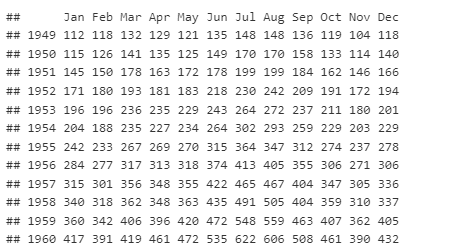
**İÇİNDEKİLER**

1. **TANIMLAYICI İSTATİSTİKLER**
   1. Özet İstatistikler………………………………………………………………………………………………3
   2. Box-Plot Grafiği………………………………………………………………………………………………3
   3. Zaman Serisi Grafiği………………………………………………………………………………………….4
2. **AYRIŞTIRMA YÖNTEMLERİ**
   1. Toplamsal Ayrıştırma Yöntemi…………………………………………………………………………...…..6
   2. Çarpımsal Ayrıştırma Yöntemi………………………………………………………………………………..7
3. **ZAMAN SERİLERİNDE REGRESYON ANALİZİ**
   1. Toplamsal Regresyon Analizi…………………………………………………………………………………8
   2. Çarpımsal Regresyon Analizi………………………………………………………………………………..10
   3. Karesel Regresyon ………………………………………………………………………………………….11
   4. Kübik Regresyon ……………………………………………………………………………………………12
   5. Lojistik Regresyon…………………………………………………………………………………………..13
4. **WINTERS ÜSTEL DÜZLEŞTİRME YÖNTEMİ**
   1. Toplamsal Winters Üstel Düzleştirme Yöntemi……………………………………………………………15
   2. Çarpımsal Winters Üstel Düzleştirme Yöntemi…………………………………………………………….17

**5. BOX-JENKINS MODELLERİ** …………………………………………….…………………..........19

**6. KAYNAKÇA**……………………………………………………………………………………………...26

**1.ÖZETLEYİCİ İSTATİSTİKLER**

Bu veri seti, 1949’dan 1960’a kadar bir ABD havayolu şirketinin aylık olarak seyahat eden müşteri toplamını açıklar. Bu veri seti R Studio kütüphanelerinden alınmıştır.

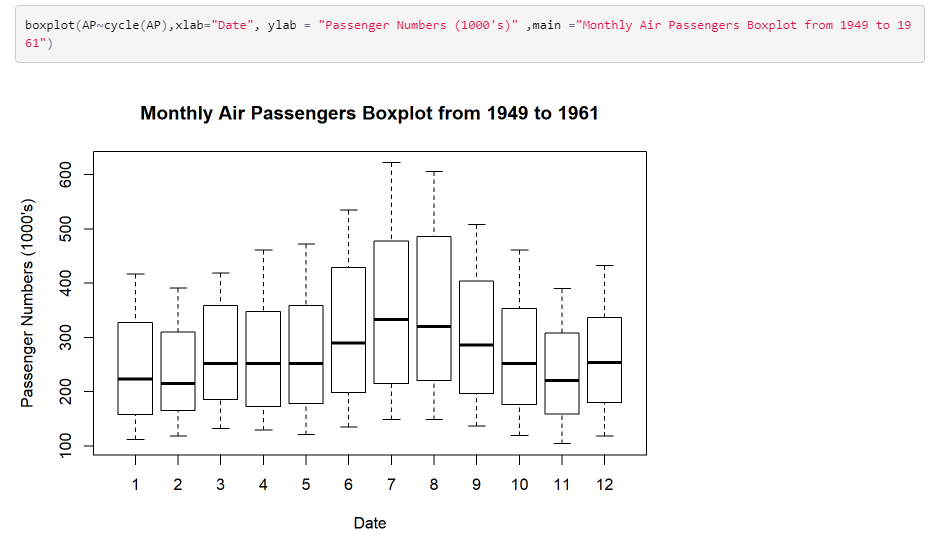
AP <- AirPassenger değişikenini temsil eder.

ap\_ts <- AirPassenger değişkenin zaman serisi formatına taşınması.

1.1 Özet İstatistikler

Zamana bağlı değişimin zaman serisi grafiği aşağıda verilmiştir.

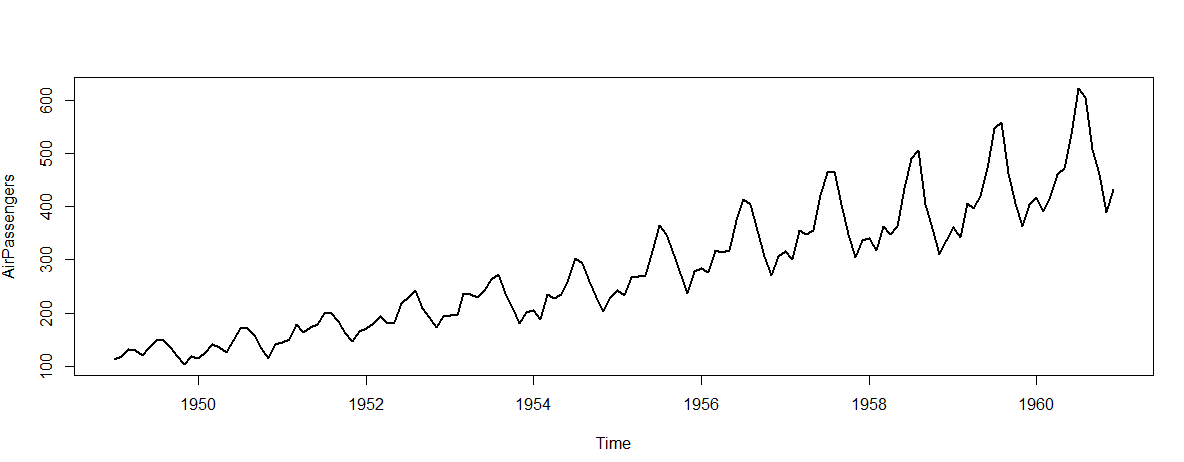
1.2 Box-Plot Grafiği



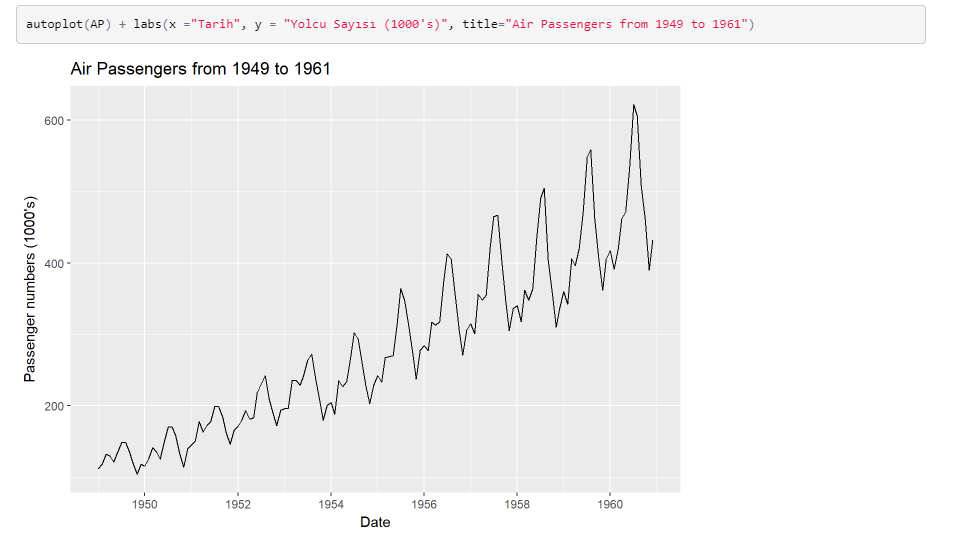
Mevsimsel efektleri görebilmek adına box-plot grafiğini inceleyelim.

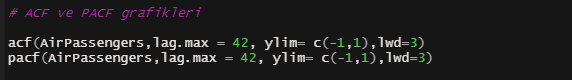
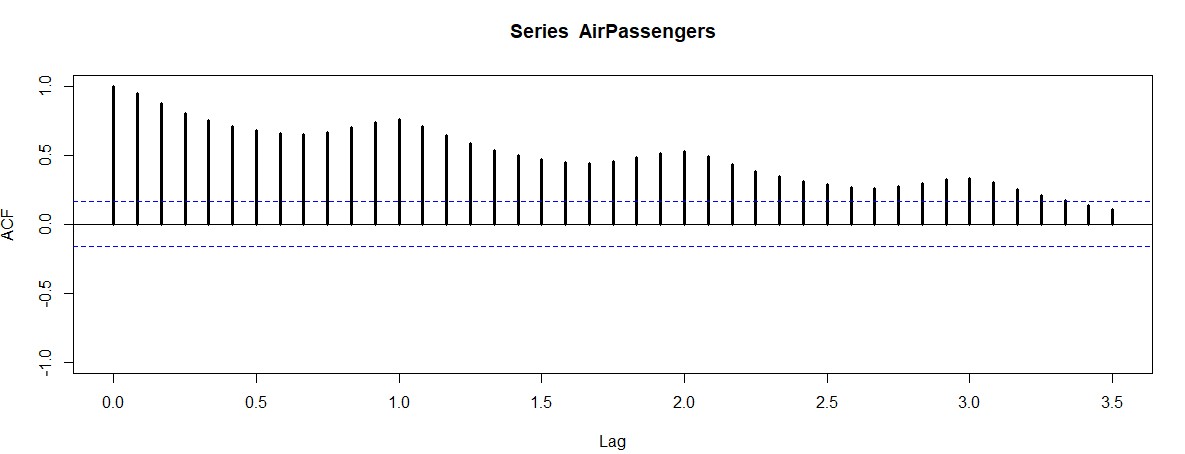
* Yolcu sayıları her yıl zamanla artıyor ve bu, belki de o dönemde uçuş yolculuğuna olan talebin artması nedeniyle artan bir doğrusal eğilimin göstergesi olabilir.
* Kutu grafiğinde, 6 ila 9. aylarda, diğer aylara göre daha yüksek ortalamalar ve daha yüksek varyanslar gözlemlenmiştir. Bu aylarda seyahat eden daha fazla yolcu vardır; bu, 12 aylık görünür bir döngü ile mevsimselliği gösterir.
* Herhangi bir aykırı değer yok gibi görünüyor ve eksik değer gözlemlenmedi. Bu nedenle veri temizleme gerekmez.

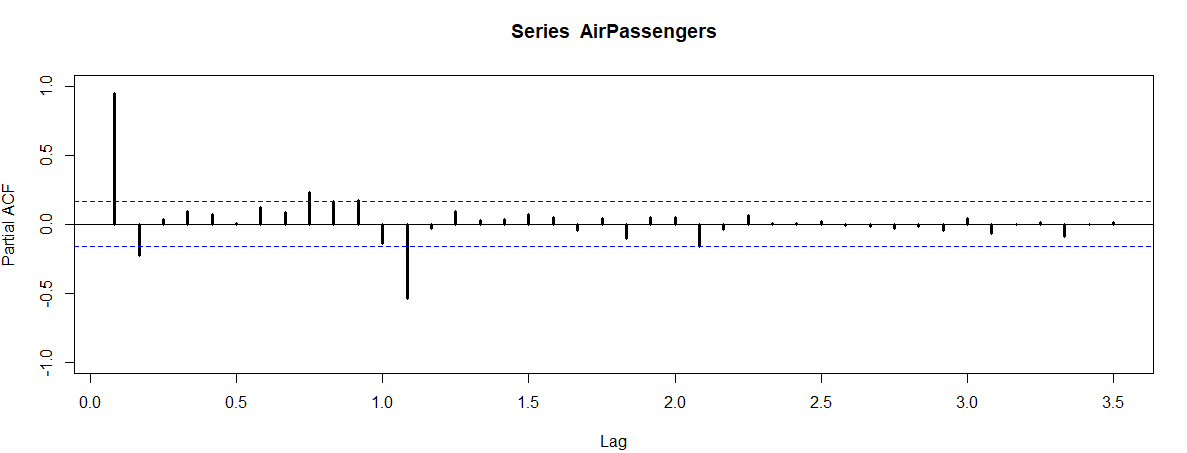
1.3 Zaman Serisi Grafiği

* Grafiğe bakıldığında dönemsel olarak artış ve azalmalar gözlemlenmiştir. Mevsimsellik olabileceğinden şüpheleniriz.

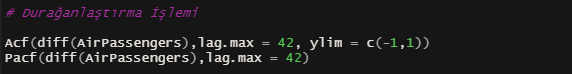
Ayrıca bu grafiği ggplot2 kütüphanesinde bulunan ggfortify paketiyle de çizdirebiliriz.

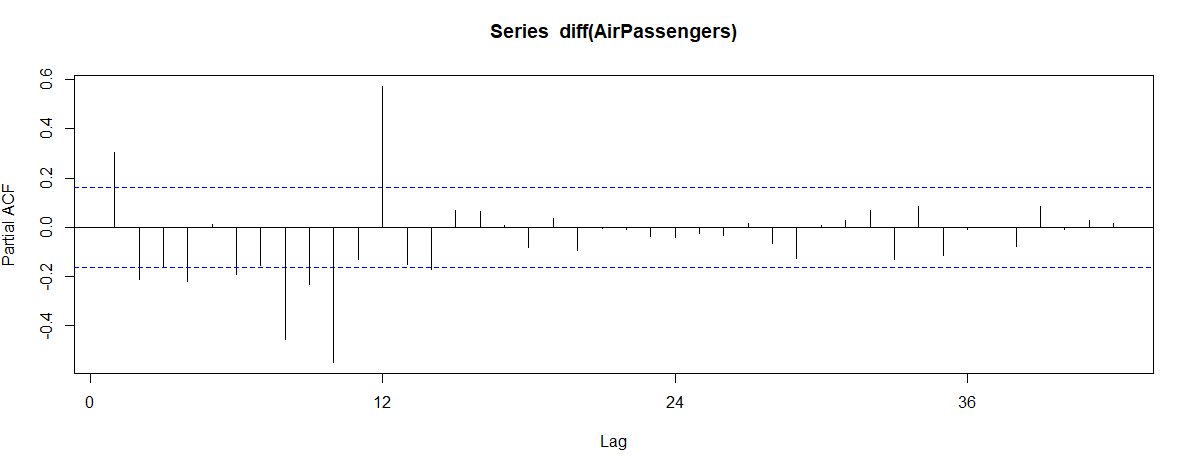


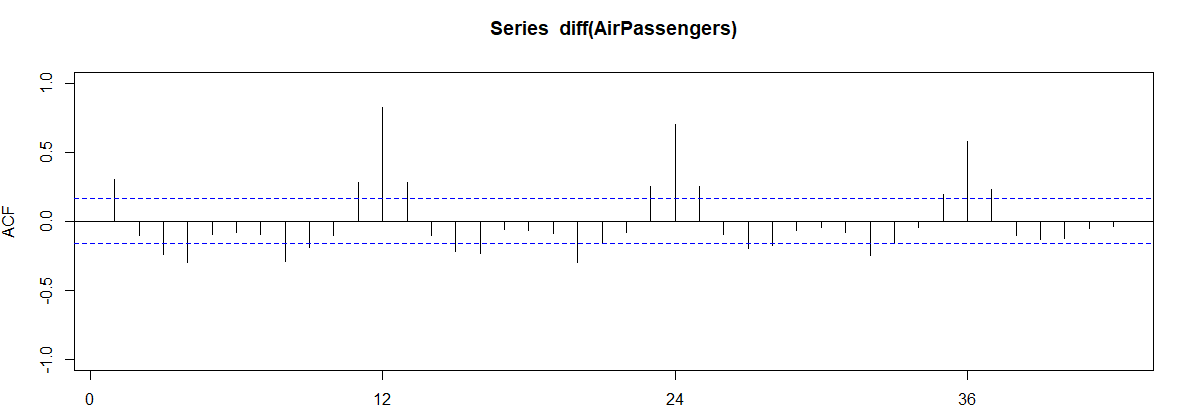




Otokorelasyon fonskiyonu yani ACF grafiğine bakıldığında ilk dört gecikmenin güven aralığı dışında olduğunu gözlemledik. Bu sebepten ötürü verimizin bir trende sahip olduğunu söyleyebiliriz. Grafikte ani dalgalanmalar olduğundan mevsimsellikten şüphelenmekteyiz. En yüksek iki gecikme arasında 12 gecikme olduğu için periyodun 12 olduğunu söyleyebiliriz. Kısmi otokorelasyon(PACF) grafiği de bu yorumlarımızı desteklemektedir.



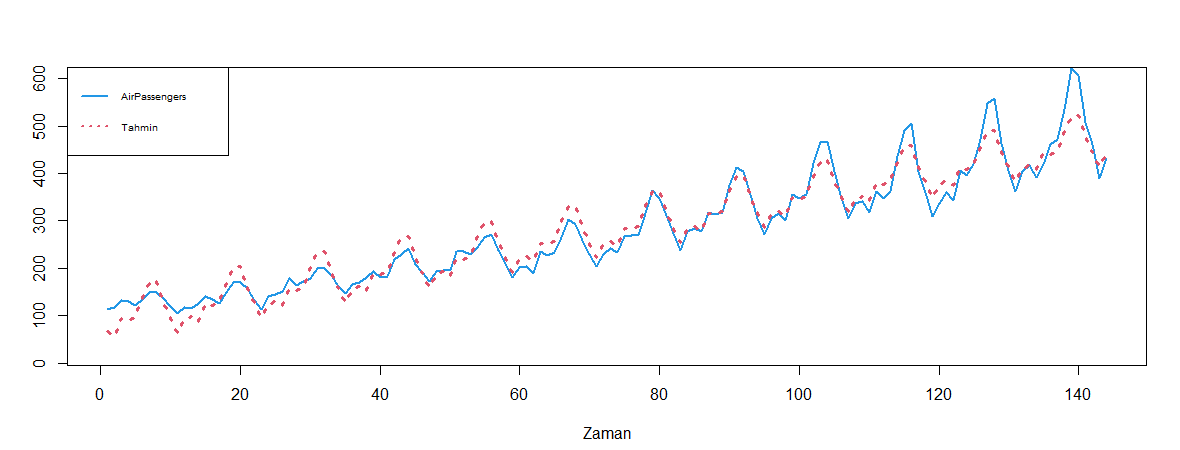
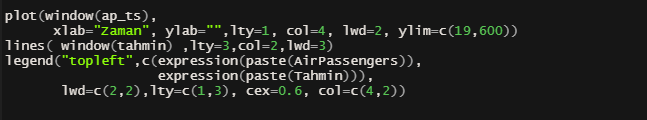




* Periyot 12 olarak bariz bir şekilde gözlemlenmiştir.
* Mevsimsellik vardır ve baskındır.

1. **AYRIŞTIRMA YÖNTEMLERİ**
   1. Toplumsal Ayrıştırma Yöntemi:



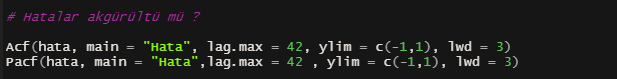
 (https://www.toppr.com/guides/fundamentals-of-business-mathematics-and-statistics/time-series-analysis/models-of-time-series-analysis/)

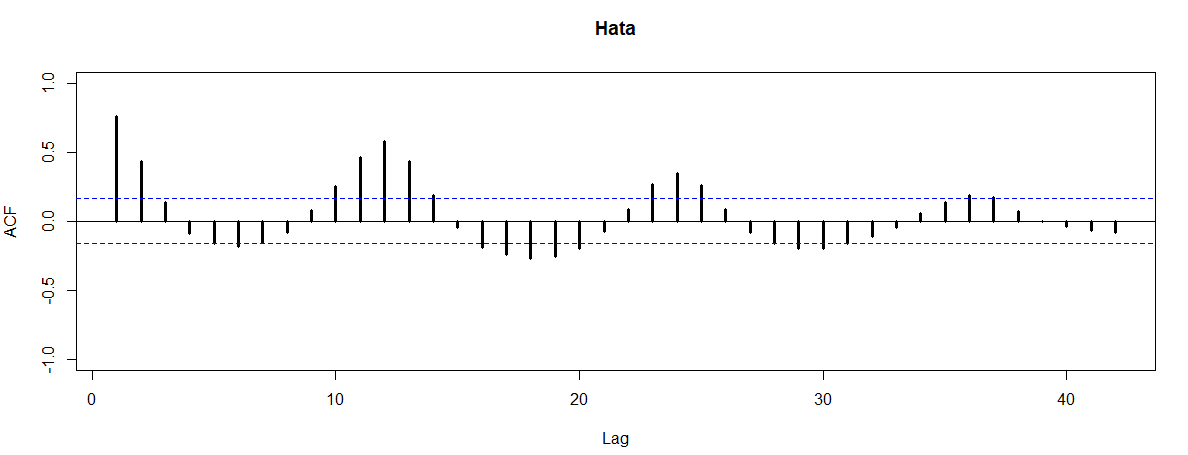
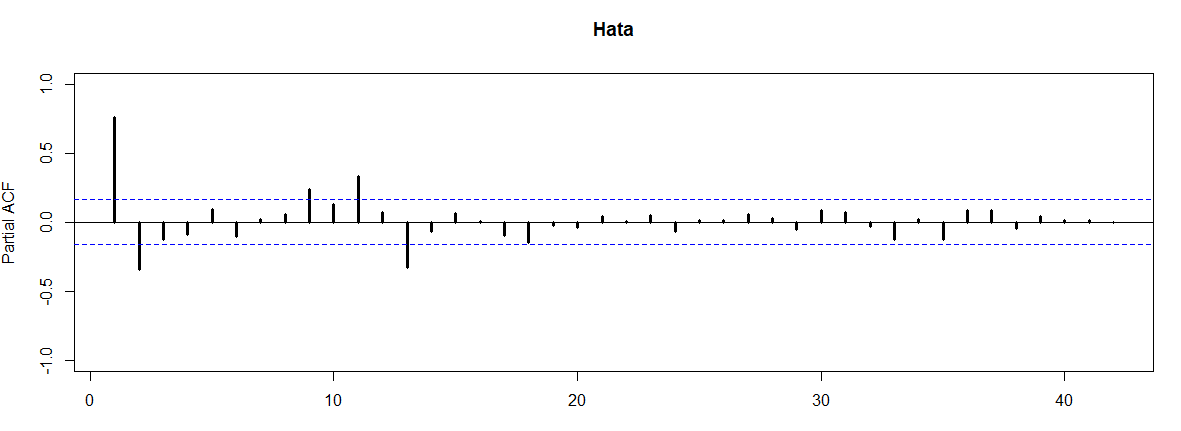
Tahmin serimiz ile orijinal seri grafiğimizi çizdiğimizde birbirine yakın değerler elde ettiğini ve modelin başarılı bir şekilde tahmin ettiğini düşünürüz. Grafiğe göre birçok verinin tahmin sınırları içinde kaldığı söylenebilir. Buradan seri için bu modelin uygun olduğu düşünülebilir ancak diğer testlere bakılması gerekir.

Hatalar akgürültü mü?

Ho: Hatalar arasında ilişki yoktur.

Hs: Hatalar arasında ilişki vardır.

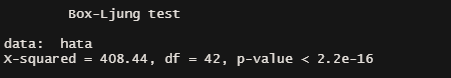




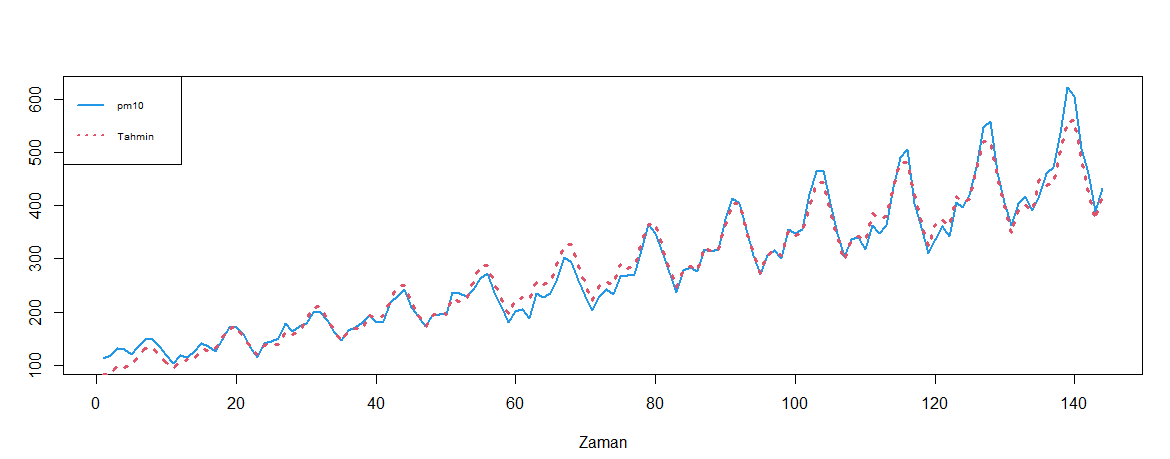
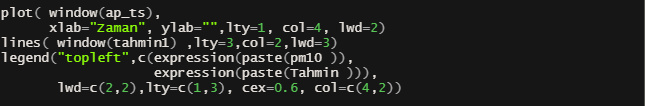
* Hataların ACF grafiğine baktığımızda sınır dışında kalan birden fazla gecikme gözlemleriz. Bu sebepten ötürü hataların akgürültü olmadığını söyleyebiliriz.

Yine de Box-Ljung Testi Uygulayalım:

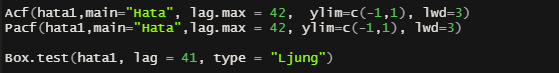


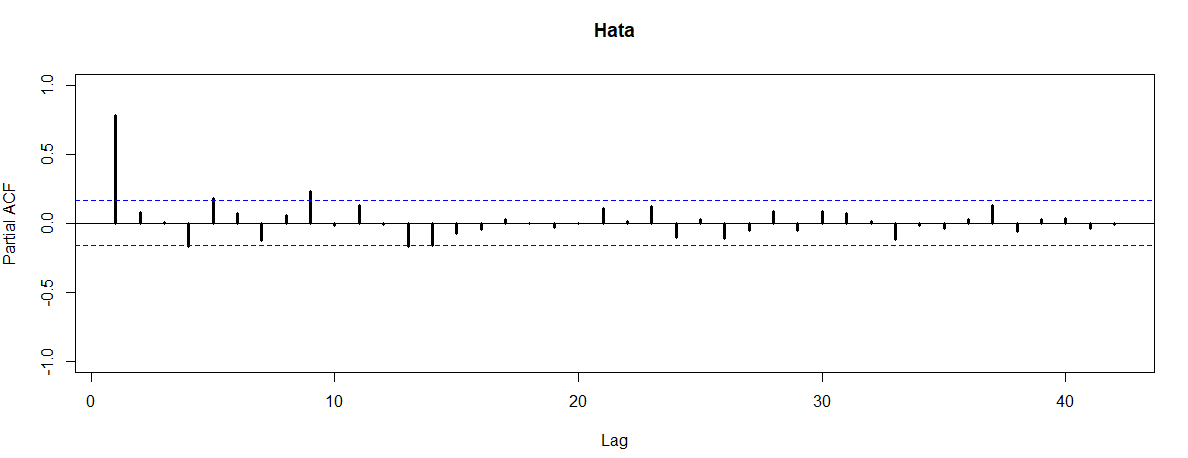
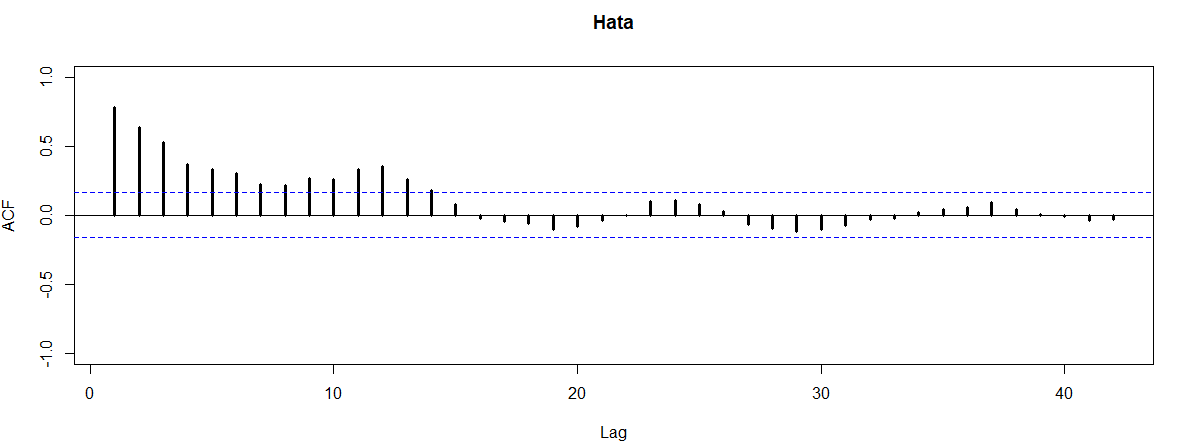
* H0 hipotezi reddedilir. Hatalar Akgürültü değildir. Bu sebepten bu modelin bu seri için uygun olmadığını söyleriz.

**2**.2 Çarpımsal Ayrıştırma Yöntemi:



Hatalar akgürültü mü?

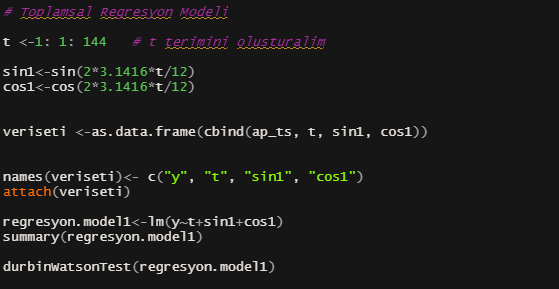




* Toplamsal ve Çarpımsal ayrıştırma yöntemlerinin hata ACF grafikleri ve Box-Ljung testleri dikkate alındığında serilerin akgürltü serisi olmadığı görülür. Bu modeller istatistiksel olarak anlamlı değildir.

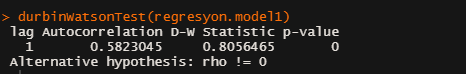
1. **ZAMAN SERİLERİNDE REGRESYON ANALİZİ**

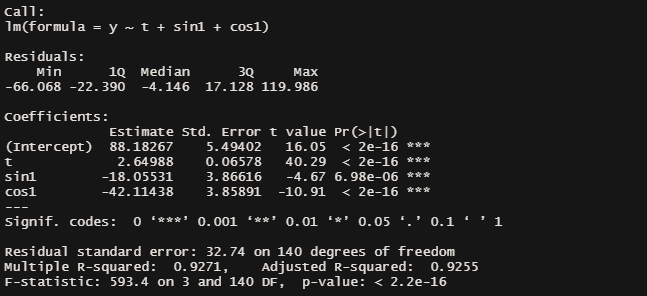
3.1 Toplamsal Regresyon Analizi

****

=Regresyon modeli anlamsızdır.

= Regresyon modeli anlamlıdır.





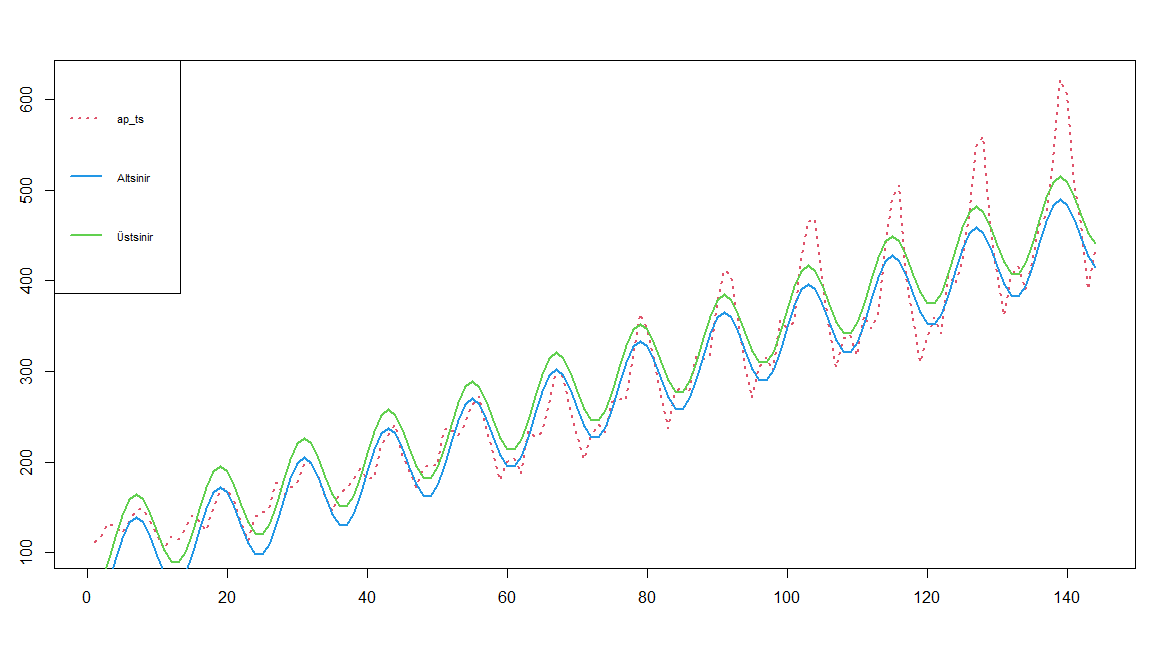
Yorum: p değeri 0.05 değerinden küçük olduğundan hipotezi reddedilir. Modelin anlamlı olduğu %95 güven düzeyinde söylenebilir.

= Modelin katsayıları anlamsızdır.

= Modelin katsayıları anlamlıdır

Yorum: sin1 ve cos1 ve t değerlerinin p değerleri <0.05 olduğu için katsayılarının anlamlı olduğunu söyleyebiliriz. Durbin-Watson test sonucunda p değeri 0,05’ten küçük olduğu için model anlamlıdır. H0 reddedilir. Böylelikle model denklemi:

= 88.1867 + 2.64988t -18.05531sin1 – 42.11438 +



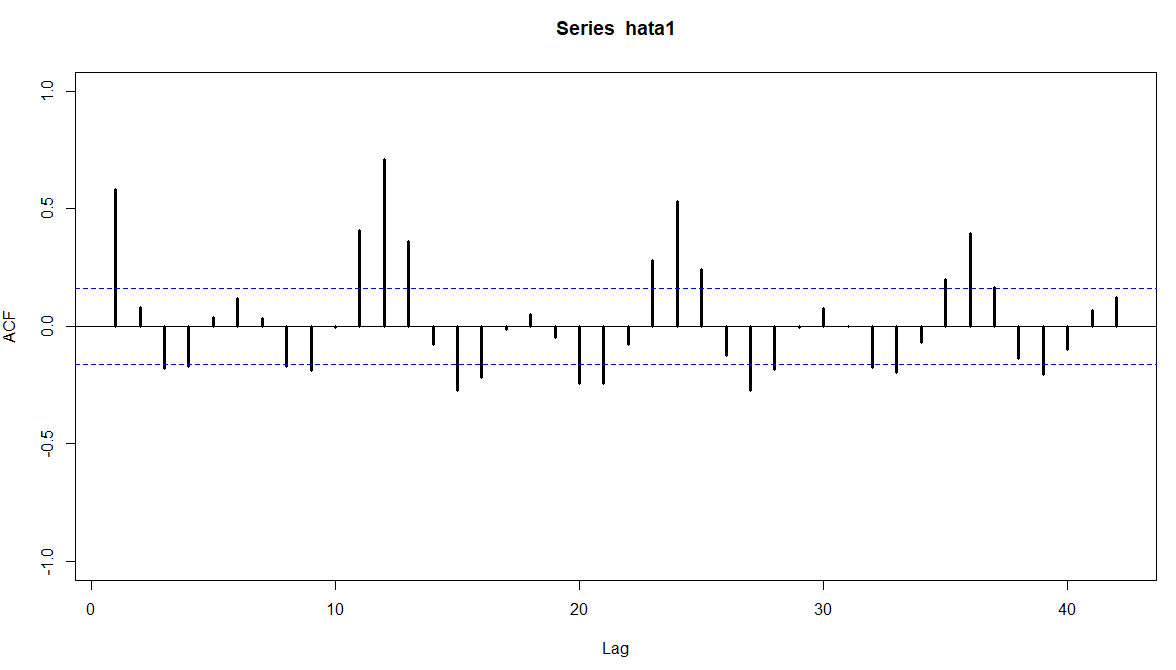
= 88.1867 + 2.64988t -18.05531sin1 – 42.11438 +

Yorum: Grafiğe bakıldığında bir çok verinin güven aralığı sınırları dışında kaldığını görürüz. Buradan hareketle bu modelin bu seri için uygulanmasının uygun olmadığı düşünülür. Nihai karar için D-W Testi uygulayalım.

: = 0

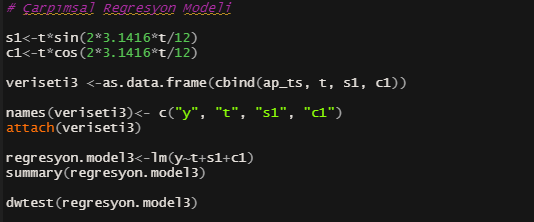
: 0

Yorum: Durbin Watson testine göre elde edilen p değeri 0.05’ten küçük olduğundan hatalar arasında özilişki olduğu %95 güven düzeyinde söylenebilir.



Yorum: Güven aralığını geçen birden fazla değer olduğundan ve önemli olduğundan hatalar arasında ilişki olduğu %95 güven düzeyinde söylenebilir. Hataların akgürültü serisi olmadığı, modelin anlamlı olmadığını söyleyebiliriz.

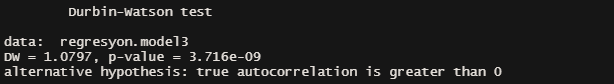
**3.2 Çarpımsal Regresyon Analizi**

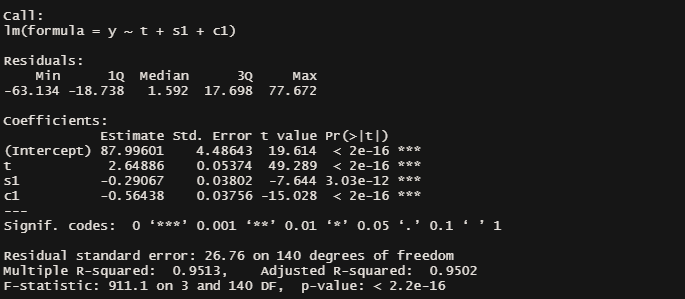


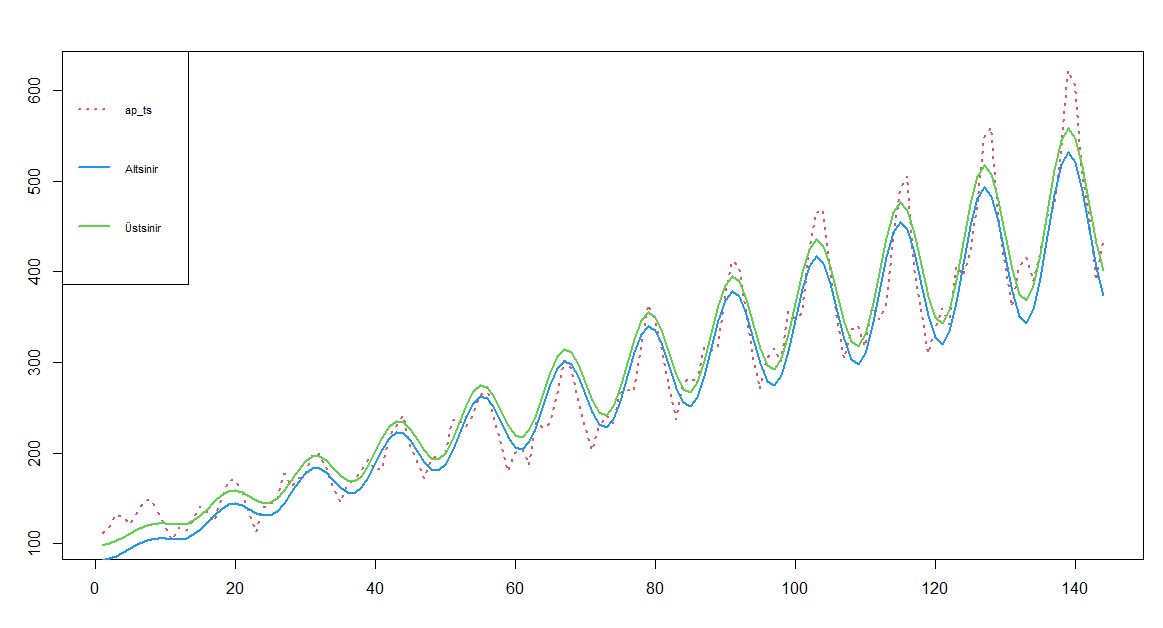
=Regresyon modeli anlamsızdır.

= Regresyon modeli anlamlıdır.

.





****

**MODEL TAHMİN GRAFİĞİ**

=Regresyon modeli anlamsızdır.

= Regresyon modeli anlamlıdır.

Yorum: p değeri 0.05 değerinden küçük olduğundan hipotezi reddedilir. Modelin anlamlı olduğu %95 güven düzeyinde söylenebilir.

= Modelin katsayıları anlamsızdır.

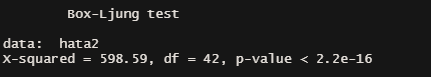
= Modelin katsayıları anlamlıdır.

Yorum: Elde edilen çıktıya göre t, cos1 ve sin1 katsayılarının anlamlı olduğu %95 güven düzeyinde söylenilebilir.

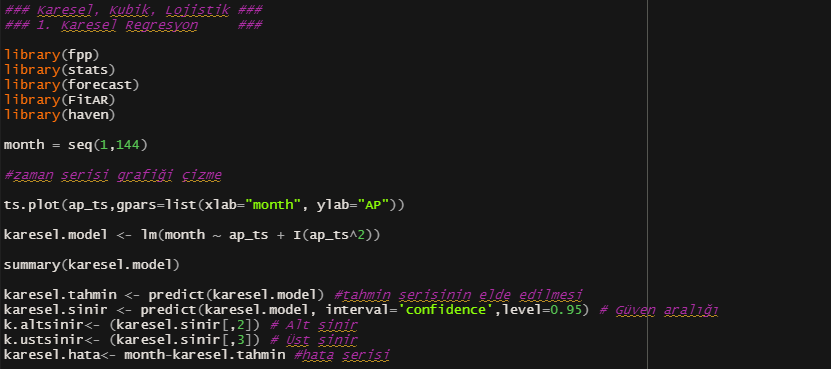
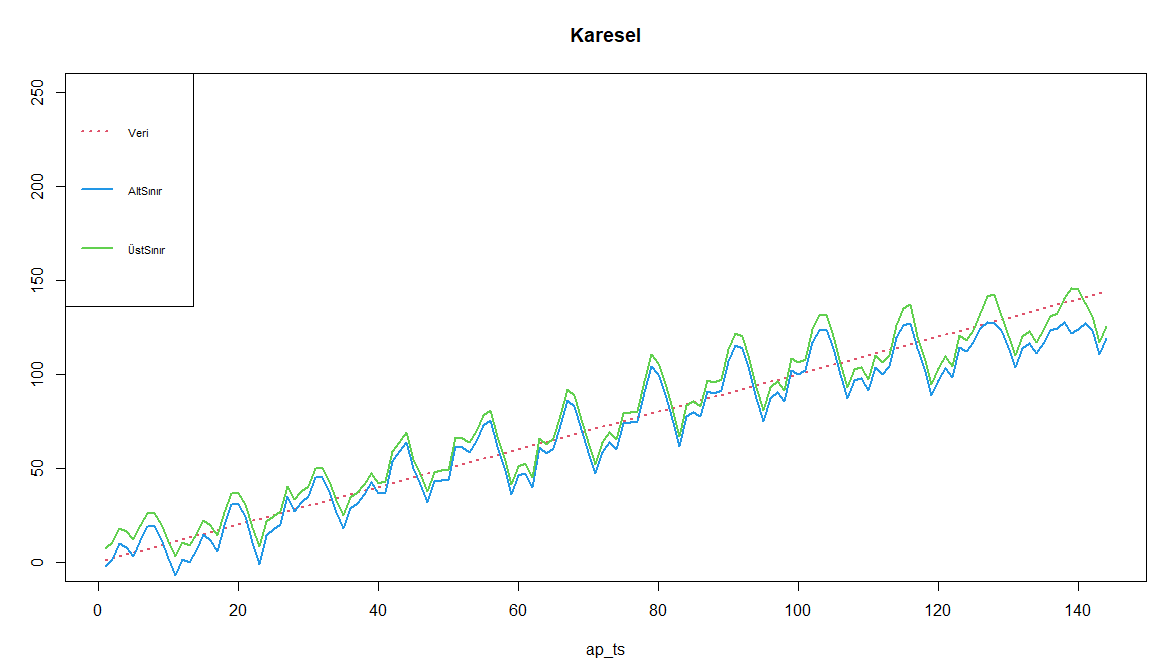
= 87.99601 + 2.64886t-0.29067sin1 – 0.56438cos1 +

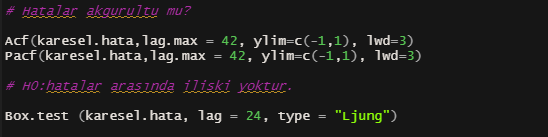
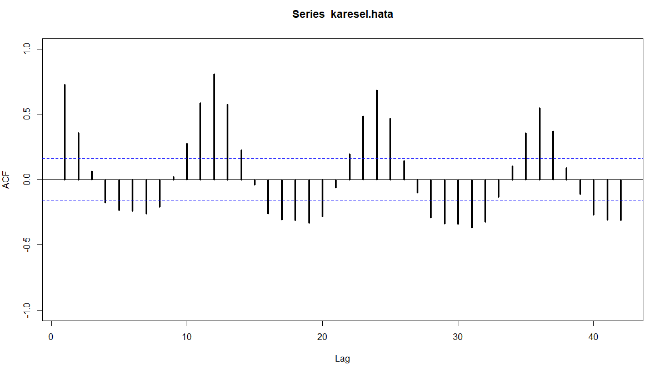
=Hatalar arasında ilişki yoktur.

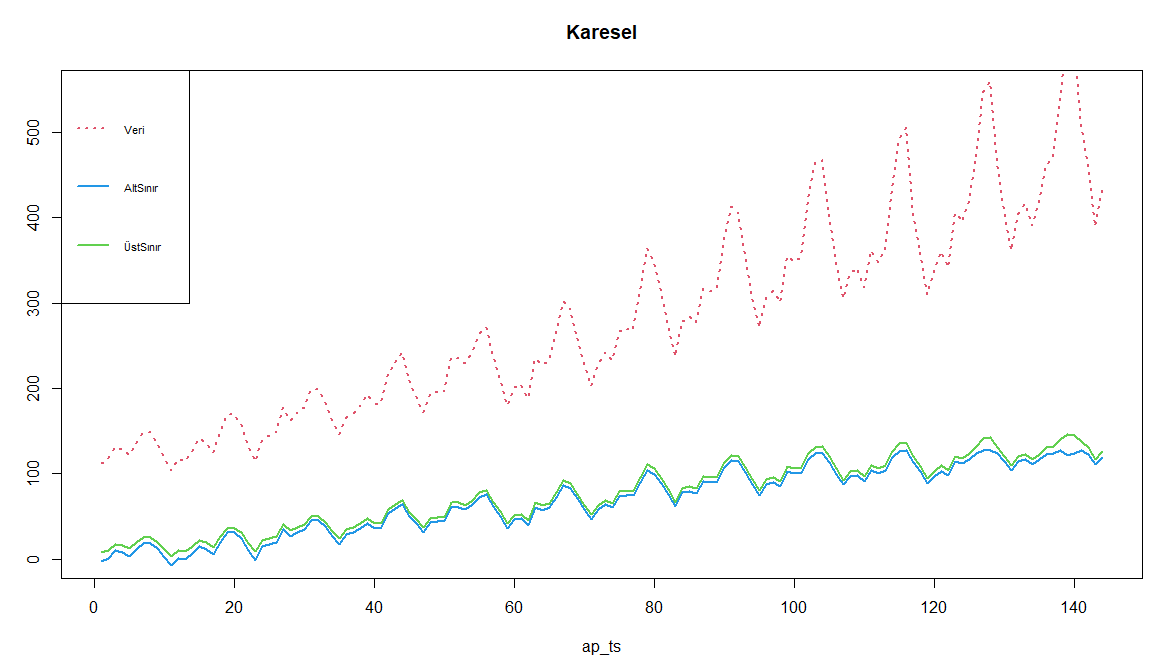
=Hatalar arasında ilişki vardır.



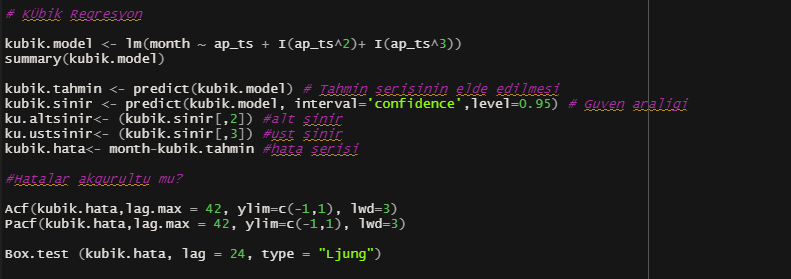
Yorum: Box-Ljung testinden elde edilen p değerinin 0.05 değerinden küçük olduğundan hatalar arasında ilişki olduğunu %95 güven düzeyinde söyleyebiliriz. Böylelikle modelin anlamlı olmadığını Hataların akgürültü serisi olmadığı, modelin anlamlı olmadığını söyleyebiliriz.

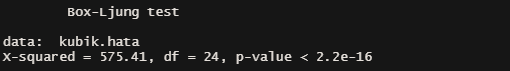
3.3 Karesel Regresyon

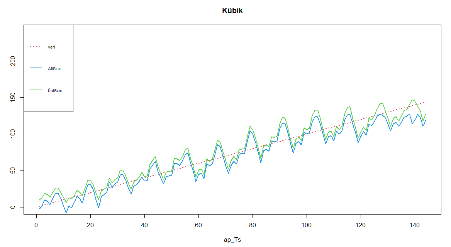
# Hatalar akgürültü mü?

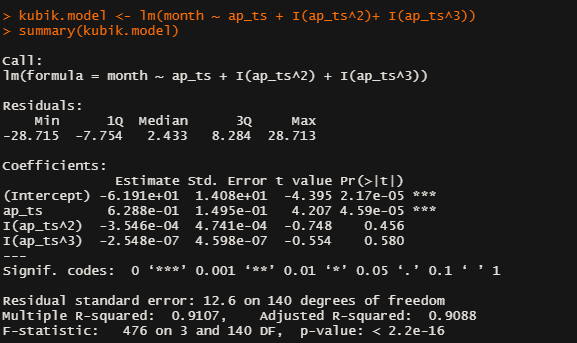
**Yorum:** Gecikmelerin önemli bir kısmı güven aralığının dışında olduğu için serimizin akgürültü serisi olduğunu söyleyemeyiz. Box-Ljung testimizin sonucu p <0,05 olduğu için hatalar arasında ilişki olduğu %95 güven düzeyinde söylenebilir. Bu model bu seri için kullanılamaz

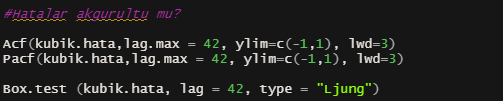
3.4 Kübik Regresyon

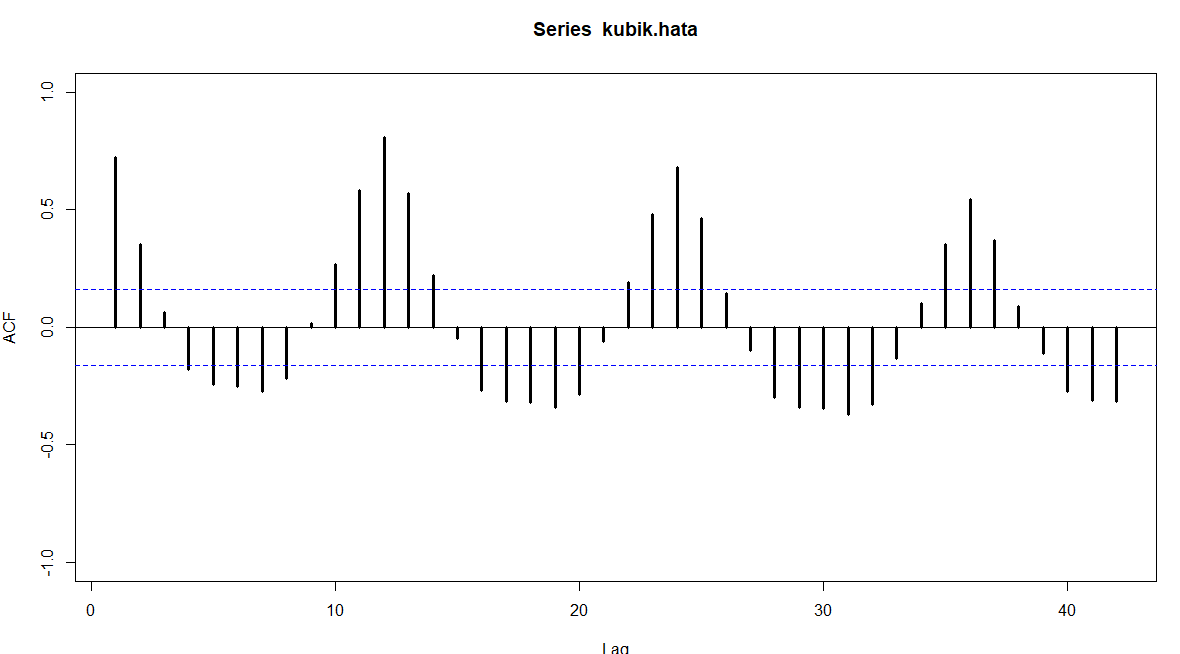
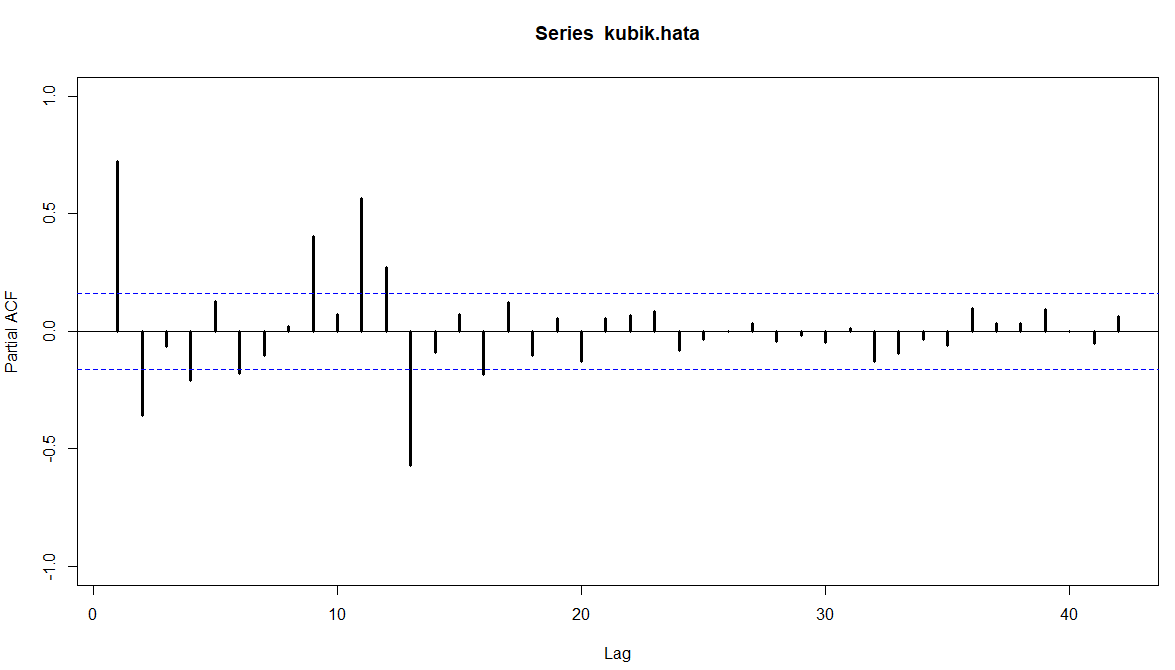




Yorum: p değeri 0,05’ten küçük olduğu için kübik model istatistiksel olarak anlamlıdır.

Yorum: Tahmin serimiz güven aralığı içinde yer almadığı için bu seri bu model ile açıklanamaz.

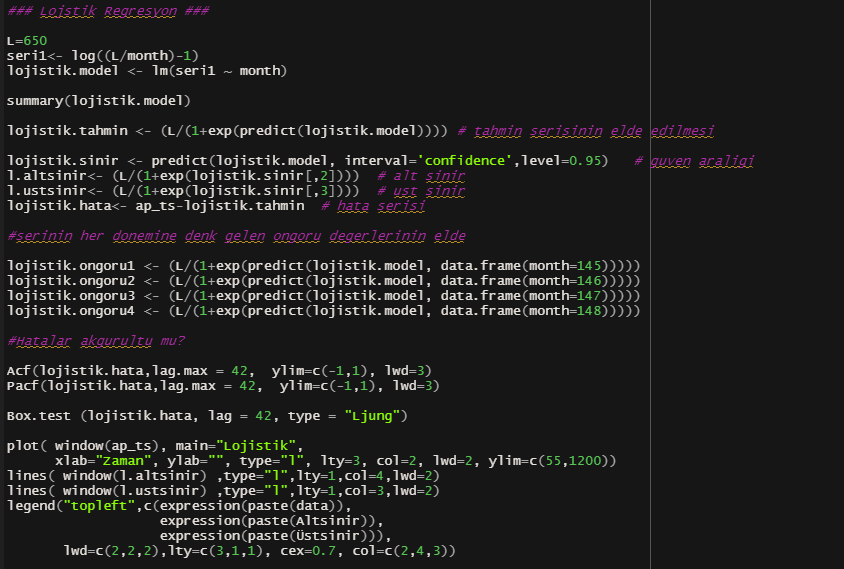


ACF ve PACF Grafikleri

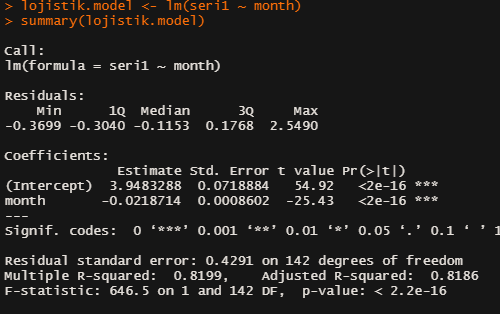
Hatalarımız güven aralığı dışında yer aldığı için akgürültü değildir. Bu model bu seriyi açıklayamaz.

3.5 Lojistik Regresyon

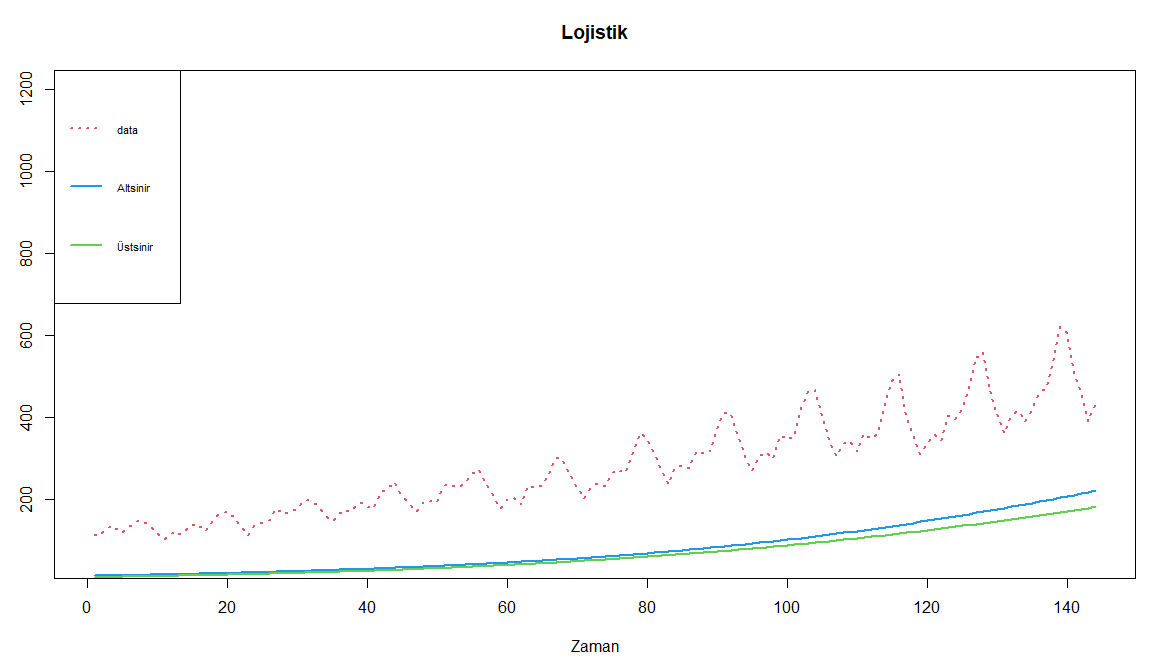
L değerini saptayabilmek için veri setimizdeki en büyük gözlemden daha büyük keyfi bir değer seçelim. Veri setimizin en büyük değeri 1.1’e baktığımızda 622 görürüz. L değerini 650 alalım.



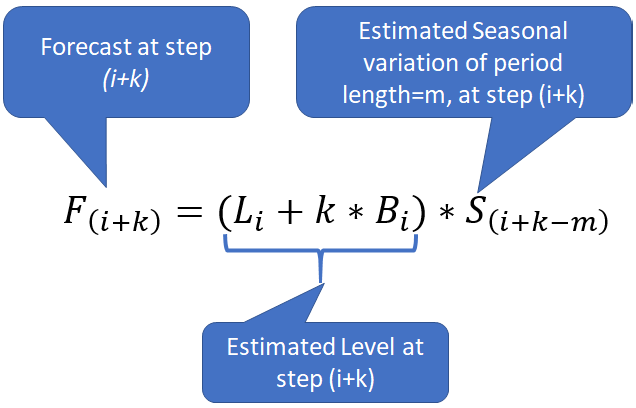
**Yorum**: p değeri 0,05’ten küçük olduğu için model istatistiksel olarak anlamlı değildir.



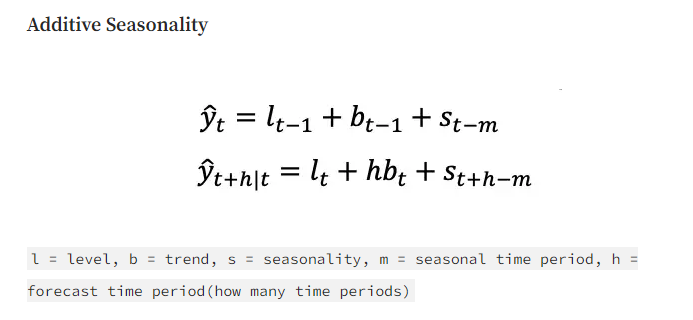
Tahmin değerleri güven aralığı içerisinde yer almadığından bu seri bu modelle tahmin edilemez.



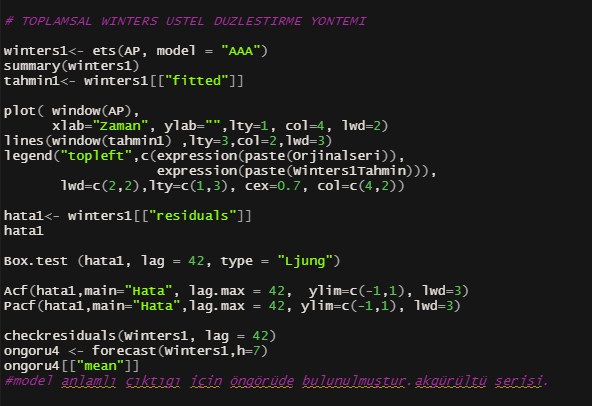
1. **WINTERS USTEL DÜZLEŞTİRME YÖNTEMİ**

****4.1 Toplamsal Winters Üstel Düzleştirme Yöntemi

**Serimizde hem trend hem de mevsimsellik mevcut. Bu sebepten Winters Üstel Düzleştirme kullanılır. (**https://timeseriesreasoning.com/contents/holt-winters-exponential-smoothing/)

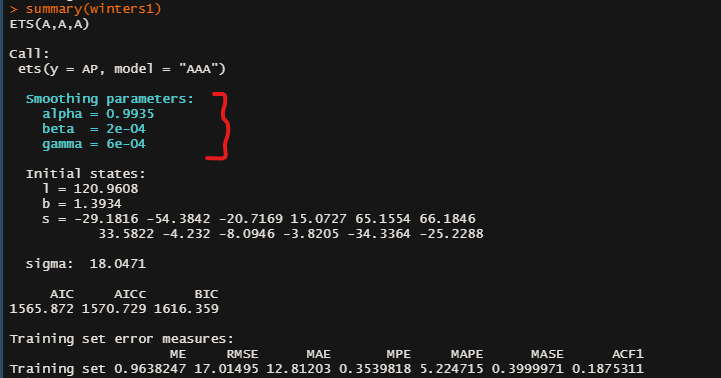
****

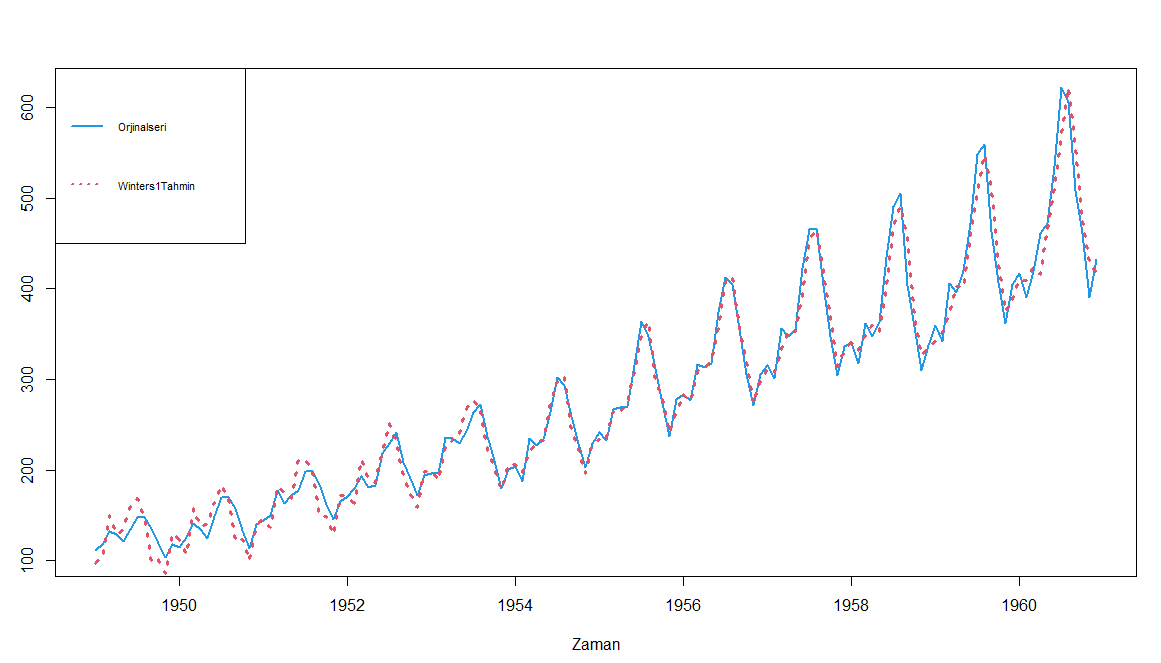
(https://otexts.com/fpp2/holt.html)

 α = 0,993

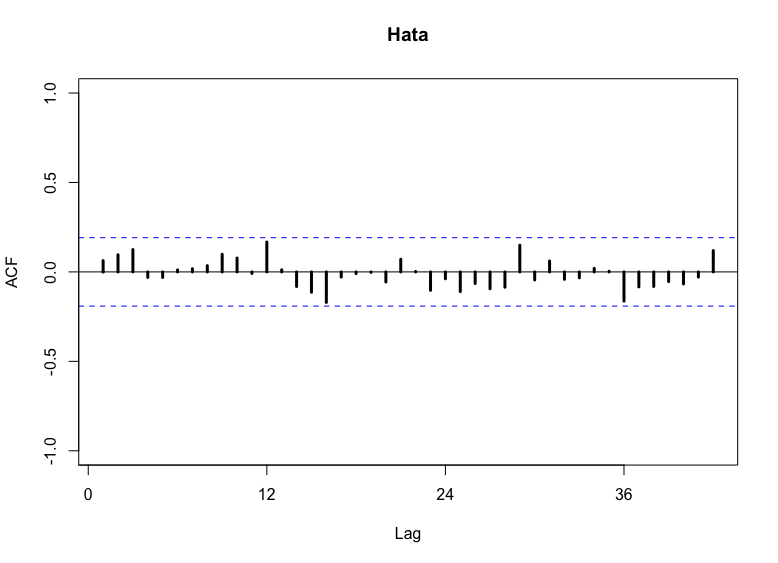
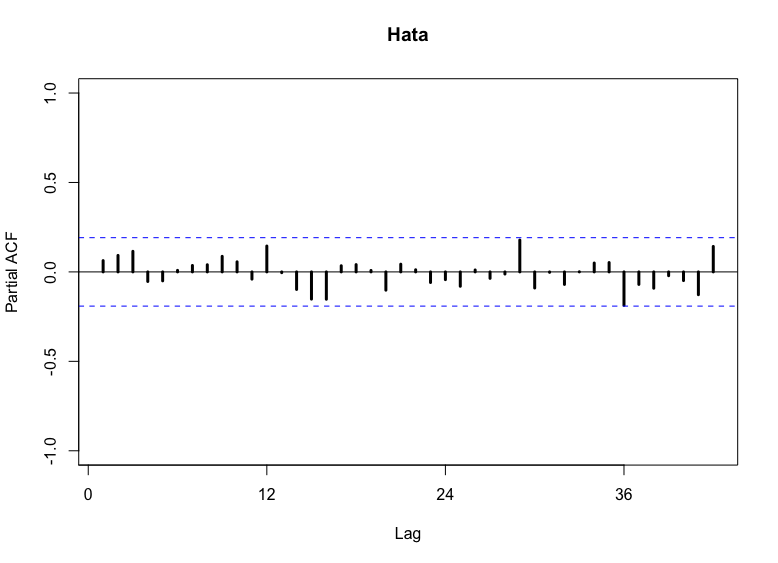
β = 0

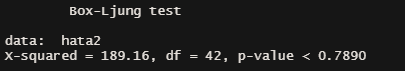
Λ = 0

* l = 120.9608
* b = 1.3934
* s değerlerimiz mevsimsel terimin başlangıç değerlerini temsil eder.
* BIC değeri 1616,359’dur.
* Winters1Tahmin serisi ile orijinal seri grafiğini çizdirdiğimizde birbirine yakın değerleri görebiliriz. Tahmin ve orijinal serinin benzerliğinden söz edebiliriz.
* ****Toplamsal Winters Yöntemi HKO’nın karekökü (RMSE) 17.01495’dir.



Yorum:Grafiğe bakıldığında orijinal seri ile toplamsal düzleştirme yöntemiyle elde edilen tahmin serisi ile arasında bir uyum olduğu söylenebilir. Fakat hata teriminin kontrol edilmesi gerekir.

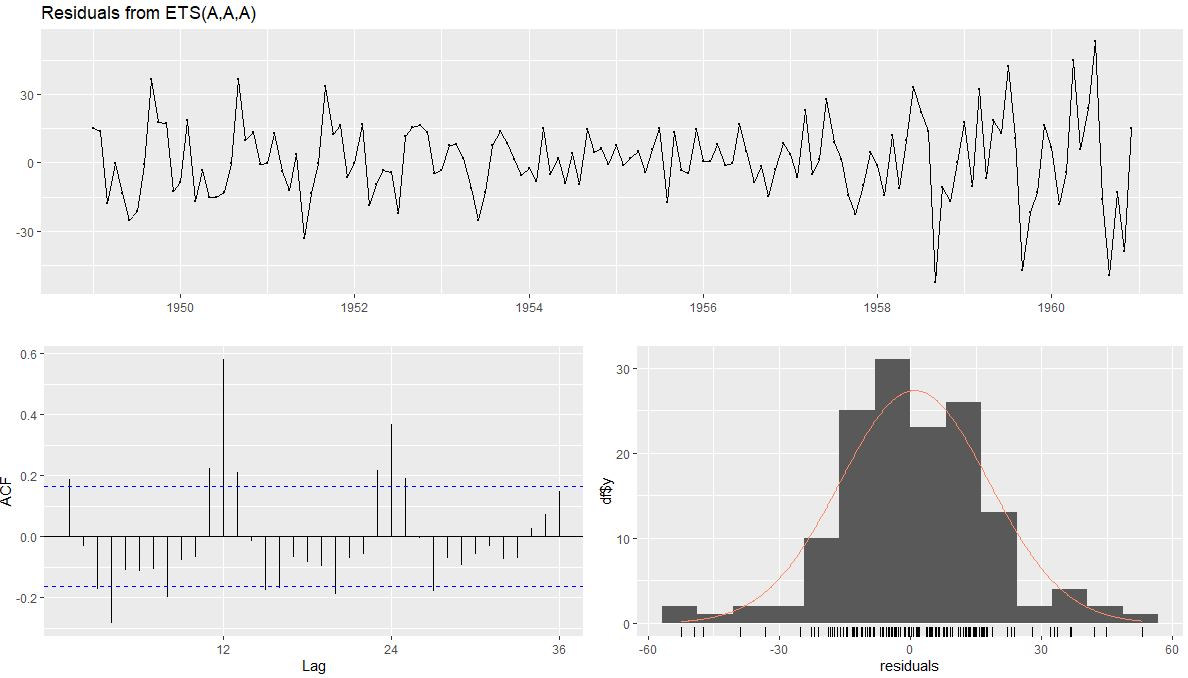




=Hatalar arasında ilişki yoktur.

=Hatalar arasında ilişki vardır.

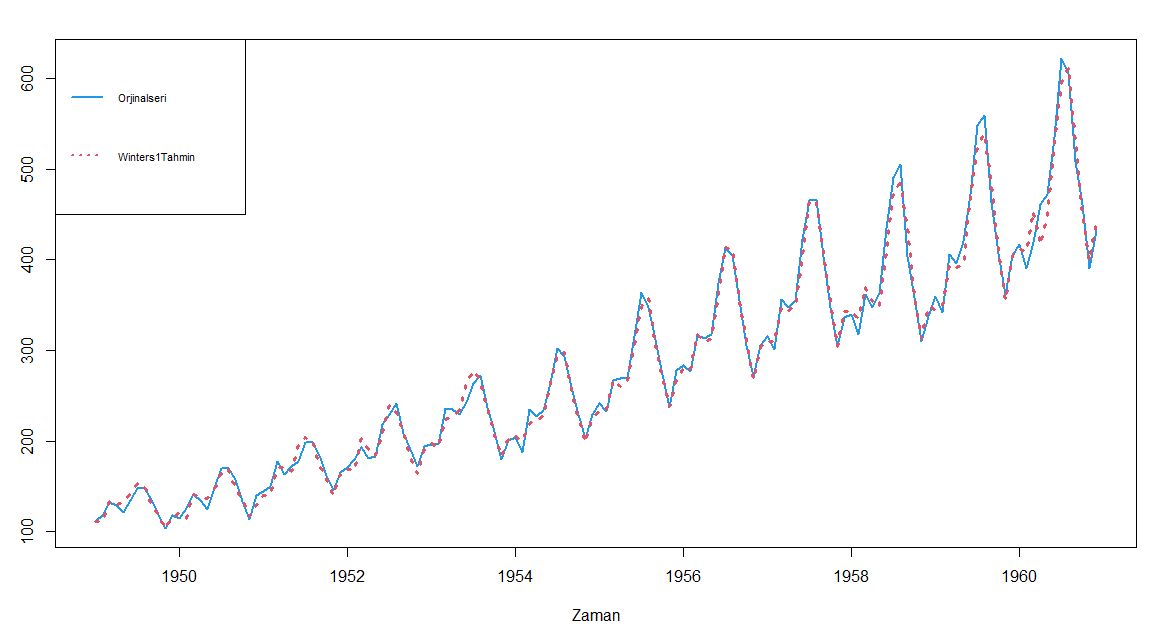
Yorum: Box-Ljung testinden elde edilen p -değerine göre hatalar arasında ilişki olmadığını %95 güven düzeyinde söylenebilir. Hata serisinin akgürültü olduğu söylenebilir. Bu model bu seriyi tahmin etmekte kullanılabilir.

****

* Kalıntıların ortalaması sıfıra yakındır ve artıklar serisinde anlamlı bir korelasyon yoktur.
* Tarihsel veriler boyunca hemen hemen aynı kaldığını ve bu nedenle artık varyansın sabit olarak ele alınabileceğini gösterir.

Yorumlanması: https://otexts.com/fpp2/residuals.html

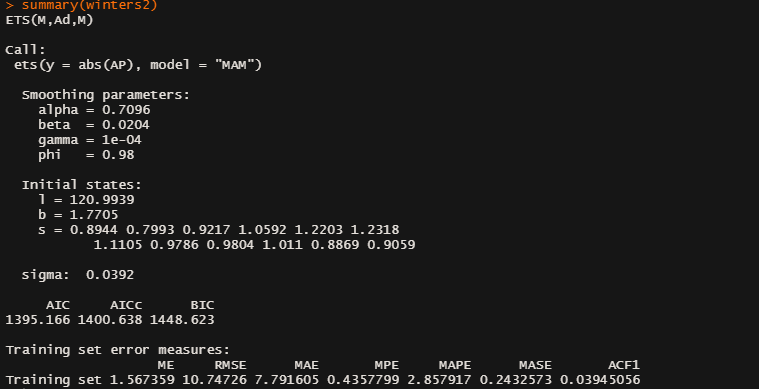
4.2 Çarpımsal Winters Üstel Düzleştirme Yöntemi

****

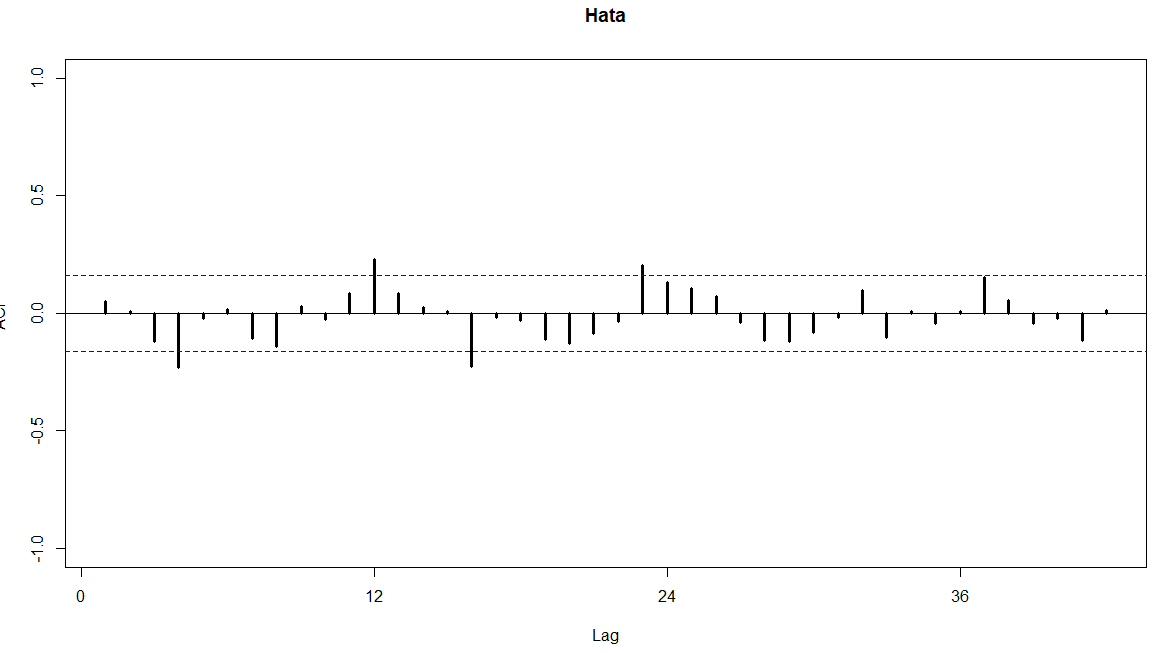
**Tahmin Grafiği**

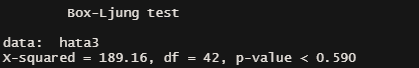
****

Yorum:Çarpımsal düzleştirme yöntemi için modeli “MAM” olarak değiştirdik.

****

* **BIC değeri = 1448.623**
* **HKO’nın karekökü =10.74**

** Yorum**: Hataların akgürültü olduğunu ve bu modelin anlamlı olduğunu %95 güven düzeyinde söyleyebiliriz.

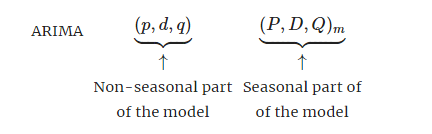


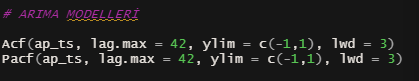
**Yorum:** Her iki modelimiz de anlamlıdır. Bu sebepten BIC değeri düşük olan yöntemi seçeriz bu da Çarpımsal modelde mevcuttur.

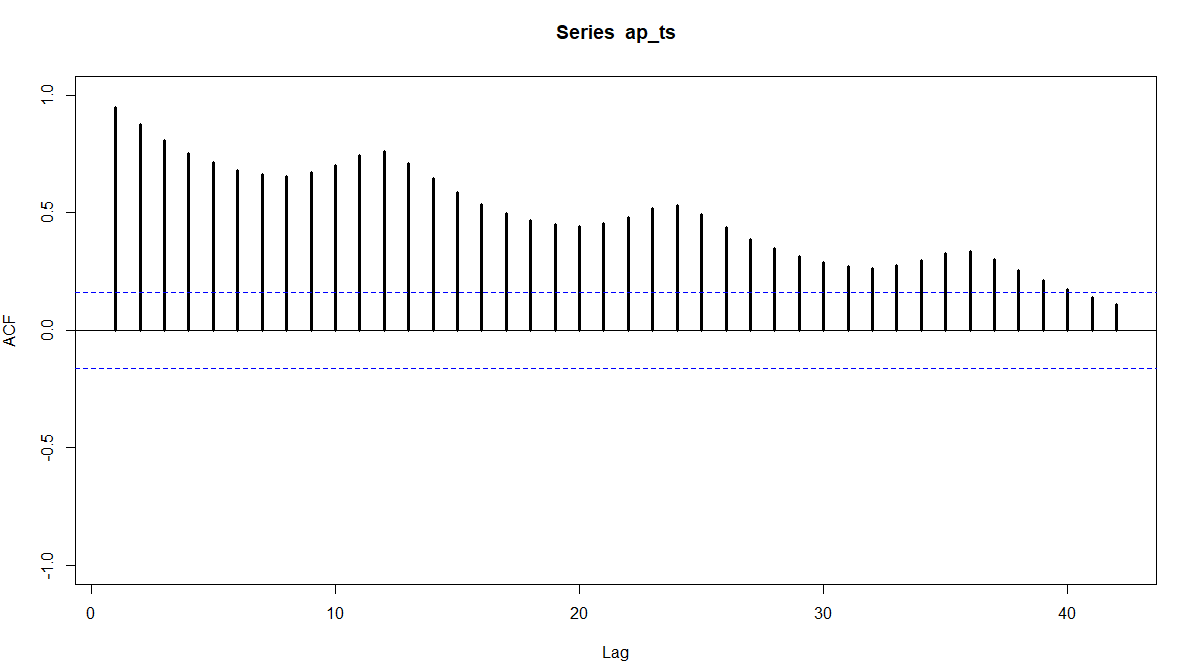
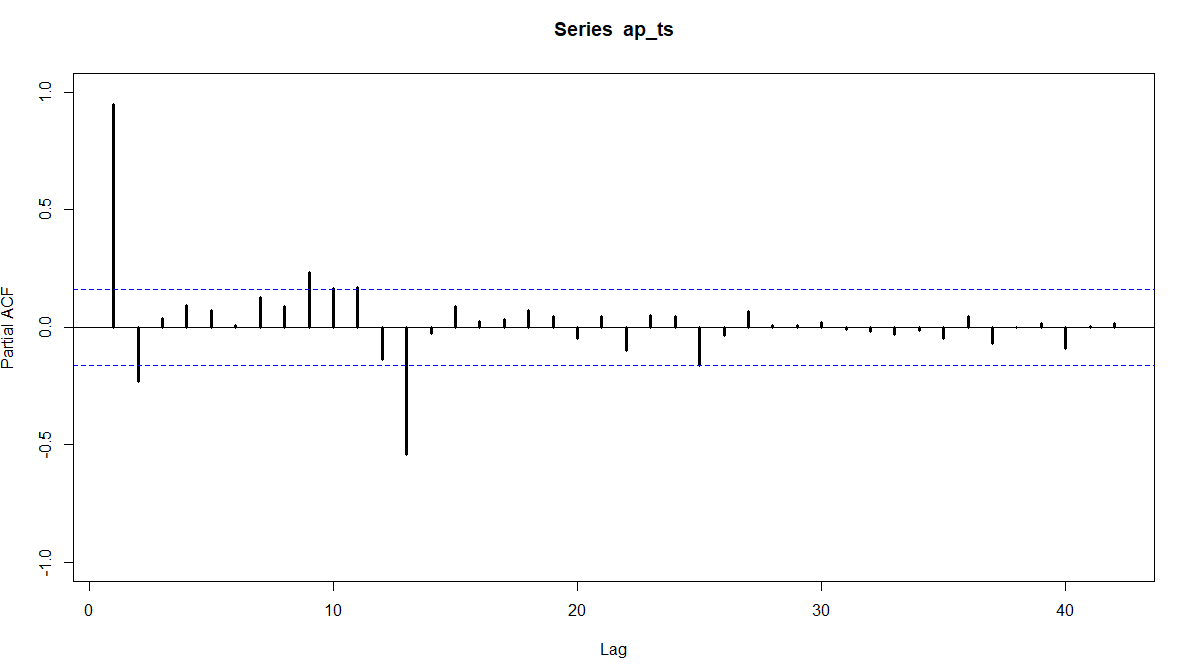
Çarpımsal Yöntemle ilk 5 ayın tahmini yapılır.



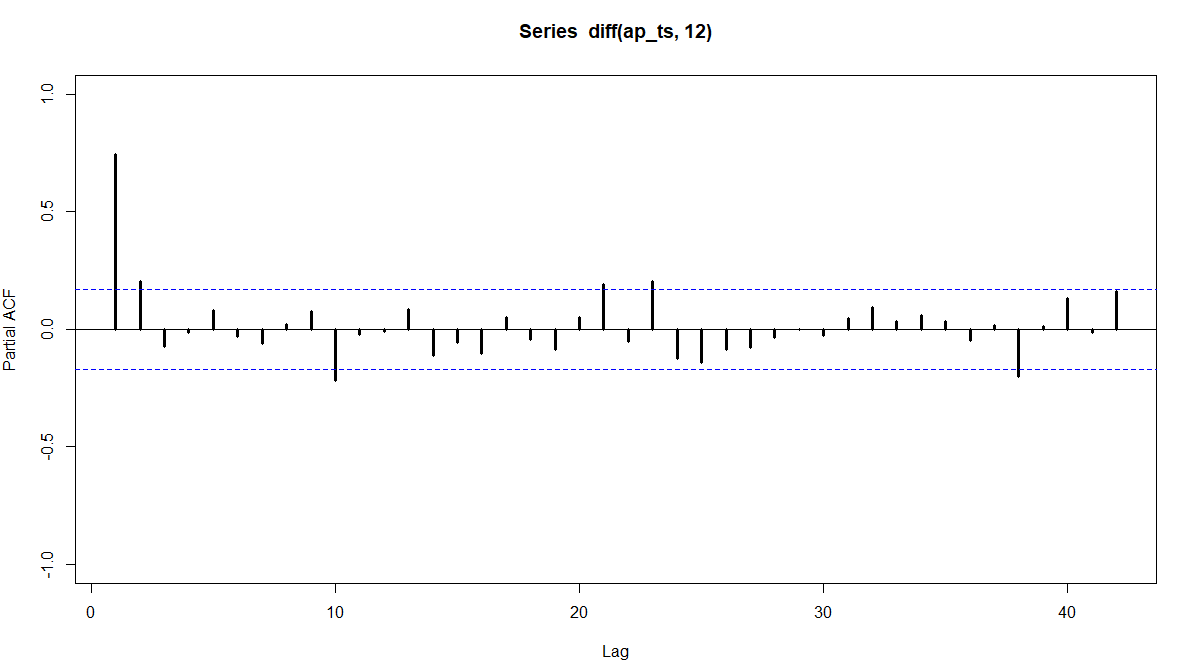
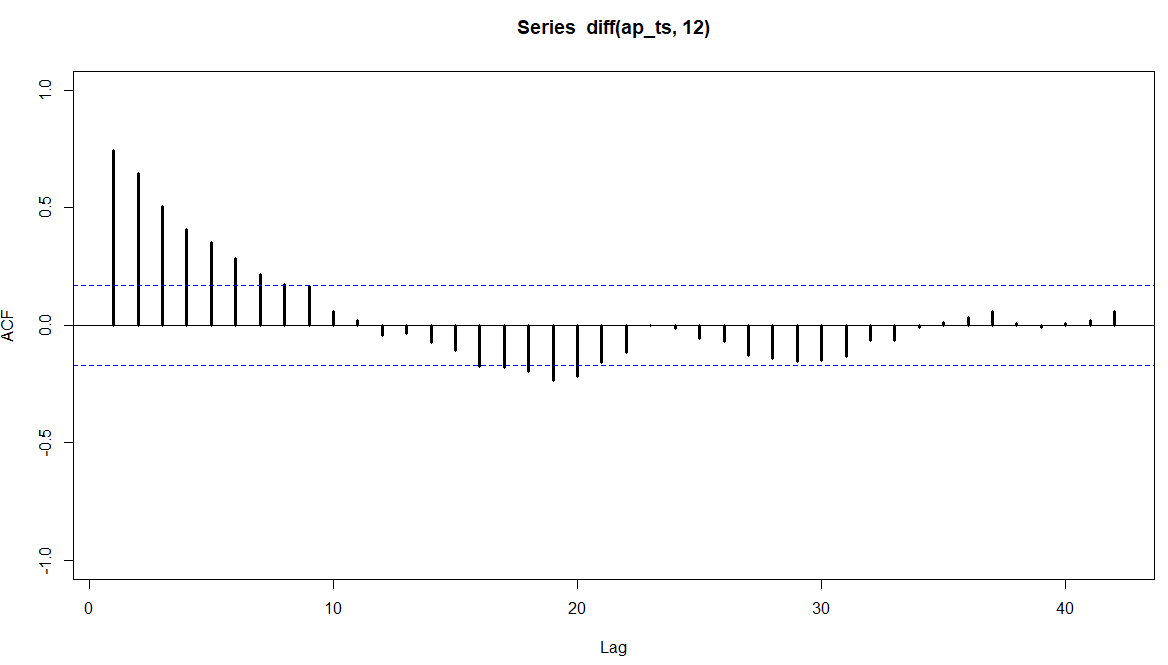
1. **BOX-JENKINS Modelleri (AutoRegressive Integrated Moving Average)**





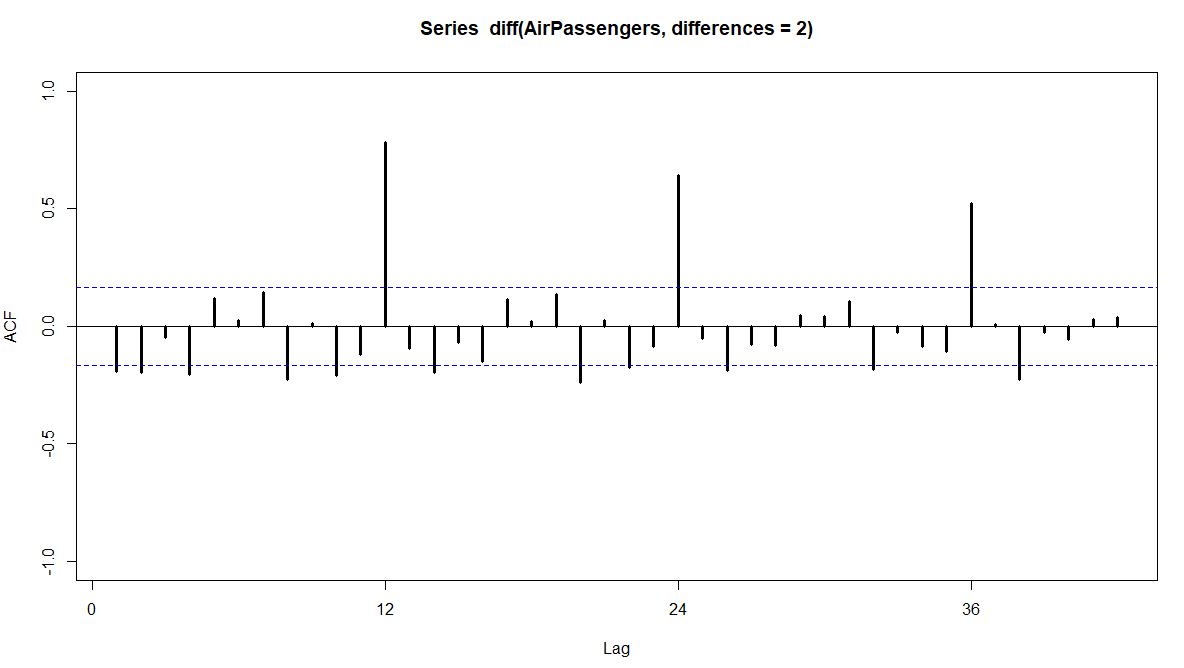


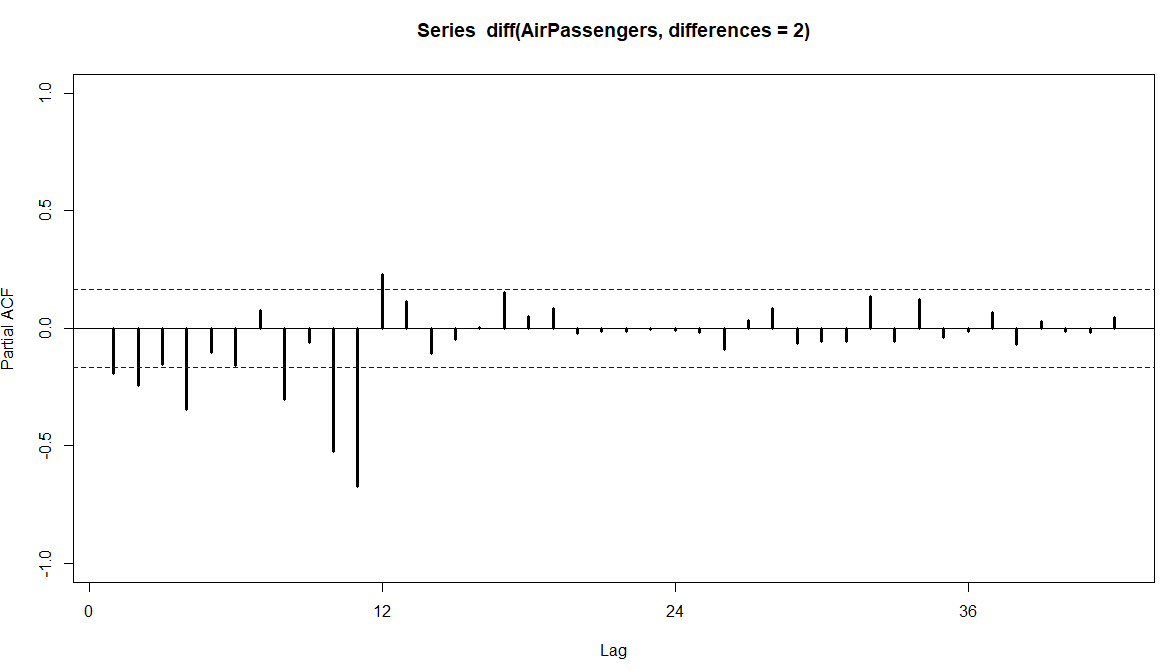
Anlamlı düzeyde serimiz durağan bir seri değildir ve mevsimsellik söz konusudur. Bu nedenle Box-Jenkins modellerine geçiş yapmadan önce seri mevsimsellikten arındıralım.



Acf(diff(AirPassengers, differences = 2),lag.max = 42, ylim = c(-1,1), lwd = 3)

Pacf(diff(AirPassengers, differences = 2),lag.max = 42, ylim = c(-1,1), lwd = 3)





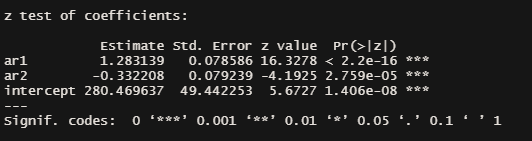
Serinin durağanlaştırılmış ACF ve PACF grafiklerine bakıldığında ACF grafiğinin PACF grafiğine göre daha yavaş azaldığı görülmektedir. Bu sebepten serimize uygun model otoregresyon modelidir. PACF grafiğinde ilk üç gecikme ait ilişki önemli olduğundan p = 2’dir.

Uygulanacak Model ->

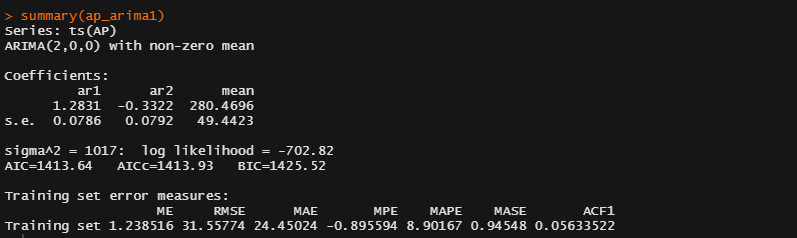
= modeli anlamsızdır.

= modeli anlamlıdır.





**Yorum:** Ar1 ve Ar2 p-values < 0.05 olduğundan ötürü H0 reddedilir. Model anlamlıdır. İşleme devam edilir.

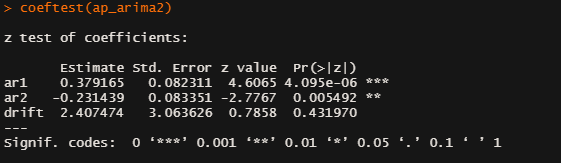


RMSE = 31.55774 , BIC = 1425.52 şeklinde elde ettik. Bu değerler en doğru modeli seçmemizde bize rehberlik edecektir.

= modeli anlamsızdır.

= modeli anlamlıdır.



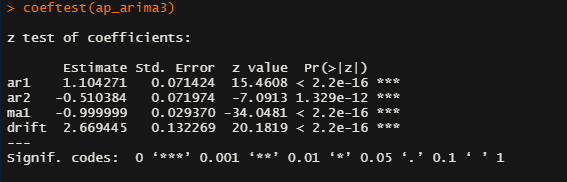


**Yorum:** ar1 anlamlı fakat ar2 anlamsız çıktı. Bu nedenle devam edilmez.

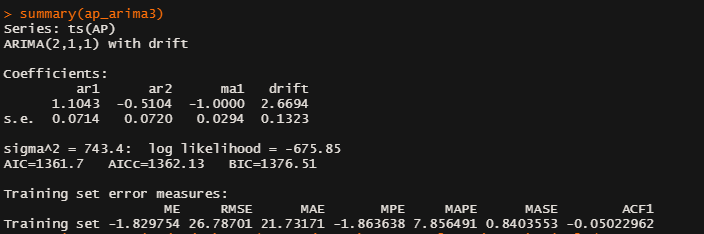
= modeli anlamsızdır.

= modeli anlamlıdır.





**Yorum:** ar1, ar2 ve ma1 değerlerli istatiksel olarak anlamlıdır.

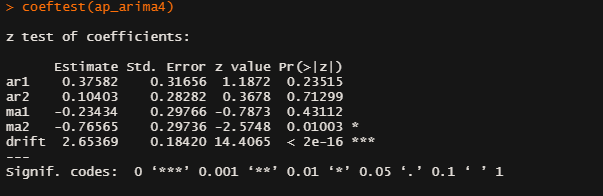


RMSE = 26.78701 , BIC = 1376,51 şeklinde elde ettik. Bu değerler en doğru modeli seçmemizde bize rehberlik edecektir.

= modeli anlamsızdır.

= modeli anlamlıdır.



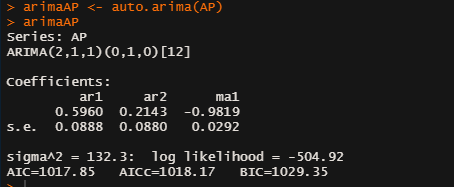


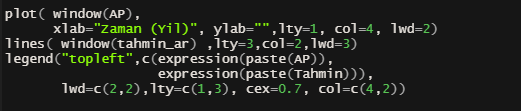
Yorum: Ar1, sar1 anlamlı ancak sar2 anlamsız çıkıyor. Bu nedenle daha fazla devam edilemez.

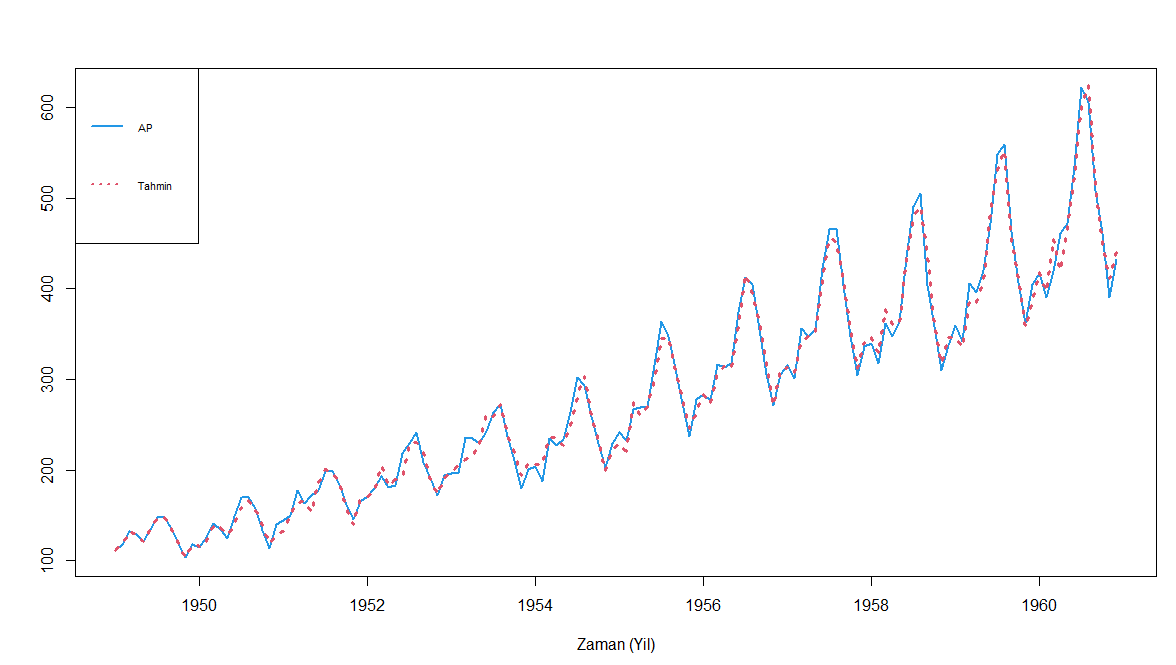
RMSE ve BIC değeri en küçük olan anlamlı modelimiz olduğundan bu model kullanılarak öngörü yaparız.

= 2.6694 + 0.1323 +

Doğru modeli bulabilmek adına, mevsimsellik de dahil olmak üzere varsayılan parametreler doğru olarak verildiğinde, en iyi modeli ve katsayıları sığdırmak için R paketindeki auto.arima fonksiyonunu da kullanabiliriz. ( Referans: <http://rstudio-pubs-static.s3.amazonaws.com/311446_08b00d63cc794e158b1f4763eb70d43a.html>)

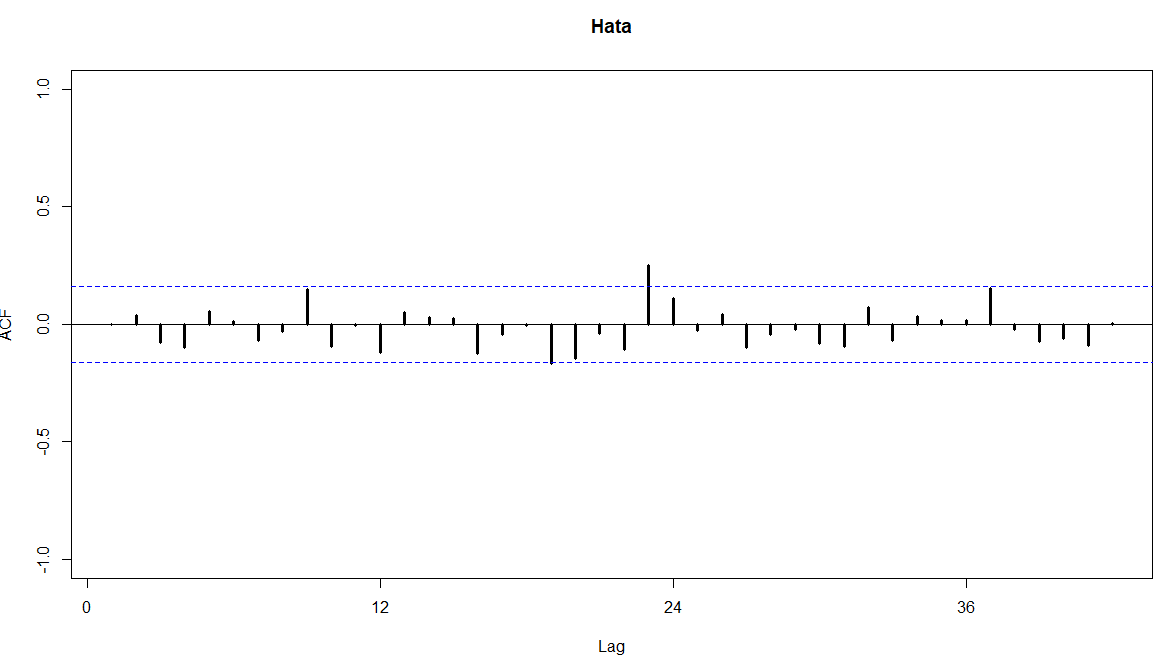
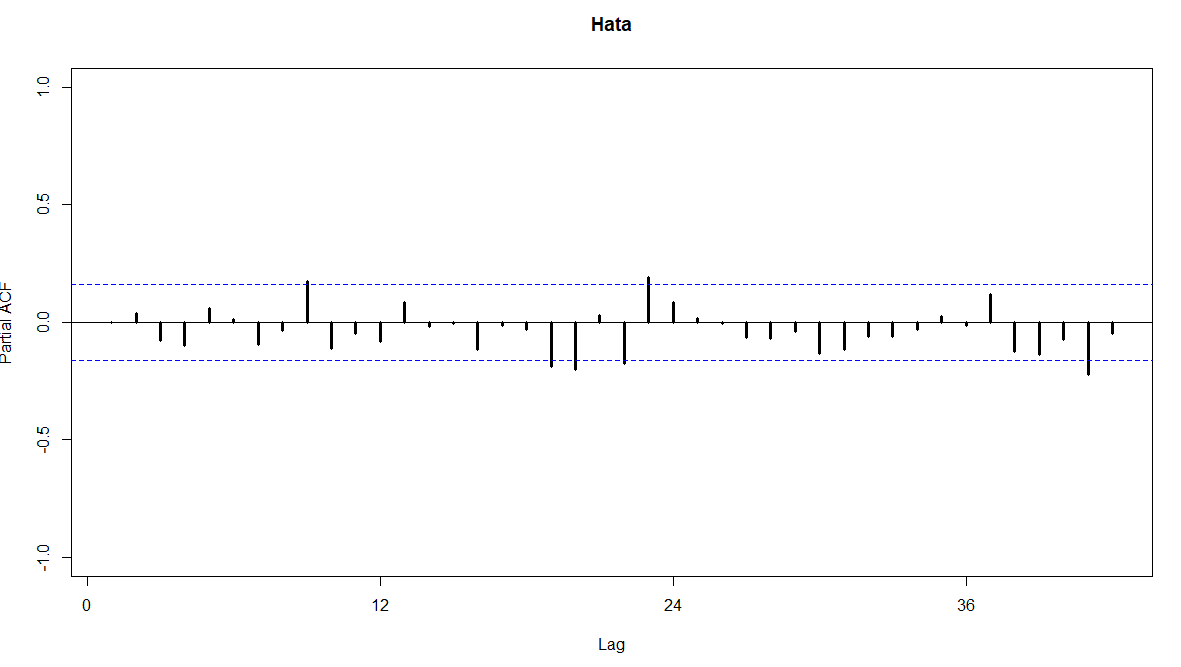




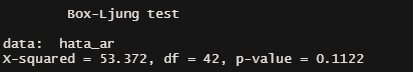


**Yorum:** Orijinal seri ile tahmin serimiz arasında oldukça güçlü bir uyum tespit ettik. Hataların akgürültü olup olmadığını test edelim.





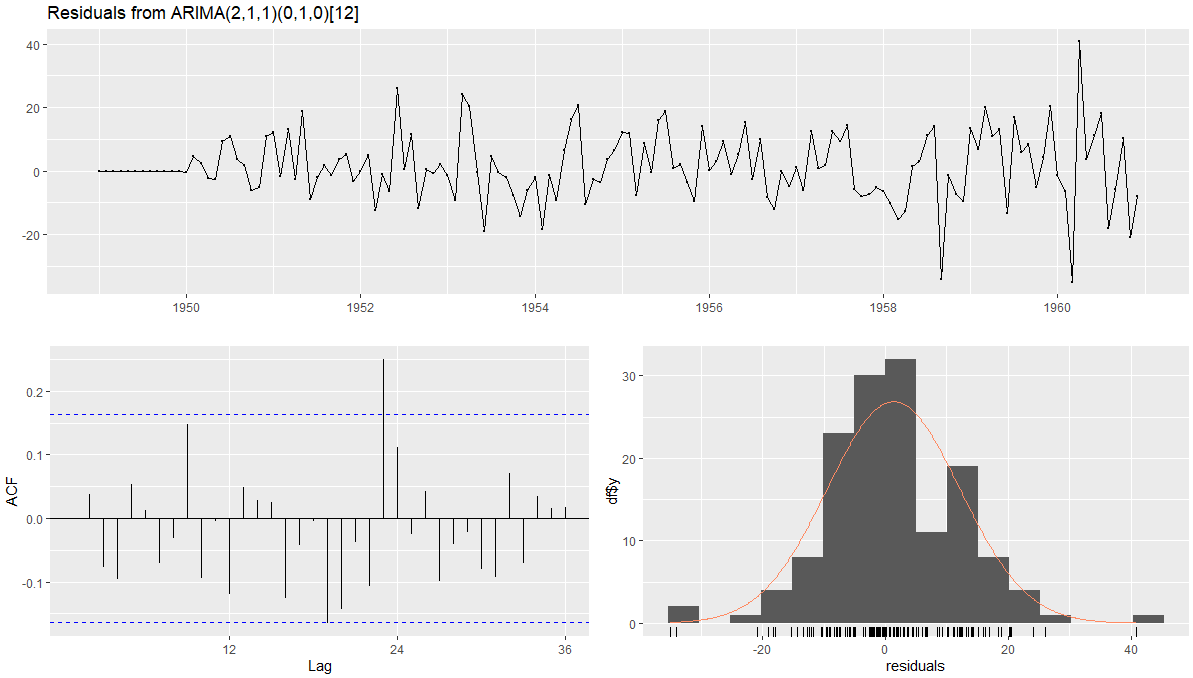
**Yorum:** ACF grafiğine baktığımızda güven aralığını geçen 1 değer saptanmıştır. Fakat önemli olup olmadığına yorum getirebilmek adına birim-kök testi yapılır.



=Hatalar arasında ilişki yoktur.

=Hatalar arasında ilişki vardır.

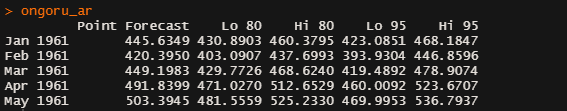
**Yorum:** Hataların ACF grafiğine baktığımızda sınırı geçen gecikmeler olduğunu gözlemledik ve Box-Ljung testinden elde edilen p-değeri 0.05 değerinden büyük olduğundan hatalar arasında ilişki olmadığını %95 güven düzeyinde söyleyebiliriz. Böylelikle modelin anlamlı olduğunu söyleriz.



**Yorum:**

* Kalıntıların ortalaması sıfıra yakındır ve artıklar serisinde anlamlı bir korelasyon yoktur.
* Tarihsel veriler boyunca hemen hemen aynı kaldığını ve bu nedenle artık varyansın sabit olarak ele alınabileceğini gösterir.





KAYNAKÇA

[https://www.researchgate.net/figure/AirPassengers-time-series\_fig1\_333645661 (18](https://www.researchgate.net/figure/AirPassengers-time-series_fig1_333645661%20(18)Aralık 2022 16:24 tarihinde erişildi)

<https://rpubs.com/emb90/137525> (260 Aralık 09:36 tarihinde erişildi)

<https://towardsdatascience.com/time-series-forecasting-in-r-with-holt-winters-16ef9ebdb6c0> (27 Aralık 19:40 tarihinde erişildi)

<https://ionides.github.io/531w20/midterm_project/project37/Midterm_project.html> (27 Aralık 19:46 tarihinde erişildi)

<https://stats.stackexchange.com/questions/285100/logistic-regression-on-time-series-data> (27 Aralık 19:46 tarihinde erişildi)

<https://machinelearningmastery.com/time-series-datasets-for-machine-learning/> (27 Aralık 19:56 tarihinde erişildi)

[https://towardsdatascience.com/how-to-remove-non-stationarity-in-time-series-forecasting- 563c05c4bfc7](https://towardsdatascience.com/how-to-remove-non-stationarity-in-time-series-forecasting-%20%20563c05c4bfc7) (27 Aralık 19:56 tarihinde erişildi)

<https://rpubs.com/neharaut05/TimeSeries_AirPassangerForecast> (27 Aralık 20:06 tarihinde erişildi)

<https://www.analyticsvidhya.com/blog/2015/12/complete-tutorial-time-series-modeling/> (27 Aralık 20:10 tarihinde erişildi)

<https://www.kaggle.com/code/mirzahasnine/amazon-stocks-prediction-time-series-analysis> (30 Aralık 13:40 tarihinde erişildi)

<https://towardsdatascience.com/interpreting-acf-and-pacf-plots-for-time-series-forecasting-af0d6db4061c> (30 Aralık 13:40 tarihinde erişildi)

<https://rdrr.io/r/datasets/AirPassengers.html> (20 Aralık 14:40 tarihinde erişildi)

<https://github.com/selva86/datasets/blob/master/AirPassengers.csv> (20 Aralık 14:40 tarihinde erişildi)

<https://github.com/raffg/air-passengers-arima> (1 Ocak 2023 18:30 tarihinde erişildi)

<https://otexts.com/fpp2/seasonal-arima.html> (1 Ocak 2023 18:30 tarihinde erişildi)

[https://otexts.com/fpp2/seasonal-arima.html (3](https://otexts.com/fpp2/seasonal-arima.html%20(3) Ocak 2022 03:51 Tarihinde erişildi)