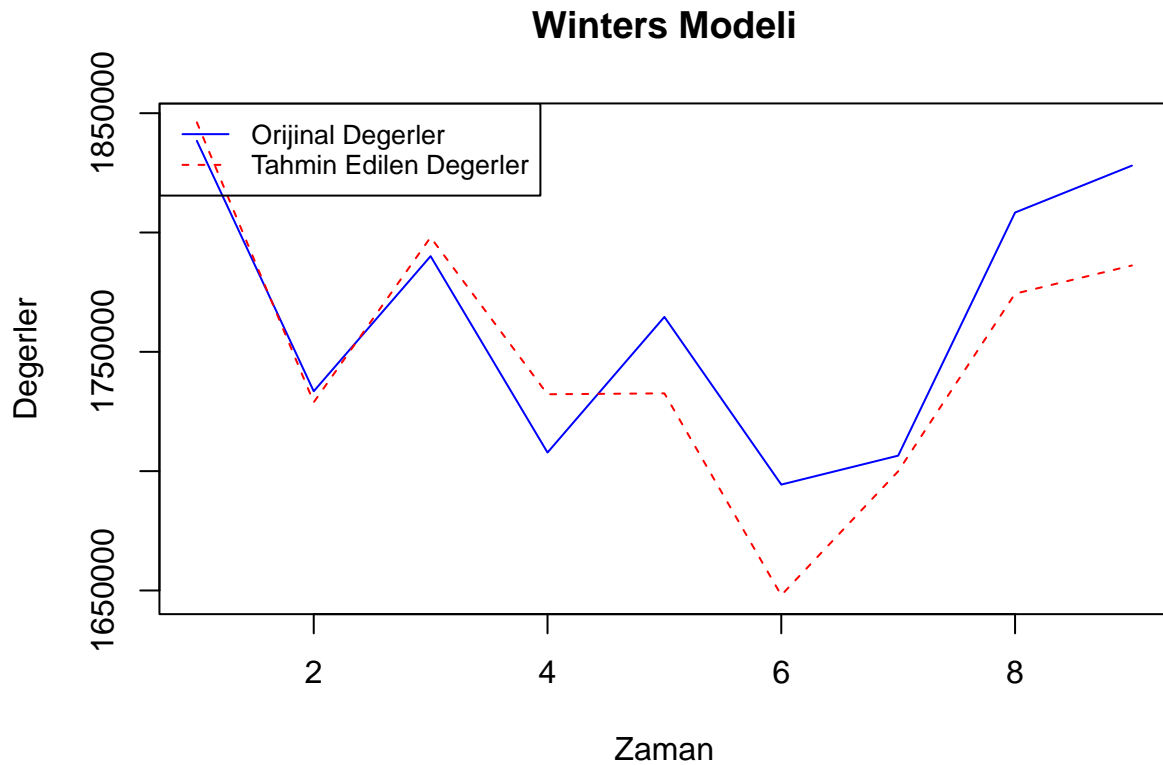
plot(original_values, col = "blue", lty = 1, ylim = range(c(original_values, forecasted_values)),
      main = "SARIMA Modeli", xlab = "Zaman", ylab = "Degerler")
lines(forecasted_values, col = "red", lty = 2)
legend("topleft", legend = c("Orijinal Degerler", "Tahmin Edilen Degerler"),
      col = c("blue", "red"), lty = c(1, 2), cex = 0.8)
```



## SARIMA Modeli



```
plot(original_values, col = "blue", lty = 1, ylim = range(c(original_values, forecasted_valueswinters))  
      main = "Winters Modeli", xlab = "Zaman", ylab = "Degerler")  
lines(forecasted_valueswinters, col = "red", lty = 2)  
legend("topleft", legend = c("Orijinal Degerler", "Tahmin Edilen Degerler"),  
      col = c("blue", "red"), lty = c(1, 2), cex = 0.8)
```



#### Ortalama Mutlak Hata (MAE)

```
mae = mean(abs(original_values - forecasted_values))
print(paste("Ortalama Mutlak Hata (MAE) SARIMA:", round(mae, 2)))
```

```
## [1] "Ortalama Mutlak Hata (MAE) SARIMA: 29086.89"
```

```
mae = mean(abs(original_values - forecasted_valueswinters))
print(paste("Ortalama Mutlak Hata (MAE) Winters:", round(mae, 2)))
```

```
## [1] "Ortalama Mutlak Hata (MAE) Winters: 22835.56"
```

#### Kök Ortalama Kare Hata (RMSE)

```
rmse = sqrt(mean((original_values - forecasted_values)^2))
print(paste("Kök Ortalama Kare Hata (RMSE) SARIMA:", round(rmse, 2)))
```

```
## [1] "Kök Ortalama Kare Hata (RMSE) SARIMA: 33503.03"
```

```
rmse = sqrt(mean((original_values - forecasted_valueswinters)^2))
print(paste("Kök Ortalama Kare Hata (RMSE) Winters:", round(rmse, 2)))
```

```
## [1] "Kök Ortalama Kare Hata (RMSE) Winters: 27624.56"
```

## Ortalama Yüzde Hata (MAPE)

```
mape <- mean(abs((original_values - forecasted_values) / original_values) * 100)
print(paste("Ortalama Yuzde Hata (MAPE) SARIMA:", round(mape, 2), "%"))

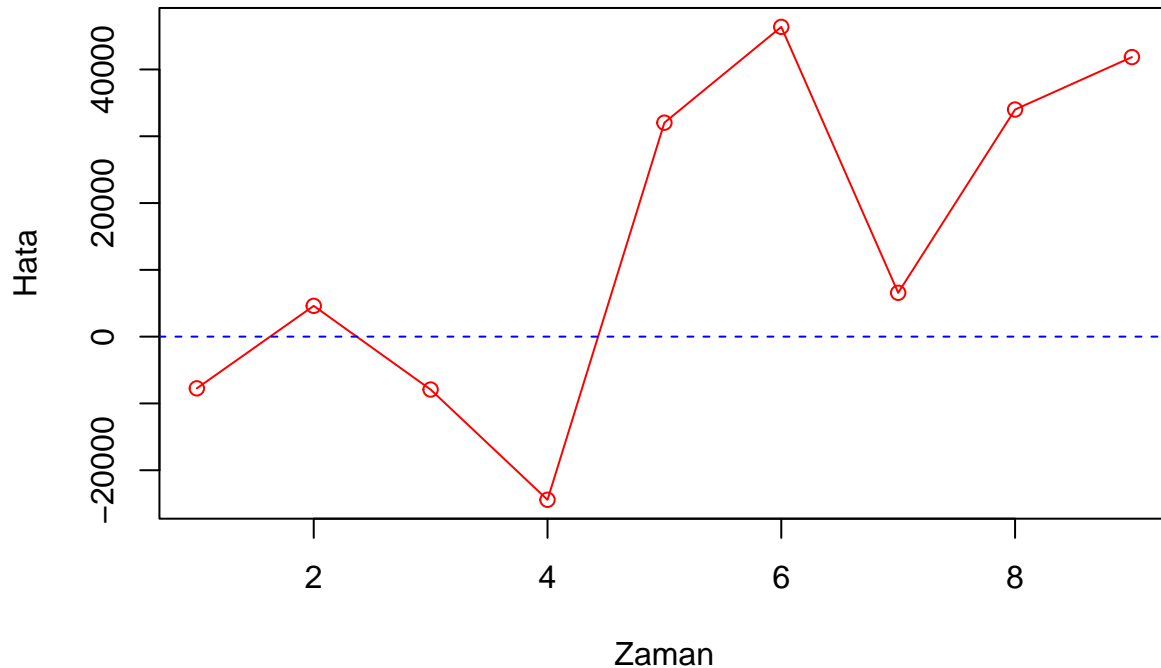
## [1] "Ortalama Yuzde Hata (MAPE) SARIMA: 1.65 %"

mape <- mean(abs((original_values - forecasted_valueswinters) / original_values) * 100)
print(paste("Ortalama Yuzde Hata (MAPE) WINTERS:", round(mape, 2), "%"))

## [1] "Ortalama Yuzde Hata (MAPE) WINTERS: 1.3 %"

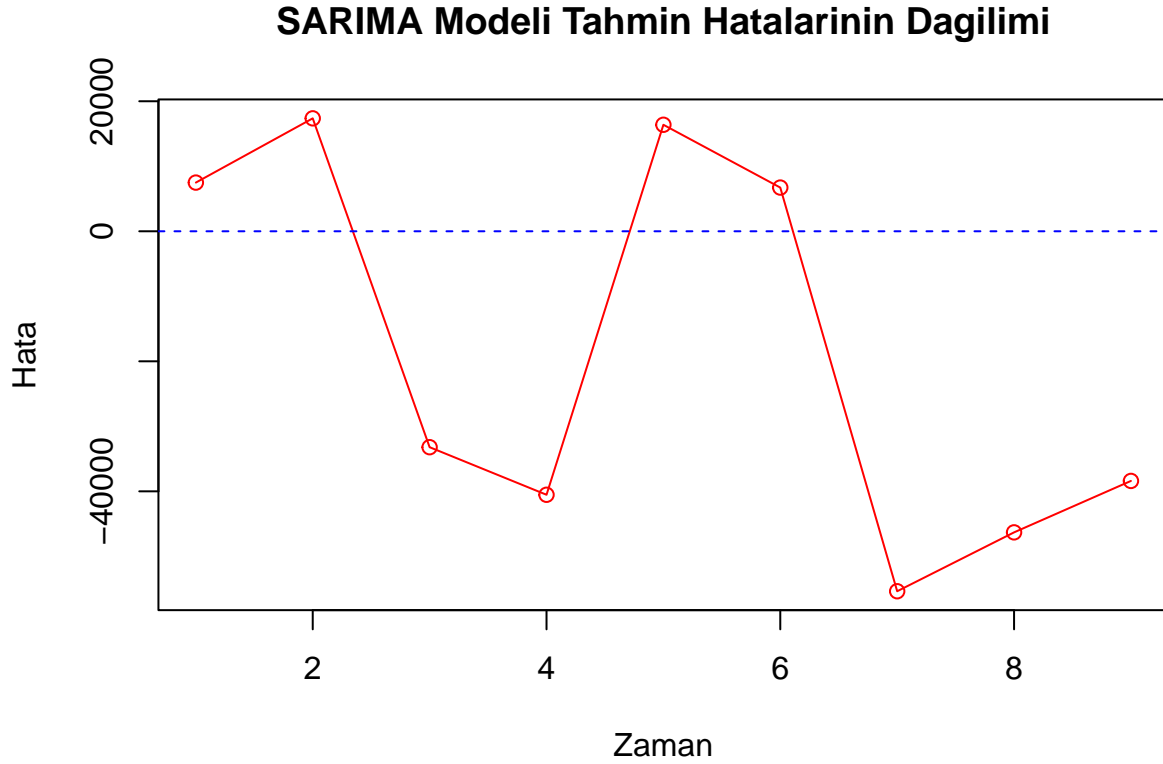
residuals <- original_values - forecasted_valueswinters
plot(residuals, type = "o", col = "red", main = "Winters Modeli Tahmin Hatalarının Dağılımı",
      xlab = "Zaman", ylab = "Hata")
abline(h = 0, col = "blue", lty = 2)
```

## Winters Modeli Tahmin Hatalarının Dağılımı



```
residuals2 <- original_values - forecasted_values

plot(residuals2, type = "o", col = "red", main = "SARIMA Modeli Tahmin Hatalarının Dağılımı",
      xlab = "Zaman", ylab = "Hata")
abline(h = 0, col = "blue", lty = 2)
```



Winters modelleri ve SARIMA modelinin farklı teorik temellere sahip olması, birinin kesinlikle diğlerinden üstün olduğı anlamına gelmez. Ancak, orijinal seri değeri ile Toplamsal Winters modelinin tahmin edilen değeri yer aldığı çizgi grafiğine bakıldığında, bu modelin SARIMA modeline göre daha uyumlu bir grafik sunduğı söylenebilir. Ayrıca, MAE, RMSE ve MAPE gibi hata metrik değeri incelendiğinde, Toplamsal Winters modelinin daha düşük hata değeri sahip olduğı ve belirli dönemlerde SARIMA modelinden daha iyi tahminler yaptığı sonucuna varılabilir.

## Sonuç

Bu çalışma, Türkiye'nin tavuk yumurtası üretim verilerini incelemek ve bu veriler üzerinden geleceğe yönelik öngörülerde bulunmak amacıyla gerçekleştirilmiştir. Analizlerde kullanılan yöntemler, verilerin trend ve mevsimsel bileşenlerini anlamada etkili olmuş, en uygun modelin belirlenmesine olanak tanımıştır. Toplamsal Winters modeli, Çarpımsal Winters modeli ve SARIMA(0,1,1)(1,1,0) modeli incelenen veri seti ile uyumlu bulunmuştur. Sonuçlar, Toplamsal Winters modelinin, daha düşük hata metrikleri ve gerçek değere olan yakınlığıyla SARIMA ve Çarpımsal Winters modellerine kıyasla daha iyi tahmin performansı sunduğunu ortaya koymuştur. Bu bulgular, üretim ve gıda politikalarının planlanmasında ve kaynak tahsisinde karar alıcılara değerli bilgiler sunabilir. Gelecekte, daha karmaşık modellerin ve ek veri kaynaklarının dahil edilmesiyle tahmin doğruluğı daha da artırılabilir.