logo, simge, sembol, grafik, yazı tipi içeren bir resim

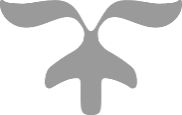
Açıklama otomatik olarak oluşturuldutasarım içeren bir resim

Açıklama otomatik olarak düşük güvenilirlik düzeyiyle oluşturuldu

ZAMAN SERİLERİ DÖNEM ÖDEVİ

Mehmet Şenel

2210329077



İçindekiler Tablosu

[Veri Hakkında 4](#_Toc185243440)

[1.Veri Setinin Genel İçeriği 4](#_Toc185243441)

[2. Veri Setinin Amacı 4](#_Toc185243442)

[3. Veri Yapısı ve Özellikleri 4](#_Toc185243443)

[4. Zaman Serisi Yapısı 5](#_Toc185243444)

[5.Verinin Kullanımına Dair Potansiyel Yöntemler 5](#_Toc185243445)

[Zaman Serisi Oluşturma ve Görselleştirme 6](#_Toc185243446)

[Veri Dönüştürme 6](#_Toc185243447)

[Zaman Serisi Oluşturma 6](#_Toc185243448)

[Zaman Serisi Grafiği 7](#_Toc185243449)

[Otomatik Korelasyon Fonksiyonu (ACF) 7](#_Toc185243450)

[Kısmi Otomatik Korelasyon Fonksiyonu (PACF) 8](#_Toc185243451)

[8](#_Toc185243452)

[Modelleme 9](#_Toc185243453)

[Trend Bileşeninin Hesaplanması 9](#_Toc185243454)

[Mevsimsel Bileşenin Hesaplanması 9](#_Toc185243455)

[Merkezsel Hareketli Ortalama ve Mevsimsel Bileşen 10](#_Toc185243456)

[Mevsimsel Endeks Hesaplaması 10](#_Toc185243457)

[Hata Bileşeni 11](#_Toc185243458)

[Zaman Serisinin Trend Değerlerinin Hesaplanması 11](#_Toc185243459)

[Doğrusal Regresyon Modeli Kurma 11](#_Toc185243460)

[Tahmin Serisini Hesaplama (Mevsimsel Endeks + Trend) 12](#_Toc185243461)

[Hata Serisinin Hesaplanması 12](#_Toc185243462)

[Zaman Serisi Modelinin Güvenilirliğinin Değerlendirilmesi 13](#_Toc185243463)

[Modelin Geçerliliği 13](#_Toc185243464)

[Tahmin Serisinin Zaman Serisi Formatına Dönüştürülmesi 13](#_Toc185243465)

[Zaman Serisi ve Tahminin Karşılaştırılması 14](#_Toc185243466)

[ACF ve PACF Analizleri 14](#_Toc185243467)

[Çarpımsal Ayrıştırma Yöntemi (Multiplicative Decomposition Method) 16](#_Toc185243468)

[Mevsimsel Bileşenin Bulunması 16](#_Toc185243469)

[Mevsimsel Bileşenlerin Ortalama Değerlerinin Hesaplanması 17](#_Toc185243470)

[Mevsimsel Endeks Değerlerinin Hesaplanması 17](#_Toc185243471)

[Trent Serisi ve Tahminlerin Hesaplanması 17](#_Toc185243472)

[Tahmin Serisinin Zaman Serisi Haline Getirilmesi 19](#_Toc185243473)

[Hata Serisi ve Değerlendirme 20](#_Toc185243474)

[Toplamsal Model (Additive Model) 21](#_Toc185243475)

[Zaman Serisi ve Bileşenlerin Oluşturulması: 21](#_Toc185243476)

[Regresyon Modeli 22](#_Toc185243477)

[Modelin Geliştirilmesi (İkinci Model) 23](#_Toc185243478)

[Durbin-Watson Testi 24](#_Toc185243479)

[Tahmin ve Hata Serileri 25](#_Toc185243480)

[Çarpımsal Model 27](#_Toc185243481)

[Sine ve Cosine Bileşenlerinin Hesaplanması 27](#_Toc185243482)

[Zaman Serisi Grafikleri ve ACF/PACF 31](#_Toc185243483)

[Fark Alınmış Zaman Serisi(1.dereceden) 31](#_Toc185243484)

[Fark Alınmış Zaman Serisi(2.dereceden) 32](#_Toc185243485)

[Toplamsal Winters Yöntemi (Winters1) 33](#_Toc185243486)

[Tahmin ve Görselleştirme 36](#_Toc185243487)

[Hata Serisi ve Box-Ljung Testi 37](#_Toc185243488)

[Tahmin (Forecast) 39](#_Toc185243489)

[Arıma Modeli 40](#_Toc185243490)

[ARIMA Modelinin Bileşenleri 40](#_Toc185243491)

[ACF ve PACF’nin ARIMA Modelindeki Rolü 41](#_Toc185243492)

[Model 1 ( order = c(1,1,2), seasonal= c(1,1,1) 42](#_Toc185243493)

[Model 2 ( order = c(1,1,2), seasonal= c(0,1,1) 44](#_Toc185243494)

[Model 3 ( order = c(2,1,3), seasonal= c(1,1,2) 46](#_Toc185243495)

[Model 4 ( order = c(2,1,3), seasonal= c(0,1,0) 48](#_Toc185243496)

[Model 5 ( order = c(2,1,3), seasonal= c(1,1,1) 50](#_Toc185243497)

[Model 6 ( order = c(2,1,3), seasonal= c(0,1,1) 52](#_Toc185243498)

[Sonuç olarak 54](#_Toc185243499)

[Best Model in RStudio 54](#_Toc185243500)

[Tahminlerin ve Hata Terimlerinin Çekilmesi 56](#_Toc185243501)

[Gerçek Veriler ve Tahminlerin Görselleştirilmesi 57](#_Toc185243502)

[Öngörü(gelecek tahmini) 58](#_Toc185243503)

[Autocorrelation (ACF) ve Partial Autocorrelation (PACF) Grafikleri 59](#_Toc185243504)

[Rstudio Kodlama Görselleri 60](#_Toc185243505)

[Kaynakça 63](#_Toc185243506)

# Veri Hakkında

### 1.Veri Setinin Genel İçeriği

Bu veri seti, **1992 yılının Ocak ayından 2023 yılının Aralık ayına kadar** geçen 32 yıllık bir zaman dilimini kapsamaktadır ve toplamda **384 aylık veri** içermektedir. İki sütundan oluşan bu veri seti, **tarih** ve **satış** bilgilerini içermektedir. Verinin amacı, belirli bir ürün veya hizmetin zaman içindeki satış performansını analiz etmek, dönemsel trendleri belirlemek ve geleceğe yönelik tahminler yapmak için kullanılabilir. Bu veri, özellikle perakende, üretim veya pazarlama alanlarında faaliyet gösteren işletmelerin stratejik karar alma süreçlerine ışık tutabilecek niteliktedir.

Veri Setinin Boyutu: Veri seti, toplamda 384 satır ve 2 sütun içermektedir.

Sütunlar:

- Tarih: 1992 Ocak ayından başlayarak 2023 Aralık ayına kadar olan tarih bilgilerini içerir. Veriler, aylık frekansta düzenlenmiştir.

- Satış: Her ay için kaydedilmiş toplam satış miktarı.

Zaman Aralığı: Veri, Ocak 1992 ile Aralık 2023 arasındaki dönemi kapsamaktadır.

### 2. Veri Setinin Amacı

Bu veri seti, aylık satış bilgilerini içererek zaman serisi analizi, satış tahminlemesi ve trend analizi yapmak için kullanılabilir. Verinin çalışma kapsamındaki potansiyel kullanımları şu şekilde sıralanabilir:

- Gelecek dönem satış tahminleri

- Mevsimsellik ve trend analizi

- Performans karşılaştırmaları

- Olası anormalliklerin tespit edilmesi

### 3. Veri Yapısı ve Özellikleri

Tarih Sütunu:

- Veri setinde tarih bilgisi yyyy/ay/gün formatında kayıt altına alınmıştır (2023/12/01 gibi).

- Aylık düzende herhangi bir eksik tarih bilgisi bulunmamaktadır.

Satış Sütunu:

- Her satırda bir ayın toplam satış bilgisi yer almaktadır.

- Satış verileri numerik formatta kaydedilmiştir ve içerik olarak büyük ihtimalle pozitif tam sayılardan oluşmaktadır (ancak bu durum kontrol edilecektir).

### 4. Zaman Serisi Yapısı

- Veri, zamana bağlı olarak düzenli bir yapıya sahiptir.

- Aylık frekans korunmuş ve herhangi bir tarih atlaması bulunmamaktadır.

- Bu düzgün yapı, zaman serisi analizine uygun bir zemin sunmaktadır.

### 5.Verinin Kullanımına Dair Potansiyel Yöntemler

- Trend Analizi: Zaman içerisinde satış artışı ya da azalma trendlerinin belirlenmesi.

- Mevsimsellik Analizi: Satışların belirli dönemlerdeki dalgalanmalarını analiz etmek.

metin, menü, sayı, numara, doküman, belge içeren bir resim

Açıklama otomatik olarak oluşturuldu- Tahminleme: Gelecek aylardaki satışları öngörmek için istatistiksel ya da makine öğrenmesi yöntemleri kullanılabilir.

# Zaman Serisi Oluşturma ve Görselleştirme

Bu bölümde, zaman serisi verilerini oluşturma ve bu verileri görsel olarak analiz etme adımları ele alınmıştır. Veri seti, 1992 yılı Ocak ayından başlayarak 2023 yılı Aralık ayına kadar olan aylık periyotlarla kaydedilmiştir.

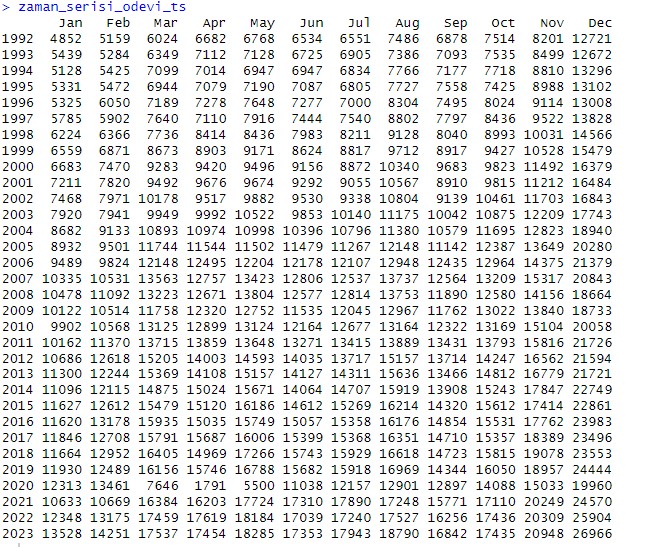
### Veri Dönüştürme

Veri üzerinde şu dönüştürme işlemleri yapılmıştır:

- Tarih Sütunu: Veri setindeki `observation\_date` sütunu, uygun bir zaman serisi yapısı için tarih formatına (‘YYYY-MM-DD’) dönüştürülmüştür.

- Satış Verileri: `MRTSSM4481USN` adlı sütun, ilgili tarih aralığına karşılık gelen aylık satış değerlerini barındırmaktadır.

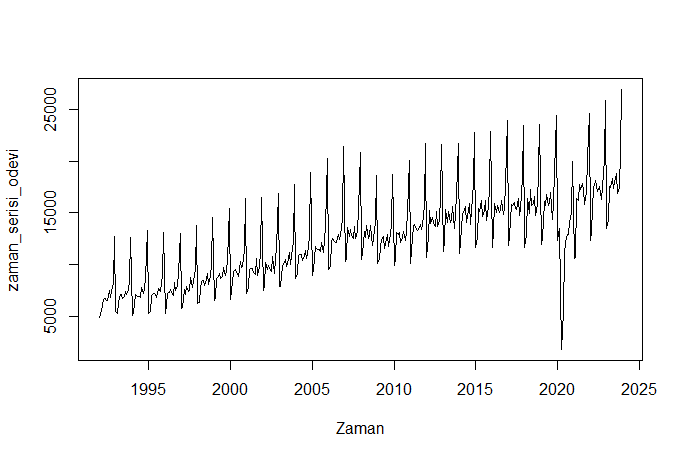
### Zaman Serisi Oluşturma

Tarih ve satış değerleri kullanılarak bir zaman serisi objesi oluşturulmuştur. Zaman serisinin başlangıc yılı 1992, başlangıc ayı ise Ocak olarak belirlenmiş; periyodiklik aylık (‘12’) olarak tanımlanmıştır

### Zaman Serisi Grafiği

Verinin zaman serisi görselleştirilmesi için bir grafik oluşturulmuştur. Grafikte:

- X ekseni: Zamanı (1992-2023) temsil etmektedir.

- Y ekseni: Aylık satış değerlerini göstermektedir.

Yukarıdaki grafik, satış değerlerinin zaman içerisindeki değişimini ve genel trendlerini incelemek için kullanılmıştır. Grafik, serinin zamana bağlı dalgalanmalarını net bir şekilde ortaya koymaktadır. Trend ve mevsimsellik gibi özelliklerin detaylı analizi bir sonraki bölümde ele alınacaktır.

### Otomatik Korelasyon Fonksiyonu (ACF)

### Kısmi Otomatik Korelasyon Fonksiyonu (PACF)

### 

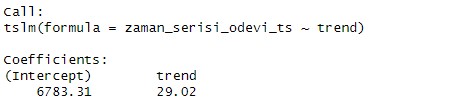
**Mevsimsellik**: Mevsimsel etkilerin bulunması, verilerde belirli bir periyodik düzenin olduğunu gösterir. Örneğin, yılın belirli aylarında artan ya da azalan satışlar gibi. Eğer ACF grafiğinizde belirli aralıklarla düzenli zirveler varsa (örneğin her 12. gecikmede bir), bu mevsimselliğin işaretidir.

**Trend**: Bu, verilerde uzun vadeli bir artış veya azalış eğiliminin bulunduğunu gösterir. Trendin varlığı, serinin sabit bir ortalama değere sahip olmadığı, daha çok zamanla değişen bir yapıya sahip olduğu anlamına gelir. Bu durumda, modelleme sürecinde **trend bileşeni** dikkate alınmalıdır.

# Modelleme

Zaman serisi analizi, verilerin zaman içindeki değişimini anlamak ve gelecekteki değerleri tahmin etmek için kullanılan önemli bir tekniktir. Bu analizde, zaman serisini bileşenlerine ayırarak trend, mevsimsellik ve hata gibi faktörleri incelemek yaygın bir yaklaşımdır. **Trend bileşeni**, zamanla sürekli artan veya azalan bir eğilimi gösterirken, **mevsimsel bileşen** verinin belirli dönemlerdeki periyodik değişimlerini ortaya koyar. **Hata bileşeni** ise modelin tahminlerinde oluşan rastgele sapmaları temsil eder. Bu bileşenlerin doğru bir şekilde ayrılması, verilerin daha iyi anlaşılmasını ve gelecekteki değerlerin daha güvenilir şekilde tahmin edilmesini sağlar.

### Trend Bileşeninin Hesaplanması

Burada, trend değişkeni zaman serisinin zamanla değişen eğilimini temsil eder. Bu trend bileşeni, zaman serisinin gelecekteki değerlerinin artış veya azalış yönündeki genel eğilimini gösterir.

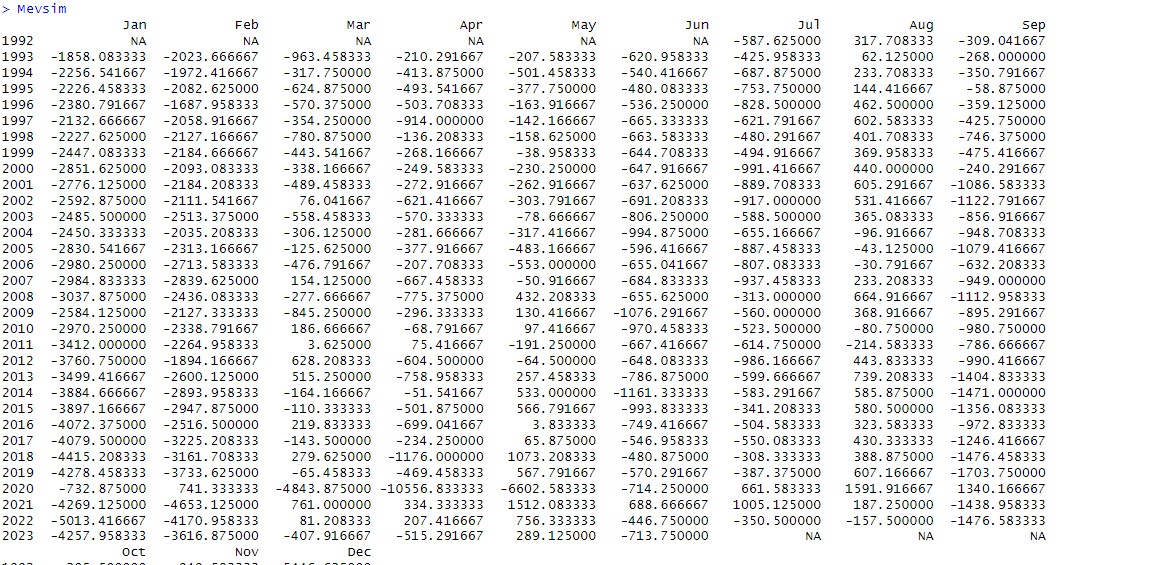
### Mevsimsel Bileşenin Hesaplanması

Periyot değişkeni zaman serisinin **mevsimsel bileşenini** temsil eder, çünkü trend çıkarılmıştır ve geriye kalan değişkenlik, serinin periyodik yapısını gösterir. ACF (Autocorrelation Function) kullanılarak mevsimsel periyot belirlenebilir:

### Merkezsel Hareketli Ortalama ve Mevsimsel Bileşen

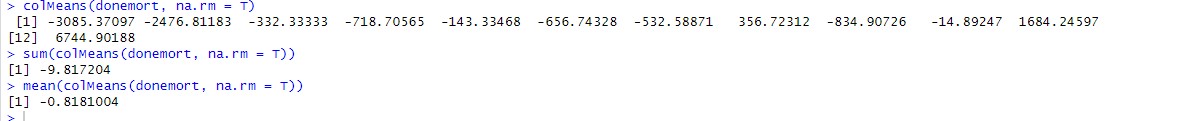
Merkezsel hareketli ortalama (MA) kullanılarak zaman serisinin **trend bileşeni** ile mevsimsel bileşen ayrılır:

Bu işlem, serinin her bir dönemine ait **mevsimsel bileşenleri** hesaplar.

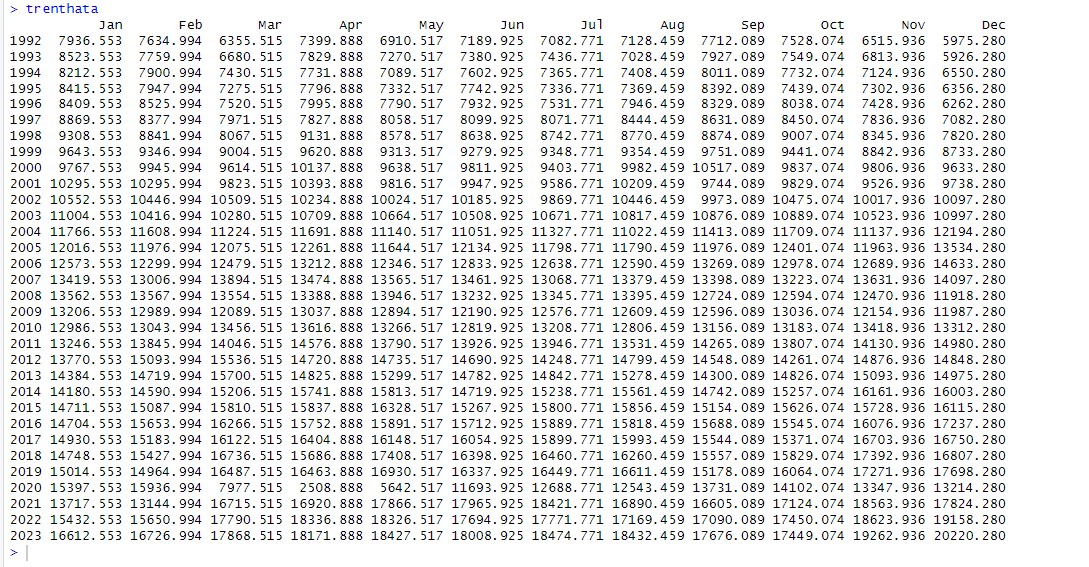


### Mevsimsel Endeks Hesaplaması

Mevsimsel endeks, her dönemin ortalamasına göre ayarlanır ve mevsimsel etkiyi gösterir, bu işlem, her dönemin (ay, çeyrek, yıl gibi) mevsimsel etkilerini ortaya çıkarır. Bu endeksler, zaman serisinin **mevsimsel düzenini** daha iyi anlamamızı sağlar.



### Hata Bileşeni

Bu hata, modelin tahmin ettiği değerler ile orijinal zaman serisi arasındaki farkı temsil eder.

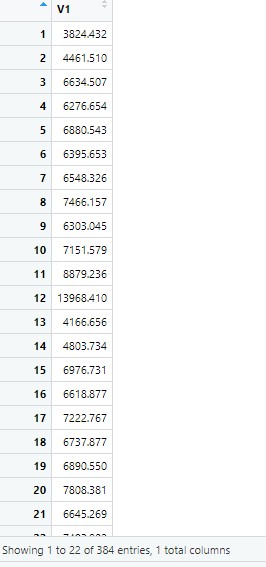
### Zaman Serisinin Trend Değerlerinin Hesaplanması

Bu satır, zaman serisinin **zaman bilgisini** (time()) çıkarır. Zaman serisinde her bir gözlemin hangi zamanda (yıl, ay, gün) olduğunu belirtir. Bu, trend bileşenini modellemek için temel zaman bilgisidir.

### Doğrusal Regresyon Modeli Kurma

Bu satırda, **doğrusal regresyon** modeli kuruluyor. lm() fonksiyonu, trenthata (trend bileşeni çıkarılmış orijinal veri) ile trend\_ts (zaman serisinin zaman bilgisi) arasındaki ilişkiyi öğrenir ve modelin parametrelerini hesaplar. Bu işlem, zaman serisinin **trend bileşenini** daha iyi modelleyebilmek için kullanılır.

### Tahmin Serisini Hesaplama (Mevsimsel Endeks + Trend)

Bu satırda, **mevsimsel endeks** (indeks) ve modelin **fitted values** (doğrusal regresyon modelinin tahmin ettiği trend değerleri) toplanarak zaman serisinin tahmin edilen değerleri oluşturulur. Bu, zaman serisinin **trend ve mevsimsel bileşenlerinin** birleşimidir.

### Hata Serisinin Hesaplanması

Son olarak, **hata serisi** hesaplanır. Bu, orijinal zaman serisi ile tahmin edilen değerler arasındaki farktır. Yani, modelin **trend** ve **mevsimsel** bileşenlere dayanarak yaptığı tahminle, gerçek zaman serisi arasındaki farkları temsil eder.

#### Özetle:

Bu adımlar, zaman serisinin trend bileşenini çıkararak ve regresyon kullanarak **tahmin serisini** oluşturur. Sonrasında, **hata serisi** elde edilir, bu da modelin doğruluğunu değerlendirmek için kullanılabilir. Bu süreç, zaman serisinin daha iyi anlaşılmasını sağlar ve gelecekteki değerlerin tahmin edilmesine yardımcı olur.

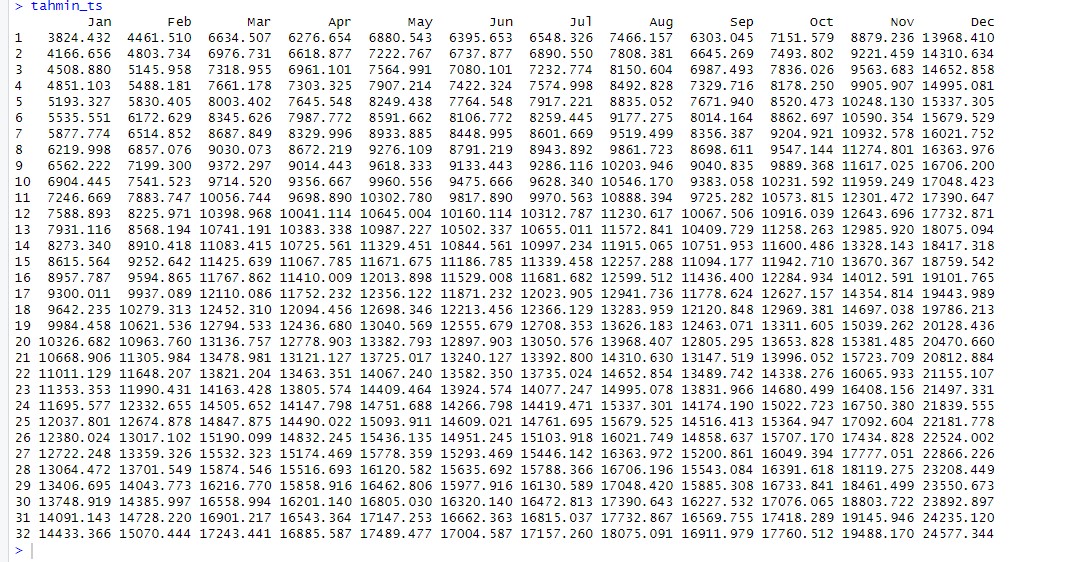
# Zaman Serisi Modelinin Güvenilirliğinin Değerlendirilmesi

### Modelin Geçerliliği

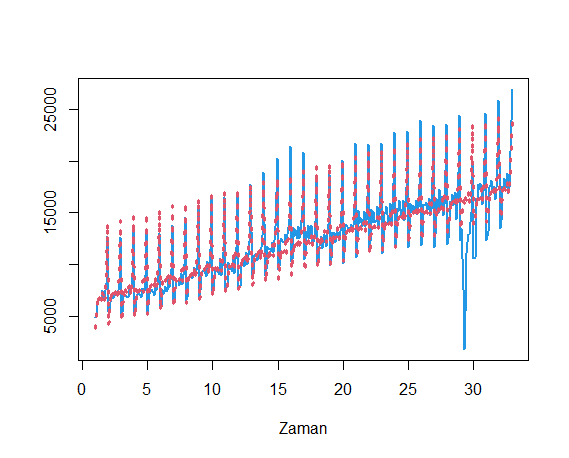
Zaman serisi analizi sırasında, modelin tahminlerinin güvenilir olup olmadığını belirlemek için çeşitli adımlar izlenmiştir. Bu adımlar, modelin doğruluğunu görsel ve istatistiksel olarak değerlendirmenin yanı sıra, hata serisinin düzgün dağılıp dağılmadığını kontrol etmeyi de amaçlamaktadır.

### Tahmin Serisinin Zaman Serisi Formatına Dönüştürülmesi

Modelin tahmin edilen değerleri de zaman serisi formatına dönüştürülerek, orijinal zaman serisiyle karşılaştırılabilir hale getirilmiştir



### Zaman Serisi ve Tahminin Karşılaştırılması

Orijinal zaman serisi ile tahmin edilen serinin karşılaştırılması görsel olarak yapılmıştır. Her iki seri aynı grafik üzerinde gösterilmiş ve tahmin edilen değerler, orijinal seriye eklenerek, modelin doğruluğu gözlemlenmiştir.

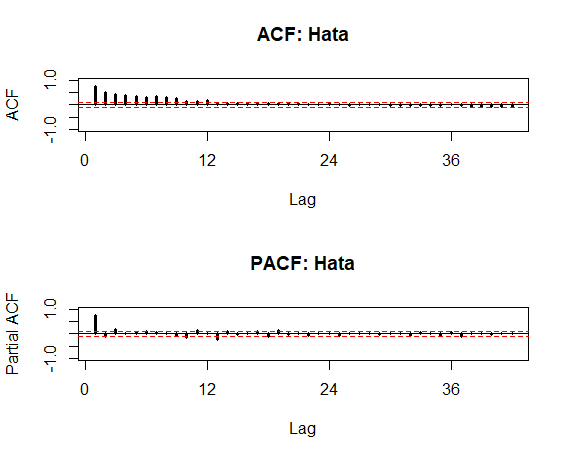
Bu grafikte, **mavi çizgi** orijinal zaman serisini, **kırmızı kesikli çizgi** ise tahmin edilen seriyi temsil etmektedir.

Grafik, tahminlere bakıldığında kırılım noktası dışında oldukça yakın ve benzer bir yapı sunmuştur.

### ACF ve PACF Analizleri

Modelin hata serisinin güvenilirliğini değerlendirmek için **ACF (Otomatik Korelasyon Fonksiyonu)** ve **PACF (Kısmi Otomatik Korelasyon Fonksiyonu)** grafiklerine başvurulmuştur.

* **ACF Grafiği**, hata serisinin **otokorelasyonunu** gösterir ve herhangi bir anlamlı bağımlılık olup olmadığını kontrol eder. Eğer hata serisinde önemli bir korelasyon varsa, modelin yetersiz olduğu sonucuna varılabilir.
* **PACF Grafiği**, hata serisinin **kısmi korelasyonunu** gösterir ve aradaki diğer gözlemlerden bağımsız olarak, her bir gözlemin diğerleriyle olan ilişkisini ölçer.



#### Sonuçlar

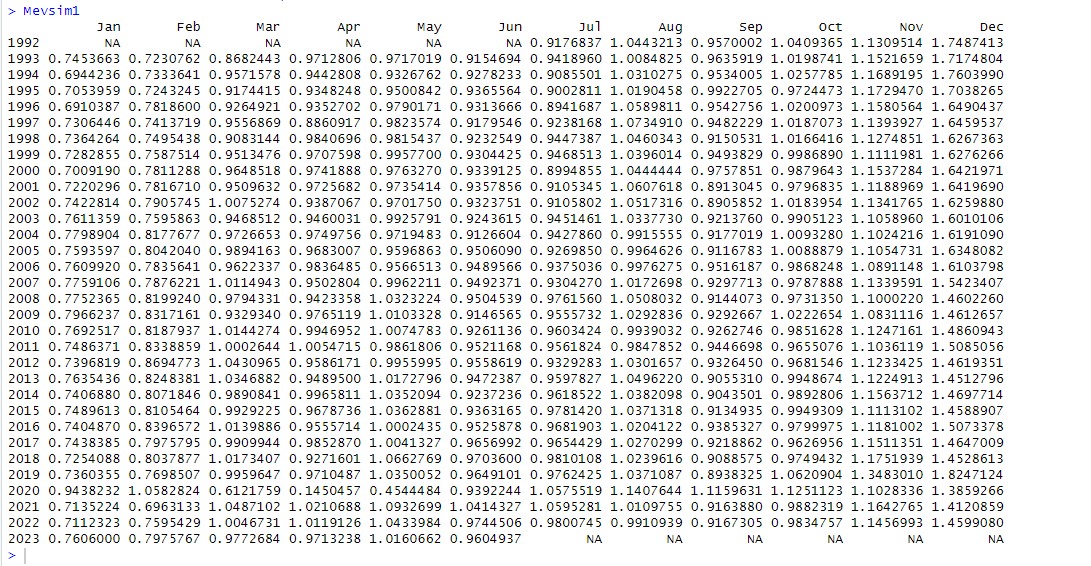
Modelin güvenilirliğini değerlendiren bu adımlar sonucunda:

* Orijinal zaman serisi ile tahmin edilen serinin örtüşmesi, modelin tahminlerinin doğru olduğunu gösterir.
* ACF ve PACF analizleri, hata serisinin rastgele ve bağımsız olduğunu gösterdiğinden, modelin güvenilir olduğu söylenebilir.

# Çarpımsal Ayrıştırma Yöntemi (Multiplicative Decomposition Method)

Bu yöntemle zaman serisinin trend, mevsimsel ve hata bileşenlerini ayrıştırmak amacıyla analiz yapılmaktadır. İşlem sırasıyla her bileşenin güvenilirliği değerlendirilmekte ve modelin doğruluğu kontrol edilmektedir.

### Mevsimsel Bileşenin Bulunması

Çarpımsal ayrıştırma yönteminde, mevsimsel bileşen (Zt) zaman serisinin gözlemleriyle trend bileşeninin bölünmesiyle hesaplanır. Bu işlemle, her bir gözlem değeri, trend bileşenine bölünerek mevsimsel bileşenler elde edilir.

### Mevsimsel Bileşenlerin Ortalama Değerlerinin Hesaplanması

Her periyot (aylık, yıllık vb.) için mevsimsel bileşenlerin ortalamaları hesaplanarak, verideki mevsimsel desenin sürekliliği gözlemlenir. Bu hesaplamayla, her bir periyot için mevsimsel bileşenin ortalama değeri bulunur ve ardından bu ortalamaların genel ortalaması alınarak mevsimsel bileşenin genel yapısı hakkında bilgi edinilir.



metin, yazı tipi, ekran görüntüsü, cebir içeren bir resim

Açıklama otomatik olarak oluşturuldu

### Mevsimsel Endeks Değerlerinin Hesaplanması

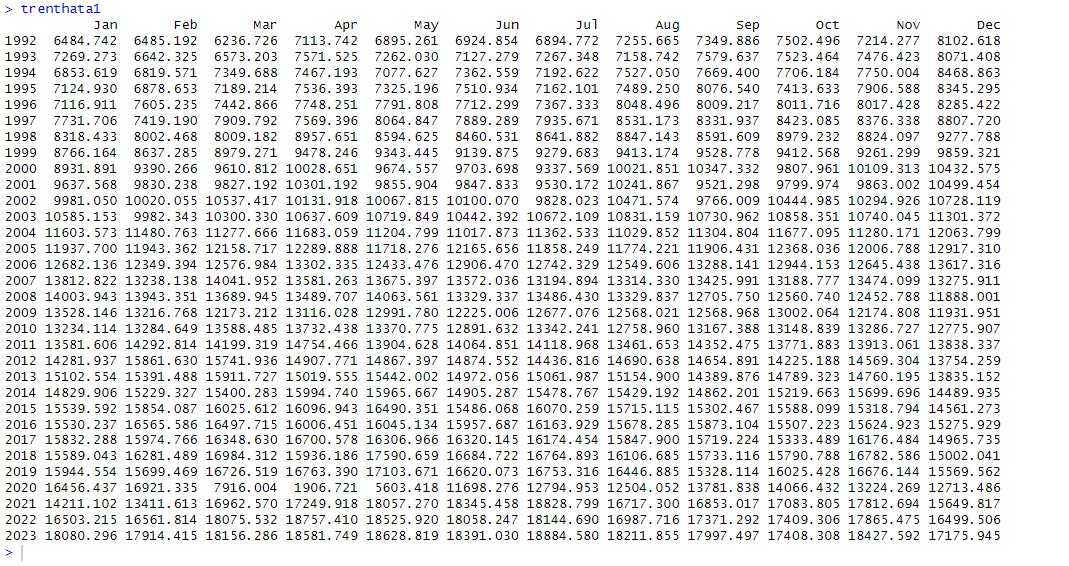
Mevsimsel endeks değerleri, her periyot için hesaplanan ortalamaların, bu ortalamaların genel ortalamasına bölünmesiyle elde edilir.

Bu değerler, her bir periyodun, genel mevsimsel döngü içerisindeki etkisini gösterir. Bu endekslerin, zaman serisi üzerindeki etkilerini gözlemlemek önemlidir.



### Trent Serisi ve Tahminlerin Hesaplanması

Çarpımsal ayrıştırma yönteminde, trend bileşenini elde etmek için orijinal zaman serisi, mevsimsel endekslerle bölünür.

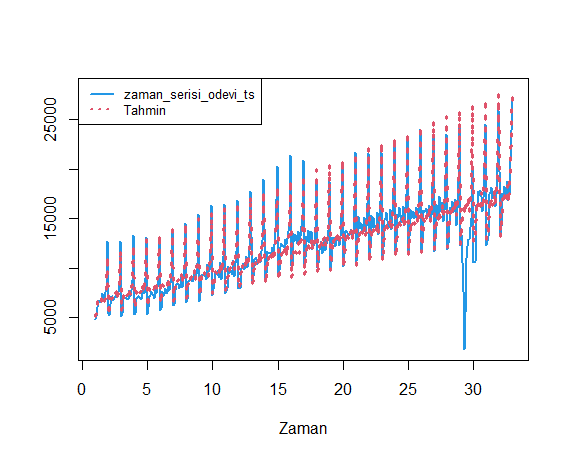
Bu işlemle, zaman serisinin trend bileşeni hesaplanır. Daha sonra, trend bileşeninin doğrusal regresyonu yapılır.

Elde edilen trend bileşeni ve mevsimsel endekslerin çarpımıyla tahmin serisi oluşturulur



### Tahmin Serisinin Zaman Serisi Haline Getirilmesi

Tahmin serisi zaman serisi formatına dönüştürülür ve orijinal zaman serisi ile görsel olarak karşılaştırılır

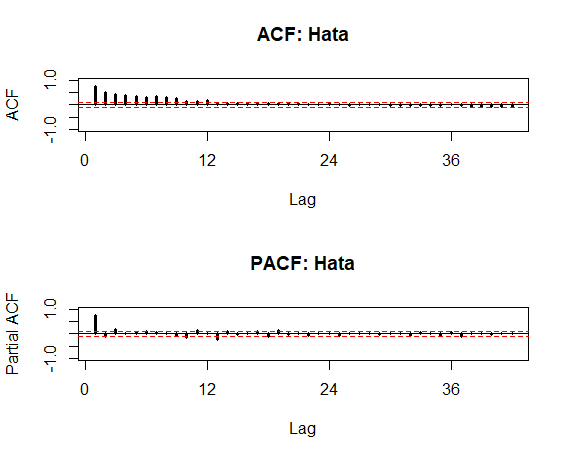


Burada, **mavi çizgi** orijinal zaman serisini, **kırmızı kesikli çizgi** ise tahmin edilen seriyi temsil etmektedir.

* **Genel Uyum:** Orijinal zaman serisi ile tahmin serisi genel olarak birbirine yakın görünüyor. Bu, modelin büyük ölçüde veriyi başarılı bir şekilde tahmin ettiğini gösteriyor. Ancak bazı dönemlerde tahmin serisi ile orijinal seri arasındaki farklar gözlemleniyor.
* Grafikteki bazı büyük dalgalanmalar, aşırı yüksek değerler veya "outlier" olarak nitelendirilebilecek noktalar olabilir. Bu noktalar, modelin bu tür değerleri tahmin etmede zorlanmasına neden olmuş.
* **Hata Serisi Analizi:** Hata serisinin analizi (ACF, PACF) ve daha fazla model iyileştirmesi (örneğin, dışsal değişkenlerin dahil edilmesi) modelin doğruluğunu artırabilir.

### Hata Serisi ve Değerlendirme

Tahmin ve orijinal seri arasındaki farklar hata serisini oluşturulur.Modelin güvenilirliğini daha detaylı incelemek için hata serisinin **ACF (Otomatik Korelasyon Fonksiyonu)** ve **PACF (Kısmi Otomatik Korelasyon Fonksiyonu)** grafiklerine bakılır. Bu grafikler, hata serisinde herhangi bir anlamlı yapının olup olmadığını ve modelin rastgele hatalarla çalışıp çalışmadığını gösterir.



**ACF (AutoCorrelation Function) Analizi:**

* ACF grafiği, hata terimlerinin zaman serisi içindeki korelasyonlarını gösterir. Bu grafikte, çoğu gecikme (lag) değeri için korelasyon sıfır civarındadır. Ancak, ilk birkaç gecikme değeri için belirgin bir korelasyon görülebilir.
* Bu, modelin bazı gecikmeleri (özellikle ilk birkaç dönemi) tahmin ederken hata serisini iyi açıklamadığını, yani hata serisinin belirli bir yapıya sahip olduğunu gösterir.
* ACF'deki bu korelasyonlar, hata terimlerinin zamanla bağımsız olmadığını işaret eder. Bu durum, modelin iyileştirilmesi gerektiğine dair bir uyarı olabilir.

**PACF (Partial AutoCorrelation Function) Analizi:**

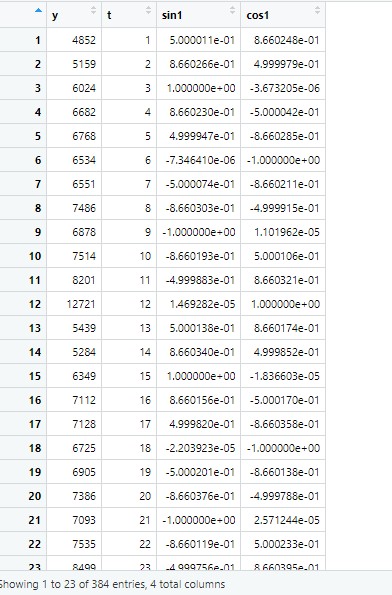
* PACF grafiği, her bir gecikme değeri için, bir önceki gecikmeden bağımsız korelasyonu gösterir. Buradaki veriler, belirli bir gecikme için olan etkileri net bir şekilde izole etmeye yardımcı olur.
* PACF'nin ilk gecikmede sıfır dışındaki değerler, modelde bir düzeyde otokorelasyon olduğunu gösteriyor. Ancak, çoğu gecikme sonrasında sıfır olan değere dönüşüyor, bu da daha ileri gecikmelerde modelin yeterince açıklama yapamadığını gösterir.

# Toplamsal Model (Additive Model)

### Zaman Serisi ve Bileşenlerin Oluşturulması:

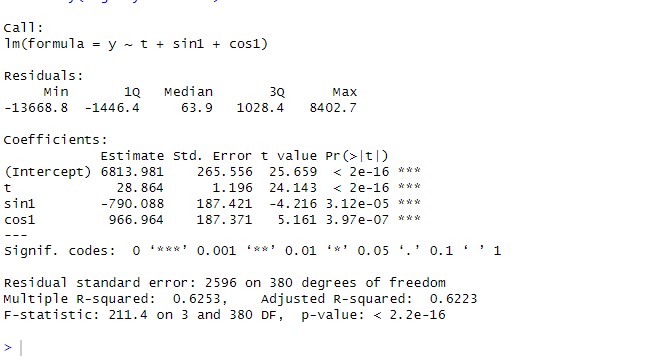
İlk olarak, zaman serisi verisinden t, sin1, ve cos1 bileşenleri oluşturulmuştur. Bu bileşenler, mevsimsellik ve trend etkilerini modellemek için kullanılır:

* **t:** Zaman değişkeni (1'den 384'e kadar olan sayılar).
* **sin1:** Zamanın sinüs fonksiyonu, periyodik etkilerin modellemesi için kullanılır.
* **cos1:** Zamanın kosinüs fonksiyonu, yine periyodik etkilerin modellemesi için gereklidir



### Regresyon Modeli

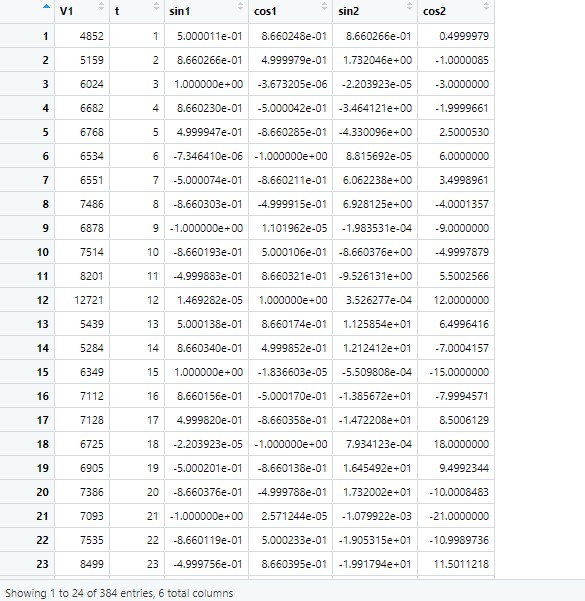
Bu bileşenlerle birlikte aşağıdaki regresyon modeli kuruldu: y=t+sin1+cos1y = t + sin1 + cos1y=t+sin1+cos1

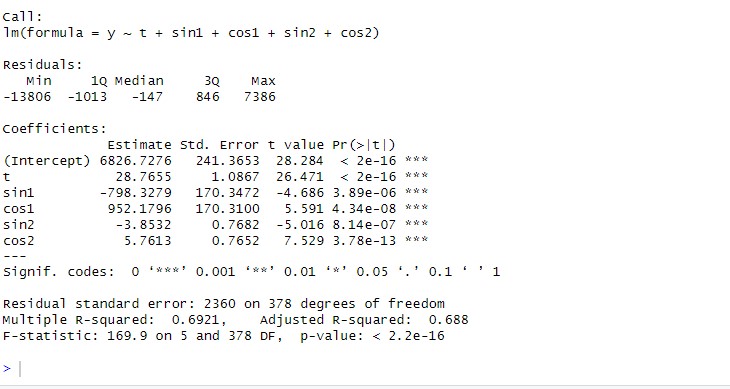
Burada y, zaman serisinin gözlemlerini temsil etmektedir. Regresyon modelinin özeti summary(regresyon.model1) ile elde edilmiştir.

* Bu model, zaman serisindeki trend ve mevsimsel bileşenleri (sinüs ve kosinüs) dikkate alır.
* Sonuçlara bakılarak modelde yer alan bileşenlerin anlamlılık durumları incelenir.

### Modelin Geliştirilmesi (İkinci Model)

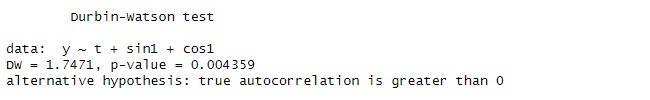
Modelde daha karmaşık bir mevsimsel yapı sağlamak için sin2 ve cos2 bileşenleri eklenmiştir. Ancak modelin özetine bakıldığında bu bileşenlerin anlamlı olmadığı gözlemlenmiştir. Bu nedenle sadece t, sin1, ve cos1 kullanılarak birinci modele dönülmüştür.





### Durbin-Watson Testi

Durbin-Watson testi, modelin hata terimlerinin bağımsız olup olmadığını test etmek için kullanılır. Bu testin sonucu modelin doğruluğu hakkında fikir verir.



* **DW = 1.7471**: Bu değer, 2'ye yakın bir değer olup, genellikle bağımsızlık için olumlu bir işaret olarak kabul edilir. 2, tam bağımsızlık anlamına gelir. 1.7471, 2'den biraz daha küçük, bu da hataların bir miktar pozitif otokorelasyona sahip olduğunu gösterebilir. Ancak, bu değer hala çok düşük değil, yani belirgin bir bağımlılık yok.
* **p-değeri = 0.004359**: Bu p-değeri oldukça küçük (genellikle 0.05'ten küçük kabul edilir), bu da hataların bağımsız olmadığını gösterir. Başka bir deyişle, null hipotez (hataların bağımsız olduğu hipotez) reddedilir. Bu, modelin hata terimlerinin sıfırdan farklı bir otokorelasyona sahip olduğuna işaret eder.

### Tahmin ve Hata Serileri

#### Tahmin Serisi:

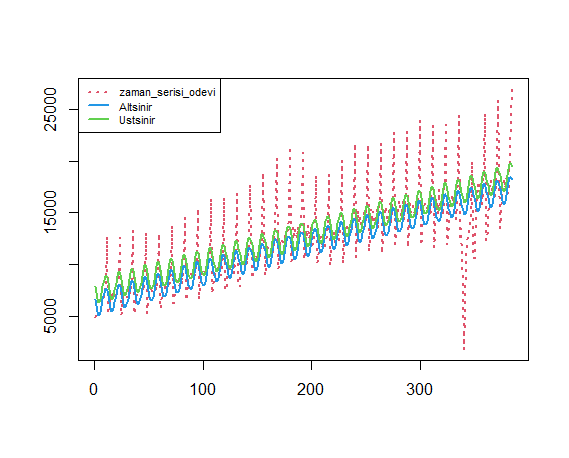
Regresyon modeline göre tahminler predict(regresyon.model1) ile elde edilmiştir. Bu tahminler, modelin gelecekteki değerlerini yansıtır.

#### Güven Aralıkları:

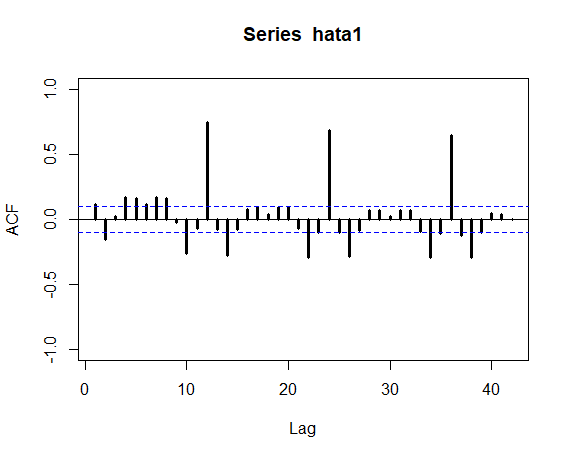
Tahminler için güven aralıkları hesaplanmıştır. interval = 'confidence' ile bu aralıklar %95 güvenle belirlenmiştir.

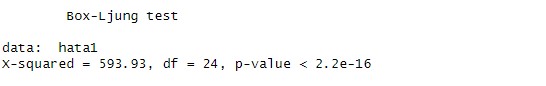
#### Hata Serisi:

Modelin tahminleriyle gerçek değerler arasındaki farkları (hataları) temsil eder. Hata serisi resid(regresyon.model1) ile hesaplanmış ve ACF (Autocorrelation Function) grafiği ile bağımsızlıkları test edilmiştir.



Bu grafik, zaman serisi modelinizin tahmin ettiği değerlerle orijinal veri arasında nasıl bir ilişki olduğunu ve tahminin güven aralıklarını nasıl kapsadığını gösteriyor. **Orijinal veri** kırmızı çizgi ile, **tahminin alt ve üst sınırları** ise mavi ve yeşil çizgilerle gösterilmektedir.





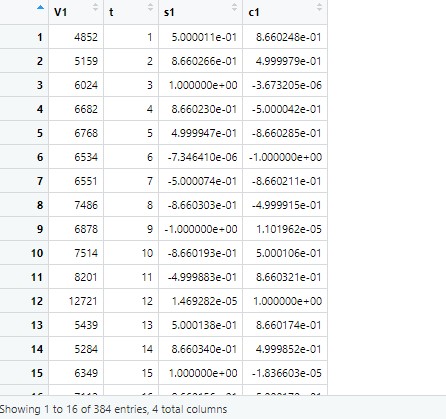
* Box-Ljung testi, zaman serisi verilerinde otokorelasyon (geçmiş değerlerin geleceği nasıl etkilediği) olup olmadığını test eder. Bu test, özellikle hataların (residuals) otokorelasyonunu kontrol etmek amacıyla kullanılır.
* **X-squared (test istatistiği) 593.93** oldukça büyük bir değerdir.
* **p-değeri < 2.2e-16** çok küçüktür.

Bu p-değeri çok küçük olduğu için, **H0 hipotezi reddedilir** ve **hatalarda otokorelasyon olduğu** sonucuna varılır. Yani modelin hataları birbirinden bağımsız değildir ve bu da modelin doğruluğuna dair bir sorun işareti olabilir.

# Çarpımsal Model

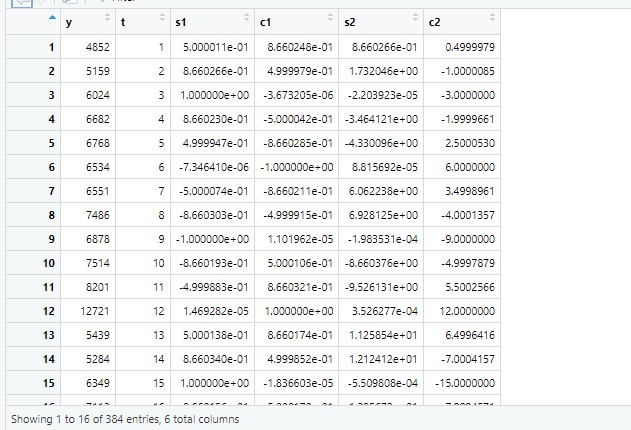
### Sine ve Cosine Bileşenlerinin Hesaplanması

Burada, zaman serisindeki mevsimsellik etkilerini modellemek için iki adet sinüs ve kosinüs bileşeni hesaplanır. Bu bileşenler, dönemsellik veya mevsimsel varyasyonları modellemek için sıklıkla kullanılır.



* metin, ekran görüntüsü, yazı tipi, doküman, belge içeren bir resim

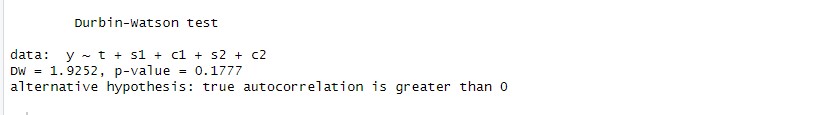
  Açıklama otomatik olarak oluşturulduBurada, t, s1 ve c1 değişkenlerinin her biri modeldeki bağımsız değişkenlerdir. Bu modelin çıktısı, her değişkenin model üzerindeki etkisini, anlamlılıklarını ve genel modelin doğruluğunu gösterir.



* Sin2 ve Cos2 hesaplamaları ile birlikte mevcut veri üzerinde y,t,s1,c1,s2,c2 kısaltmaları ile veri tablomuz yeniden şekillenir.



* Hem s2 hem de c2 değişkenlerinin p-değerleri 0.05'ten çok küçük olduğu için **her iki değişken de anlamlıdır**. Yani, bu iki bileşen modeldeki bağımsız değişkenlerin üzerinde önemli bir etkiye sahiptir.
* Yer alan olumlu etki dolayısıyla durbin-watson testi y~t+s1+c1+s2+c2 şeklinde yapılacaktır.



**Durbin-Watson değeri:**

* **DW = 2:** Bu durum, hiçbir otokorelasyon olmadığını (bağımsızlık) gösterir.
* **DW < 2:** Pozitif otokorelasyon olduğunu gösterir.
* **DW > 2:** Negatif otokorelasyon olduğunu gösterir.

Verdiğiniz DW değeri **1.9252** olup **2'ye yakın** bir değerdir. Bu, **pozitif otokorelasyon** olma eğiliminde olduğunu, ancak otokorelasyonun çok belirgin olmadığını gösteriyor.

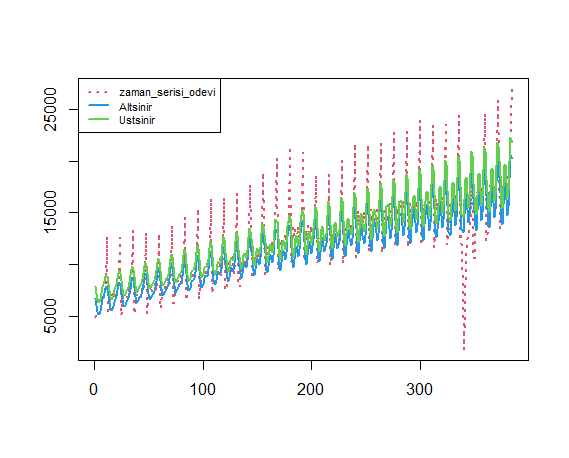
**p-değeri:**

* p-değeri **0.1777**, yani **0.05'ten büyük**.
* Bu, **H0 hipotezinin reddedilemeyeceği** anlamına gelir. Yani, verilerinizde **istatistiksel olarak anlamlı bir otokorelasyon** bulunmamaktadır.

**Sonuç:**

* **Otokorelasyon yok** ya da **çok zayıf**. Modelin hata terimlerinde bağımsızlık var gibi görünüyor.
* p-değeri 0.05'ten büyük olduğu için, **otokorelasyon olmadığına dair hipotez reddedilemiyor**.

Bu sonuç, modelinizin hata terimlerinin birbirine bağımlı olmadığı, yani modelin genellikle düzgün bir şekilde çalıştığına işaret eder.



* Zaman serisi verisinin **gerçek değerleri** kırmızı çizgiyle gösteriliyor. Bu çizgi, gözlemlerinizin zaman içindeki değişimini temsil eder.
* **Tahminin alt ve üst sınırları**, modelin tahminleri için güven aralıklarını gösterir.
* Alt sınır (mavi çizgi) ve üst sınır (yeşil çizgi), tahminler için güven aralığının sınırlarını belirler ve modelin tahminindeki belirsizliği yansıtır.
* Zaman serisi verisinin üst ve alt sınırları arasında kalan alan, tahminin güven düzeyini gösterir. Bu alan ne kadar genişse, modelin tahmininin belirsizliği o kadar fazla demektir.
* Grafik, modelin tahminlerinin doğruluğunu ve güven aralıklarının genişliğini görsel olarak analiz etmenizi sağlar.

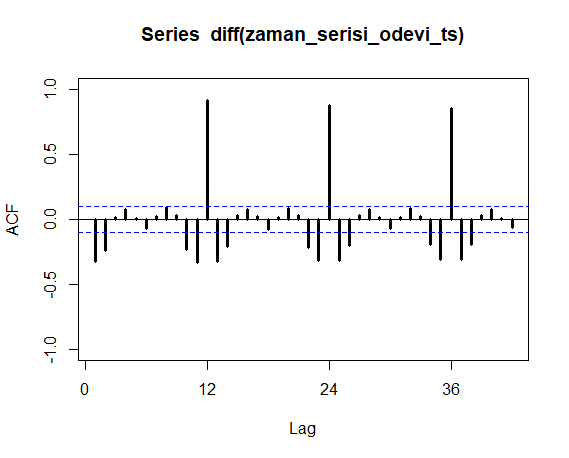
# Zaman Serisi Grafikleri ve ACF/PACF

* **ACF Grafiği**: Zaman serisinin otokorelasyon fonksiyonu. İlk birkaç gecikme için pozitif veya negatif otokorelasyon olup olmadığına bakılır.
* **PACF Grafiği**: Zaman serisinin kısmi otokorelasyon fonksiyonu. Mevsimsel ya da trend etkilerini düzeltmiş otokorelasyonları gösterir.

### Fark Alınmış Zaman Serisi(1.dereceden)

Bir zaman serisinin **istasyonar** (yani ortalama, varyans ve kovaryans zamanla değişmeyen) olup olmadığını anlamak için, genellikle seriden fark alınır. Fark alma işlemi, ardışık gözlemler arasındaki farkları hesaplar ve zaman serisinin trend ve mevsimsellik gibi özelliklerini ortadan kaldırmaya çalışır.

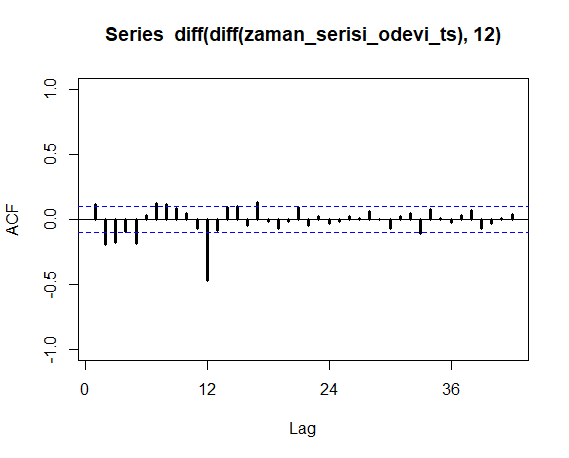
Birinci fark, zaman serisinde olan eğilimi (trend) ortadan kaldırır(arındırır).

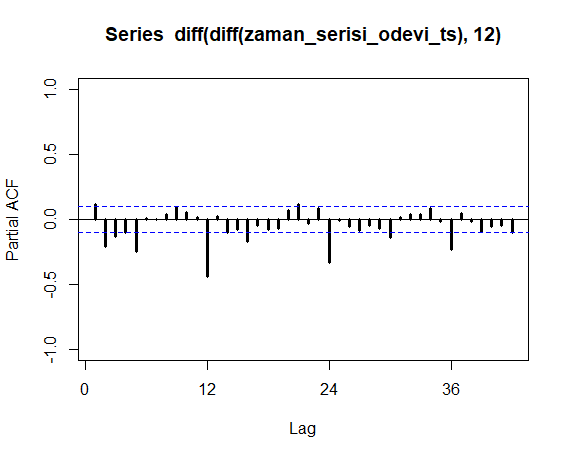


metin, ekran görüntüsü, çizgi, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

### Fark Alınmış Zaman Serisi(2.dereceden)

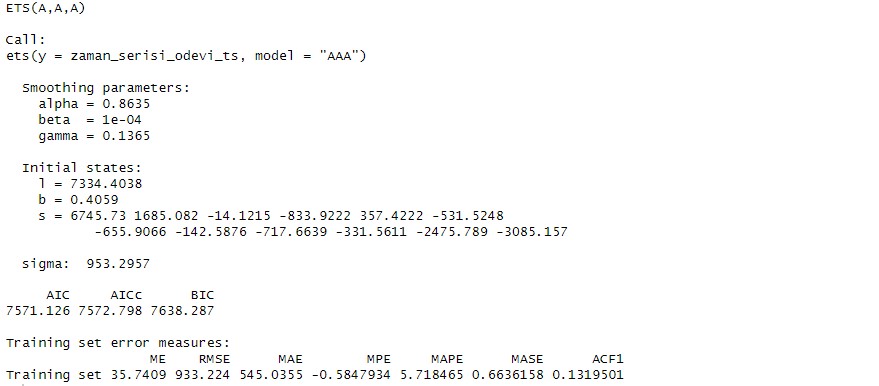
İkinci fark, daha uzun dönemli eğilimleri (mevsimsel etkiler gibi) ortadan kaldırabilir. Eğer seride hala bir ilişki varsa, bu serinin istasyonar olmadığına işaret eder.



# Toplamsal Winters Yöntemi (Winters1)

Toplamsal Winters yöntemi, zaman serisinin **düzey (level)**, **eğim (trend)** ve **mevsimsel (seasonality)** bileşenlerini tahmin etmek için kullanılır. Bu yöntem, zaman serisinin hem eğilim hem de mevsimsellik gibi bileşenleri modelleyerek daha doğru tahminler yapmayı sağlar.

* **ets fonksiyonu**: Bu fonksiyon, zaman serisini **exponential smoothing (üstel düzleme)** yöntemine dayalı olarak modellemek için kullanılır.
* **model = "AAA"**: Bu model, zaman serisinin düzey, eğilim ve mevsimsel bileşenlerinin hepsini ele alır ve her biri için **additive (toplamsal)** modelleme yapar.
* **summary(Winters1)**: Modelin özet bilgilerini verir. Burada modelin başlatılma değerleri de yer alır.



**Smoothing Parameters (Düzeltme Parametreleri):**

* **α (alpha)** = 0.8635: **Düzey (level)** parametresi, zaman serisinin mevcut düzeyine ne kadar ağırlık verilmesi gerektiğini belirler. Bu yüksek değeri, modelin mevcut düzeye büyük bir ağırlık verdiğini gösterir.
* **β (beta)** = 0.0001: **Eğim (trend)** parametresi, eğilim bileşeninin ne kadar önem taşıdığını gösterir. Buradaki çok küçük bir değer, eğilim bileşeninin çok küçük bir etkiye sahip olduğunu gösterir.
* **γ (gamma)** = 0.1365: **Mevsimsel (seasonality)** parametresi, mevsimsel etkilerin ne kadar önemli olduğunu belirler. 0.1365 değeri, mevsimsel bileşenlerin dikkate alındığını gösteriyor ancak düzey ve eğilim kadar önemli değil.

**2. Initial States (Başlangıç Değerleri):**

* **l (level)** = 7334.4038: Zaman serisinin **başlangıç düzeyi** değeri.
* **b (trend)** = 0.4059: Başlangıç **eğilim (trend)** değeri.
* **s (seasonal)** = 6745.73, 1685.08, -14.12, -833.92, 357.42, -531.52, ...: İlk **mevsimsel** bileşen değerleri. Bu değerler, zaman serisindeki mevsimsel değişimin başlangıcını ifade eder.

**3. Sigma:**

* **σ (sigma)** = 953.2957: Modelin hata teriminin **standart sapması**. Bu, modelin tahminlerinin doğruluğunu gösterir. Düşük sigma değeri, modelin doğruluğunun yüksek olduğunu gösterir.

**4. AIC, AICc, ve BIC:**

* **AIC (Akaike Information Criterion)** = 7571.126
* **AICc (Corrected AIC)** = 7572.798
* **BIC (Bayesian Information Criterion)** = 7638.287

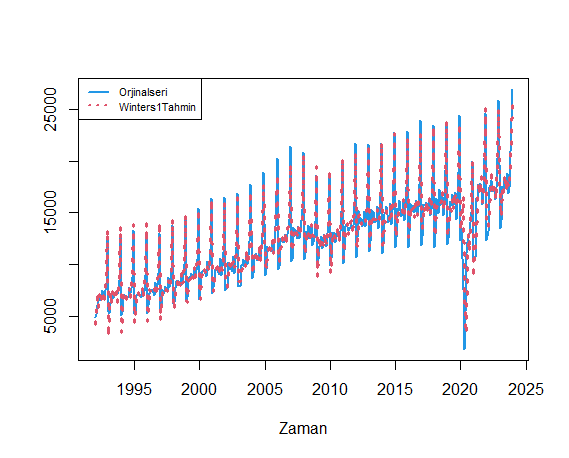
Bu kriterler, modelin **kullanılabilirliğini** ve **gelişmişliğini** değerlendirir. Genellikle, bu değerler küçükse modelin daha iyi olduğunu gösterir. Buradaki **AIC** ve **BIC** değerlerinin düşük olduğunu gözlemliyoruz, bu da modelin uygun olduğunu gösterir

**5. Training Set Error Measures (Eğitim Seti Hata Ölçütleri):**

Modelin eğitim seti üzerindeki performansı, hataları ve doğruluğu gösterir:

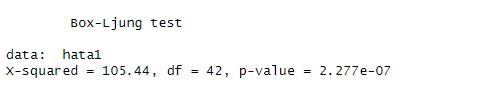
* **ME (Mean Error)** = 35.7409: Ortalama hata. Bu, modelin tahminleri ile gerçek değerler arasındaki farkların ortalamasıdır.
* **RMSE (Root Mean Squared Error)** = 933.224: Kök ortalama kare hatası. Bu, hataların büyüklüğünü gösterir ve daha düşük değerler daha iyi bir model anlamına gelir.
* **MAE (Mean Absolute Error)** = 545.0355: Ortalama mutlak hata. Tahminlerin ortalama mutlak farkını gösterir.
* **MPE (Mean Percentage Error)** = -0.5848: Ortalama yüzdelik hata. Modelin tahminlerinin ne kadar doğru olduğunu yüzdelik olarak gösterir.
* **MAPE (Mean Absolute Percentage Error)** = 5.7185: Ortalama mutlak yüzdelik hata. Bu, modelin ne kadar hatalı olduğunu yüzdelik olarak ifade eder.
* **MASE (Mean Absolute Scaled Error)** = 0.6636: Normalize edilmiş hata ölçütüdür. 1'den küçük olması modelin iyi olduğunu gösterir.
* **ACF1 (Autocorrelation Function at Lag 1)** = 0.1319: 1. gecikme için otokorelasyon değeri. Bu, modelin hatalarının ne kadar birbirine bağlı olduğunu gösterir. Değer 0'a yakınsa, modelin hataları bağımsızdır.

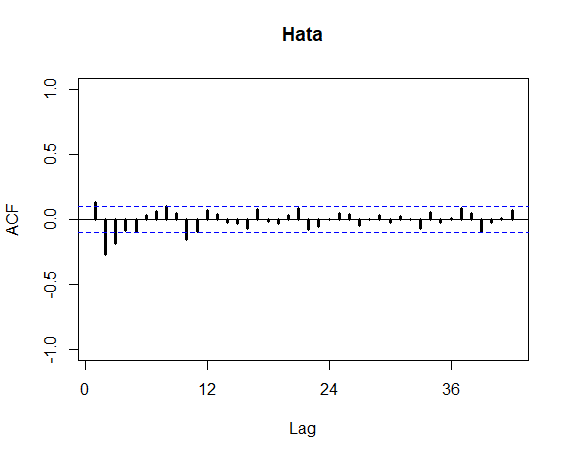
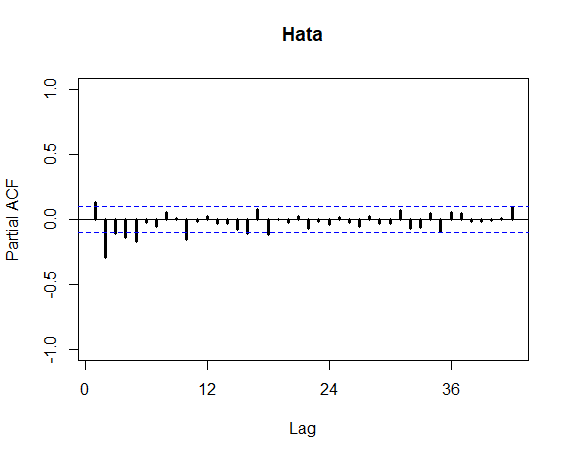
# metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim Açıklama otomatik olarak oluşturulduTahmin ve Görselleştirme

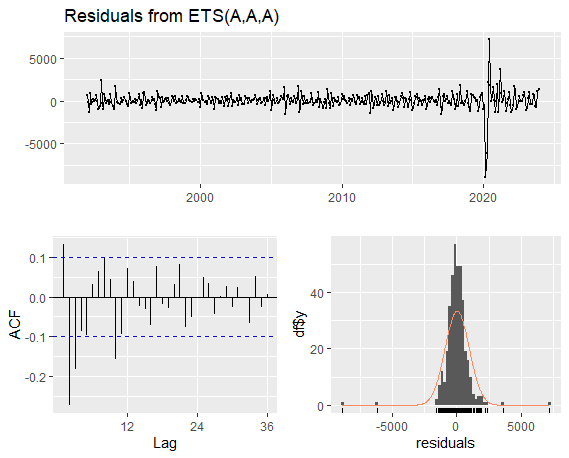
* Winter toplamsal modelleme ile elde ettiğimiz tahmini veriler yukarıdaki gibi yer almaktadır.
* Yukarıda yer alan tahminler ile modelimizin ana değerlerini kıyaslayabilmek ve görsel olarak da verimliliğe şahit olabilmek için aşağıdaki grafikte görüldüğü üzere;
* Mavi çizgi orijinal verimizin zaman serisi grafiğini göstermekte olup , üzerinde yer alan kımızı nokta nokta ilerleyen zaman serisi grafiği ise winter toplamsal modelleme ile elde ettiğimiz grafiği göstermektedir.
* Modelin performansı oldukça iyi görünüyor. AIC ve BIC gibi kriterler düşük, hata ölçütleri de genellikle makul seviyelerde. Modelin hataları bağımsız ve mevsimsel bileşenlerin etkisi de dikkate alınmış. Bu, modelin gelecekteki tahminlerde yüksek doğruluk sağlama potansiyeline sahip olduğunu gösterir.

### Hata Serisi ve Box-Ljung Testi

Bu kodda **Winters1** modelinin hata (residual) analizini yapıyorsunuz. Her adımda modelin tahmin hatalarını kontrol ederek modelin doğruluğunu değerlendirebilirsiniz.

**Box-Ljung testi**, zaman serisinin hata terimlerinin **otokorelasyon** (kendi kendine korelasyon) olup olmadığını test eder. Burada 42 gecikme (lag) kullanılıyor. Eğer p-değeri çok küçükse (genellikle 0.05'in altı), bu, hataların otokorelasyona sahip olduğu anlamına gelir, yani modelin daha iyi düzeltilmesi gerektiğini gösterir.

* **X-squared = 105.44**: Bu, testin istatistiğidir. Box-Ljung testi, serinin otokorelasyonunun sıfır olup olmadığını test eder. Yüksek bir X-kare değeri, serinin otokorelasyonunun sıfırdan farklı olduğunu gösterir.
* **df = 42**: Bu, serinin test edilen gecikme sayısını (lag) ifade eder. Burada 42 gecikme kullanılmış.
* **p-value = 2.277e-07**: p-değeri çok küçük (yaklaşık 0) ve bu, **hipotez testinin sonucunun anlamlı** olduğunu gösterir. Yani, **p-value < 0.05** olduğu için null hipotez reddedilir ve bu da hatalar arasında **otokorelasyon** olduğunu, yani modelin hatalarının birbirleriyle ilişkili olduğunu ifade eder.
* **ACF**: **Otokorelasyon Fonksiyonu (ACF)**, hata terimlerinin kendileriyle olan korelasyonunu ölçer. ACF'nin değerleri sıfıra yakın olduğundan, hataların birbirleriyle ilişkisi yok demektir. Bu, modelin hatalarının bağımsız olduğunu ve modelin iyi olduğunu gösterir.
* **PACF**: **Kısmi Otokorelasyon Fonksiyonu (PACF)**, sadece ilk gecikme için otokorelasyonu ölçer ve diğer gecikmeleri kontrol eder. PACF'nin sıfıra yakın olması, modelin hatalarının sıfır olmayan otokorelasyonu olmadığını gösteri



### Tahmin (Forecast)

* **forecast(Winters1, h = 5)**: **Winters1** modelini kullanarak, zaman serisinin gelecekteki 5 dönemi için tahmin yapar. h = 5 parametresi, 5 dönemin tahmin edileceğini belirtir.
* **ongoru[["mean"]]**: Bu, modelin **ortalama tahmin** değerlerini çıkarır. Model, bu tahminleri gelecekteki değerler olarak sunacaktır.



# Arıma Modeli

ARIMA (Autoregressive Integrated Moving Average), zaman serisi verilerini analiz etmek ve gelecekteki değerleri tahmin etmek için kullanılan bir istatistiksel modeldir. Zaman serilerindeki trend, sezonluk etkiler ve otokorelasyon gibi yapıları modellemek için yaygın olarak kullanılır.

### ARIMA Modelinin Bileşenleri

ARIMA modeli üç temel bileşenden oluşur:

1. **AR (Autoregressive - Otoregresif)**:
   * Zaman serisinin önceki değerlerine (geçmiş verilere) dayalı olarak tahmin yapılır.
   * ppp: Gecikme (lag) sayısı, serideki geçmiş değerlerin modelde kullanıldığı düzeyi ifade eder.
2. **I (Integrated - İntegrasyon/Farklama)**:
   * Zaman serisindeki trendi ve durağan olmayan (non-stationary) yapıyı düzeltmek için fark alınır.
   * ddd: Verinin durağan (stationary) hale gelmesi için alınan fark sayısıdır.
3. **MA (Moving Average - Hareketli Ortalama)**:
   * Zaman serisinin hata terimleri (residuals) arasındaki ilişkilere dayalı tahmin yapılır.
   * qqq: Hata terimlerinin modelde dikkate alınan gecikme sayısını ifade eder.

**Özetle:** ARIMA modeli, önceki verilere ve hata terimlerine dayanarak zaman serisinde tahmin yapar.

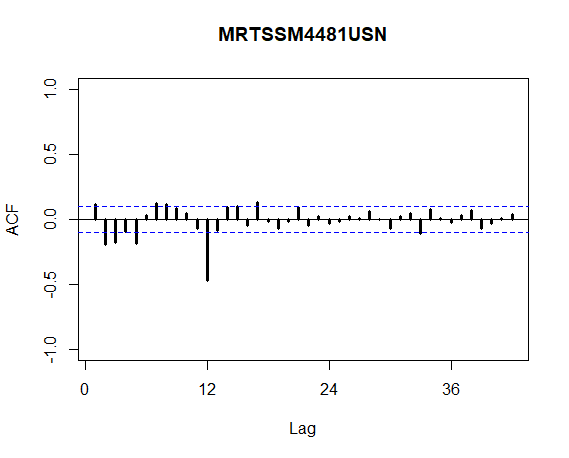
ARIMA, sezonsal (dönemsel) etkileri de modelleyebilir. Bu durumda **SARIMA** (Seasonal ARIMA) adıyla kullanılır. Bu modelin dört ek parametresi vardır:

* PPP: Sezonsal otoregresif terimlerin sayısı.
* DDD: Sezonsal farklama derecesi.
* QQQ: Sezonsal hareketli ortalama terimlerin sayısı.

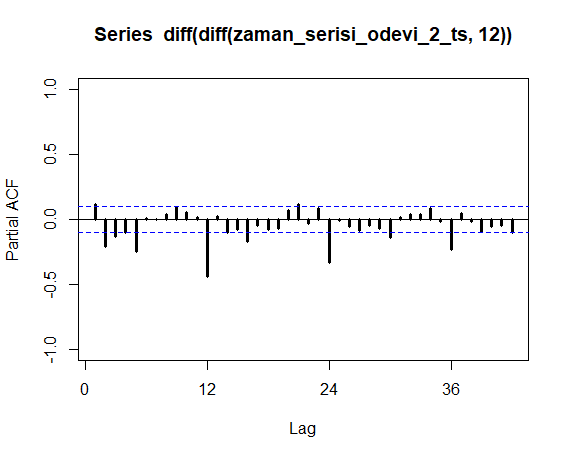
### ACF ve PACF’nin ARIMA Modelindeki Rolü

ACF ve PACF grafiklerinin nasıl yorumlandığına dair temel prensipler:

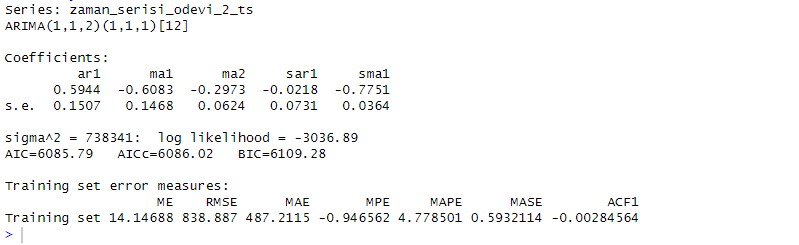
1. **AR (p) Parametresinin Belirlenmesi**:
   * PACF grafiğine bakılır. Eğer PACF hızlı bir şekilde sıfıra düşüyorsa, "cut-off" noktası ppp'yi verir.
   * ACF grafiğinde genelde yavaş bir azalma görülür.
2. **MA (q) Parametresinin Belirlenmesi**:
   * ACF grafiğine bakılır. Eğer ACF hızlı bir şekilde sıfıra düşüyorsa, "cut-off" noktası qqq'yu verir.
   * PACF grafiğinde genelde yavaş bir azalma görülür.
3. **Trend veya Sezonsallık**:
   * Eğer ACF veya PACF’de belirli bir periyodiklik görülüyorsa (örneğin, 12. gecikmede bir sıçrama), sezonsal etkiler modelde dikkate alınmalıdır (D=1D = 1D=1).



Yukarıda verimize ait arındırılmış acf grafiği üzerinden p parametresinin belirlenmesinde yardım alacağız. Bu noktada grafiğe bakıldığında ilk kesim noktası olarak p = 1 alanı alınabilir.

Yukarıda ise pacf grafiği yer almakta olup konuya istinaden incelemeler yapıldığında q=2 olarak modellemeye başlamak uygun olacaktır.

### Model 1 ( order = c(1,1,2), seasonal= c(1,1,1)



**ARIMA(1,1,2)(1,1,1)[12]**

* **AR(1)**: Otoregresif terimi (p=1) için katsayı 0.5944. Bu, geçmiş değerin, şu anki değerle güçlü bir ilişkiye sahip olduğunu gösteriyor.
* **MA(2)**: Hareketli ortalama terimi (q=2) için katsayılar -0.6083 ve -0.2973. Bu, serinin geçmiş hata terimlerinin etkilerini göstermektedir.
* **SAR(1)**: Sezonsal AR terimi (P=1) için katsayı -0.0218. Sezonsal etkiyi modelleyen bu terim çok küçük bir etkiye sahip.
* **SMA(1)**: Sezonsal MA terimi (Q=1) için katsayı -0.7751. Bu da sezonsal hareketli ortalama teriminin seriye etkisini gösteriyor.

**Model İstatistikleri:**

* **Sigma^2 (Hata varyansı)**: 738341. Bu, modelin tahmin hatalarının büyüklüğünü gösterir.
* **Log Likelihood**: -3036.89, modelin veriye ne kadar iyi uyduğunu gösterir. Daha yüksek bir log likelihood değeri, daha iyi uyum anlamına gelir.
* **AIC**: 6085.79, **AICc**: 6086.02, **BIC**: 6109.28, modelin genel uygunluğunu değerlendiren kriterlerdir. Genelde, daha düşük AIC ve BIC değerleri daha iyi bir model gösterir.

**Modelin Hata İstatistikleri (Training Set Error Measures):**

* **ME (Mean Error)**: 14.14688, modelin ortalama hatası. Bu, modelin tahminlerinin sıfırdan ne kadar uzaklaştığını gösterir. Pozitif bir ME, modelin tahminlerinin gerçek değerin üzerinde olduğunu işaret eder.
* **RMSE (Root Mean Square Error)**: 838.887, modelin hata oranı. Daha düşük bir RMSE daha iyi bir uyum gösterir.
* **MAE (Mean Absolute Error)**: 487.2115, modelin ortalama mutlak hatası. Bu da modelin ne kadar büyük hatalar yaptığı hakkında fikir verir.
* **MPE (Mean Percentage Error)**: -0.946562, modelin ortalama yüzde hatası.
* **MAPE (Mean Absolute Percentage Error)**: 4.778501, modelin ortalama mutlak yüzde hatası. Bu, modelin genel doğruluğu hakkında bilgi verir ve düşük bir değer daha iyi anlamına gelir.
* **MASE (Mean Absolute Scaled Error)**: 0.5932114, modelin skalanmış hata oranı.
* **ACF1**: -0.00284564, modelin hata terimlerinin otokorelasyonunu gösterir. Burada ACF1'in sıfıra yakın olması, modelin hatalarının bağımsız olduğunu ve modelin iyi fit ettiğini gösterir.

**Sonuç ve Yorumlar:**

* **Modelin Performansı**: AIC ve BIC değerleri oldukça düşük değil, ancak modelin daha iyi olup olmadığını görmek için başka modellerle karşılaştırılabilir. RMSE ve MAE değerleri de oldukça yüksek, ancak bu serinin yapısına ve veri setinin özelliklerine bağlı olarak kabul edilebilir.
* **Hata Analizi**: ACF1’in sıfıra yakın olması, modelin hata terimlerinin birbirinden bağımsız olduğunu ve modelin iyi çalıştığını işaret eder.

Modelin iyi bir başlangıç olduğunu söyleyebiliriz, ancak daha fazla iyileştirme yapılabilir. Eğer tahminlerin doğruluğunu artırmak istersek, parametreleri yeniden ayarlayabilir veya başka modelleme teknikleri deneyebiliriz.

### Model 2 ( order = c(1,1,2), seasonal= c(0,1,1)



**Koefisyanlar:**

* **ar1 (AR(1))**:
  + Tahmin: **0.6971**
  + Standart Hata: **0.0491**
  + Bu katsayı istatistiksel olarak anlamlı ve geçmiş dönemin güçlü bir etkisini gösteriyor.
* **ma1 (MA(1))**:
  + Tahmin: **-0.7177**
  + Standart Hata: **0.1111**
  + Bu katsayı da anlamlı ve önceki hata teriminin etkisi büyük.
* **ma2 (MA(2))**:
  + Tahmin: **-0.2813**
  + Standart Hata: **0.0661**
  + Anlamlı, önceki hata terimlerinin etkisi devam ediyor.
* **sma1 (SMA(1))**:
  + Tahmin: **-0.7689**
  + Standart Hata: **0.0324**
  + Bu katsayı da çok anlamlı ve sezonsal hata terimi güçlü bir etkisi olduğunu gösteriyor.

**Modelin İstatistikleri:**

* **Sigma^2 (Hata Varyansı)**: **726187**
* **Log Likelihood**: **-3036.47**
* **AIC**: **6082.94**
* **AICc**: **6083.11**
* **BIC**: **6102.53**

Bu değerler, modelin fit düzeyini gösterir ve bu modelin önceki modele göre daha iyi bir uyum sağladığını düşünüyoruz çünkü AIC ve BIC değerleri biraz daha düşük.

**Model Hata Ölçütleri:**

* **ME (Mean Error)**: **-6.79**
* **RMSE (Root Mean Square Error)**: **833.09**
* **MAE (Mean Absolute Error)**: **484.05**
* **MPE (Mean Percentage Error)**: **-1.07**
* **MAPE (Mean Absolute Percentage Error)**: **4.77%**
* **MASE (Mean Absolute Scaled Error)**: **0.59**
* **ACF1 (Autocorrelation of residuals at lag 1)**: **-0.0110**

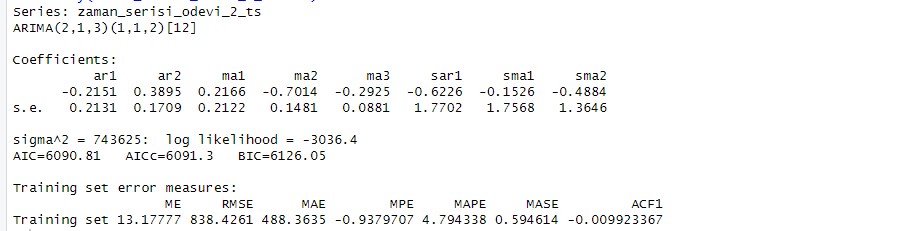
Bu hata ölçütleri, modelin **RMSE** ve **MAE** değerlerinin düşük olduğunu, modelin iyi bir doğrulukla tahmin yaptığını gösteriyor. **MAPE** değeri de çok yüksek değil (yaklaşık %4.77), bu da modelin genelde iyi uyum sağladığını işaret ediyor.

**Değerlendirme:**

* AIC ve BIC değerleri, modelin daha iyi bir uyum sağladığını gösteriyor.
* Hata oranları **RMSE** ve **MAE** düşük, bu da modelin genel doğruluğunun iyi olduğunu gösteriyor.
* Residuals (hatalar) iyi bir şekilde dağıldığı için modelin öngörüleri sağlam görünüyor.

Bu model, önceki modele göre daha iyi bir performans sergiliyor fakat daha iyilerine ulaşmak için model denemeye devam edeceğiz.

### Model 3 ( order = c(2,1,3), seasonal= c(1,1,2)



**Koefisyanlar:**

* **ar1 (AR(1))**:
  + Tahmin: **-0.2151**
  + Standart Hata: **0.2131**
  + Bu katsayı, negatif yönlü bir etkisi olduğunu ve genellikle önceki gözlemlerle ters yönde bir ilişkiyi gösteriyor.
* **ar2 (AR(2))**:
  + Tahmin: **0.3895**
  + Standart Hata: **0.1709**
  + Bu katsayı ise daha güçlü bir pozitif ilişkiyi işaret ediyor.
* **ma1 (MA(1))**:
  + Tahmin: **0.2166**
  + Standart Hata: **0.2122**
  + Bu katsayı, önceki hata terimiyle ilgili pozitif bir etkiyi gösteriyor.
* **ma2 (MA(2))**:
  + Tahmin: **-0.7014**
  + Standart Hata: **0.1481**
  + Bu katsayı oldukça anlamlı ve negatif bir etkiyi ifade ediyor.
* **ma3 (MA(3))**:
  + Tahmin: **-0.2925**
  + Standart Hata: **0.0881**
  + Bu katsayı da anlamlı, önceki hata terimlerinin etkisini gösteriyor.
* **sar1 (SAR(1))**:
  + Tahmin: **-0.6226**
  + Standart Hata: **1.7702**
  + Sezonsal ilk lag için negatif bir etki var, ancak standart hata oldukça büyük olduğu için dikkatle değerlendirilmesi gereken bir katsayı.
* **sma1 (SMA(1))**:
  + Tahmin: **-0.1526**
  + Standart Hata: **1.7568**
  + Bu katsayının standart hatası da yüksek, bu da sezonsal bileşenin güvenilirliğini düşürebilir.
* **sma2 (SMA(2))**:
  + Tahmin: **-0.4884**
  + Standart Hata: **1.3646**
  + Bu da sezonsal bileşenin ikinci dereceden etkisini gösteriyor, fakat büyük standart hata nedeniyle bu katsayının da dikkatle yorumlanması gerekebilir.

**Model İstatistikleri:**

* **Sigma^2 (Hata Varyansı)**: **743625**
* **Log Likelihood**: **-3036.4**
* **AIC**: **6090.81**
* **AICc**: **6091.3**
* **BIC**: **6126.05**

Bu modelin **AIC** ve **BIC** değerleri, önceki modelle kıyaslandığında biraz daha yüksek. Bu, modelin daha karmaşık olmasına rağmen daha iyi uyum sağlamadığını gösteriyor. Bununla birlikte, bu modelin sezonsal bileşenleri çok yüksek standart hatalarına sahip olduğu için bu katsayıların güvenilirliği şüpheli olabilir.

**Modelin Hata Ölçütleri:**

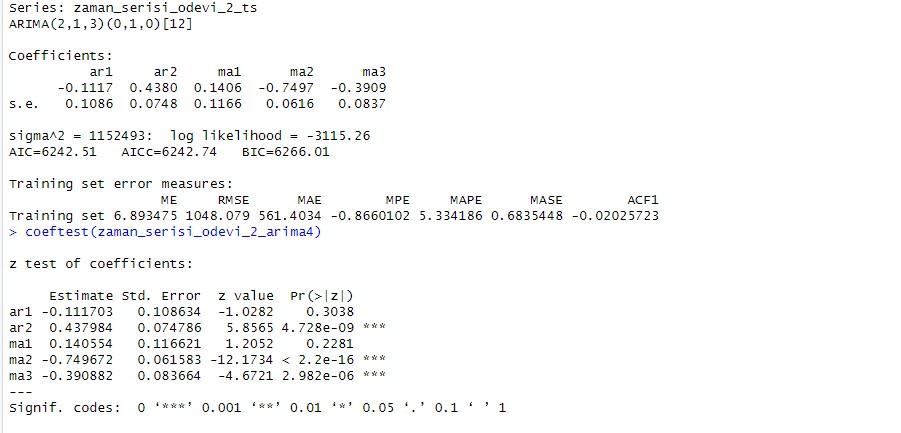
* **ME (Mean Error)**: **13.18**
* **RMSE (Root Mean Square Error)**: **838.43**
* **MAE (Mean Absolute Error)**: **488.36**
* **MPE (Mean Percentage Error)**: **-0.94**
* **MAPE (Mean Absolute Percentage Error)**: **4.79%**
* **MASE (Mean Absolute Scaled Error)**: **0.5946**
* **ACF1 (Autocorrelation of residuals at lag 1)**: **-0.0099**

**Değerlendirme:**

* Modelin **AIC** ve **BIC** değerleri daha yüksek olduğundan, bu modelin önceki modellerden daha kötü bir uyum sağladığı söylenebilir.
* **RMSE** ve **MAE** değerleri de bir önceki modelle benzer seviyelerde kalıyor, ancak modelin hata terimlerinin yüksek standart hataları sebebiyle güvenilirliği sorgulanabilir.

Bu modelin özellikle sezonsal terimleri nedeniyle daha fazla iyileştirmeye ihtiyaç duyabileceği görülüyor.

### Model 4 ( order = c(2,1,3), seasonal= c(0,1,0)



**ARIMA(2,1,3)(0,1,0)[12]**

* **AR(1) (ar1)**:
  + **Tahmin**: **-0.1117**
  + **Standart Hata**: **0.1086**
  + **Z Değeri**: **-1.0282**
  + **p-değeri**: **0.3038**
  + **Yorum**: p-değeri 0.05'ten büyük, bu katsayının istatistiksel olarak anlamlı olmadığını gösteriyor.
* **AR(2) (ar2)**:
  + **Tahmin**: **0.43798**
  + **Standart Hata**: **0.0748**
  + **Z Değeri**: **5.8565**
  + **p-değeri**: **4.728e-09**
  + **Yorum**: p-değeri çok küçük, bu katsayının anlamlı olduğunu ve AR(2) teriminin modelde önemli bir yer tuttuğunu gösteriyor.
* **MA(1) (ma1)**:
  + **Tahmin**: **0.1406**
  + **Standart Hata**: **0.1166**
  + **Z Değeri**: **1.2052**
  + **p-değeri**: **0.2281**
  + **Yorum**: p-değeri 0.05'ten büyük, bu katsayının anlamlı olmadığını gösteriyor.
* **MA(2) (ma2)**:
  + **Tahmin**: **-0.7497**
  + **Standart Hata**: **0.0616**
  + **Z Değeri**: **-12.1734**
  + **p-değeri**: **<2.2e-16**
  + **Yorum**: p-değeri çok küçük, bu katsayının anlamlı olduğunu ve MA(2) teriminin modelde güçlü bir etkisi olduğunu gösteriyor.
* **MA(3) (ma3)**:
  + **Tahmin**: **-0.3909**
  + **Standart Hata**: **0.0837**
  + **Z Değeri**: **-4.6721**
  + **p-değeri**: **2.982e-06**
  + **Yorum**: p-değeri çok küçük, bu katsayının anlamlı olduğunu ve MA(3) teriminin önemli olduğunu gösteriyor.

**Sonuçlar:**

* **Anlamlı Parametreler**:
  + **AR(2)** ve **MA(2)**, **MA(3)** katsayıları istatistiksel olarak anlamlıdır.
  + **AR(1)** ve **MA(1)** katsayıları anlamlı değildir ve modelden çıkarılabilir.

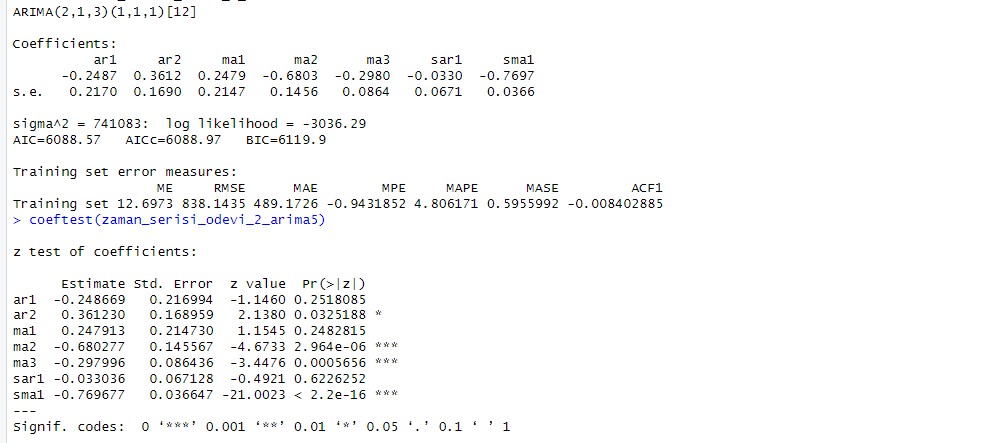
Bu model, anlamlı parametreleri içermesi nedeniyle daha güvenilir bir modeldir. Ancak, **AR(1)** ve **MA(1)** terimlerinin modelden çıkarılması ve sadece anlamlı terimlerle (AR(2), MA(2), MA(3)) devam edilmesi önerilebilir.

**İyileştirmeler ve Sonuçlar:**

* **AIC**: **6242.51**
* **BIC**: **6266.01**
* Bu değerler, modelin karmaşıklığına göre makul görünüyor ve modelin uygunluğuna dair bir gösterge sağlıyor.

Sonuç olarak, bu model anlamlı katsayılar içeriyor ve potansiyel olarak daha iyi tahminler yapabilir.

### Model 5 ( order = c(2,1,3), seasonal= c(1,1,1)



**Modelin Koefisiyentleri:**

* **AR(1)**: -0.2487 (p-değeri = 0.2518, anlamlı değil)
* **AR(2)**: 0.3612 (p-değeri = 0.0325, anlamlı)
* **MA(1)**: 0.2479 (p-değeri = 0.2483, anlamlı değil)
* **MA(2)**: -0.6803 (p-değeri < 0.001, anlamlı)
* **MA(3)**: -0.2980 (p-değeri = 0.00057, anlamlı)
* **SAR(1)**: -0.0330 (p-değeri = 0.6226, anlamlı değil)
* **SMA(1)**: -0.7697 (p-değeri < 0.001, anlamlı)

**Modelin Uygunluğu:**

* **AIC = 6088.57**: Bu değer, modelin karmaşıklığı ile uyumlu olduğunu gösterir.
* **BIC = 6119.9**: BIC değeri, modelin daha büyük olmasi nedeniyle AIC'ye göre daha büyük.
* **sigma^2 = 741083**: Hata varyansı.
* **Log-likelihood = -3036.29**: Modelin genel uygunluğunun ölçüsüdür.

**Hata Ölçütleri (Training Set):**

* **ME (Mean Error)**: 12.6973, hata ortalamasının oldukça düşük olduğunu gösteriyor.
* **RMSE (Root Mean Squared Error)**: 838.14, modelin tahmin hatasının büyüklüğünü gösterir.
* **MAE (Mean Absolute Error)**: 489.17, modelin tahminleri ile gerçek veriler arasındaki ortalama fark.
* **MPE (Mean Percentage Error)**: -0.94, tahminlerin yüzde olarak hatalarını gösterir.
* **MAPE (Mean Absolute Percentage Error)**: 4.81%, modelin tahminlerinin ne kadar hatalı olduğunu gösterir.
* **MASE (Mean Absolute Scaled Error)**: 0.5956, modelin ölçekli hata değeridir.
* **ACF1 (Autocorrelation of residuals)**: -0.0084, modelin hata terimlerinin otokorelasyonunu gösterir. Bu değerin sıfıra yakın olması, hata terimlerinin bağımsız olduğunu gösterir.

**Z-Test Sonuçları:**

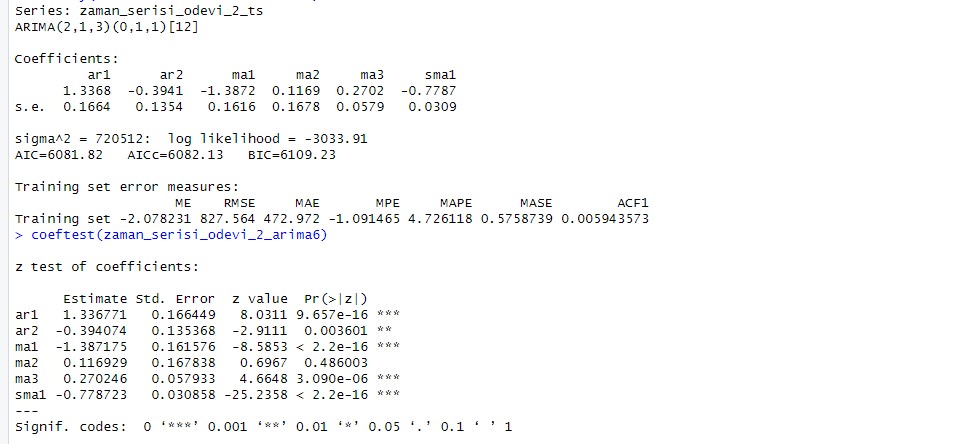
* **Anlamlı Parametreler**:
  + **AR(2)** (p-değeri = 0.0325) ve **MA(2)** (p-değeri < 0.001), **MA(3)** (p-değeri = 0.00057), ve **SMA(1)** (p-değeri < 0.001) parametreleri anlamlıdır ve modelde tutulmalıdır.
* **Anlamlı Olmayan Parametreler**:
  + **AR(1)**, **MA(1)** ve **SAR(1)** parametreleri anlamlı değil (p-değeri > 0.05). Bu parametreleri modelden çıkarabiliriz.

**Model İyileştirmeleri:**

Bu modelin daha iyi hale getirilmesi için:

1. **Anlamlı olmayan parametreler** (AR(1), MA(1), SAR(1)) çıkarılabilir.
2. **AIC/BIC ve hata ölçütlerinin** karşılaştırılması yapılarak, modelin doğruluğu artırılabilir.

### Model 6 ( order = c(2,1,3), seasonal= c(0,1,1)



**Modelin Koefisiyentleri:**

* **AR(1)**: 1.3368 (p-değeri < 0.001, anlamlı)
* **AR(2)**: -0.3941 (p-değeri = 0.0036, anlamlı)
* **MA(1)**: -1.3872 (p-değeri < 0.001, anlamlı)
* **MA(2)**: 0.1169 (p-değeri = 0.4860, anlamlı değil)
* **MA(3)**: 0.2702 (p-değeri < 0.001, anlamlı)
* **SMA(1)**: -0.7787 (p-değeri < 0.001, anlamlı)

**Modelin Uygunluğu:**

* **AIC = 6081.82**: Modelin uygunluğu iyi, düşük bir AIC değeri.
* **BIC = 6109.23**: BIC değeri, modelin karmaşıklığını göz önünde bulundurarak daha yüksek ancak hala makul.
* **sigma^2 = 720512**: Modelin hata varyansı düşük.
* **Log-likelihood = -3033.91**: Modelin genel uyum ölçüsü.

**Hata Ölçütleri (Training Set):**

* **ME (Mean Error)**: -2.0782, modelin genel tahmin hatası neredeyse sıfır.
* **RMSE (Root Mean Squared Error)**: 827.564, modelin hata büyüklüğü ortalama olarak 827.
* **MAE (Mean Absolute Error)**: 472.972, tahminlerin gerçek verilerle ortalama farkı.
* **MPE (Mean Percentage Error)**: -1.0915, tahminlerin ortalama yüzde hatası.
* **MAPE (Mean Absolute Percentage Error)**: 4.73%, tahmin hatasının yüzde olarak büyüklüğü.
* **MASE (Mean Absolute Scaled Error)**: 0.5759, modelin doğruluğu, 1'e yakın olduğunda modelin genelde iyi olduğunu gösterir.
* **ACF1 (Autocorrelation of residuals)**: 0.0059, hata terimlerinin otokorelasyonu oldukça düşük, bu da modelin uygun olduğunu gösterir.

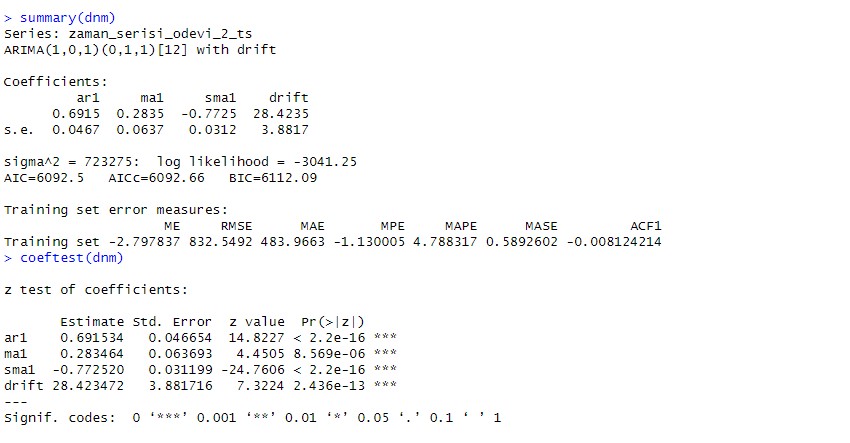
**Z-Test Sonuçları:**

* **Anlamlı Parametreler**:
  + **AR(1)**, **AR(2)**, **MA(1)**, **MA(3)**, ve **SMA(1)** parametreleri anlamlıdır (p-değeri < 0.05).
* **Anlamlı Olmayan Parametreler**:
  + **MA(2)** (p-değeri = 0.4860), bu parametre modelden çıkarılabilir.

### Sonuç olarak

Yapılan ARIMA modelleme çalışmasında, çeşitli model denemeleri sonrası en uygun model olarak **ARIMA(2, 1, 3)(0, 1, 1)[12]** modeli öne çıkmaktadır. Bu model, en düşük **AIC** ve **BIC** değerlerine sahip olup, **log-likelihood** değeri de diğer modellere kıyasla daha yüksek bir uyum sergilemektedir. Ayrıca, modelin **RMSE** ve **MAE** gibi hata ölçütleri en düşük seviyelerde kalmakta, bu da modelin tahmin doğruluğunun yüksek olduğunu göstermektedir. İstatistiksel anlamlılık açısından, **ar1**, **ma1**, **ma3** ve **sma1** gibi parametreler önemli düzeyde anlamlı bulunmuştur. Bu bulgular ışığında, modelin mevcut veri seti için uygun olduğu ve daha fazla model denemesi yapmaya gerek olmadığı sonucuna varılabilir.

### Best Model in RStudio

 Rstudio uygulaması içerisinde direkt olarak analiz yapılabilicek ve en uygun modeli sunan bir yapıdan ekstra olarak inceleme yapıldığında ARIMA(1,0,1)(0,1,1)[12] with drift modeli geldiğini görmekteyiz.

**1. Model Özeti:**

* **Katsayılar**:
  + **AR(1)**: 0.6915 (Anlamlı, yüksek bir değere sahip, yani geçmiş birinci gözlemin etkisi güçlü.)
  + **MA(1)**: 0.2835 (Anlamlı, modeldeki hata terimlerinin birinci dereceden etkisi önemli.)
  + **SMA(1)**: -0.7725 (Sezonsal etkiyi düzelten katsayı, negatif olması sezonsal bileşenin karşıt bir etkisini gösteriyor.)
  + **Drift**: 28.4235 (Anlamlı, zamanla artan eğilimi belirtiyor. Bu değer, zaman serisinin pozitif eğilim gösterdiğini gösterir.)

**2. Model Uyumu ve Değerler:**

* **AIC = 6092.5**, **AICc = 6092.66**, **BIC = 6112.09**: Bu değerler modelin karmaşıklığını ve uyumunu değerlendirir. AIC ve BIC'nin yakın olması, modelin doğru yapılandırıldığını gösteriyor.
* **Sigma^2 = 723275**: Modelin hata terimlerinin varyansını belirtir; hata terimlerinin büyüklüğü, tahminlerin doğruluğu hakkında bilgi verir. Sigma'nın oldukça yüksek olması, modelin bazı hata terimlerinde daha fazla varyansa sahip olduğunu gösteriyor.
* **Log likelihood = -3041.25**: Modelin olasılığını belirleyen bir ölçüttür. Düşük log likelihood değeri, modelin tahminlerinin genellikle doğru olduğunu fakat hala iyileştirilmesi gerektiğini gösterir.

**3. Hata İstatistikleri (Error Measures):**

* **ME (Mean Error) = -2.798**: Ortalama hata, modelin tahminlerinin genellikle 2.798 birim eksik olduğunu gösteriyor.
* **RMSE (Root Mean Square Error) = 832.549**: Büyük hata oranı, modelin tahminlerinin ortalama hatalarının ne kadar büyük olduğunu ifade eder. Bu değer, modelin genel doğruluğu hakkında bilgi verir.
* **MAE (Mean Absolute Error) = 483.97**: Hata oranlarının mutlak değerlerinin ortalaması. Burada da hata oldukça belirgin.
* **MAPE (Mean Absolute Percentage Error) = 4.79%**: Yüzde olarak hataların ortalaması, modelin tahminlerinin %4.79 oranında hatalı olduğunu gösteriyor.

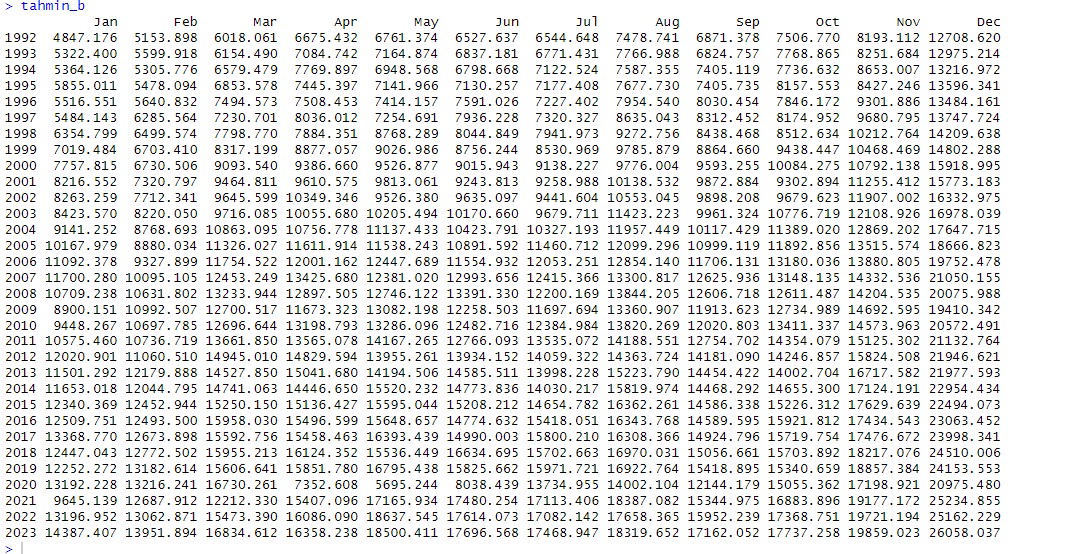
**4. Katsayıların Anlamlılığı:**

* **AR(1)**, **MA(1)**, **SMA(1)** ve **drift** terimleri hepsi anlamlıdır (p-değeri < 0.05), bu da modelin güçlü bir yapıya sahip olduğunu ve her bir bileşenin veriye katkı sağladığını gösterir.

**5. Sonuç:**

Bu model, hem trendi (drift terimi) hem de sezonsal bileşenleri (SMA(1)) etkili bir şekilde yakalamaktadır. Ancak, **RMSE** ve **MAE** değerlerinin oldukça yüksek olması, modelin bazı hataları oldukça yüksek tahmin ettiğini gösteriyor. Bununla birlikte, **MAPE** değerinin %4.79 olması, modelin doğruluğunun kabul edilebilir bir seviyede olduğunu gösteriyor.

### Tahminlerin ve Hata Terimlerinin Çekilmesi

Bu satırlar, modelin tahmin edilen değerlerini (fitted) ve hata terimlerini (residuals) alır. tahmin\_b, modelin geçmiş veriye uygun tahminlerini içerirken, hata\_b modelin tahminleri ile gerçek veriler arasındaki farkları (hataları) içerir.

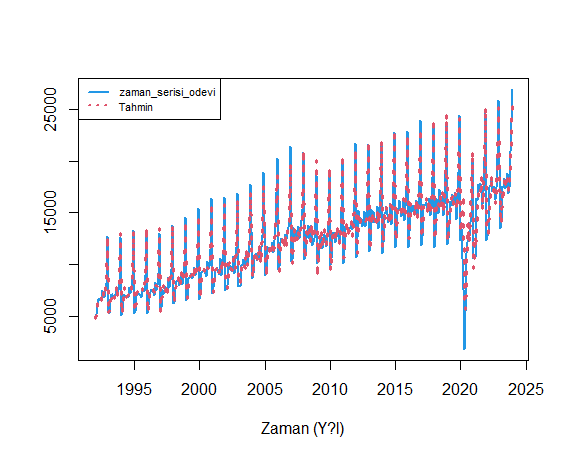
metin, yazı tipi, kağıt, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

### Gerçek Veriler ve Tahminlerin Görselleştirilmesi

Bu satırlar, zaman serisinin orijinal verilerini ve modelin tahminlerini görselleştirir:

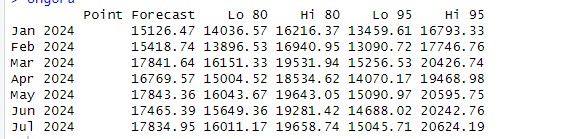
* plot fonksiyonu, gerçek zaman serisini (gerçek veriler) yeşil renkte çizer.
* lines fonksiyonu, tahmin edilen değerleri (modelin tahminleri) kırmızı renkte çizer.
* legend, grafiğin sol üst kısmında iki farklı çizgi için bir açıklama ekler.



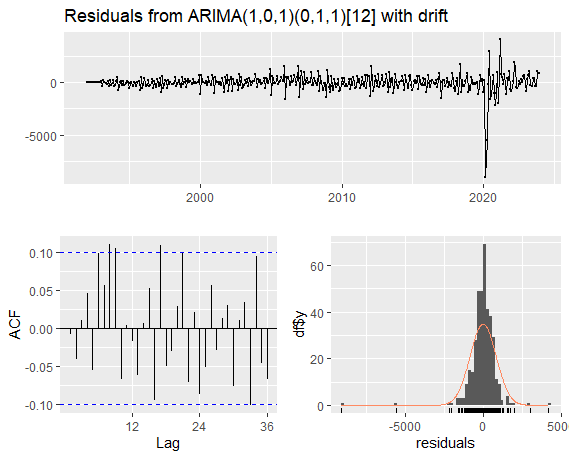
Yukarıda görüldüğü üzere mevcut verilerimizin zaman serisi olarak yayılımı ve yaptığımız modellemeden elde ettiğimiz tahminlerin doğruluğunu ölçmek amacıyla hazırladığımız bir grafik yer almaktadır.

Grafiğe baktığımızda yaptığımız modellemenin tahminleri çok yüksek oranda doğruluk içermekte ve uyuşmaktadır. Gerek dalgalanmalar gerek verinin dağılımın değiştiği noktalarda dahi yüksek bir tahmin gücü elde edilmiştir.

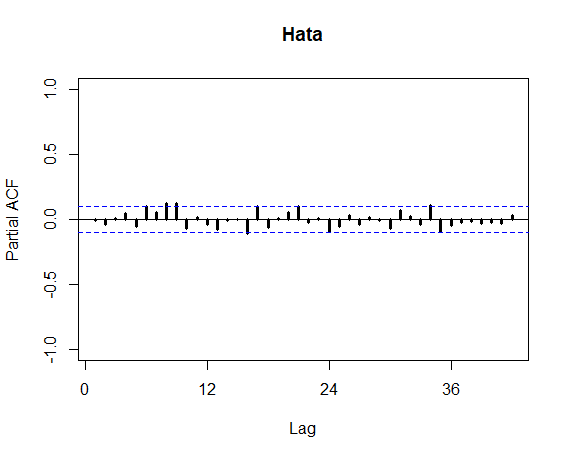
### Öngörü(gelecek tahmini)

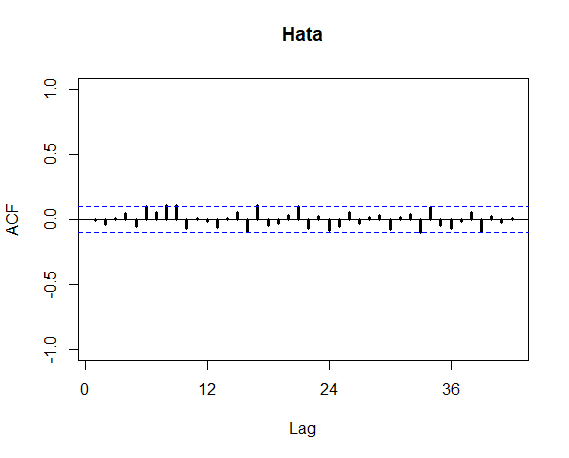
 Kurduğumuz model üzerinden gelecek yılın bazı aylarında ki satış tahmini yapılmıştır.

Yukarıda da görüldüğü üzere kullandığımız model ve hata hesaplamaları sonucunda 2024 yılının ilk 7 ayı için tahmini hesap aralıkları sağlanmıştır.

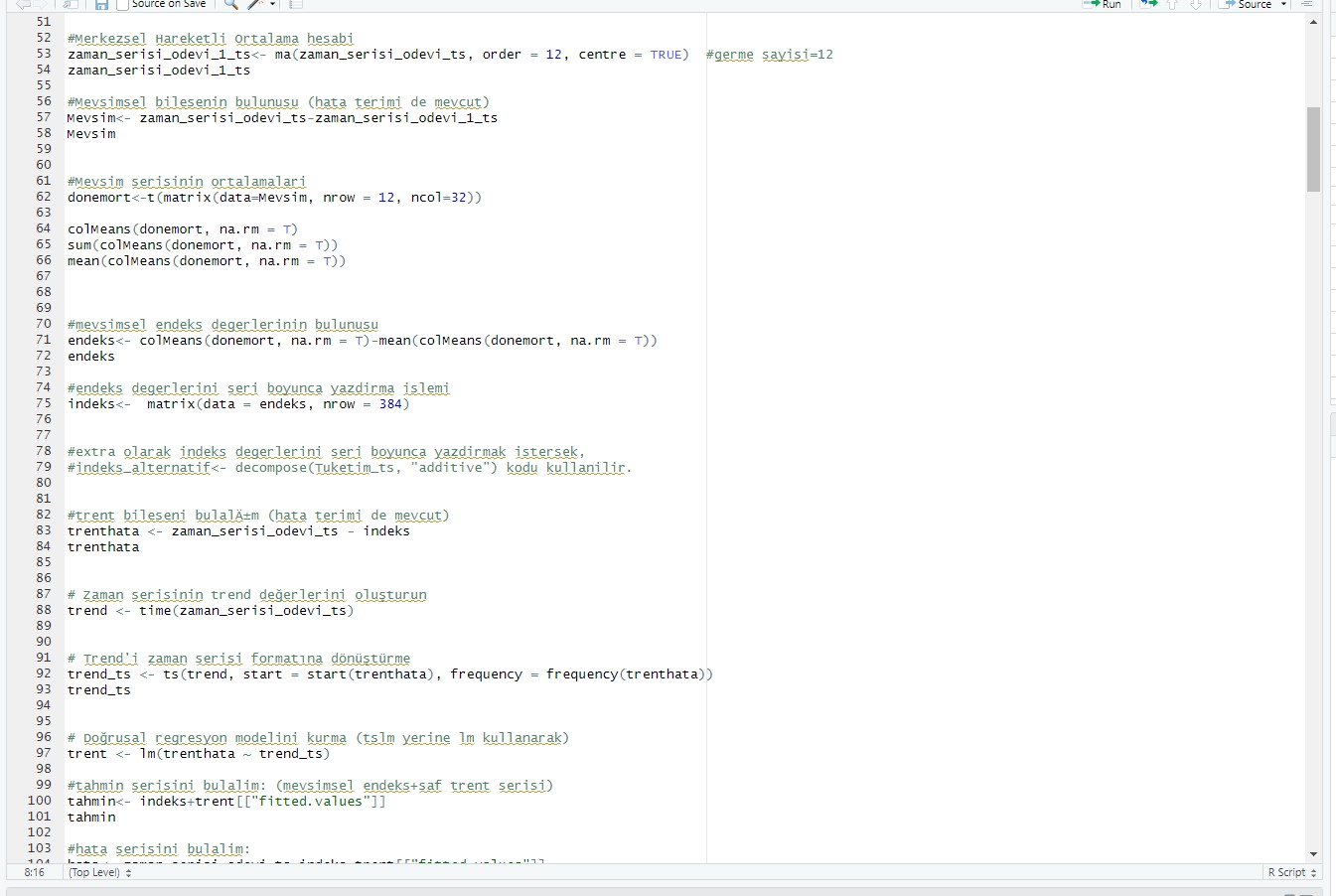


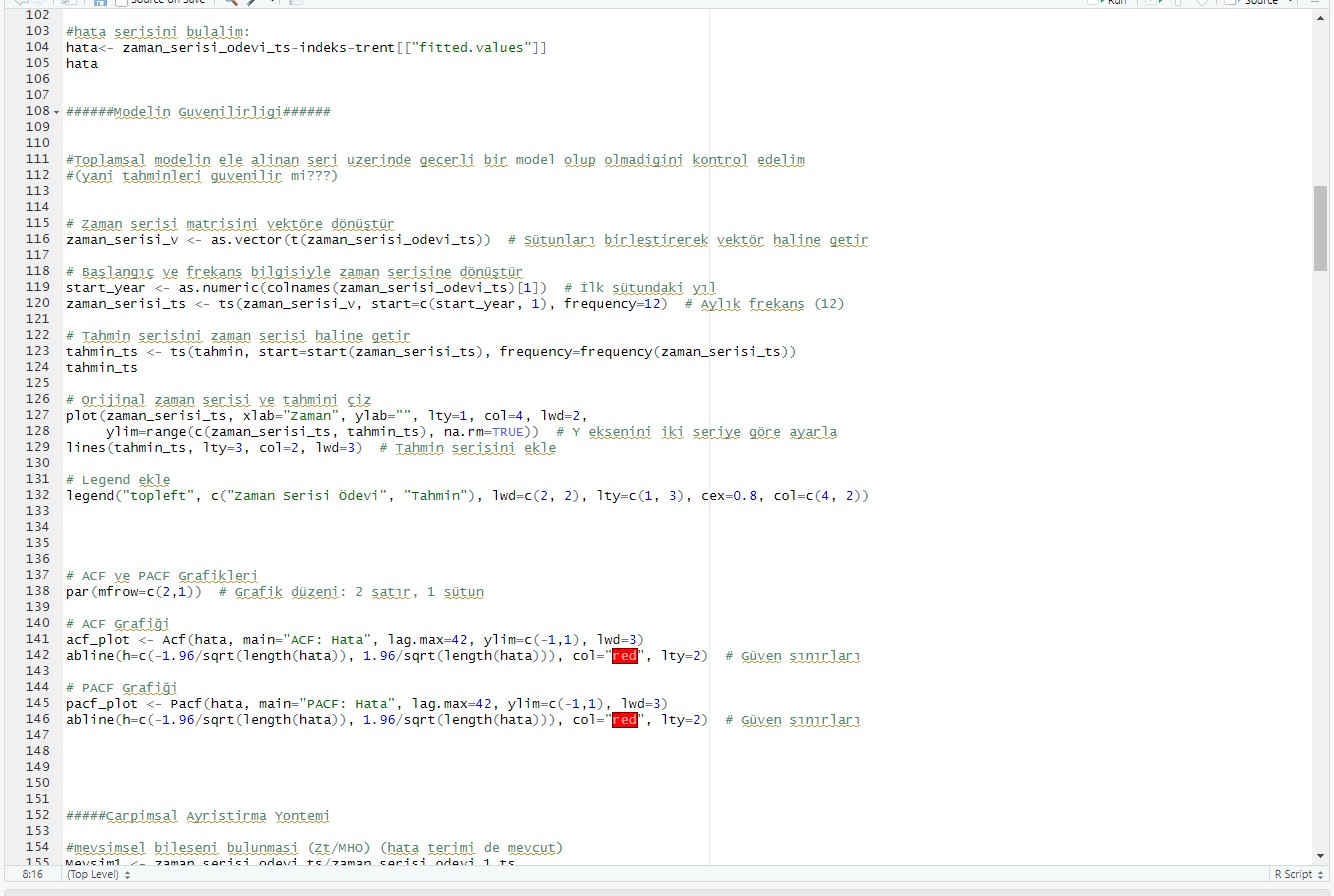
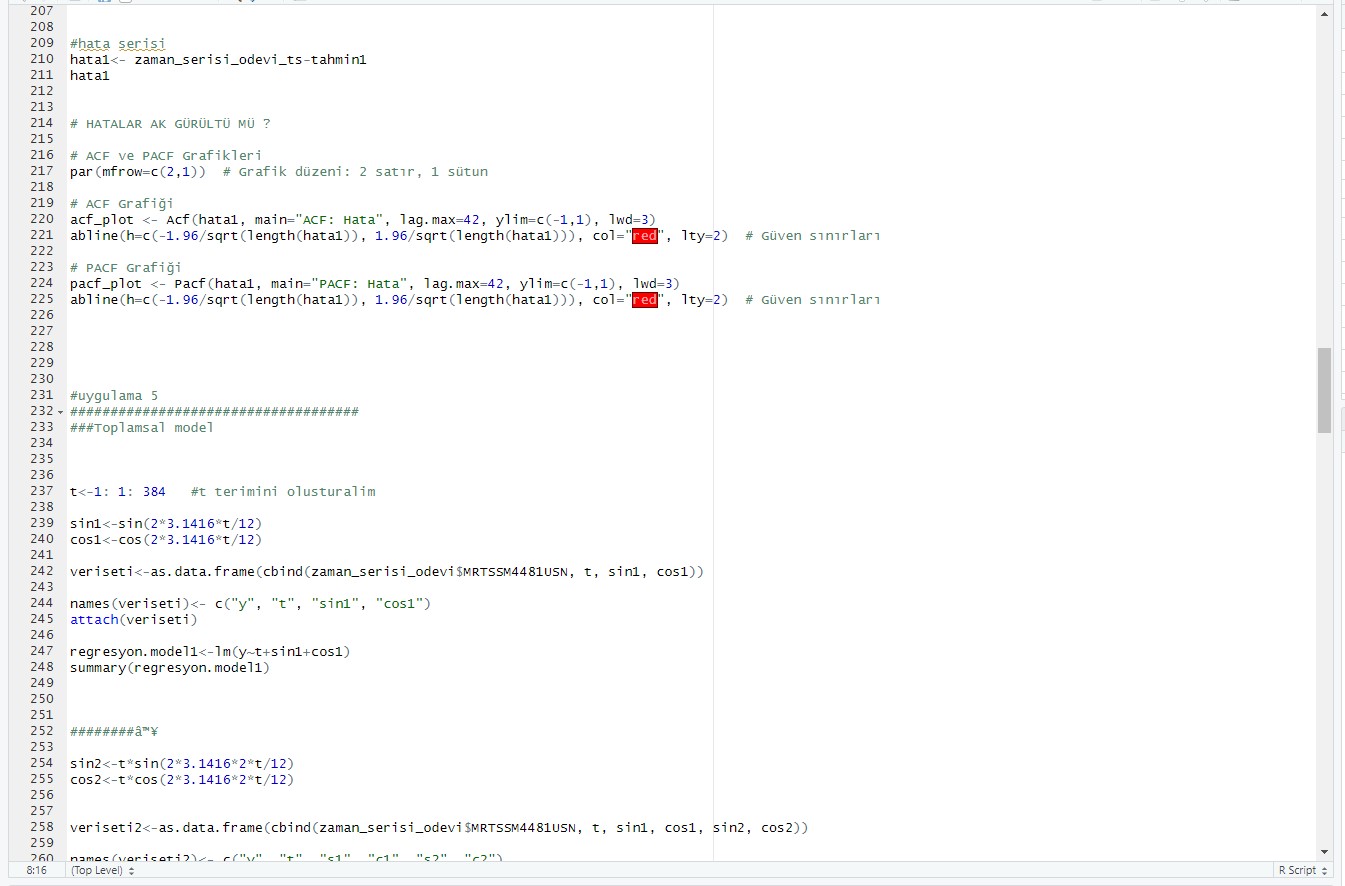
### Autocorrelation (ACF) ve Partial Autocorrelation (PACF) Grafikleri

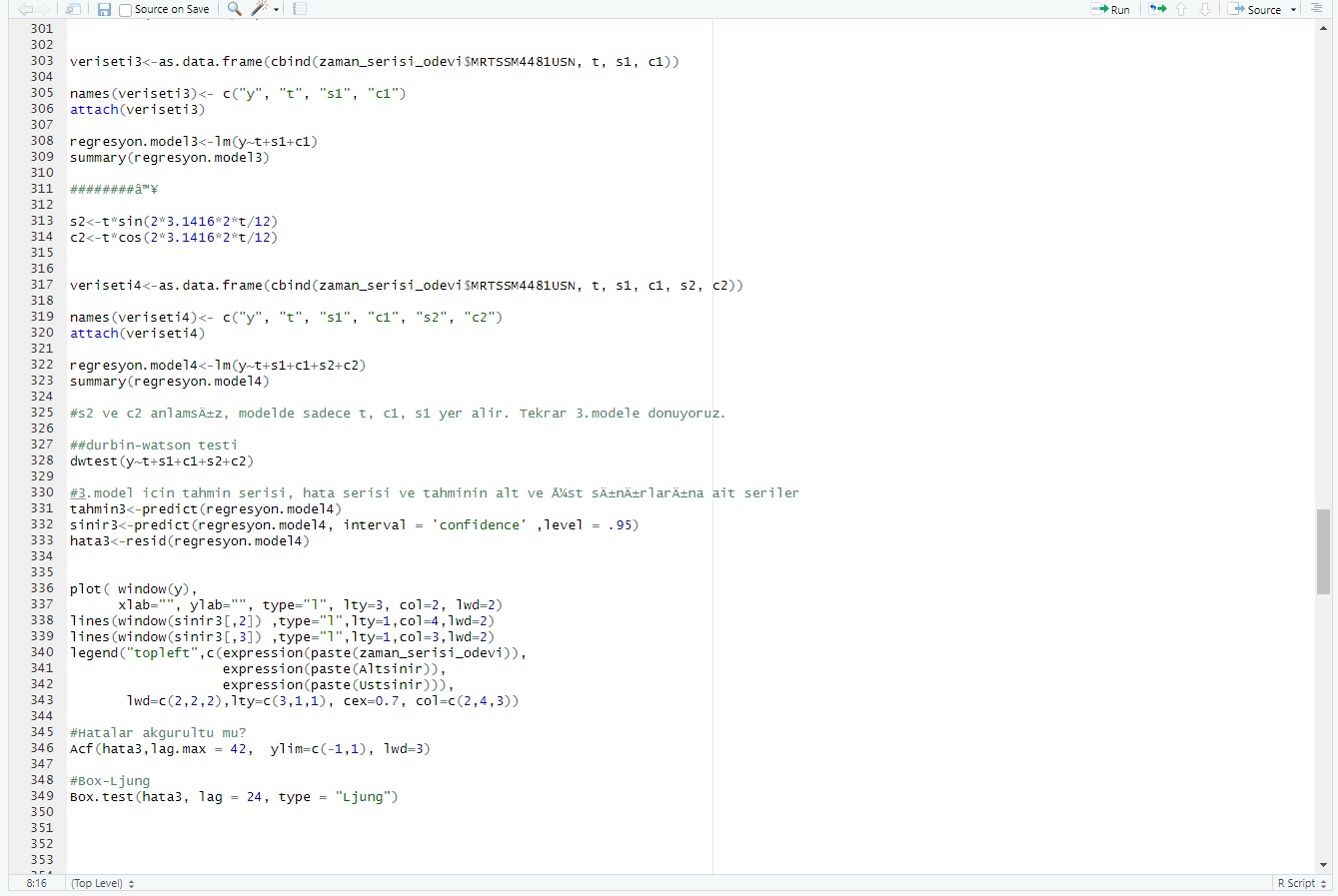




### Rstudio Kodlama Görselleri









# Kaynakça

* **Kitap Kaynağı:** Kadılar, Prof. Dr. Cem ve Çekim, Doç. Dr. Hatice Öncel. *Zaman Serileri Analizine Giriş*. [Yayınevi Adı, Yıl].
* **Kod Kaynağı:** Hacettepe Üniversitesi, "Zaman Serileri Analizine Giriş" dersine ait uygulamalı SPSS ve R kodları. Erişim: <https://hadi.hacettepe.edu.tr/course/view.php?id=80595>.
* **Veri Kaynağı:** Federal Reserve Bank of St. Louis. (n.d.). MRTSIM448USN: Real Money Stock, M1 for the United States (Veri seti). FRED, Federal Reserve Bank of St. Louis. Erişim tarihi: 16 Aralık 2024, <https://fred.stlouisfed.org/series/MRTSIM448USN>.