1. **Veri Tanımı**

**Günlük Web Sitesi Ziyaret Durumu**

Bu veri seti statforecasting.com adlı istatistiksel tahmin öğretim notlarını içeren web sitesindeki çeşitli ziyaret trafiği için 5 yılı kapsayacak şekilde günlük zaman serisi verilerini içermektedir. Değişkenler; günlük sayfa yükleme sayıları, tekil ziyaretçiler, ilk kez gelen ziyaretçiler ve web sitesine geri dönen ziyaretçilerdir.

Veri setimiz, 14 Eylül 2014 ile 19 Ağustos 2020 arasındaki tarih aralığından oluşan 2167 satır 8 sütundan oluşmaktadır.

Ziyaret, tanımlandığı şekilde aynı kullanıcı tarafından belirli bir günde sitedeki bir veya daha fazla sayfaya yapılan giriş akışı olarak tanımlanır. Aynı IP adresini paylaşan birden fazla kişi (örneğin bir bilgisayar laboratuvarında) tek bir kullanıcı olarak kabul edilir, bu nedenle gerçek kullanıcı sayısı bir dereceye kadar eksik sayılabilir. Son 6 saat içinde aynı IP adresinden bir giriş gelmemişse, ziyaret "benzersiz(unique)" olarak sınıflandırılır. Siteyi tekrardan ziyaret eden kişiler(Returning.Visits), kabul ettikleri takdirde çerezler tarafından tanımlanır. Diğerleri ilk kez gelen ziyaretçiler(First.Time.Visits) olarak sınıflandırılır, bu nedenle benzersiz ziyaretçi sayısı(Unique.Visits), tanım gereği geri gelen ve ilk kez gelen ziyaretçilerin sayılarının toplamıdır. Veriyle ilgili zaman serisi analizlerine başlamadan önce veride eksik değer olup olmadığı kontrol edilmiştir ve kontrollerimiz sonucu eksik değer olmadığı sonucuna ulaşılmıştır.

**Değişkenler**

**Row:** Satır Numaraları.

**Day:** Text formatında Günler.

**Day.Of.Week:** Haftanın Günleri.(1:Pazar(Sunday), 2:Pazartesi(Monday), ...)

**Date:** Tarih(Ay\Gün\Yıl).

**Page.Loads:** Günlük Sayfa Yüklenmeleri.

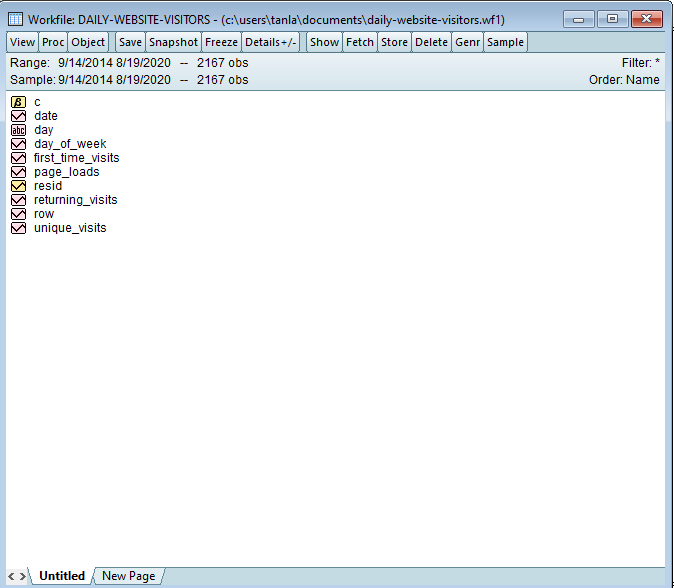
**Unique.Visits:** Benzersiz ziyaretçi sayısı. İlk kez gelen ziyaretçi sayısı ve siteyi tekrardan ziyaret eden kişi sayısının toplamıdır. (IP adreslerinden 6 saatten fazla bir süredir herhangi bir sayfaya ulaşılmayan günlük ziyaretçi sayısı)

**First.Time.Visits:** İlk kez gelen ziyaretçi sayısı.

**Returning.Visits:** Siteyi tekrardan ziyaret eden kişiler.

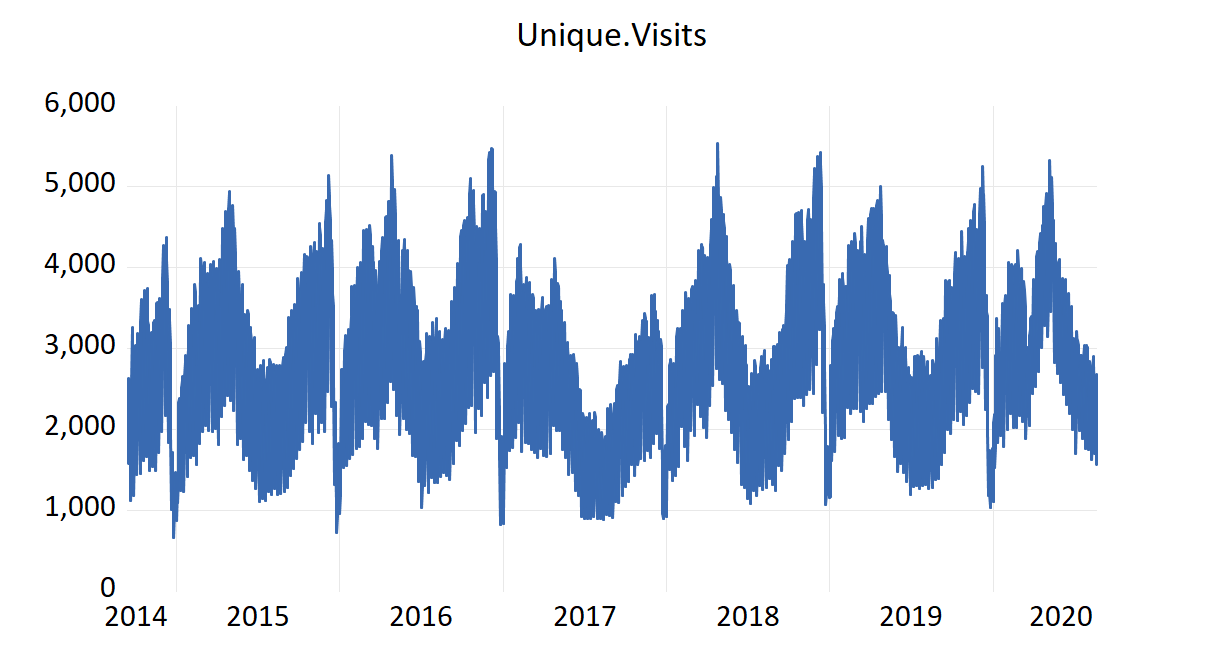
**Veri Linki:**

https://www.kaggle.com/datasets/bobnau/daily-website-visitors



Veri setimizdeki eksik gözlem varlığını analizlerimize başlamadan önce sorguladık ve eksik değer olmadığını gördük. Sonrasında verimizi EViews’da içe aktardık ve analizimize hazırladık.

**Çizgi Grafiği**



Bu grafiğe baktığımız zaman sarmal bir yapıda olduğunu görüyoruz. Bu tür grafikler mevsimselliğin göstergesidir. Şu an için grafikten hareketle mevsimsellik var diyebiliriz. Grafiğe baktığımızda verimiz durağan gözükmektedir. Daha sonrasında analizler yardımıyla daha net bir karara varacağız.

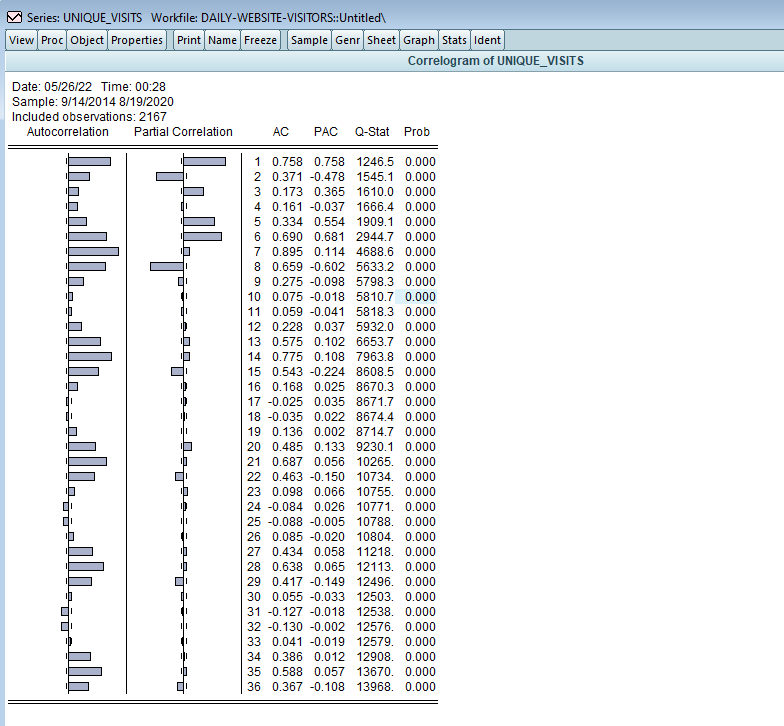
**Not:** Analizlerimizin tümünde **alfa = 0.05** alınmıştır.

1. **Verimiz Zaman Serisi Mi?**

Verimizin zaman serisi olup olmadığına yani gözlemlerin ilişkili olup olmadığına korelasyon ile bakıyoruz.

**H0 : Gözlemler ilişkisizdir. Veri zaman serisi değildir.**

**H1 : Gözlemler ilişkilidir. Veri zaman serisidir.**



Korelogramımızı oluşturduk. EViews’un default önerisiyle gecikme değerimizi 36’da bıraktık.

Prob değerleri 0.05’ten küçük olduğu için gözlemler ilişkilidir, H0 hiptoezi reddedilir. Veri setimiz bir zaman serisidir / gözlemler ilişkilidir.

1. **Otokorelasyon Durumu**

**H0 : Otokorelasyon yoktur.**

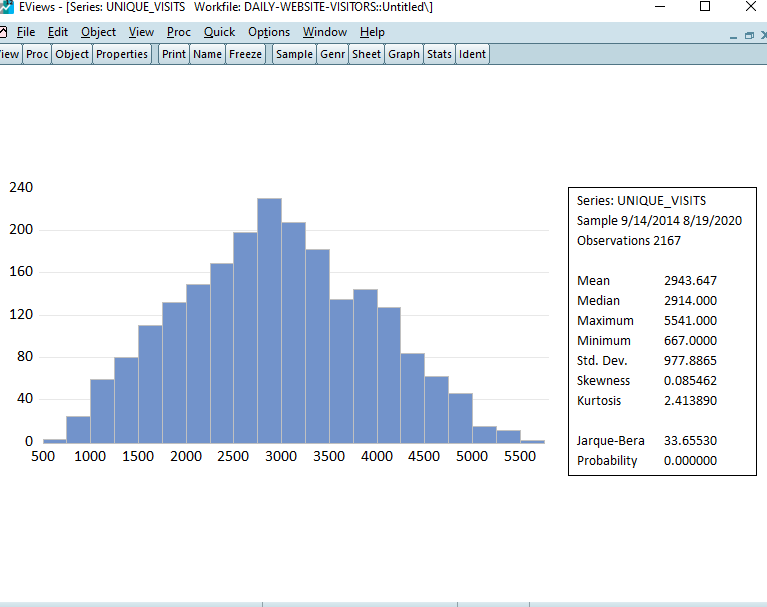
**H1 : Otokorelasyon vardır.**

PAC kısmında alfa değeri 0.05’ten büyük olan değerler olduğu için H0 reddedilir. Otokorelasyon bu seri için vardır.

1. **Veri Normal Dağılıyor Mu?**

**H0: Veri normal dağılmaktadır.**

**H1: Veri normal dağılmamaktadır.**



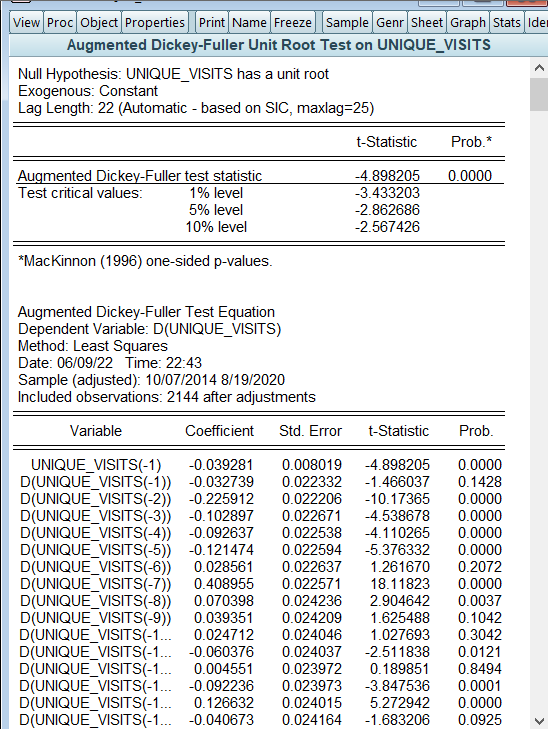
Jarque Bera Testi ile karar verilir. Jarque Bera probability değeri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilir. Verimiz normal dağılmamaktadır.

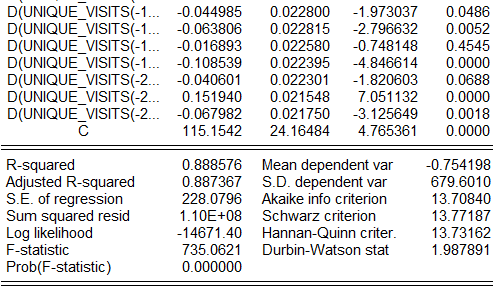
Çarpıklık ve basıklık durumunu inceleyecek olursak;  
**Skewness değeri** (0.085462) 0’dan büyük olduğu için sağa çarpıktır. **Kurtosis değeri** (2.413) için 3’ten küçük olduğu için basıklık vardır; fakat grafikten de görüleceği üzere verimiz normale çok yakın bir dağılıma sahiptir. Bu nedenle dönüşüm yapmadan devam etmeyi tercih ediyoruz.

**E) Durağanlık Testi**

**H0 : Seride birim kök içerir ve seri durağan değildir. Stokastik trend vardır.**

**H1 : Seride birim kök içermez ve seri durağandır. Stokastik trend yoktur.**





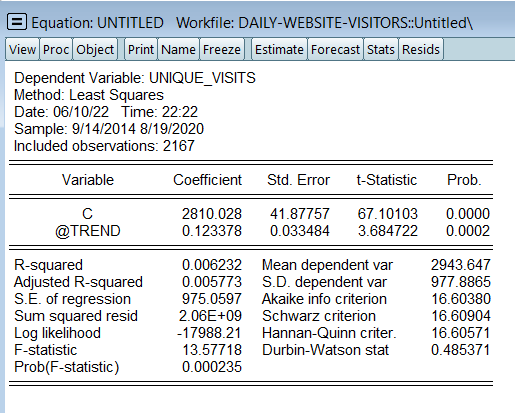
Augmented Dickey-Fuller testindeki prob değeri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilir. Seride birim kök yoktur ve seri durağandır. Stokastik trend yoktur.

Bu noktada, verimiz zaten durağan olduğu için durağanlaştırma için fark alma işlemlerini yapmayacağız.

**F) Trend Var Mı?**

**ls (unique\_visits) c @trend**

ls: least squares



**Yorum: @trend katsayısının prob değeri alfadan (0,05) küçük olduğu için H0 hipotezi reddedilir: Evet, deterministik trend vardır.**

**B1 katsayısı için hipotez:**

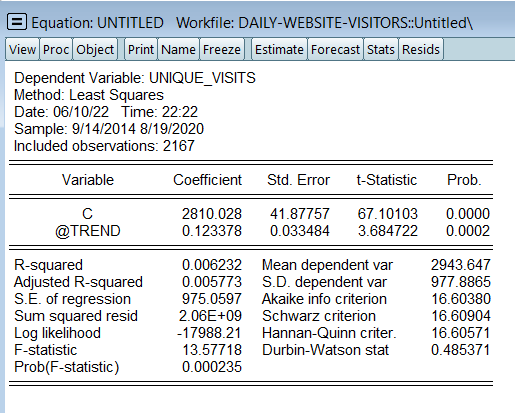
H0: B1 katsayısı anlamsızdır. Deterministik trend yoktur.

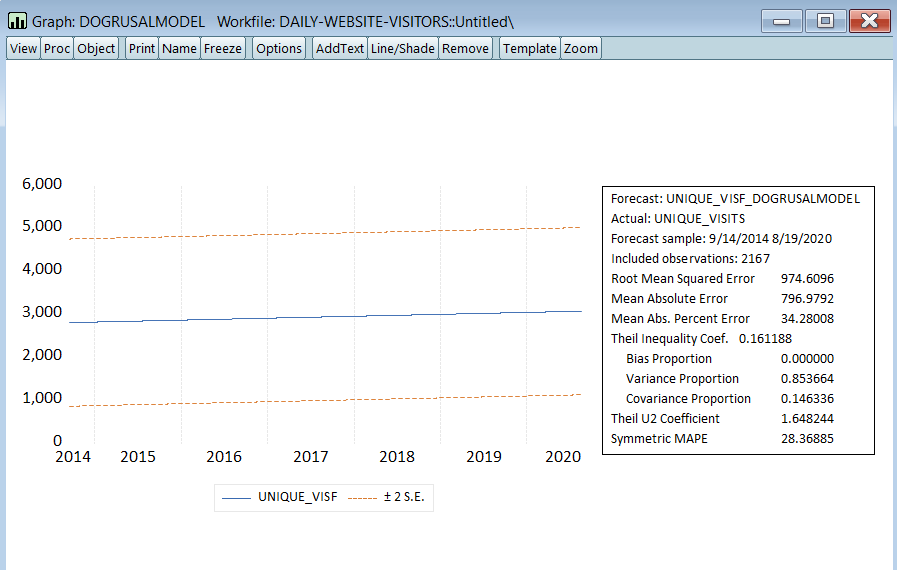
H1: B1 katsayısı anlamlıdır. Deterministik trend vardır.

**G) Deterministik Trendin Modellenmesi**

1. **Doğrusal Trend Modeli**

**ls (unique\_visits) c @trend**





**Kurulan Model:**

Yt = 2810.028 + 0.123t

**B0 katsayısı için hipotez:**

H0: B0 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B0 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

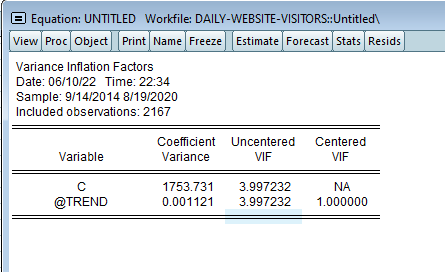
**B1 katsayısı için hipotez:**

H0: B1 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B1 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar : Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

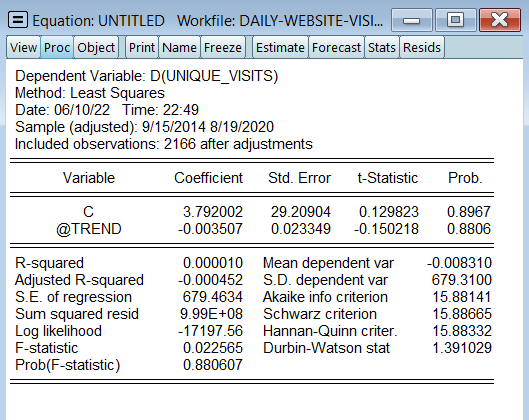
**Sonuç : B1 katsayısını baz alıyoruz. B1 katsayısı, yani trend katsayısı anlamlıdır, bu model uygun model olarak kullanılabilir.**

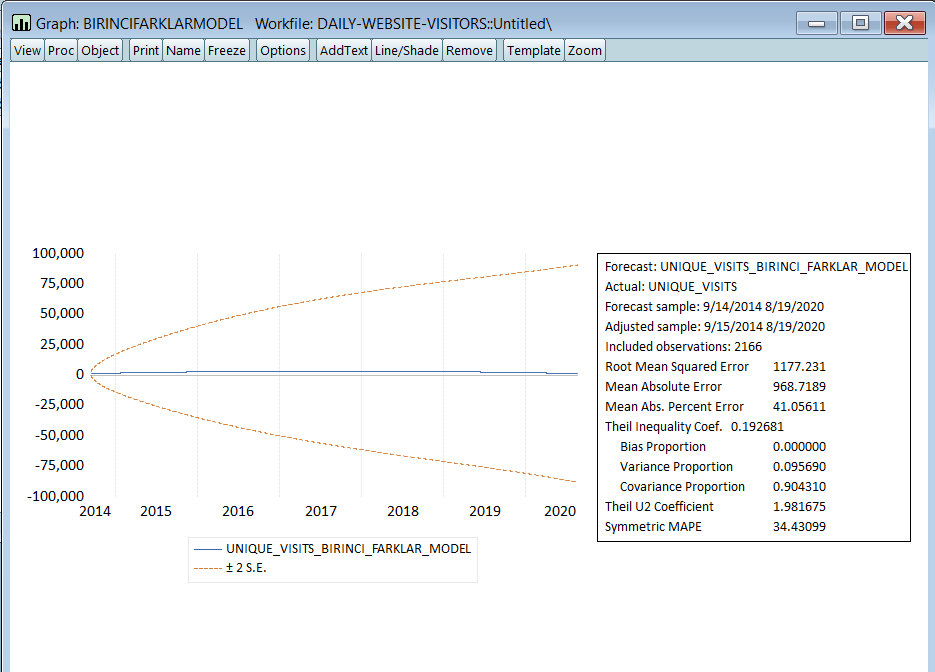


VIF değeri 5 ya da 10’dan büyükse çoklu bağlantı göstergesidir. Çoklu bağlantı varsa ilgili katsayılar bu nedenle zayıf bir şekilde kestiriliyor demektir. Verimizin değerlerini incelediğimiz zaman böyle bir problemle karşılaşılmadığını görüyoruz.

1. **Birinci Farklar Trend Modeli**

**ls d(unique\_visits) c @trend**

****



**Kurulan Model:**

**Yt = 3.792 - 0.0035t**

**B0 katsayısı için hipotez:**

H0: B0 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B0 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri > 0,05 olduğu için H0 reddedilemez. Deterministik trend yoktur.**

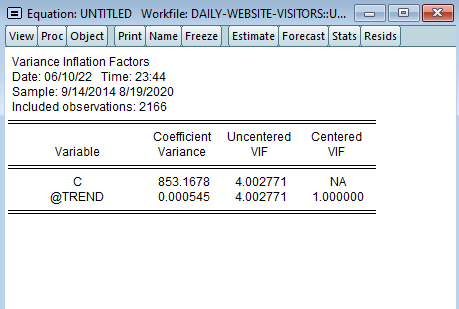
**B1 katsayısı için hipotez:**

H0: B1 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B1 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar : Prob değeri > 0,05 olduğu için H0 reddedilemez. Deterministik trend yoktur.**

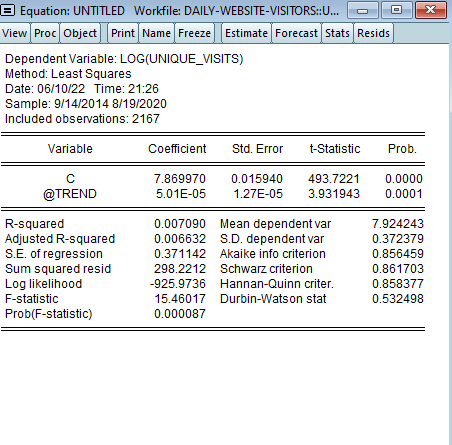
**Sonuç : B1 katsayısını baz alıyoruz. B1 katsayısı, yani trend katsayısı anlamlı değildir., bu model uygun model olarak kullanılamaz.**

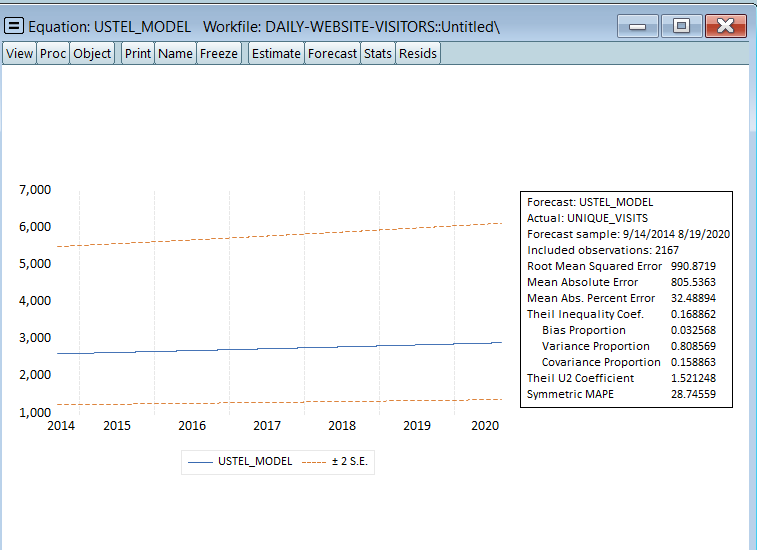


VIF değerlerini incelediğimizde 5’i geçmediği için modelde çoklu bağlantı probleminden söz edemeyiz.

1. **Üstel Trend Modeli**

**ls log(unique\_visits) c @trend**

****



**Kurulan Model:**

**ln(Yt) = 7.87+ 0.0000501t**

**B0 katsayısı için hipotez:**

H0: B0 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B0 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

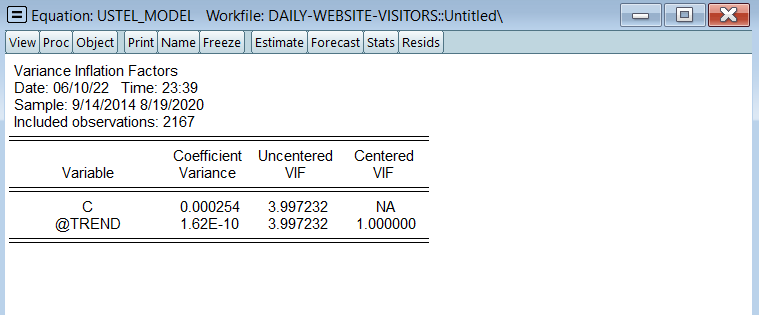
**B1 katsayısı için hipotez:**

H0: B1 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B1 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar : Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

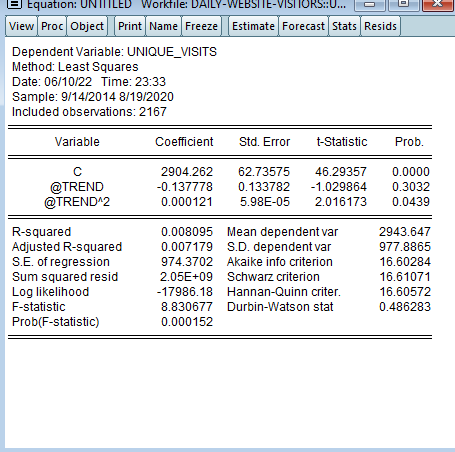
**Sonuç : B1 katsayısını baz alıyoruz. B1 katsayısı, yani trend katsayısı anlamlıdır, bu model uygun model olarak kullanılabilir.**

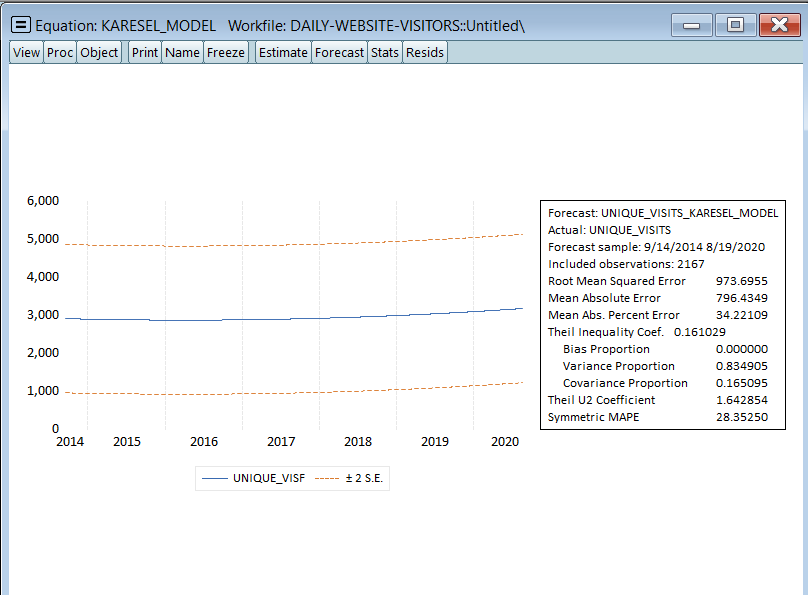


VIF değeri 5 ya da 10’dan büyükse çoklu bağlantı göstergesidir.

1. **Karesel Trend Modeli**

**ls (unique\_visits) c @trend @trend^2**

****

****

**Kurulan Model:**

**Yt = 2904.262 −0.137778t + 0.000121t^2**

**B0 katsayısı için hipotez:**

H0: B0 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B0 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

**B1 katsayısı için hipotez:**

H0: B1 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B1 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar : Prob değeri > 0,05 olduğu için H0 reddedilemez. Deterministik trend yoktur..**

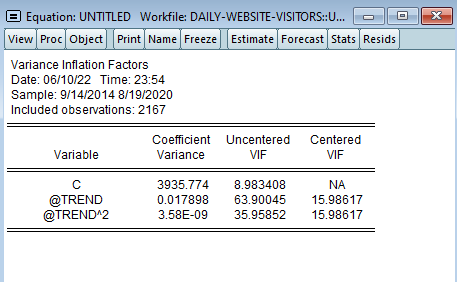
**B2 katsayısı için hipotez:**

H0: B2 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B2 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

**Sonuç : B2 katsayısını baz alıyoruz. B2 katsayısı, yani trend^2 katsayısı anlamlıdır, bu model uygun model olarak kullanılabilir.**

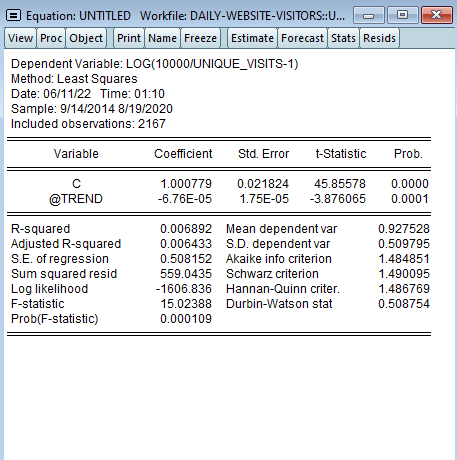
****

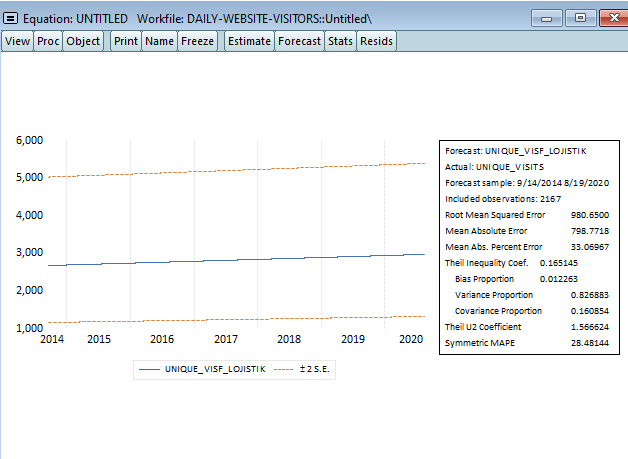
VIF değerleri 5 ve 10’dan yüksek çıktığı için çoklu bağlantı probleminden söz edilebilir. Bu model için çoklu bağlantı problemi vardır.

1. **Lojistik Trend Modeli**

Serimizdeki en büyük değer 5541 olduğu için, bu değerden büyük bir değer olan 10000’i seçiyoruz.

**ls log(10000/unique\_visits -1) c @trend**

****

****

**Kurulan Model:**

**Yt = 1.000779 − 0.0000676t**

**B0 katsayısı için hipotez:**

H0: B0 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B0 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

**B1 katsayısı için hipotez:**

H0: B1 katsayısı anlamsızdır. Deterministik trend yoktur.

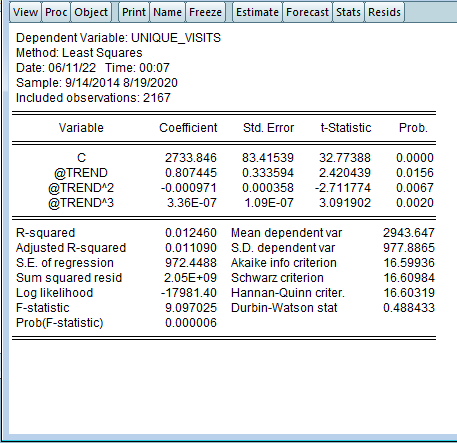
H1: B1 katsayısı anlamlıdır. Deterministik trend vardır.

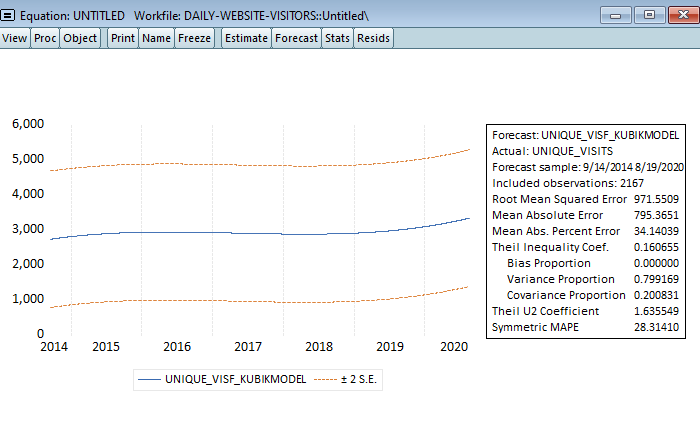
**Karar : Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

**Sonuç : B0 ve B1 katsayıları anlamlıdır. B1 katsayısını baz alıyoruz. B1 katsayısı, yani trend katsayısı anlamlıdır, bu model uygun model olarak kullanılabilir.**

1. **Kübik Trend Modeli**

**ls unique\_visits c @trend @trend^2 @trend^3**

****

****

**Kurulan Model:**

**Yt = 2733.846 + 0.807445t - 0.000971t^2 + 0.000000336t^3**

**B0 katsayısı için hipotez:**

H0: B0 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B0 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

**B1 katsayısı için hipotez:**

H0: B1 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B1 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar : Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

**B2 katsayısı için hipotez:**

H0: B2 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B2 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

**B3 katsayısı için hipotez:**

H0: B3 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B3 katsayısı anlamlıdır. Deterministik trend vardır.

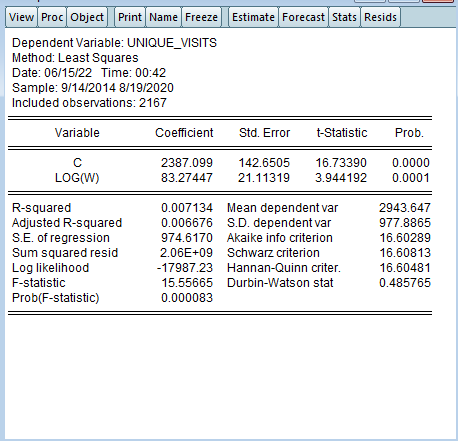
**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

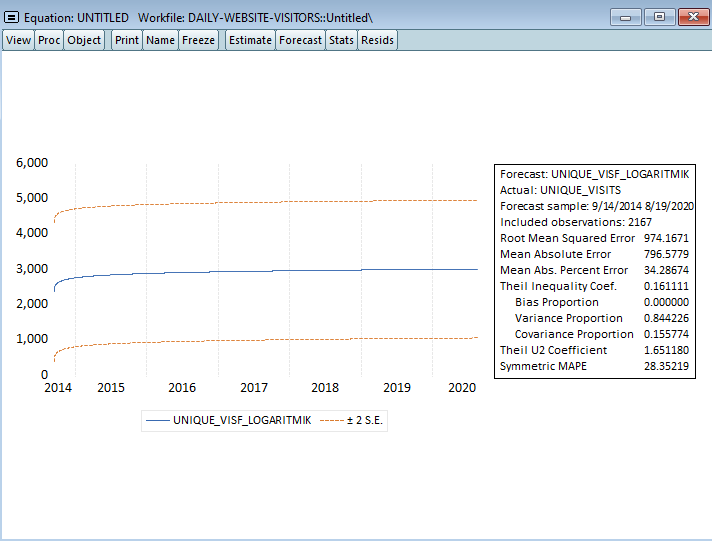
**Sonuç : B0, B1, B2 ve B3 katsayıları anlamlı olduğu için bu model geçerli olarak kullanılabilir.**

1. **Logaritmik Trend Modeli**

**series a=@trend+1**

**ls unique\_visits c log(a)**

****

****

**Kurulan Model:**

**Yt = 2387.099 + 83.27447ln(t)**

**B0 katsayısı için hipotez:**

H0: B0 katsayısı anlamsızdır. Deterministik trend yoktur.

H1: B0 katsayısı anlamlıdır. Deterministik trend vardır.

**Karar: Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

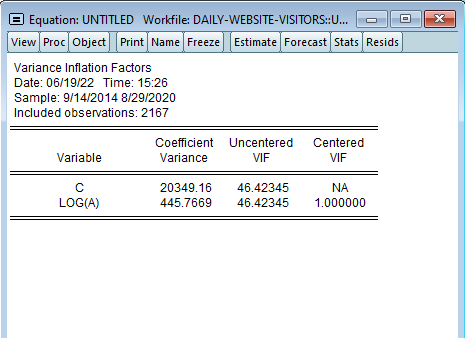
**B1 katsayısı için hipotez:**

H0: B1 katsayısı anlamsızdır. Deterministik trend yoktur.

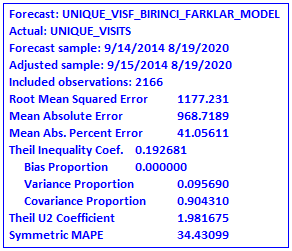
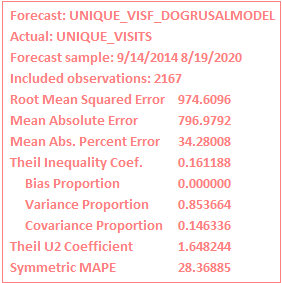
H1: B1 katsayısı anlamlıdır. Deterministik trend vardır.

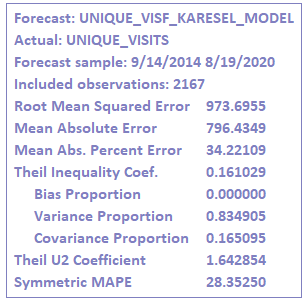
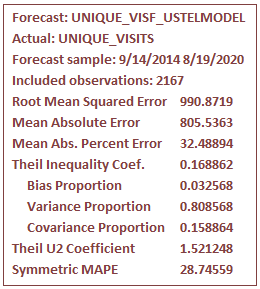
**Karar : Prob değeri < 0,05 olduğu için H0 reddedilir. Deterministik trend vardır.**

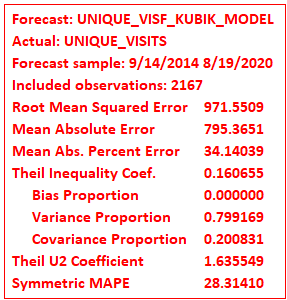
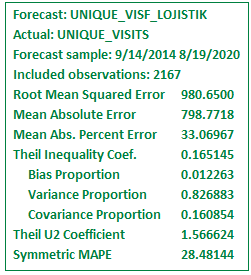
**Sonuç : B0 ve B1 katsayıları anlamlıdır. B1 katsayısını baz alıyoruz. B1 katsayısı, yani trend katsayısı anlamlıdır, bu model uygun model olarak kullanılabilir.**

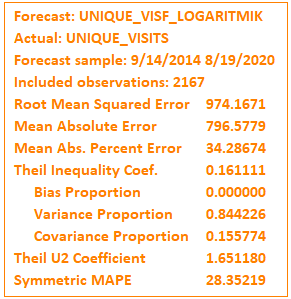
****

**Modellerin Karşılaştırılması**







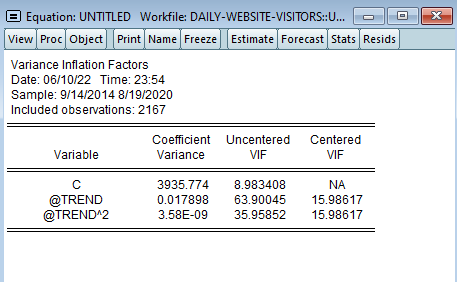


**SONUÇ:**

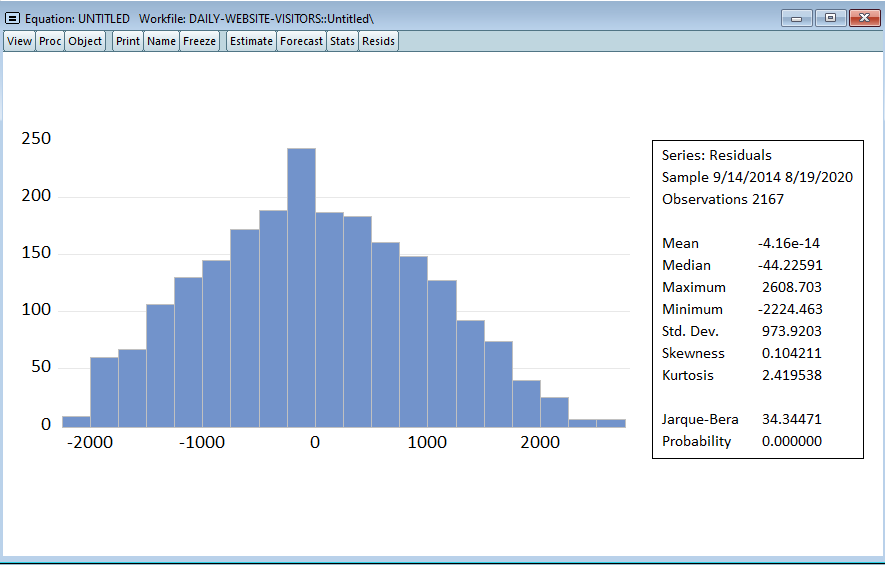
R^2 olarak en büyük olan ve RMSE değeri olarak en küçük model olan **Karesel Model**’in seçilmesine karar verilmiştir.

**Seçilen Karesel Trend Modelinde:**

Yukarıda da bahsettiğimiz gibi seçtiğimiz karesel modelde çoklu bağlantı problemi vardır.

****

**Hatalar Normal Dağılıyor Mu?**



Prob. değeri 0.05'den küçük olduğundan:

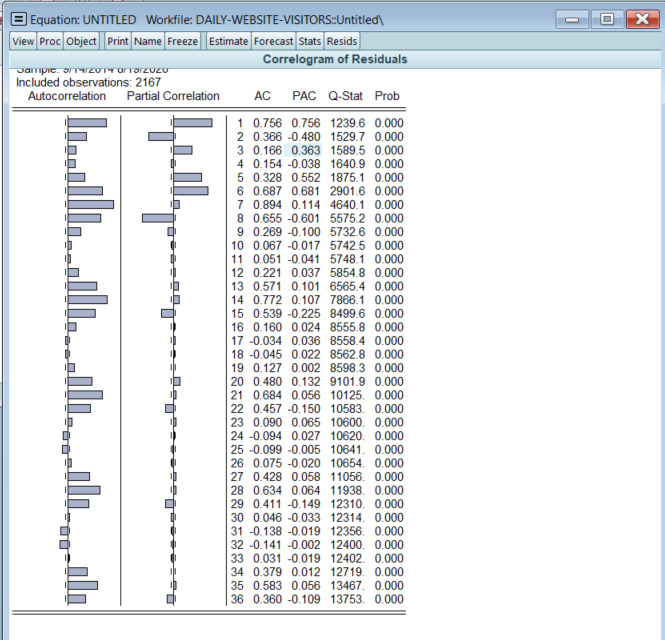
**H0: Artıklar normal dağılıyor.**

**H1: Artıklar normal dağılmıyor.**

**Karar:** Prob değeri 0.05'den küçüktür. Yokluk hipotezi reddedilmiştir. Artıklar normal dağılmıyor.

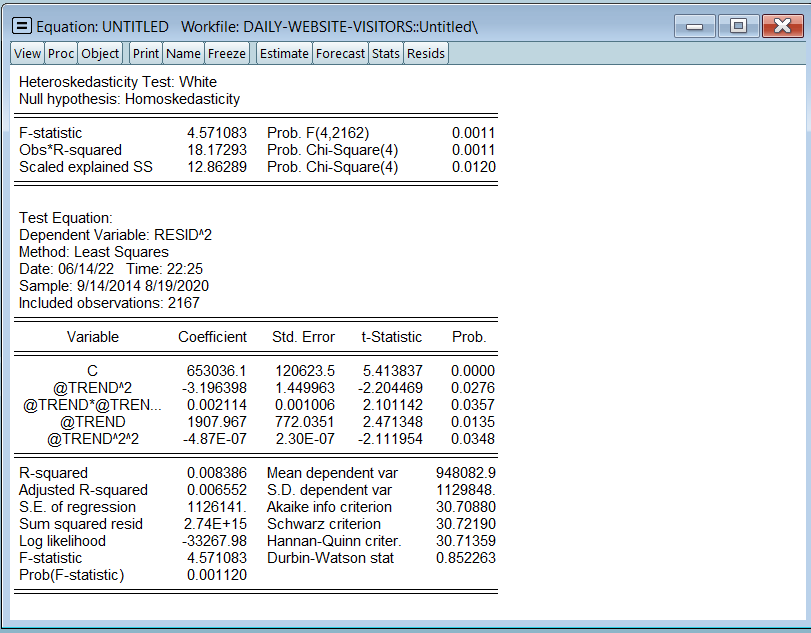
Histogramdan da görülebildiği gibi artıklar hafif sağa çarpıktır.

**Otokorelasyon problemi var mı ?**



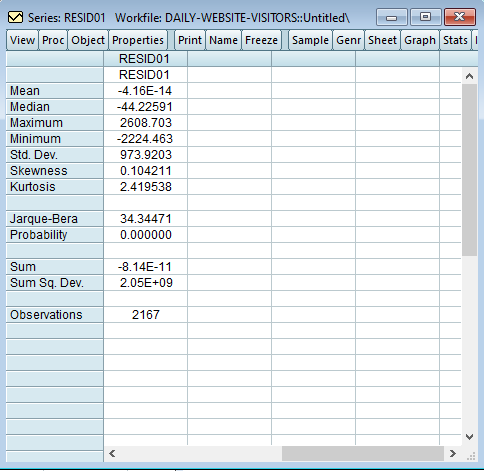
Partial Correlation bölümünde güven aralığını aşan değerler vardır. Seride otokorelasyon olduğunu söyleyebiliriz.

**Değişen varyans problemi var mı?**



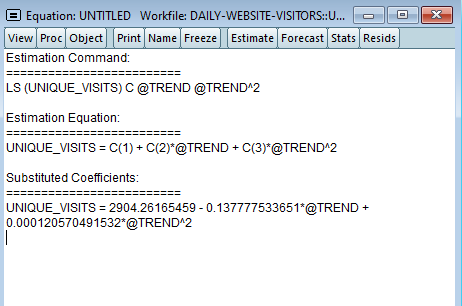
Probability değeri 0.05’ten küçük olduğundan değişen varyans problemi vardır.

**0 Ortalama varyans problemi var mı?**



Verinin ortalaması 0’a yakın olduğundan dolayı bu varsayım sağlanmaktadır.

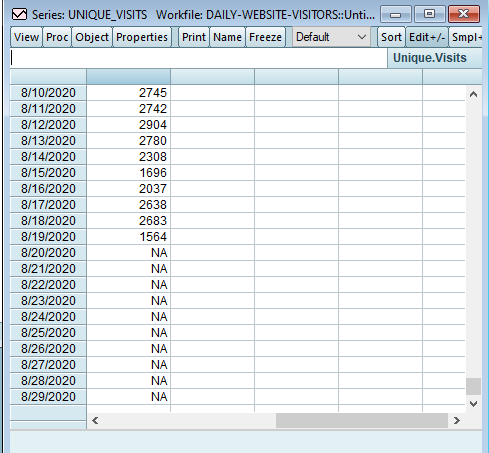
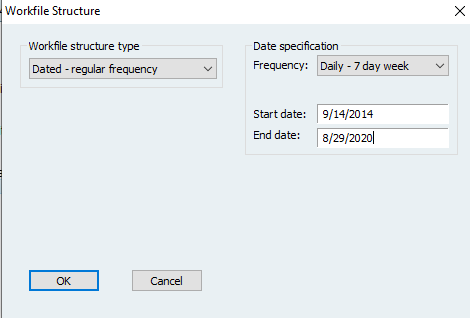
Aşağıda karesel trend modelinin denklemini ve temsil ettiği katsayı değerlerini görebilirsiniz.



**Üstel Düzleştirme**

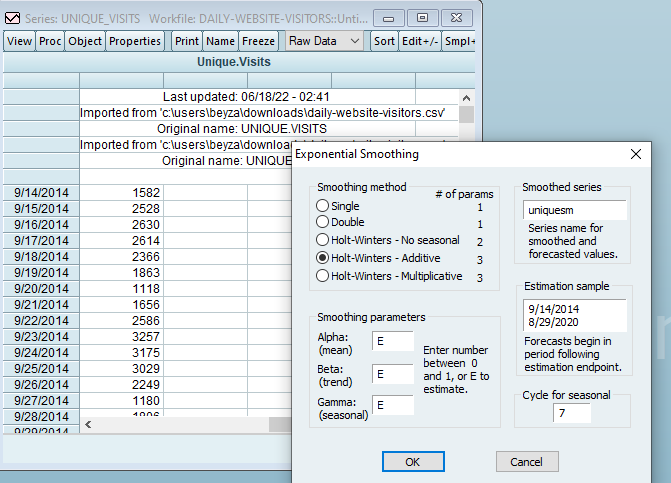
Üstel düzleştirme yöntemi ile veri setimizdeki son değişim ve sıçramaları baz alarak öngörüleri devamlı güncelleyebiliriz.

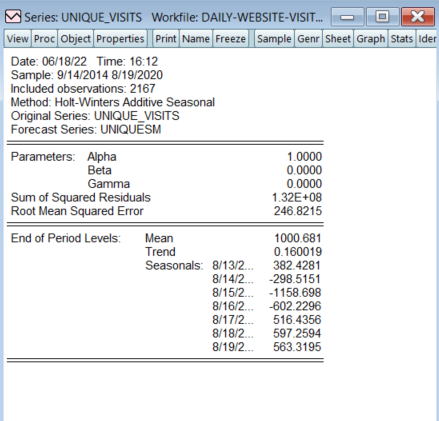
Öncelikle verimizi tahmin etmek istediğimiz tarihlere göre ayarlayarak genişletiyoruz. Biz on günü tahmin etmek istedik bu yüzden end date tarihini 08/19/2020 iken 08/29/2020 olarak ayarlıyoruz. Böylece gözlem sayımızı 2167’den 2177’ye çıkarmış olduk.

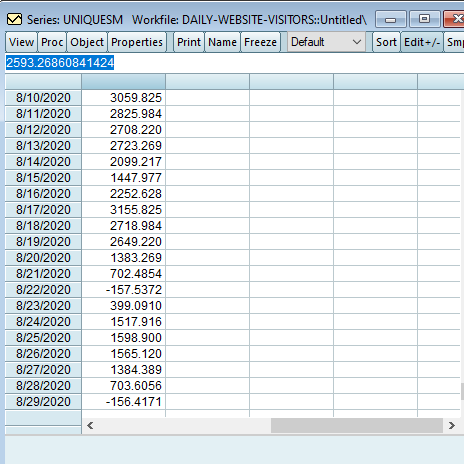
****

Bizim verimizin yapısını göz önüne aldığımız zaman Holt-Winters-Additive ve Holt-Winters\_Multiplicative yöntemlerini deneyip aralarında bir karara varacağız.

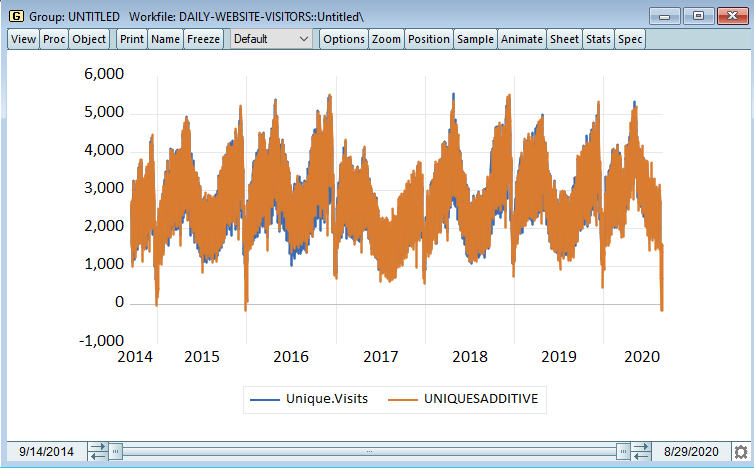
**HOLT-WINTERS-ADDITIVE:**

****

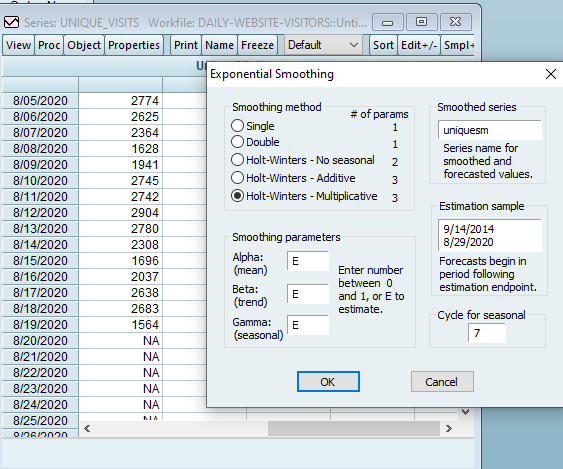
****

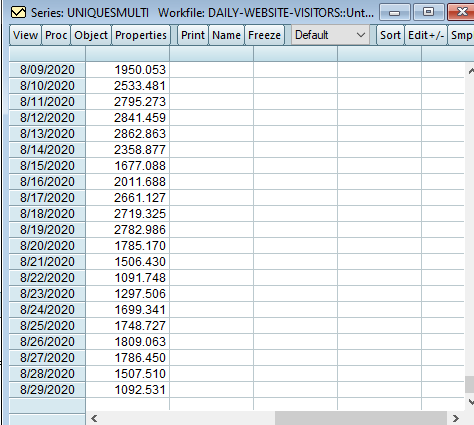
****

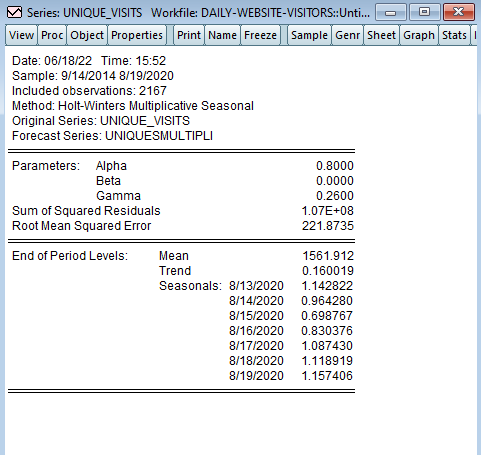
Aşağıdaki grafikte bu yöntemin orijinal veriyle oldukça uyumlu olduğunu gözlemliyoruz.

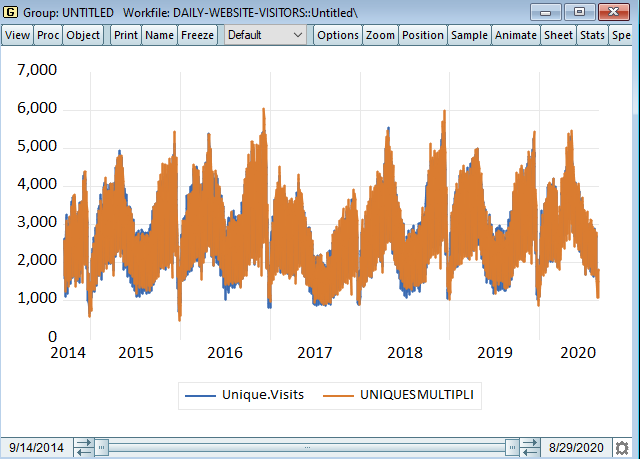
****

**HOLT-WINTERS-Multiplicative:**

****

****

****

****

Yine aynı şekilde bu grafikte HOLT-WINTERS-Multiplicative yönteminin de iyi uyum sağladığını görüyoruz.

**Karar :** ModellerinRMSE değerlerini karşılaştırıyoruz. HOLT-WINTERS-Multiplicative modelinde hata oranı(RMSE) daha küçük olduğu için HOLT-WINTERS-Multiplicative modelini seçiyoruz.

**AYRIŞTIRMA İŞLEMLERİ**

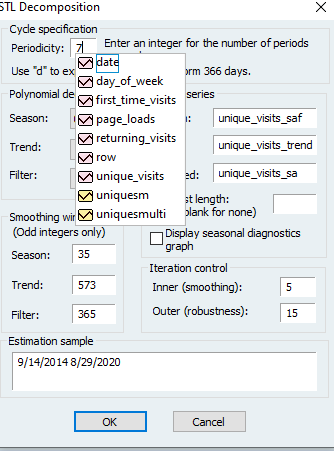
Daha önceden yaptığımız uygulamada 10 günü tahmin etmek için gözlemlerimize NA değerlerini eklemiştik. Şimdi ise verimizi mevsimsellikten ve trendden arındırmak için STL algoritmasıyla ayrıştırma işlemini yapacağız.

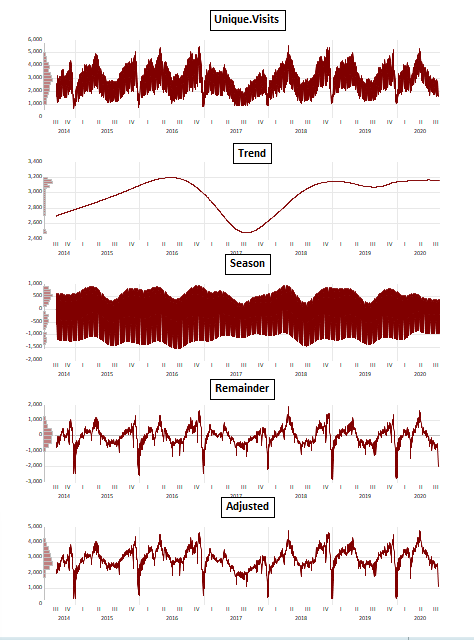
“LOESS (STL) algoritması kullanılarak yapılan Mevsimsellik-Trend ayrıştırması, ekonomik ve çevresel analizlerde sıklıkla kullanılan sağlam bir zaman serisi ayrıştırma yöntemidir.

STL yöntemi, bir zaman serisini trend, mevsimsel ve kalan bileşenlere ayrıştırmak için yerel olarak yerleştirilmiş regresyon modellerini kullanır.” **(1)**

STL'yi herhangi bir veri kümesine uygulayabilirsiniz, ancak anlamlı sonuçlar yalnızca verilerde yinelenen bir zamansal model varsa döndürülür.

STL algoritması, iki döngüde LOESS kullanarak zaman serilerinde yumuşatma gerçekleştirir.

****

****

Yukarıdaki grafiklerde zaman serimizin ayrıştırılmış bileşenlerini görebilirsiniz. Trend ve Mevsimselliğe ek olarak, “Remainder” grafiği verilerde mevcut olan gürültü miktarını gösterir.

Remainder’da sıfıra yakın değerler, mevsimsel ve trend bileşenlerinin zaman serilerini tanımlamada doğru olduğunu gösterirken, daha büyük kalan değerler gürültünün varlığını gösterir.

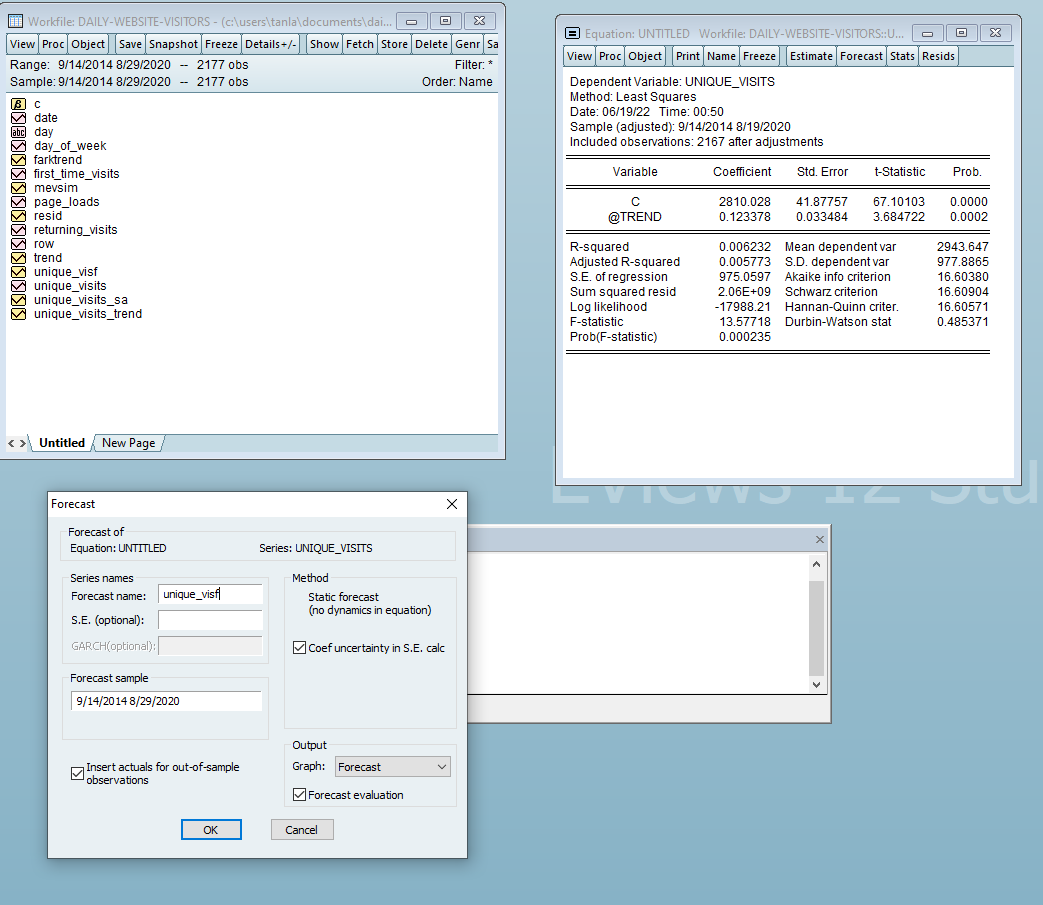
Remainder bileşenini, verilerdeki aykırı değerleri belirlemek için de kullanabilirsiniz.

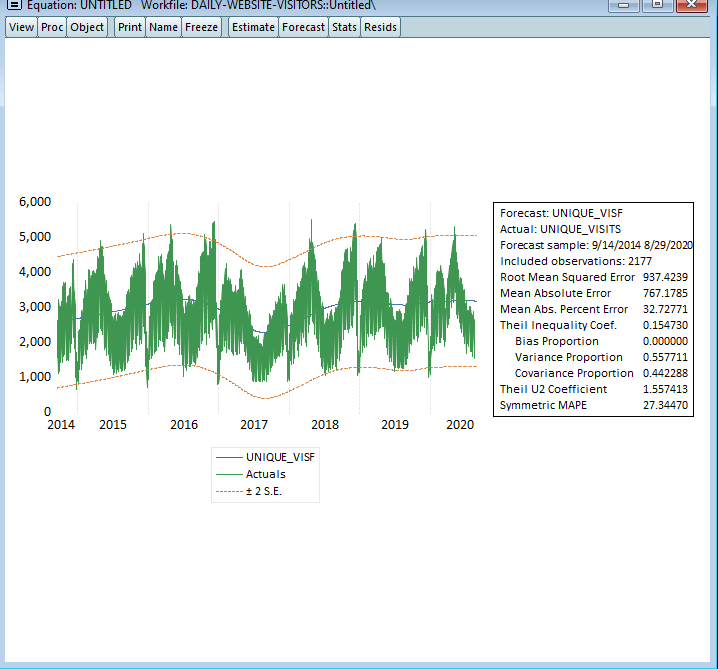
STL algoritması, zaman serimizi bileşenlerine ayırarak her bir bileşeni yeni değişkenler olarak environment’ımıza atadı. Şimdi manuel olarak seriyi kendimiz ayrıştıracağız.

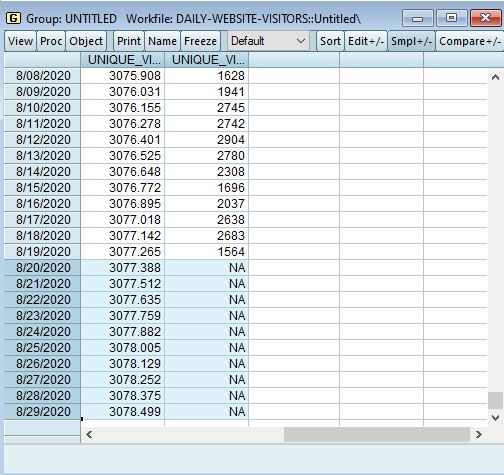
ÇARPIMSAL MODEL: **Tahmin = Trend . Mevsim . Hata** olduğunu unutmayalım.

**Serinin Trendden Ayrıştırılarak Forecastlenmesi**

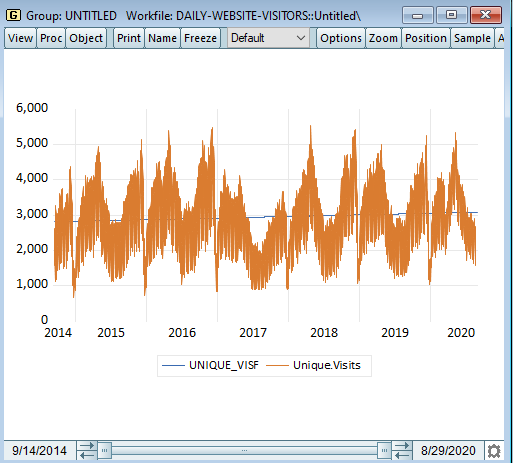
Aşağıda “unique\_visf” değişkeni oluşturarak sonraki 10 günü tahminledik:

****

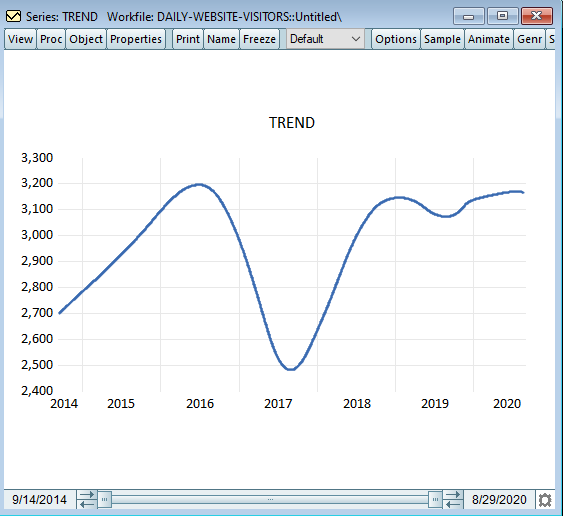
****

****

Trendin çizgi grafiğine baktığımızda hafif bir pozitif trend olduğunu görebiliriz:

****

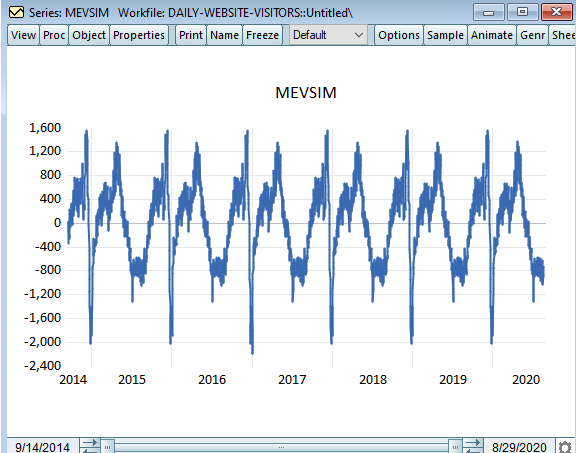
Sonrasında orijinal veriden trendi çıkararak saf trend verisini “trend” (serinin trendden ayrıştırılmış hali) isimli değişkende depoladık.



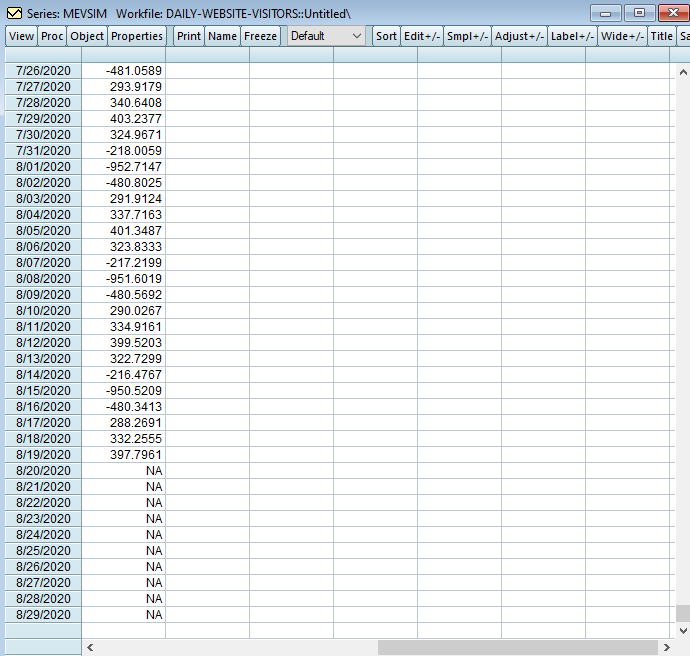
Not: Trend grafiğinin üstte yaptığımız STL ayrıştırma yönteminin sonucunda oluşan unique\_visits\_trend’le aynı olduğunu görebiliriz.

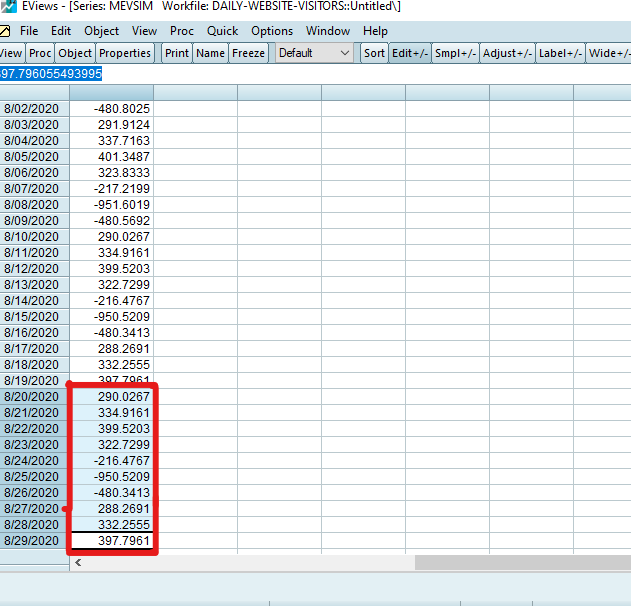
**Serinin Mevsimsellikten Ayrıştırılması**

Ayrıştırmayı yaparken, “Mevsim” adı altında orijinal veriden mevsimsellikten ayrıştırılmış halinin farkını depoladık:

****

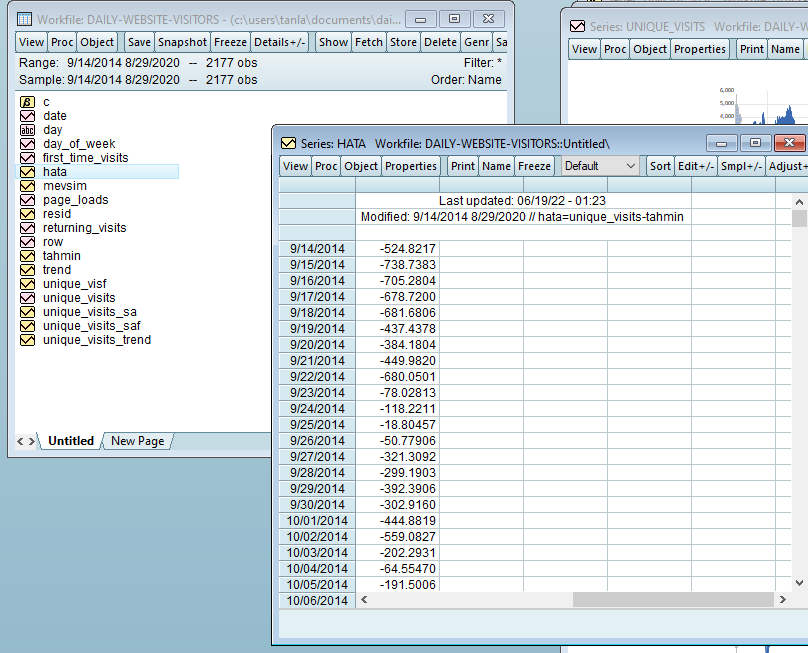
Oluşturduğumuz mevsim serisinde tahmin edilecek 10 günü (NA) bir önceki 10 günün değerleriyle dolduruyoruz:

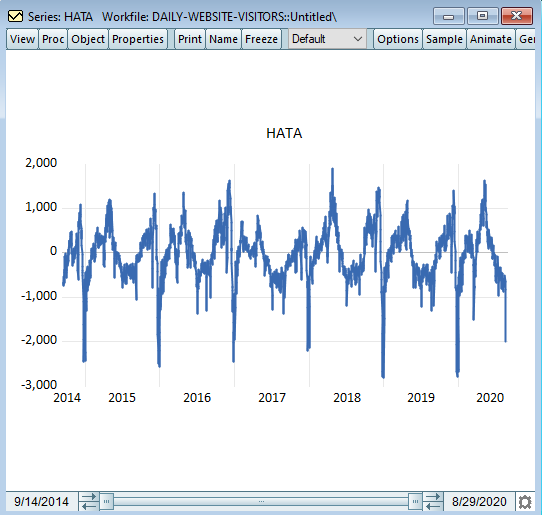
****

****

**Hata Serisini Elde Etme**

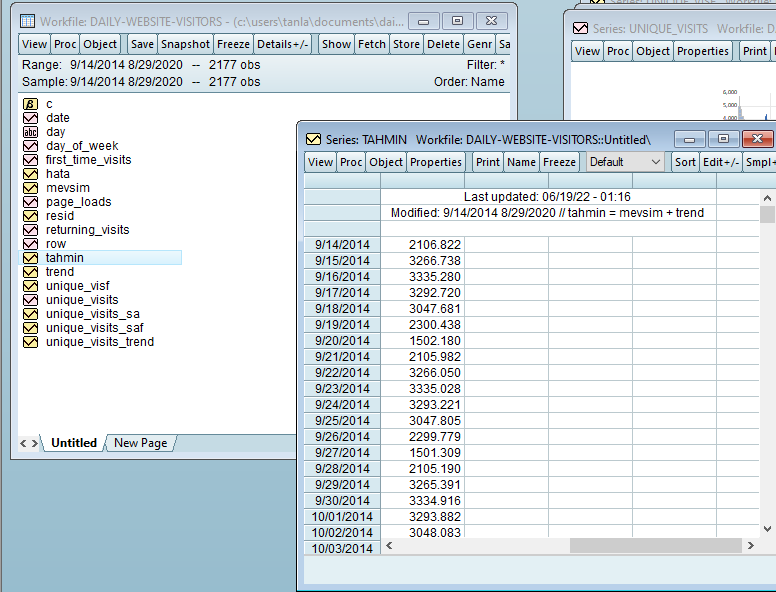
**series hata=unique\_visits-tahmin**

****

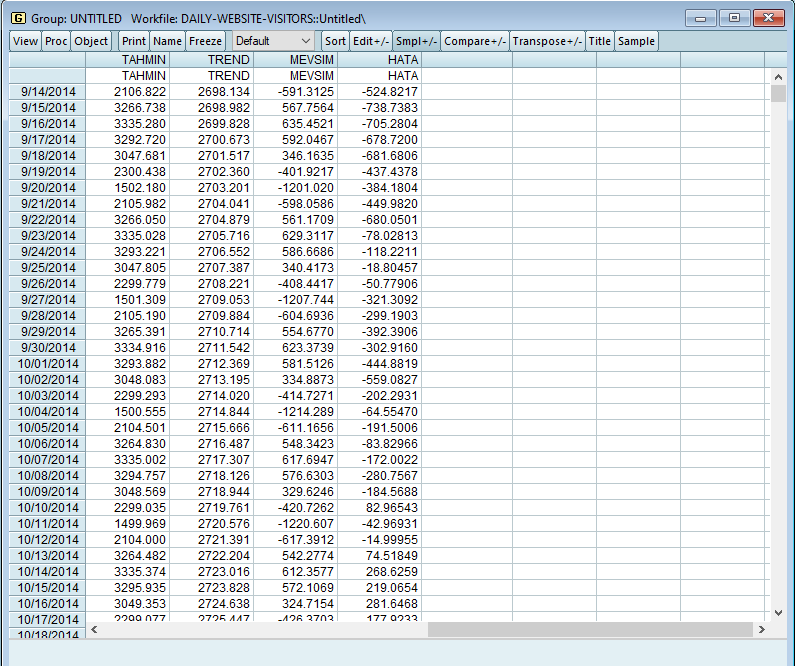
****

**Tahmin’i Elde Etme**

**series hata=unique\_visits-tahmin**

****

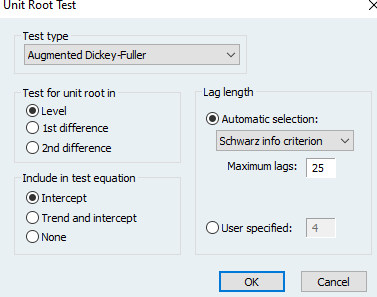
**Hepsinin Beraber Görünümü:**

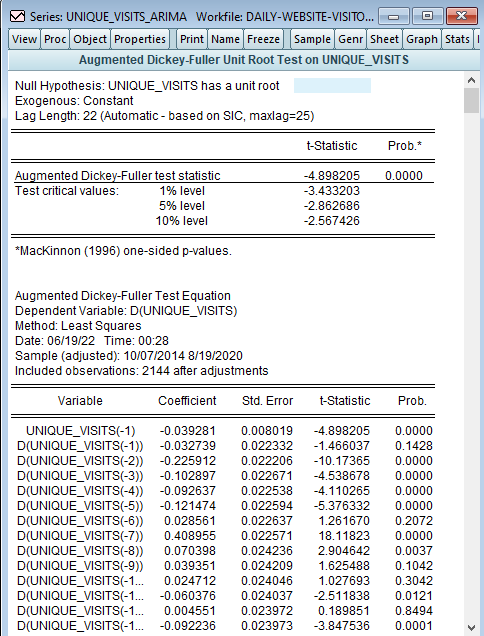
****

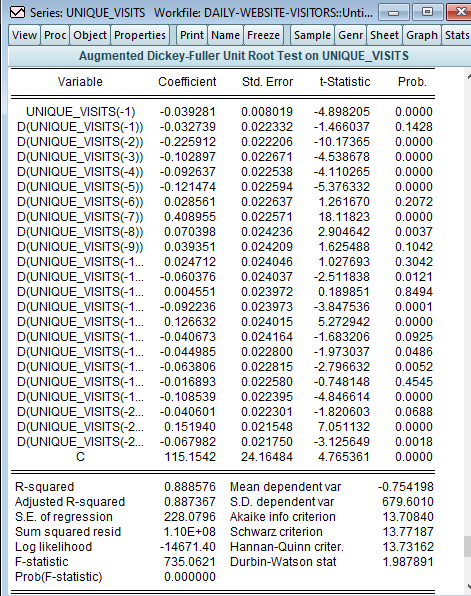
**BOX-JENKINS METODOLOJİSİ**

**Mevsimsel Box-Jenkins Modelleri : (SAR) (SMA) (SARMA/SARIMA)**

Verimiz mevsimsellik içermektedir. Aynı zamanda günlük verilerden oluşmaktadır. SARIMA modelleri serimiz için uygun modellerdir.

****

****

****

Durbin-Watsons Test istatistiğine göre değerler 1.7 ile 2.3 arasında yer aldığından dolayı otokorelasyon yoktur diyebiliriz.

Prob değeri alfa değeri olan 0.05’ten küçüktür . Ayrıca serimiz durağandır. Buradan d = 0 diyebiliriz.

**SARIMA (p,d,q) (P,D,Q)s**

**P: mevsimsel otoregresyon (SAR) modelinin derecesi,**

**D: mevsimsel fark alma işlemi sayısı,**

**Q: mevsimsel hareketli ortalama (SMA) modelinin derecesi**

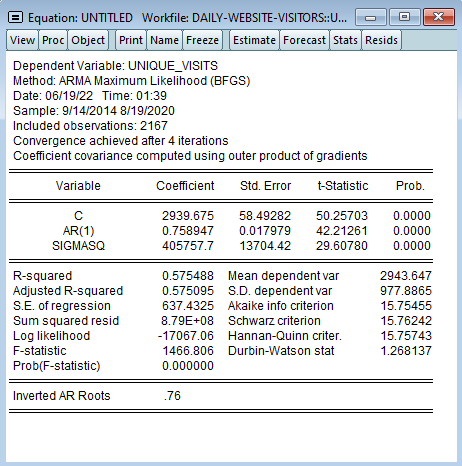
**s: periyot**

**Serimiz durağan olduğu için bu adımlarda fark alınmayacaktır(d=0).**

**Mevsimsel Otoregresyon Modelleri (SAR)**

**SAR(1) = SARIMA(1,0,0)(1,0,0)**

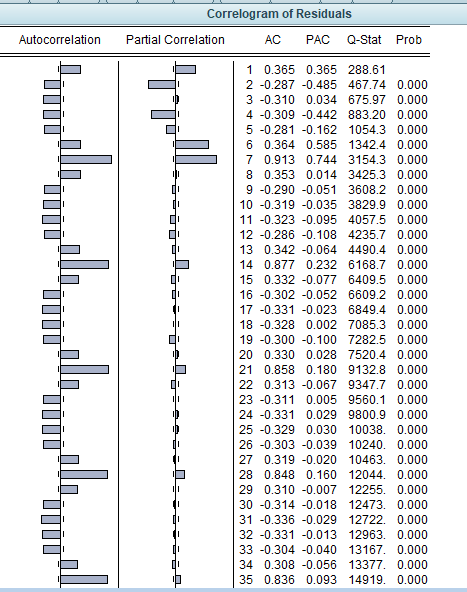
ls unique\_visits c sar(1)



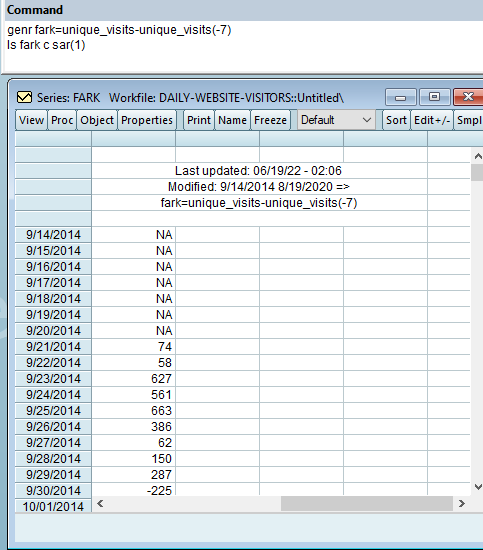
**H0 :Model anlamlıdır**

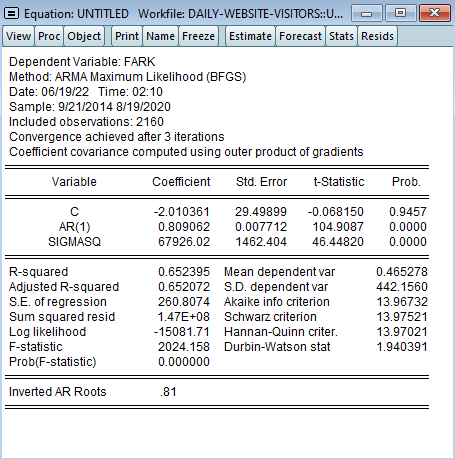
**H1 : Model anlamsızdır.**

Prob değerleri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilemez. Inverted AR roots değerinin 1 değerine yakın olması çok iyi değil ve bizim verimizde 1 değerine yakın. Kolegrama bakalım.



Partial Correlation kısmına baktığımda değerlerin güven aralıklarının çok değişmediğini görebiliyoruz. Bu yüzden bu model sorunumuzu çözmedi.Fark almamız gerekiyor.. Periyodumuz 7 olduğundan farkı 7 olarak aldık ve tekrardan kodumuzu yazıp partial correlationa bakalım.





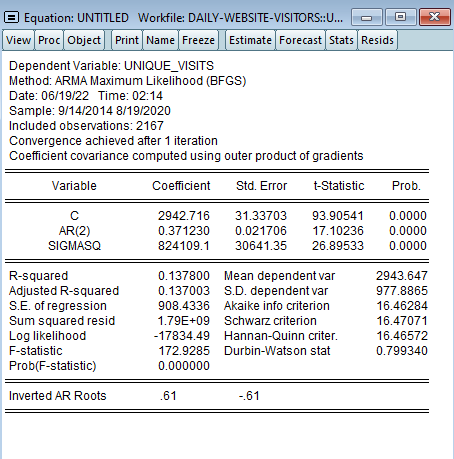
**H0 :Model anlamlıdır**

**H1 : Model anlamsızdır.**

Prob değeri alfa değeri olan 0.05’ten büyük olduğundan H0 reddedilir. Model anlamsızdır.

**SAR(2) = SARIMA(2,0,0)(2,1,0)**

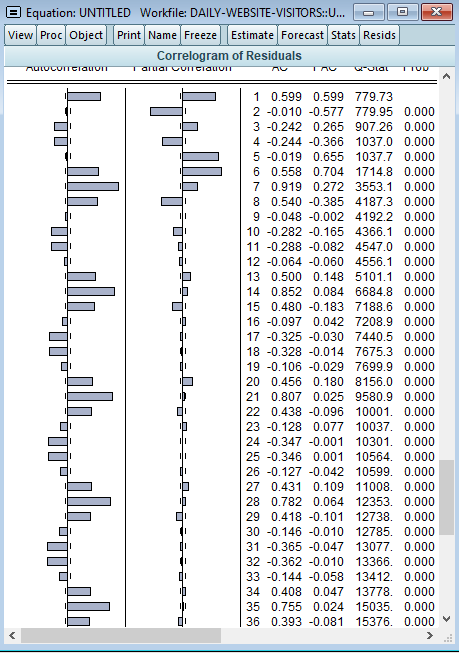
ls unique\_visits c sar(2)



**H0 :Model anlamlıdır**

**H1 : Model anlamsızdır.**

Prob değerleri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilemez. Model anlamlıdır.Inverted AR roots değerinin 1 değerine yakın olması çok iyi değil ancak bizim verimizde 1 değerine çok yakın değil. Kolegrama bakalım.

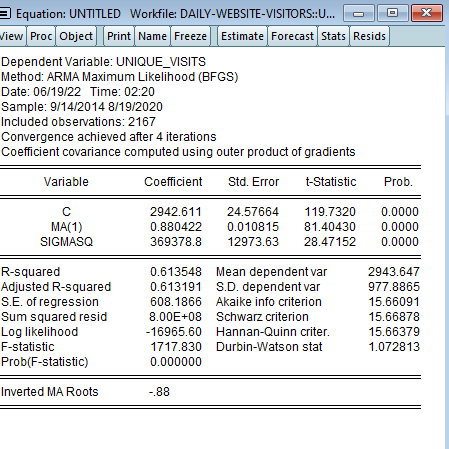


Güven sınırlarını aşan değerler hala var. Bu modelimiz de uygun bir model gibi gözükmüyor.

**Mevsimsel Hareketli Ortalama Modelleri (SMA)**

**SMA(1) = SARIMA(0,0,1)(0,1,1)**

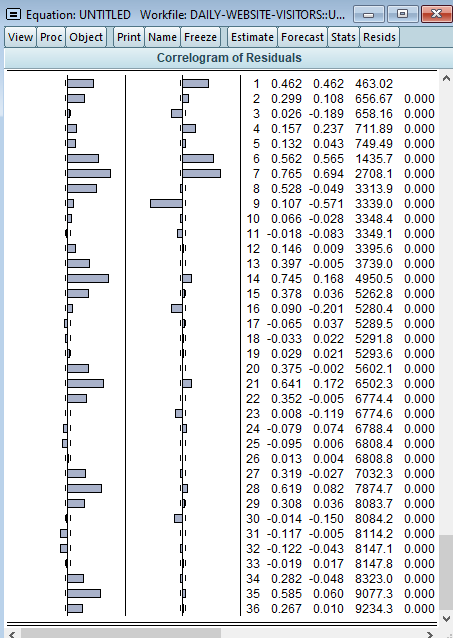
ls unique\_visits c sma(1)

****

**H0 :Model anlamlıdır**

**H1 : Model anlamsızdır.**

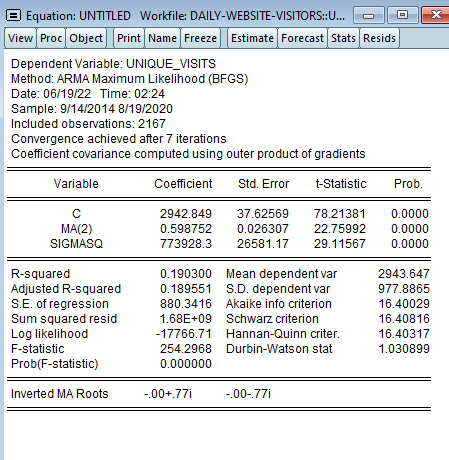
Prob değerleri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilemez.Model anlamlıdır. Kolegrama bakalım.



Güven aralığı sorunları hala çözülmedi. En sonunda en uygun modele karşılaştırma yaparak karar vereceğiz.

**SMA(2) = SARIMA(0,0,)(0,1,2)**

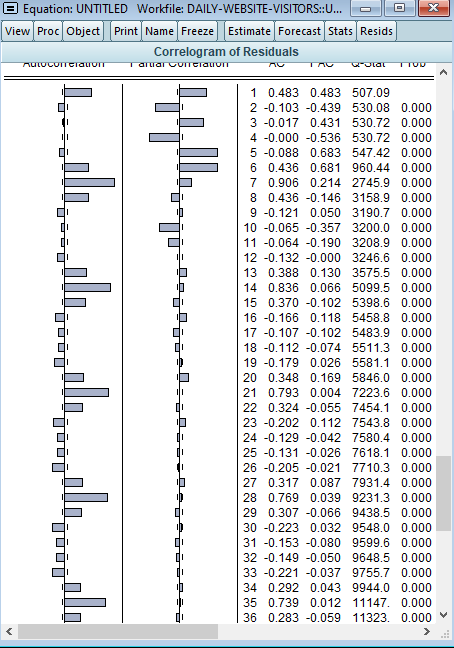
ls unique\_visits c sma(2)

****

**H0 :Model anlamlıdır**

**H1 : Model anlamsızdır.**

Prob değerleri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilemez.Model anlamlıdır. Kolegrama bakalım.

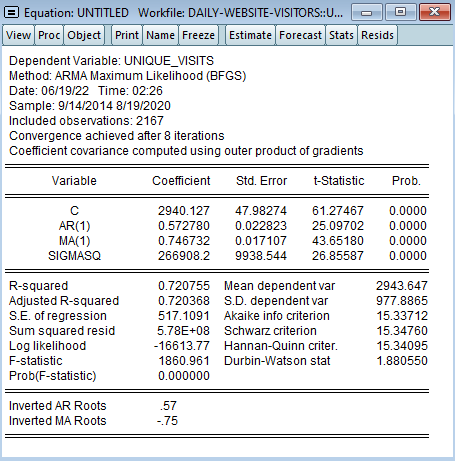
****

Bu model de uygun bir model olarak gözükmüyor, en sonunda inceleyeceğiz.

**Mevsimsel Otoregresif Hareketli Ortalama Modelleri (SARMA/SARIMA)**

**SARIMA(1,0,1)(1,1,1)= SAR(1) + SMA(1)**

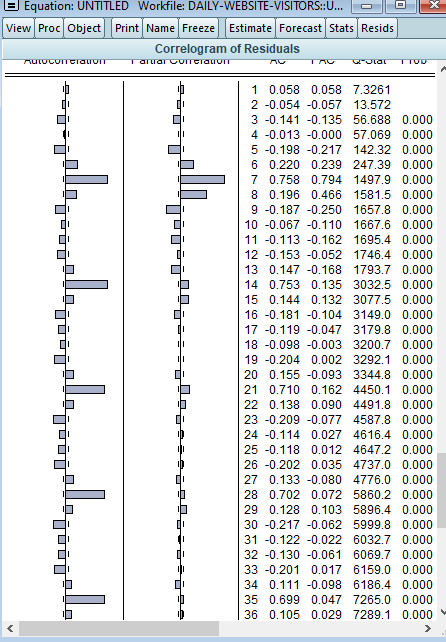
**ls unique\_visits c sar(1) sma(1)**

****

**H0 :Model anlamlıdır**

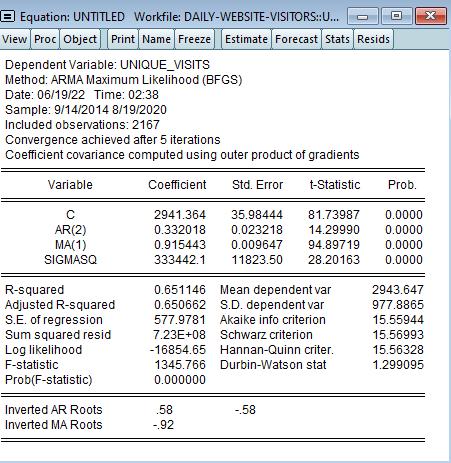
**H1 : Model anlamsızdır.**

Prob değerleri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilemez.Model anlamlıdır. Kolegrama bakalım.

****

Modeldeki birçok güven aralığı sorununun giderilmiş olduğunu görüyoruz fakat hala mevcut.

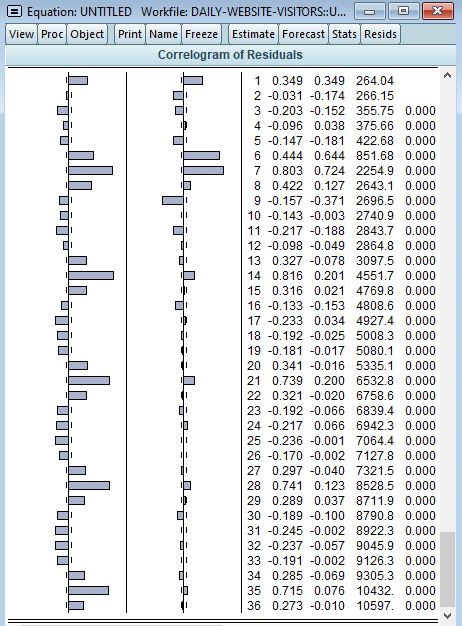
**SARIMA(2,0,1)(2,1,1)= SAR(2) + SMA(1)**



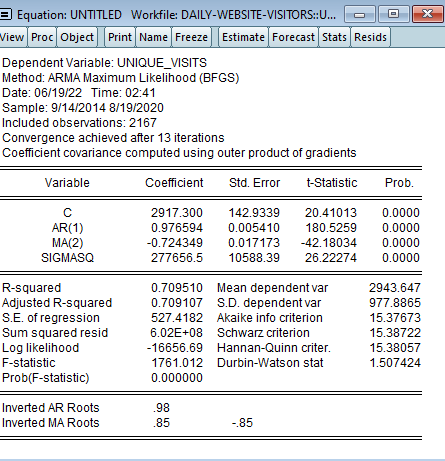
**H0 :Model anlamlıdır**

**H1 : Model anlamsızdır.**

Prob değerleri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilemez.Model anlamlıdır. Kolegrama bakalım.



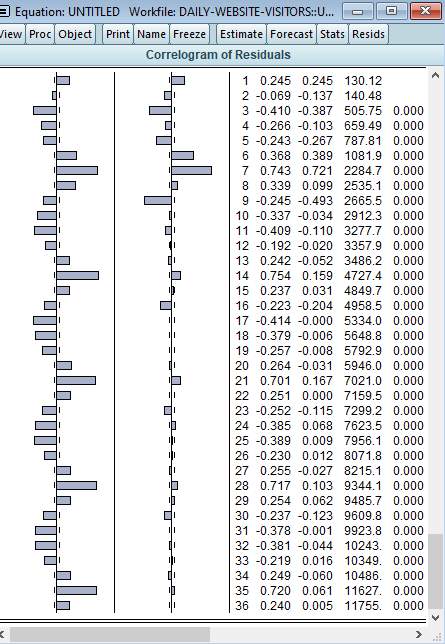
**SARIMA(1,0,2)(1,1,2)= SAR(1) + SMA(2)**

****

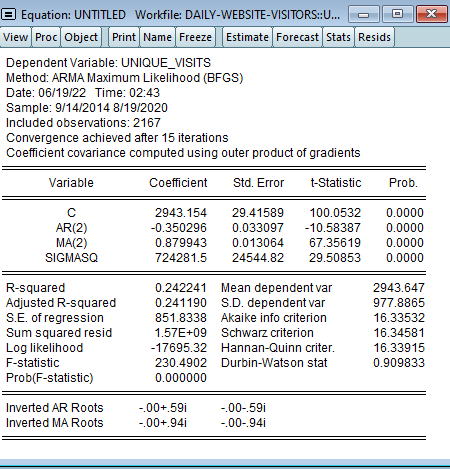
**H0 :Model anlamlıdır**

**H1 : Model anlamsızdır.**

Prob değerleri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilemez.Model anlamlıdır. Kolegrama bakalım.



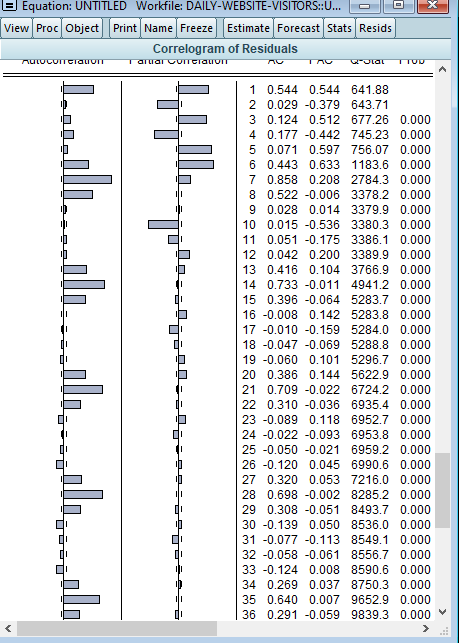
**SARIMA(2,0,2)(2,1,2)= SAR(2) + SMA(2)**

****

**H0 :Model anlamlıdır**

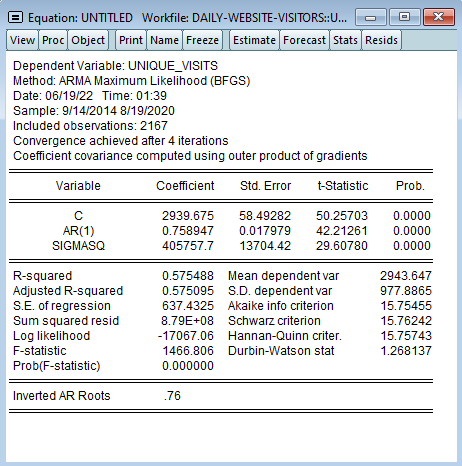
**H1 : Model anlamsızdır.**

Prob değerleri alfa değeri olan 0.05’ten küçük olduğundan H0 reddedilemez.Model anlamlıdır. Kolegrama bakalım.

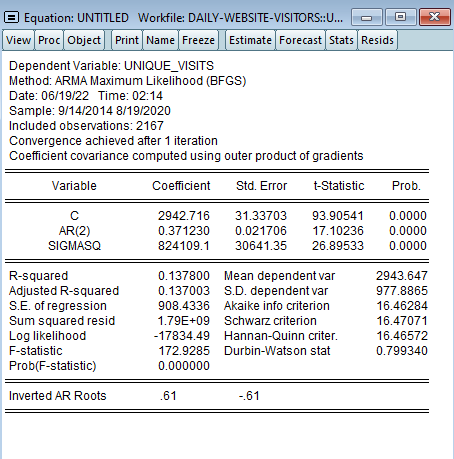


**UYGUN MODEL SEÇİMİ**

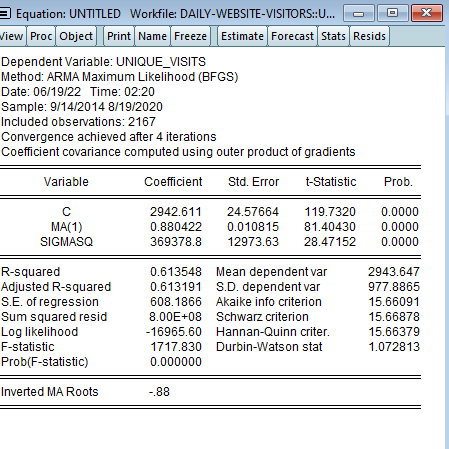
**sar(1)**



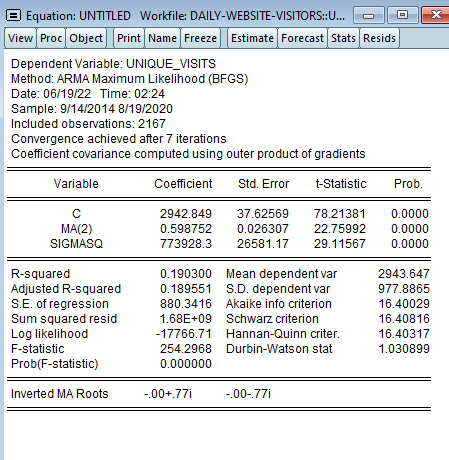
**sar(2)**



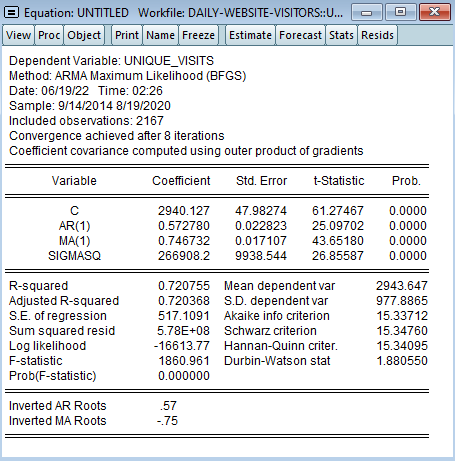
**sma(1)**

****

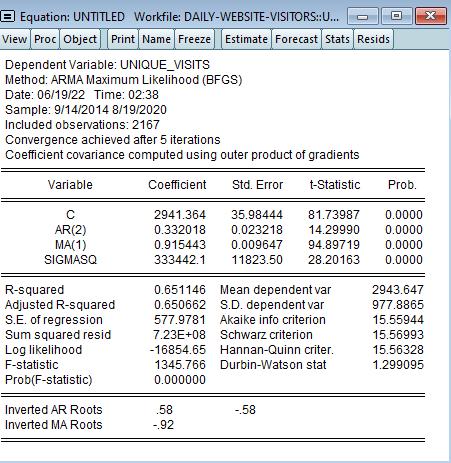
**sma(2)**

****

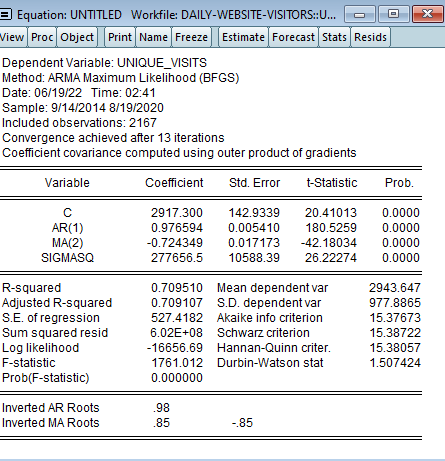
**sar(1) sma(1)**

****

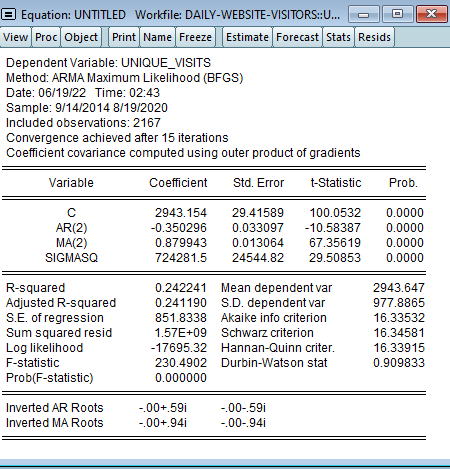
**sar(2) sma(1)**



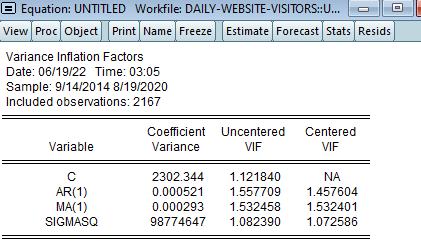
**sar(1) sma(2)**

****

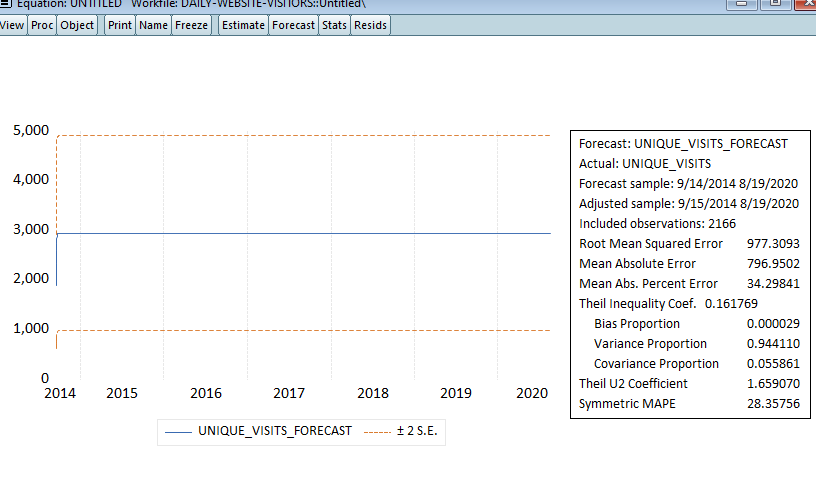
**sar(2) sma(2)**

****

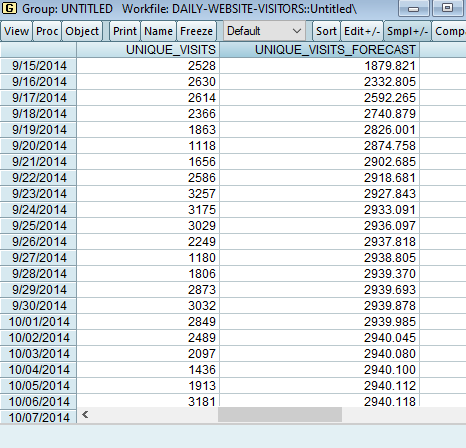
Kurduğum tüm modellerin ve değişkenlerin R^2leri ve değişkenlerin güven aralıkları incelendiğinde en uygun model SARIMA(1,0,1)(1,1,1)= SAR(1) + SMA(1) olarak gözüküyor. Bu modelin VIF değerlerine bakalım.

****

VIF değerleri 1 ila 10 arasında. Çoklu bağlantı sorunumuz yok gözüküyor. Bir de forecasti inceleyelim.

****

Hem serimizi hem de tahmini beraber görüntülediğimizde;



Seçtiğimiz modele göre siteye giren günlük ziyaretçi sayısını verileri göz önünde bulundurarak çeşitli uygulamalarla tahmin etmiş olduk.

**KAYNAKÇA**

[**https://doc.arcgis.com/en/insights/latest/analyze/stl.htm#:~:text=Seasonal%2DTrend%20decomposition%20using%20LOESS%20(STL)%20is%20a%20robust,%2C%20seasonal%2C%20and%20remainder%20components**](https://doc.arcgis.com/en/insights/latest/analyze/stl.htm#:~:text=Seasonal%2DTrend%20decomposition%20using%20LOESS%20(STL)%20is%20a%20robust,%2C%20seasonal%2C%20and%20remainder%20components)**. (1)**