Zaman Serileri - Temel Konular

Zaman Serileri Analizi
Ekonometrik Modelleme ve Zaman Serileri Analizi

Dr. Ömer Kara[1]

[1]İktisat Bölümü
Eskişehir Osmangazi Üniversitesi

14 Nisan 2021

Taslak

Motivasyon

Bu bölümde sırasıyla aşağıdaki konular incelenecektir.

- Veri türleri ve özelliklerinin gözden geçirilmesi: yatay-kesit verisi, zaman serileri verisi, havuzlanmış yatay-kesit verisi ve panel veri
- Zaman serileri verisinin özellikleri ve stokastik süreç
- Zaman serileri verisinin hazırlanmasında kullanılan teknikler
- Zaman serileri örüntüleri: trend, mevsimsellik ve döngüsellik
- Zaman serileri verisi için örnekler
- Zaman serilerinin dekompozizasyonu: klasik dekompozizasyon
- Zaman serilerinde bağımlılığın ölçülmesi: örneklem otokovaryansı, örneklem otokorelasyonu ve örneklem otokorelasyon fonksiyonu
- Stokastik bir sürece örnek olarak pür rassal süreç

Veri Türleri ve Özellikleri

- Zaman serileri verisinin özelliklerini detaylı olarak incelemeden önce diğer veri türlerini hatırlamamız faydalı olacaktır.
- Ekonometrik analizlerde temel olarak kullanılan 4 farklı veri türü vardır.
    - Yatay-kesit verisi (Cross-sectional data)
    - Zaman Serileri verisi (Time series data)
    - Havuzlanmış yatay-kesit verisi (Pooled cross-section data)
    - Panel veri (Panel Data)

Yatay-Kesit Verisi

- Yatay-kesit verisi değişken(ler)e ait verilen zamanın belirli bir kesitinde farklı birimlerden oluşan veri türüdür.
- Şekil 1'deki veri tablosu bireylerin özelliklerini gösteren yatay-kesit verisine bir örnektir.

TABLE 1.1 A Cross-Sectional Data Set on Wages and Other Individual Characteristics

obsno	wage	educ	exper	female	married
1	3.10	11	2	1	0
2	3.24	12	22	1	1
3	3.00	11	2	0	0
4	6.00	8	44	0	1
5	5.30	12	7	0	1
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
525	11.56	16	5	0	1
526	3.50	14	5	1	0

Şekil 1: Yatay-Kesit Verisi Örneği 1

Kaynak: Wooldridge (2016)

Yatay-Kesit Verisi

- Şekil 2'deki veri tablosu ülkelerin ekonomik büyüme oranlarını ve ülke özelliklerini gösteren yatay-kesit verisine bir başka örnektir.

TABLE 1.2 A Data Set on Economic Growth Rates and Country Characteristics

obsno	country	gpcrgdp	govcons60	second60
1	Argentina	0.89	9	32
2	Austria	3.32	16	50
3	Belgium	2.56	13	69
4	Bolivia	1.24	18	12
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
61	Zimbabwe	2.30	17	6

Şekil 2: Yatay-Kesit Verisi Örneği 2

Kaynak: Wooldridge (2016)

Zaman Serileri Verisi

- Zaman Serileri verisi değişken(ler)e ait verilen aynı birimin farklı zamanlarından oluşan veri türüdür.
- Şekil 3'deki veri tablosu Porto Riko'daki minimum maaş, işsizlik ve benzer istatistikleri gösteren zaman serileri verisine bir örnektir.

TABLE 1.3 Minimum Wage, Unemployment, and Related Data for Puerto Rico

obsno	year	avgmin	avgcov	prunemp	prgnp
1	1950	0.20	20.1	15.4	878.7
2	1951	0.21	20.7	16.0	925.0
3	1952	0.23	22.6	14.8	1015.9
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
37	1986	3.35	58.1	18.9	4281.6
38	1987	3.35	58.2	16.8	4496.7

Şekil 3: Zaman Serileri Verisi Örneği 1

Kaynak: Wooldridge (2016)

Havuzlanmış Yatay-Kesit

- Havuzlanmış yatay-kesit verisi değişken(ler)e ait verilen farklı zamanlarındaki yatay-kesit verilerinin birleştirilmesiyle oluşan veri türüdür.
- Şekil 4'deki veri tablosu iki farklı yıldaki havuzlanmış (bir araya getirilmiş) ev fiyatlarını gösteren havuzlanmış yatay-kesit verisine bir örnektir.

TABLE 1.4 Pooled Cross Sections: Two Years of Housing Prices

obsno	year	hprice	proptax	sqrft	bdrms	bthrms
1	1993	85,500	42	1600	3	2.0
2	1993	67,300	36	1440	3	2.5
3	1993	134,000	38	2000	4	2.5
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
250	1993	243,600	41	2600	4	3.0
251	1995	65,000	16	1250	2	1.0
252	1995	182,400	20	2200	4	2.0
253	1995	97,500	15	1540	3	2.0
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
520	1995	57,200	16	1100	2	1.5

Şekil 4: Havuzlanmış Yatay-Kesit Verisi Örneği

Kaynak: Wooldridge (2016)

Panel Veri

- Panel veri değişken(ler)e ait verilen farklı birimlerin farklı zamanlarından oluşan veri türüdür.
- Şekil 5'deki veri tablosu iki farklı yıldaki suç istatistiklerini gösteren panel veriye bir örnektir.

TABLE 1.5 A Two-Year Panel Data Set on City Crime Statistics

obsno	city	year	murders	population	unem	police
1	1	1986	5	350,000	8.7	440
2	1	1990	8	359,200	7.2	471
3	2	1986	2	64,300	5.4	75
4	2	1990	1	65,100	5.5	75
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
297	149	1986	10	260,700	9.6	286
298	149	1990	6	245,000	9.8	334
299	150	1986	25	543,000	4.3	520
300	150	1990	32	546,200	5.2	493

Şekil 5: Panel Veri Örneği

Kaynak: Wooldridge (2016)

Zaman Serileri Verisinin Özellikleri

- Bir zaman serisi değişkeni, zamana göre indekslenmiş bir gözlem veya ölçüm dizisi olarak tanımlanabilir.
  - Örneğin, y_t bir zaman serisidir. Burada zaman indeksi t'nin ($t = 1, 2, \ldots, n$) ayrık olduğu varsayılır ve n gözlem sayısıdır.
- Zaman serilerinde veriler, yatay-kesit verisinden farklı olarak genellikle eskiden yeniye belli bir zaman sıralaması izlemektedir.
  - Gözlemler arasındaki zaman aralıkları (zaman frekansı) düzenli veya düzensiz olabilir.
  - Biz sadece düzenli olarak ölçülen zaman serisi verilerine odaklanacağız. Örneğin: aylık, yıllık, haftalık ve günlük frekanstaki veri.

Tablo 1: Zaman Serileri Verisi - Frekanslar

Data	Frekans
Yıllık (Annual)	1
Çeyreklik (Quarterly)	4
Aylık (Monthly)	12
Haftalık (Weekly)	52.25
Günlük (Daily)	365.25
Saatlik (Hourly)	8766

Notlar: Zaman aralığı olarak 1 yıl alınmıştır.

Zaman Serileri Verisinin Özellikleri

- Şekil 6'deki veri tablosu ABD'deki enflasyon ve işsizlik oranlarını gösteren zaman serileri verisine bir başka örnektir.

TABLE 10.1 Partial Listing of Data on U.S. Inflation and Unemployment Rates, 1948–2003

Year	Inflation	Unemployment
1948	8.1	3.8
1949	−1.2	5.9
1950	1.3	5.3
1951	7.9	3.3
.	.	.
.	.	.
.	.	.
1998	1.6	4.5
1999	2.2	4.2
2000	3.4	4.0
2001	2.8	4.7
2002	1.6	5.8
2003	2.3	6.0

Şekil 6: Zaman Serileri Verisi Örneği 2

Kaynak: Wooldridge (2016)

- Zaman serileri analizinde geçmiş değerler gelecekteki değerleri etkilemektedir fakat bunun tersi geçerli değildir.
- Şekil 6'deki zaman serisi verisinde 2000 yılındaki enflasyon ilerleyen yıllardaki enflasyonu etkilerken geçmiş yıllardaki enflasyon verilerini etkileyemez.

Zaman Serileri Verisinin Özellikleri

- Yatay-kesit verisinde kullandığımız önemli varsayımlardan birisi **rassallık** varsayımıydı.
  - Rassallık varsayımına göre tahminde kullanılan n tane gözlem ilgili anakütleden rassal örnekleme yoluyla seçilmiştir. Yani gözlemler stokhastiktir (rassal), deterministik (kesin) değil.
  - Anakütleden alınan farklı bir örnek genellikle bağımlı ve bağımsız değişkenlerin farklı değerlerini içereceğinden, bu örneklemler yardımıyla ulaşılan SEKK parametre tahmin değerleri de genellikle farklılık gösterir.
  - Bu nedenle SEKK parametre tahmincileri de rassal değişkenler olarak değerlendirilir.
- Peki zaman serilerinde de rassallık varsayımını kullanabilir miyiz? Eğer kullanabilirsek, rassallığı nasıl yorumlamamız gerekir?
- Zaman serisi değişkenlerinin (enflasyon, işsizlik, gayri safi yurtiçi hasıla, BIST 100 kapanış fiyatları, vs) bir sonraki dönemde hangi değerleri alacaklarını öngöremediğimiz için bu değişkenleri rassal değişken olarak düşünebiliriz.

Stokastik Süreç

Stokastik Süreç / Zaman Serisi Süreci

Zaman (t) indeksi taşıyan rassal değişkenlerin oluşturduğu diziye/seriye **stokastik süreç** (stochastic process) ya da **zaman serisi süreci** (time series process) denir.

- Stokastik sözcüğü rassal ile aynı anlamda kullanılmaktadır.
- Mevcut bir zaman serisi, stokastik sürecin **olası bir gerçekleşmesi** olarak görülebilir.
- Zamanda geriye gidip başka bir gerçekleşme elde edemeyeceğimiz için zaman serileri tek bir gerşekleşmenin sonuçlarıdır.
- Bununla birlikte, farklı tarihsel koşullar altında ilgilendiğimiz stokastik sürecin genellikle farklı bir gerçekleşmesini elde ederiz.
- Bu nedenle, bir zaman serisi sürecinin bütün olası gerçekleşmelerinin oluşturacağı küme, zaman serisi analizinde yatay-kesit verisindeki anakütlenin rolünü üstlenecektir.

Stokastik Süreç

Stokastik Süreç: Tekil Gerçekleşme

Bir stokastik sürecin tekil gerçekleşmesi $\{y_t : t = 1, 2, \ldots, n\}$ ya da $\{y_t\}_{t=1}^{n}$ ile gösterilebilir. Bu tekil gerçekleşme aşağıdaki sonsuz serinin bir alt kümesi olarak düşünülebilir.

$$\{y_t\}_{t=-\infty}^{\infty} = \{\ldots, y_{-1}, y_0, \underbrace{y_1, y_2, \ldots, y_{n-1}, y_n}_{\{y_t\}_{t=1}^{n} \text{ gerçekleşme}}, y_{n+1}, y_{n+2}, \ldots\}$$

- Uygulamada zaman indeksi her zaman 1 değerinden başlar. Fakat, teorik olarak her hangi bir tam sayı değerini olabilir (hatta sürekli reel sayı da olabilir).
- Eğer süreç tekrarlanırsa, aynı stokastik süreç kullanılarak farklı bir gerçekleşme elde edilebilir.
- Ekonomi gibi sosyal bilimlerde, biz çoğunlukla stokastik süreçlerin tekil bir gerçekleşmesini kullanacağız.
- Şimdi, zaman serileri verisi örneklerini ve sıklıkla kullanacağımız teknik ve kavramları inceleyelim.

Zaman Serileri Verisinin Hazırlanması

- Zaman serileri analizinde, veri öncelikle analize uygun şekilde hazırlanmalıdır.
- Örneğin, zaman serileri analizde:
  - Paranın değerinden yani enflasyondan etkilenen nominal zaman serisi kullanmak yerine, fiyatların etkisinden arındırılmış reel değerler hesaplanıp kullanılmalıdır. Yani, zaman serisi için **enflasyon ayarlaması** (inflation adjustment) yapılmalıdır.
  - Nüfustan etkilenen zaman serisi kullanmak yerine, zaman serisini nüfusun etkisinden arındırmak için kişi başı değerler hesaplanıp kullanılmalıdır. Yani, zaman serisi için **nüfus ayarlaması** (population adjustment) yapılmalıdır.
  - Parasal zaman serisi verilerinin diğer ülke verileriyle karşılaştırılabilmesi için verinin uluslararası kullanımı olan ABD Doları, Euro ya da Sterlin cinsinden hesaplanıp kullanılması, yani **kur ayarlaması** yapılması, daha faydalı olacaktır.
  - y_{t-s} gibi **gecikmeli zaman serisi** değişkenleri kullanılacaksa, istenilen gecikmeler için veri tablo halinde hazırlanmalıdır.
  - Zaman serisine ait **büyüme oranı** kullanılacaksa, büyüme oranı ve istenilen gecikmeler için veri tablo halinde hazırlanmalıdır.
  - **İndeks** kullanılacaksa, indeksin baz dönemine dikkat edilmeli ve gerekiyorsa baz dönemi değiştirilmelidir. Kullanılan tüm indekslerin baz dönemi aynı olmalıdır.
  - Bazen, zaman serisinin normal dağılım yapabilmesi için **matematiksel transformasyon** ile dönüştürülmesi gerekebilir. Örneğin, logaritmik transformasyon ve Box–Cox transformasyonu.

Zaman Serileri Verisinin Grafiği

- Zaman serisi verisi hazırlandıktan sonra ilk adım verinin zamana göre grafiğini çıkartmak olmalıdır.
  - Yani, düz çizgilerle birleştirilen ardışık zaman serisi gözlemleri, gözlem zamanı t'ye göre çizilir ve grafikle gösterilir.
- Bu bölümde, zaman serileri analizinde sıklıkla kullanılan farklı frekanslardaki zaman serilerine ait grafikler karşılaştırılmalı olarak incelenecektir.
- İnceleyeceğimiz zaman serisi frekansları (artan frekansta):
  - Yıllık
  - Çeyreklik
  - Aylık
  - Günlük
- Ayrıca, zaman serileri versinin hazırlanmasında kullanılan teknikler de grafiklerle beraber kısaca incelenecektir.

Zaman Serileri Örüntüleri

- Zaman serilerine ait grafikleri incelemeden önce, zaman serileri analizde sıkça görülen zaman serisi örüntülerini (time series patterns) inceleyelim.

Trend (Trend)

Zaman serisi verisinde uzun dönemli bir artış veya azalış olduğunda bir **trend** vardır.

- Örneğin, Şekil 11'deki Kişi Başı Reel GSYH - Türkiye verisinde güçlü bir trend vardır.
- Trendin her zaman doğrusal olması gerekmez. Zaman serilerinde doğrusal olmayan trend de gözlenebilir, örneğin Şekil 17'deki Tüketici Fiyatları - Türkiye verisinde artan trend vardır.

Mevsimsellik (Seasonality)

Mevsimsel örüntü, bir zaman serisi yılın üç aylık dönemi, ayı veya haftanın günü gibi mevsimsel faktörlerden etkilendiğinde ortaya çıkar.

- Mevsimsellik her zaman sabit ve bilinen bir dönemdedir.
- Örneğin, Şekil 23'deki aylık Nottingham Ortalama Sıcaklık verisinde yaz ve kış aylarında olmak üzere mevsimsellik vardır.

Zaman Serileri Örüntüleri

Döngüsellik (Cyclical)

Zaman serisi verisinde sabit bir frekansta olmayan yükseliş ve düşüşler olduğunda döngüsellik vardır.

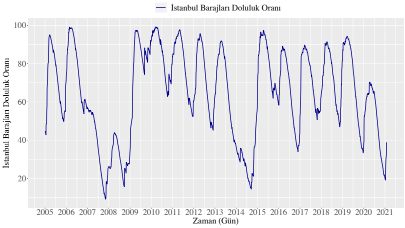
- Dalgalanmalar genellikle ekonomik koşullardan kaynaklanır ve dolayısıyla genellikle "iş döngüsü" ile ilgilidir.
- Dalgalanmaların süresi genellikle en az 2 yıldır.
- Örneğin, Şekil 11'deki Kişi Başı Reel GSYH - Türkiye verisinde döngüsellik vardır.

- Döngüsellik, sıklıkla mevsimsellikle karıştırılır.
  - Dalgalanmalar sabit bir frekanstaysa (yani dalgalanmaların sıklığı değişmiyorsa) ve zamanın bazı yönleriyle ilişkiliyse mevsimsellik vardır.
  - Dalgalanmalar sabit bir frekansta değilse döngüsellik vardır.
  - Genel olarak, döngüsel dalgalanmaların ortalama uzunluğu, mevsimsel bir örüntünün uzunluğundan daha uzundur.
  - Döngüsel dalgalanmaların büyüklüğü, mevsimsel bir örüntünün büyüklüğünden daha değişken olma eğilimindedir. Yani, döngüsel dalgalanmalarda dalga boyu dönemden döneme farklılık gösterirken, mevsimsel örüntüde dalga boyu aşağı yukarı hep aynıdır.

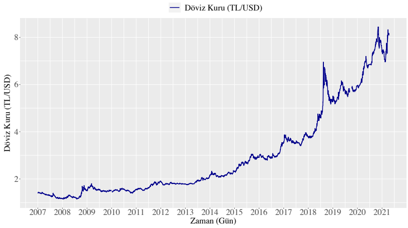
Zaman Serileri Örüntüleri

- Çoğu zaman serisi trend, mevsimsellik ve döngüsellik içerir.
- Bir tahmin yöntemi seçerken, önce verilerdeki zaman serisi örüntülerini tanımlamamız ve ardından örüntüleri doğru şekilde yakalayabilen bir yöntem seçmemiz gerekir.
- **Şekil** 7'de bu örüntülerin farklı kombinasyonlarını içeren örnekler verilmiştir.
  - **Şekil** 7**a**'daki İstanbul Barajları Doluluk Oranları verisi her yıl güçlü bir mevsimsellik göstermektedir. Ayrıca belirli dönemler tekrarlanan ve yaklaşık 2 yıl süren döngüsellik de mevcuttur. Belirgin bir trend yoktur.
  - **Şekil** 7**b**'deki Döviz Kuru (TL/USD) verisinde mevsimsellik olmamakla birlikte, belirgin bir yukarı yönlü trend vardır. Hatta, 2017 yılından itibaren artan bir trend mevcuttur denilebilir. Belirgin ve güçlü bir döngüsellik mevcuttur.
  - **Şekil** 7**c**'deki Avustralya Elektrik Üretimi verisinde güçlü bir mevsimsellikle birlikte güçlü yukarı yönlü bir trend mevcuttur. Belirgin bir döngüsellik yoktur.
  - **Şekil** 7**d**'deki Google Hisse Senedi Getiri Oranı verisinde trend, mevsimsellik ve döngüsellik yoktur. Görüldüğü gibi tahmin etmesi zor rassal dalgalanmalar mevcuttur. Bu yönüyle daha sonra Slayt 71'da göreceğimiz pür rassal sürece benzemektedir.

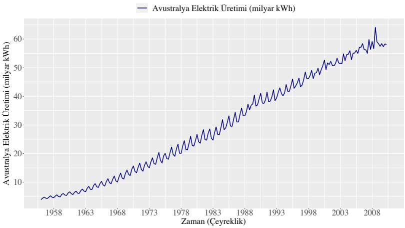
Zaman Serileri Örüntüleri



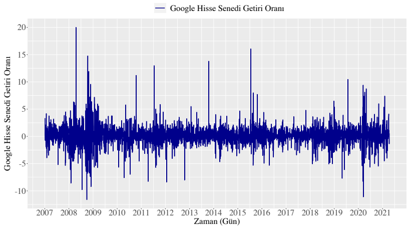
(a) İstanbul Barajları Doluluk Oranları



(b) Döviz Kuru (TL/USD)



(c) Avustralya Elektrik Üretimi

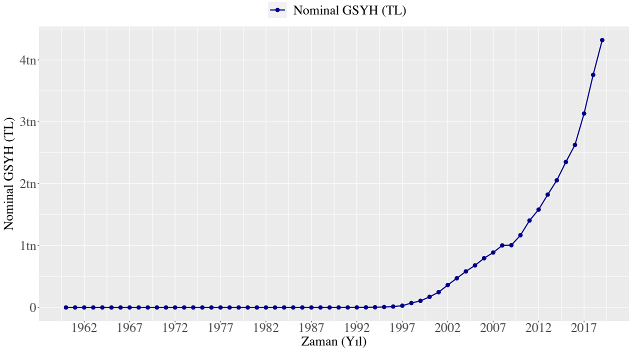


(d) Google Hisse Senedi Getiri Oranı

Şekil 7: Farklı Örüntüleri Gösteren Dört Zaman Serisi Örneği

Nominal GSYH (TL) - Türkiye

- Şekil 8'de Türkiye için Gayri Safi Yurtiçi Hasıla (GSYH) verisi cari fiyatlarla (nominal) Türk Lirası (TL) cinsinden gösterilmiştir. Yani, Nominal GSYH (TL).
- Şekil 8'de kullanılan verinin enflasyona ve nüfusa göre ayarlanmadığına, ve ayrıca TL cinsinden olduğuna, dikkat edin!

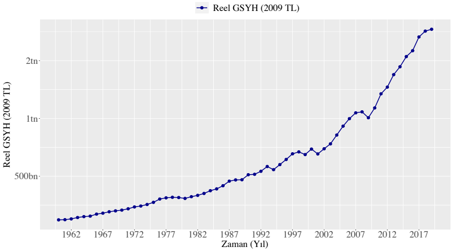


Şekil 8: Nominal GSYH (TL) - Türkiye

Kaynak: World Bank

Reel GSYH (TL) - Türkiye

- Şekil 8'de gösterilen Nominal GSYH (TL) verisi, Şekil 9'de enflasyon ayarlaması yapılmış (reel) GSYH (TL) olarak gösterilmiştir. Yani, Reel GSYH (TL).
- Şekil 9'de kullanılan verinin nüfusa göre ayarlanmadığına, ve ayrıca TL cinsinden olduğuna, dikkat edin!



Şekil 9: Reel GSYH (TL) - Türkiye

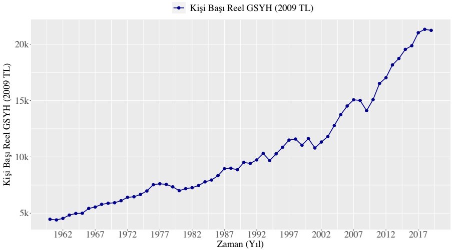
Kaynak: World Bank

- Reel GSYH şu şekilde hesaplanabilir:

$$\text{Reel GSYH}_t = \frac{\text{Nominal GSYH}_t}{\text{GSYH Deflatörü}_t} \times 100$$

Kişi Başı Reel GSYH (TL) - Türkiye

- Şekil 9'de gösterilen Reel GSYH (TL) verisi, Şekil 10'de nüfus ayarlaması yapılmış (kişi başı) Reel GSYH (TL) olarak gösterilmiştir. Yani, Kişi Başı Reel GSYH (TL).
- Şekil 10'de kullanılan verinin TL cinsinden olduğuna dikkat edin!



Şekil 10: Kişi Başı Reel GSYH (TL) - Türkiye

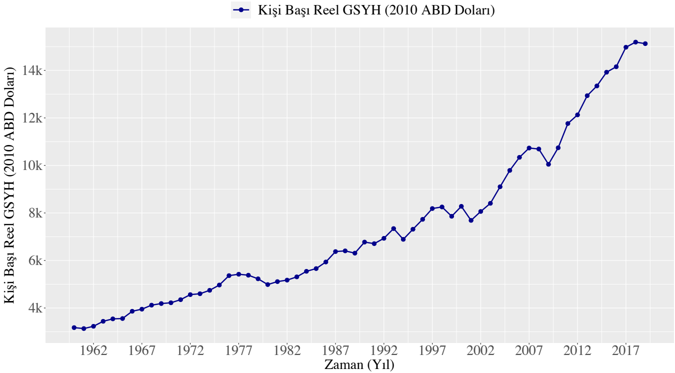
Kaynak: World Bank

- Kişi Başı Reel GSYH şu şekilde hesaplanabilir:

$$\text{Kişi Başı Reel GSYH}_t = \frac{\text{Reel GSYH}_t}{\text{Nüfus}_t}$$

Kişi Başı Reel GSYH (ABD Doları) - Türkiye

- Şekil 11'de Türkiye için Gayri Safi Yurtiçi Hasıla (GSYH) verisi enflasyona ve nüfusa göre ayarlanmıştır, ve ayrıca ABD Doları cinsindendir. Yani, Kişi Başı Reel GSYH (ABD Doları).



Şekil 11: Kişi Başı Reel GSYH (ABD Doları) - Türkiye

Kaynak: World Bank

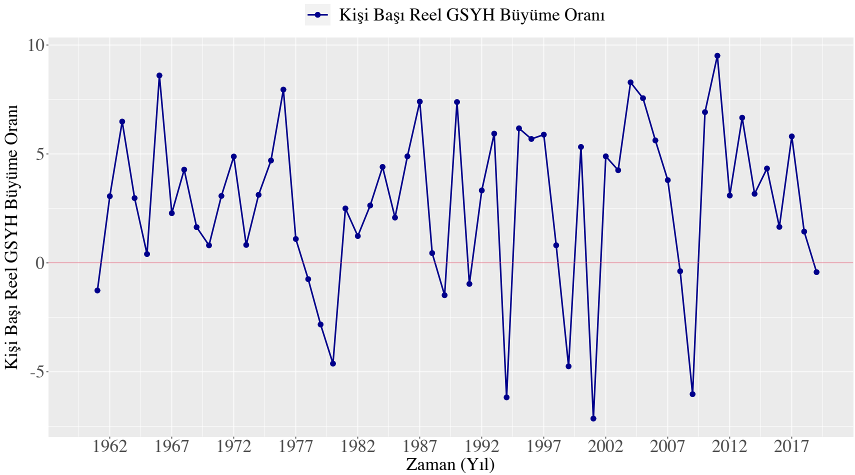
Kişi Başı Reel GSYH (ABD Doları) - Türkiye

Tablo 2: Kişi Başı Reel GSYH (ABD Doları) - Türkiye

Yıl	t	$GSYH_t$	$GSYH_{t-1}$	$GSYH_{t-2}$	$GSYH_{t-3}$	$GSYH_{t-4}$
1960	1	3175	NA	NA	NA	NA
1961	2	3135	3175	NA	NA	NA
1962	3	3230	3135	3175	NA	NA
1963	4	3440	3230	3135	3175	NA
1964	5	3542	3440	3230	3135	3175
1965	6	3557	3542	3440	3230	3135
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2016	57	14153	13924	13346	12936	12128
2017	58	14975	14153	13924	13346	12936
2018	59	15190	14975	14153	13924	13346
2019	60	15125	15190	14975	14153	13924

Notlar: World Bank datası kullanılmıştır.

Kişi Başı Reel GSYH Büyüme Oranı - Türkiye



Şekil 12: Kişi Başı Reel GSYH Büyüme Oranı - Türkiye

Kaynak: World Bank

Kişi Başı Reel GSYH Büyüme Oranı - Türkiye

Tablo 3: Kişi Başı Reel GSYH Büyüme Oranı - Türkiye

Yıl	t	GSYH_t	GR_t	GR_{t-1}	GR_{t-2}	GR_{t-3}	GR_{t-4}
1960	1	3175	NA	NA	NA	NA	NA
1961	2	3135	−1.27	NA	NA	NA	NA
1962	3	3230	3.06	−1.27	NA	NA	NA
1963	4	3440	6.49	3.06	−1.27	NA	NA
1964	5	3542	2.97	6.49	3.06	−1.27	NA
1965	6	3557	0.4	2.97	6.49	3.06	−1.27
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2018	59	15190	1.44	5.8	1.65	4.33	3.17
2019	60	15125	−0.43	1.44	5.8	1.65	4.33

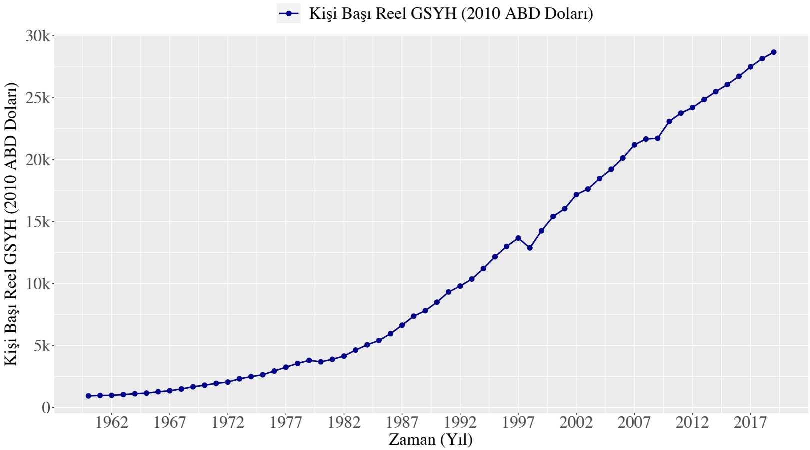
Notlar: World Bank datası kullanılmıştır.

- Büyüme oranı şu şekilde hesaplanabilir:

$$\text{GR}_t = \frac{\text{GSYH}_t - \text{GSYH}_{t-1}}{\text{GSYH}_{t-1}} \times 100$$

Kişi Başı Reel GSYH (ABD Doları) - Kore

- Şekil 11'deki Türkiye'ye ait grafik ile Kore grafiğini karşılaştıralım.

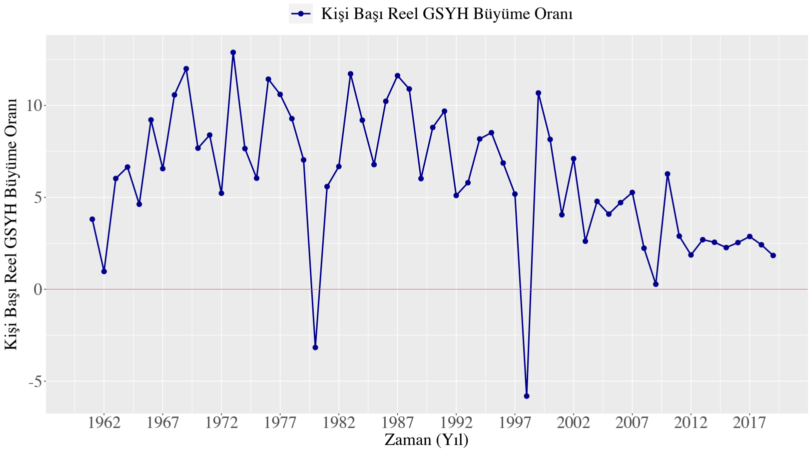


Şekil 13: Kişi Başı Reel GSYH (ABD Doları) - Kore

Kaynak: World Bank

Kişi Başı Reel GSYH Büyüme Oranı - Kore

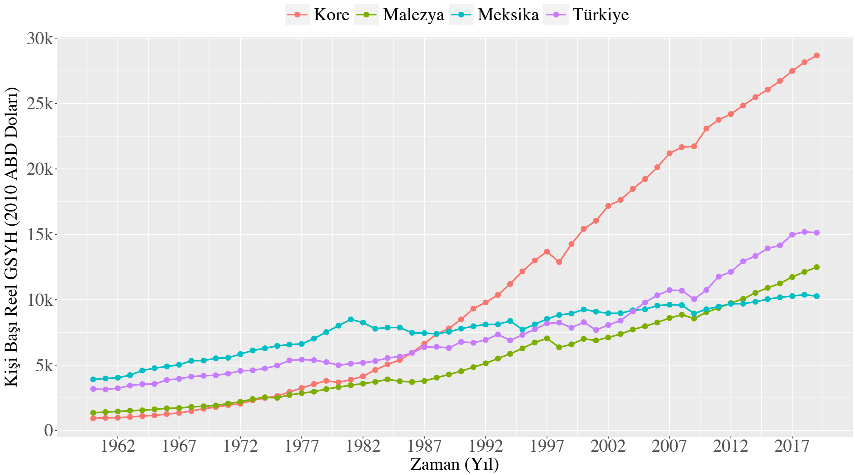
- Şekil 12'deki Türkiye'ye ait grafik ile Kore grafiğini karşılaştıralım.



Şekil 14: Kişi Başı Reel GSYH Büyüme Oranı - Kore

Kaynak: World Bank

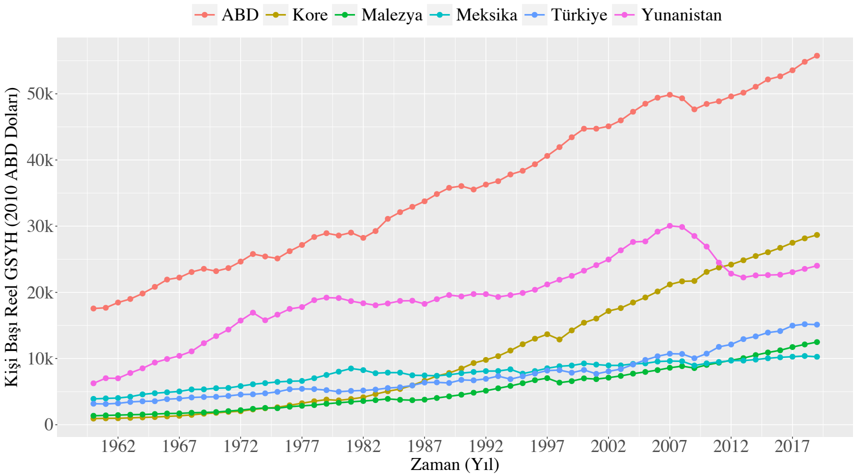
Kişi Başı Reel GSYH (ABD Doları) - Seçilmiş Ülkeler 1



Şekil 15: Kişi Başı Reel GSYH (ABD Doları) - Seçilmiş Ülkeler 1

Kaynak: World Bank

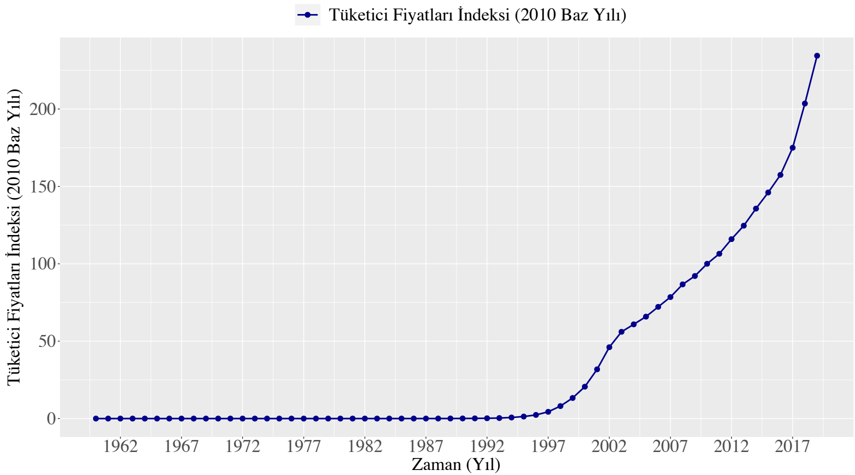
Kişi Başı Reel GSYH (ABD Doları) - Seçilmiş Ülkeler 2



Şekil 16: Kişi Başı Reel GSYH (ABD Doları) - Seçilmiş Ülkeler 2

Kaynak: World Bank

Tüketici Fiyatları İndeksi - Türkiye (2010 Baz Yılı)



Şekil 17: Tüketici Fiyatları İndeksi - Türkiye (2010 Baz Yılı)

Kaynak: World Bank

İndeksler

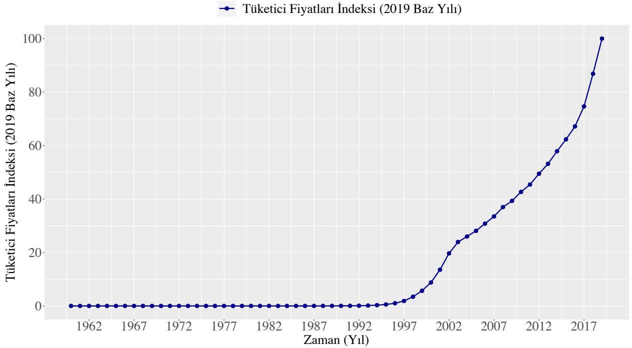
- İndeksler, makroekonometrik ve finansal uygulamalarda yaygın olarak kullanılmaktadır.
    - Örneğin, Tüketici Fiyatları İndeksi, Üretici Fiyatları İndeksi ve Endüstriyel Üretim İndeksi.
- Her indeksin belli bir baz dönemi vardır. Baz döneminde indeks değeri 100'dür.
- Bir indeksin belirli değerleri yalnızca baz dönemdeki değerle karşılaştırılarak yorumlanabilir.
    - Örneğin, baz dönemi 2010 yılı olan bir indekste, diğer dönemlerindeki değerler ancak baz dönemine göre karşılaştırılarak yorumlanabilir. Değer 2014'te 130 ise, endeksin 2010'dan 2014'e kadar %30 arttığını söyleyebiliriz.
- Aşağıdaki formülü kullanarak herhangi bir indeksin baz dönemini kolayca değiştirebiliriz.

$$\text{yeni indeks}_t = \frac{\text{eski indeks}_t}{\text{eski indeks}_{\text{yeni baz dönemi}}} \times 100$$

  burada eski $\text{indeks}_{\text{yeni baz dönemi}}$ eski indeksin yeni baz dönemindeki değeridir.
- İndeks büyüme oranları indekste kullanılan baz dönemine göre farklılık göstermez, yani her zaman aynıdır. Bakınız Şekil 19.

Tüketici Fiyatları İndeksi - Türkiye (2019 Baz Yılı)

- Şekil 17'te 2010 baz yılı kullanılarak gösterilen Tüketici Fiyatları İndeksi, Slayt 33'deki formül kullanılarak, Şekil 18'te 2019 baz yılı ile gösterilmiştir.

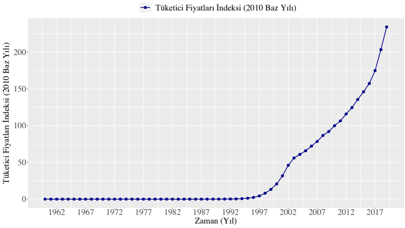


Şekil 18: Tüketici Fiyatları İndeksi - Türkiye (2019 Baz Yılı)

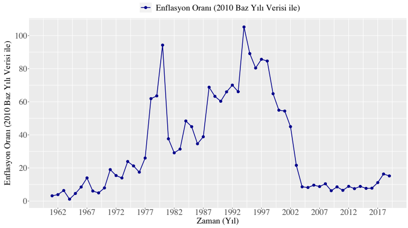
Kaynak: World Bank

- Açıkça görüldüğü gibi iki grafik, aralarındaki ölçek farkı hariç tamamen aynıdır.

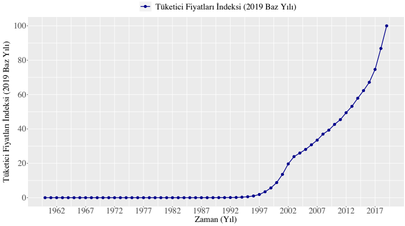
Enflasyon - Türkiye (Tüketici Fiyatları İndeksi ile)



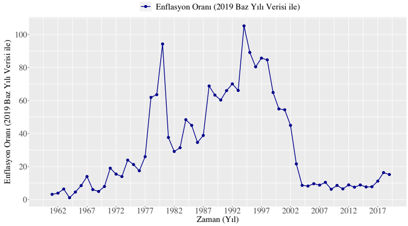
(a) Tüketici Fiyatları İndeksi (2010 Baz Yılı)



(b) Enflasyon Oranı (2010 Baz Yılı Verisi ile)



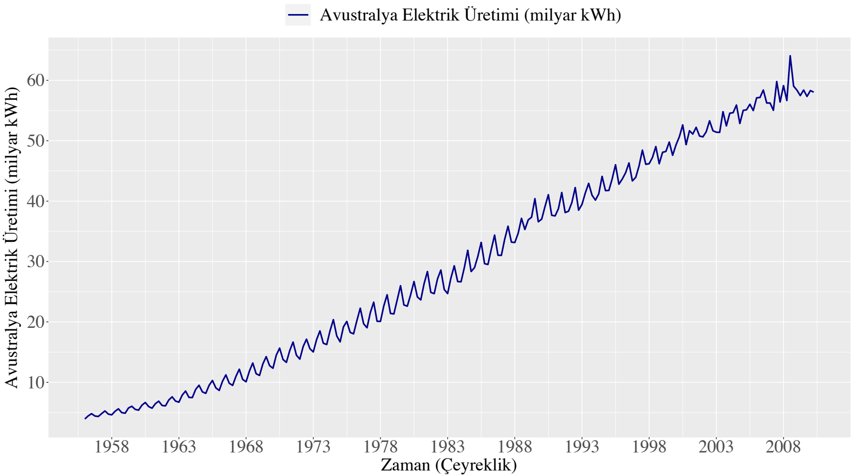
(c) Tüketici Fiyatları İndeksi (2019 Baz Yılı)



(d) Enflasyon Oranı (2019 Baz Yılı Verisi ile)

Şekil 19: Farklı Baz Yılına Sahip Verilerle Enflasyon Oranı Hesaplanması

Avustralya Elektrik Üretimi - Çeyreklik Veri



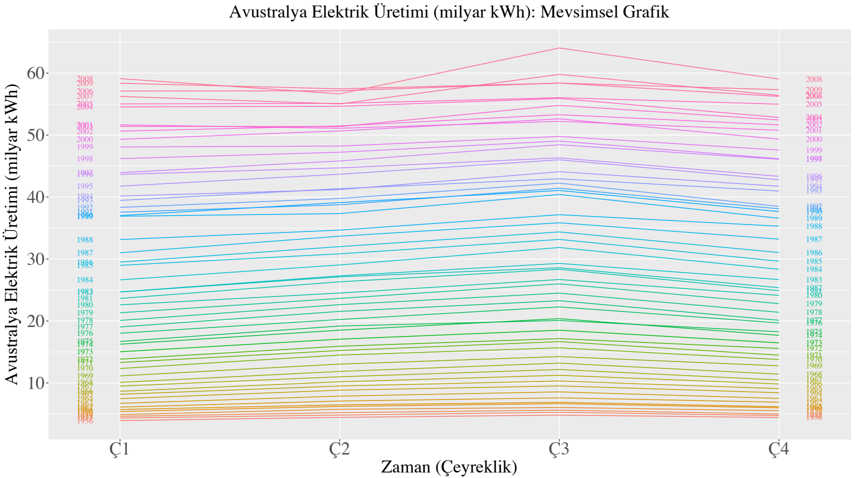
Şekil 20: Avustralya Elektrik Üretimi

Kaynak: fpp2 R paketi

Mevsimsel Grafik ve Mevsimsel Altseri Grafiği

- Zaman serisi verisindeki mevsimsel örüntünün görsel olarak belirlenmesinde **mevsimsel grafik** ve **mevsimsel altseri grafiği** kullanılabilir.
- Mevsimsel grafik, verinin mevsimlere göre grafiğe dökülmesi dışında zaman serisi grafiğine çok benzerdir.
  - Mevsimsel grafik, her mevsime ait verilerin karşılaştırma amacıyla üst üste getirildiği grafiktir.
  - Mevsimsel grafik, verinin altında yatan mevsimsel örüntünün daha net görülmesini sağlar ve özellikle mevsimsel örüntünün değiştiği yılların belirlenmesinde faydalıdır.
  - Örneğin, Şekil 20'de ve Şekil 21'de kullanılan veriler tamamen aynı olmasına rağmen Şekil 21'deki mevsimsel grafik mevsimselliğin etkisini net bir şekilde göstermektedir.
- Mevsimsel altseri grafiği, her mevsime ait verilerin ayrı ayrı zaman çizelgelerinde bir araya toplandığı grafiktir.
  - Mevsimsel altseri grafiğinde, mavi yatay çizgiler, her mevsimin ortalamasını gösterir.
  - Mevsimsel altseri grafiği, özellikle belirli mevsimlerdeki zaman içindeki değişiklikleri belirlemede çok kullanışlıdır.
  - Örneğin, Şekil 22'deki mevsimsel altseri grafiği her mevsimdeki değişikleri net bir şekilde göstermektedir.

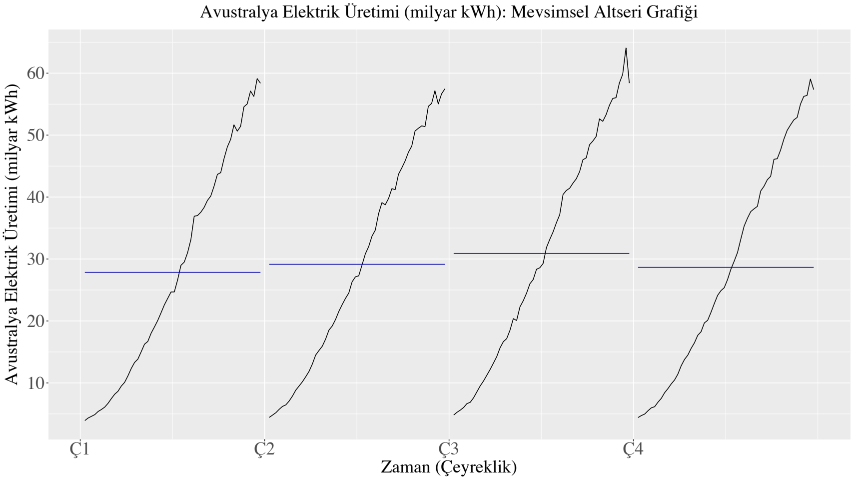
Avustralya Elektrik Üretimi: Mevsimsel Grafik



Şekil 21: Avustralya Elektrik Üretimi: Mevsimsel Grafik

Kaynak: fpp2 R paketi

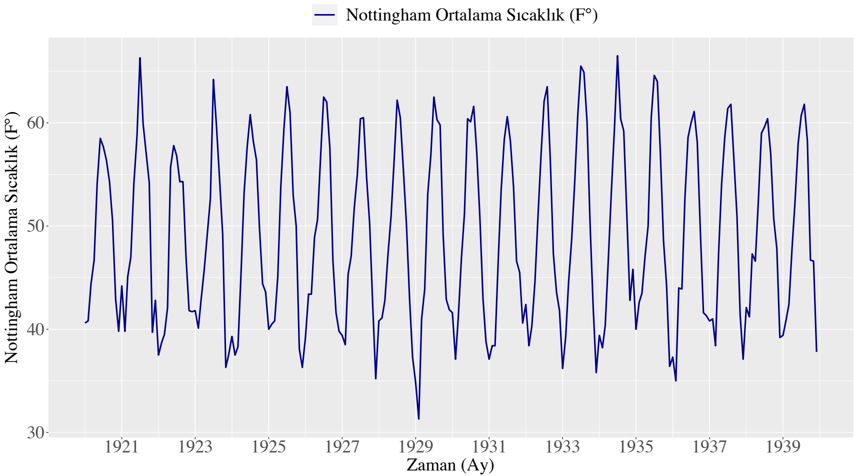
Avustralya Elektrik Üretimi: Mevsimsel Altseri Grafiği



Şekil 22: Avustralya Elektrik Üretimi: Mevsimsel Altseri Grafiği

Kaynak: fpp2 R paketi

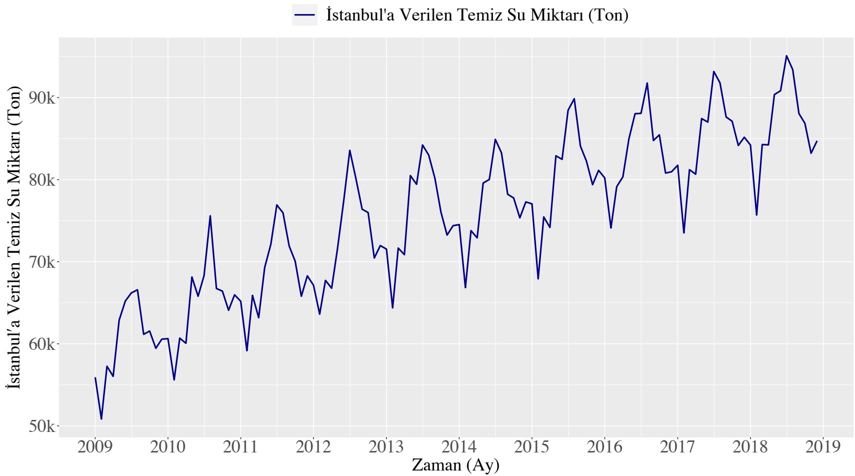
Nottingham Ortalama Sıcaklık - Aylık Veri



Şekil 23: Nottingham Ortalama Sıcaklık

Kaynak: datasets R paketi

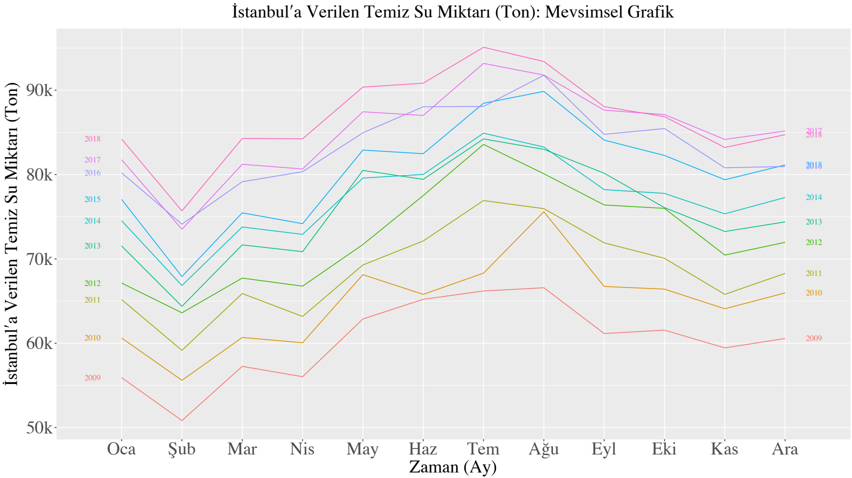
İstanbul'a Verilen Temiz Su Miktarı - Aylık Veri



Şekil 24: İstanbul'a Verilen Temiz Su Miktarı

Kaynak: İstanbul Büyükşehir Belediyesi

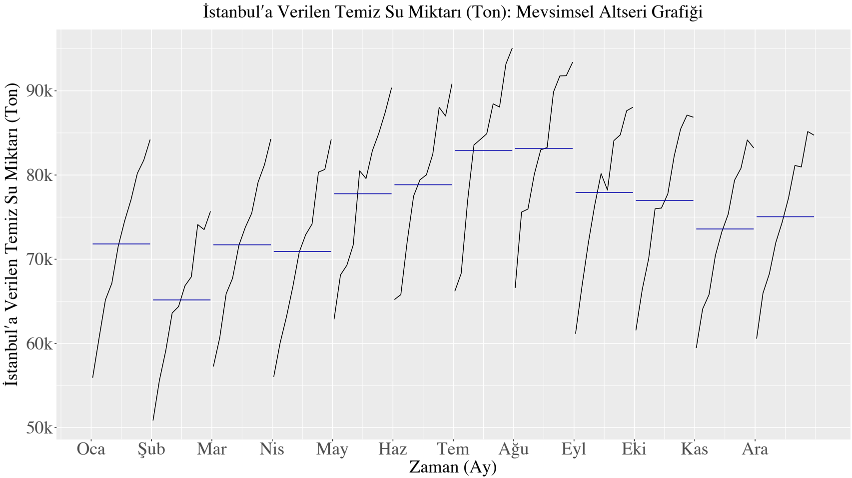
İstanbul'a Verilen Temiz Su Miktarı: Mevsimsel Grafik



Şekil 25: İstanbul'a Verilen Temiz Su Miktarı: Mevsimsel Grafik

Kaynak: İstanbul Büyükşehir Belediyesi

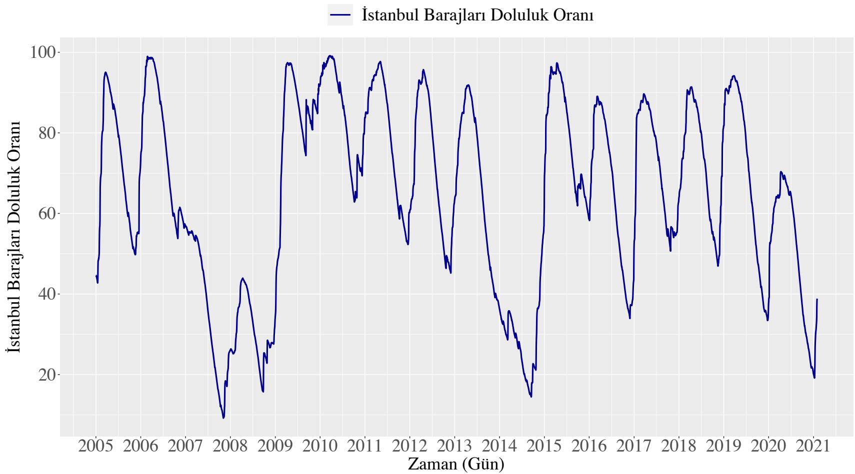
İstanbul'a Verilen Temiz Su Miktarı: Mevsimsel Altseri Grafiği



Şekil 26: İstanbul'a Verilen Temiz Su Miktarı: Mevsimsel Altseri Grafiği

Kaynak: İstanbul Büyükşehir Belediyesi

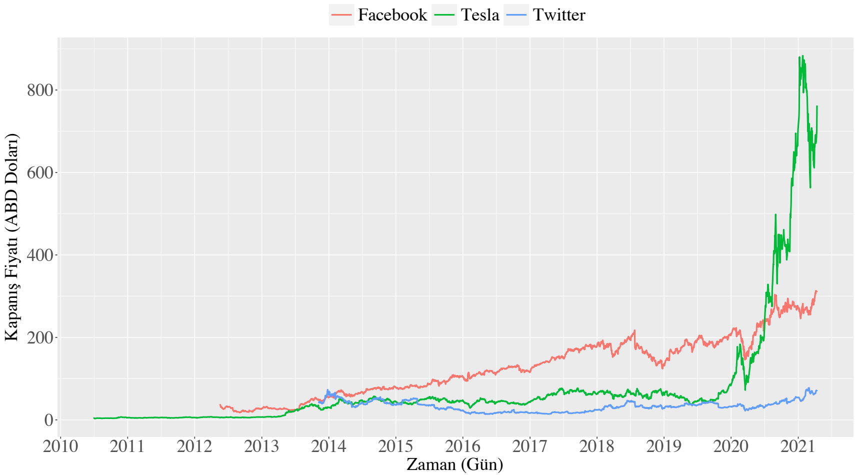
İstanbul Barajları Doluluk Oranları - Günlük Veri



Şekil 27: İstanbul Barajları Doluluk Oranları

Kaynak: İstanbul Büyükşehir Belediyesi

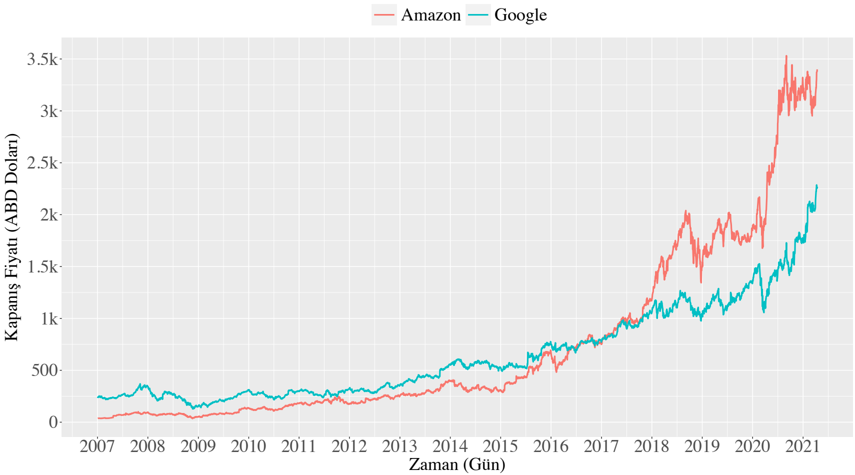
Hisse Senedi Fiyatları 1 (ABD Doları) - Günlük Veri



Şekil 28: Hisse Senedi Fiyatları 1 (ABD Doları)

Kaynak: Yahoo Finance

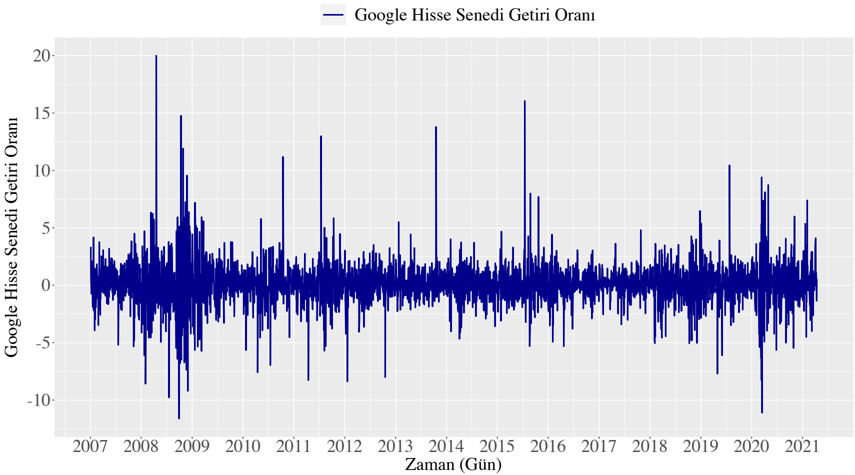
Hisse Senedi Fiyatları 2 (ABD Doları) - Günlük Veri



Şekil 29: Hisse Senedi Fiyatları 2 (ABD Doları)

Kaynak: Yahoo Finance

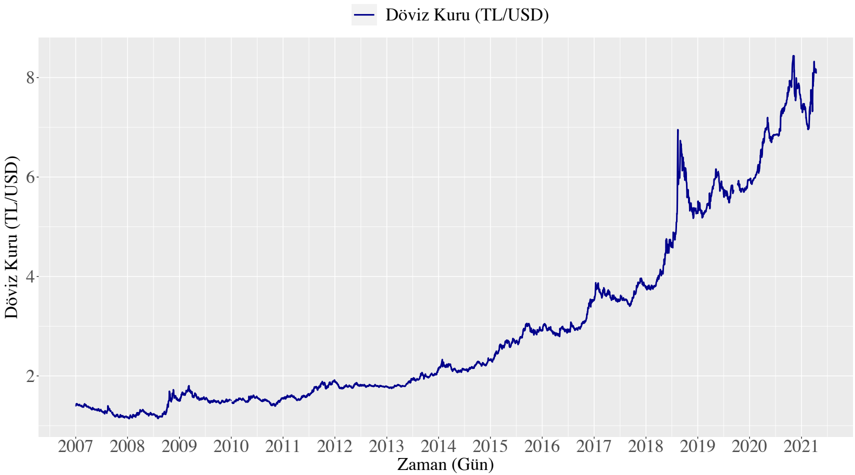
Google Hisse Senedi Getiri Oranı - Günlük Veri



Şekil 30: Google Hisse Senedi Getiri Oranı

Kaynak: Yahoo Finance

Döviz Kuru (TL/USD) - Günlük Veri



Şekil 31: Döviz Kuru (TL/USD)

Kaynak: Yahoo Finance

Matematiksel Transformasyonlar

- Birçok ekonometrik model tahmininde, hata terimlerinin normal dağıldığı (normallik varsayımı) varsayılır.
- Normallik varsayımı, güven aralıklarının oluşturulabilmesi ve hipotez testlerinin yapılabilmesi için istatistiksel bir çerçeve sağlar.
- Ekonometrik bir modeldeki bağımlı değişkeni normal dağılıma yakınsaması için **matematiksel transformasyon** ile dönüştürmek, hata terimlerini de normal dağılıma yakın (eğer zaten normal dağılım yapmıyorsa) hale getirebilir.
    - Yani, matematiksel transformasyonlarla beraber normallik varsayımı sağlanabilir.
    - Değişkenleri dönüştürmek modellerin tahmin gücünü artırabilir çünkü matematiksel transformasyonlar değişkenin varyansını stabilize edebilir ve değişken içindeki pür rassal süreç kısımlarını yok edebilir.
    - Bazı matematiksel transformasyonların kullanılması model sonuçlarının yorumlanmasında da kolaylık sağlar. Örnegin, logaritmik transformasyon.
    - Matematiksel transformasyonlar, zaman serisi haricindeki verilerde de kullanılabilir.
- Zaman serisi verilerinde, veri artan ve azalan bir varyasyon gösteriyorsa matematiksel transformasyon yararlı olabilir.
- Bu bölümde, zaman serileri analizinde kullanılan en yaygın matematiksel transformasyonlar olan **logaritmik transformasyon** ve **Box–Cox transformasyonu** ele alacağız.

Logaritmik Transformasyon

- Logaritmik transformasyon, hem normallik varsayımının sağlanmasında hem de model sonuçlarının yorumlanmasında kolaylık sağlar.
- Ekonometrik modellerde aksi belirtilmediği taktirde **doğal logaritmik transformasyon** (ln) kullanılır. Yani **e** (euler sayısı) tabanında logaritma alınır.
- Düzey formundaki $\{y_t\}_{t=1}^{n} = \{y_1, y_2, \ldots, y_t\}$ serisi için, logaritmik tranformasyon kullanılarak logaritmik formdaki $\{w_t\}_{t=1}^{n} = \{w_1, w_2, \ldots, w_t\}$ serisi elde edilebilir.
  - Burada y_t düzey formu, $\ln y_t = w_t$ ise logaritmik form olarak adlandırılır.
- Logaritmik yakınsama özelliği ile logaritmik formdaki veride oluşan değişimler, düzey formundaki veride oransal ya da yüzdesel değişim olarak yorumlanabilir.

$$\ln y_t - \ln y_{t-1} \approx (y_t - y_{t-1})/y_t \qquad \text{(logaritmik yakınsama)}$$

$$\Delta \ln y_t \approx \Delta y_t / y_t \qquad \text{(oransal değişim)}$$

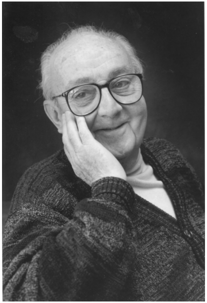
$$100 \cdot \Delta \ln y_t \approx 100 \cdot \Delta y_t / y_t \qquad \text{(100 ile çarpım)}$$

$$100 \cdot \Delta \ln y_t \approx \% \Delta y_t \qquad \text{(yüzdesel değişim)}$$

- Eğer düzey formundaki veride pozitif olmayan değerler varsa, logaritmik transformasyon uygulanamaz.

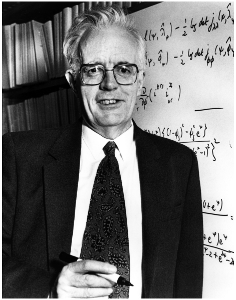
Box–Cox Transformasyonu

- Bazen, model sonuçlarının yorumlanmasında kolaylık sağlamasalarda başka transformasyonlar kullanılabilir.
  - Örneğin: verinin karesinin, küpünün ve kare-kökünün alınması.
  - Bunlar **üstel transformasyonlar** (power transformations) olarak adlandırılır çünkü genelde $w_t = y_t^p$ formunda yazılabilirler. Burada p üstel değeridir.
- Hem logaritmik transformasyon hem de üstel transformasyonu içeren kullanışlı bir transformasyon ailesi Box–Cox transformasyonudur (Box & Cox, 1964).



George E. P. Box (1919-2013)

Kaynak: Wikipedia



David Cox (1924 –)

Kaynak: Wikipedia

Box–Cox Transformasyonu

Box–Cox Transformasyonu

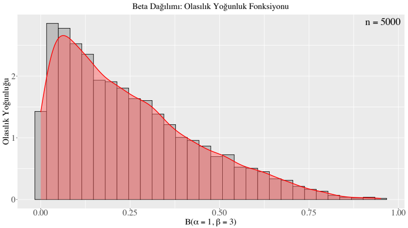
Düzey formundaki $\{y_t\}_{t=1}^{n} = \{y_1, y_2, \ldots, y_t\}$ serisi, optimal λ parametre değeri kullanılarak dönüştürülmüş formdaki $\{w_t\}_{t=1}^{n} = \{w_1, w_2, \ldots, w_t\}$ serisi aşağıdaki gibi elde edilebilir.

$$w_t = \begin{cases} \ln y_t & \text{eğer } \lambda = 0, \\ \frac{y_t^{\lambda} - 1}{\lambda} & \text{eğer } \lambda \neq 0. \end{cases}$$

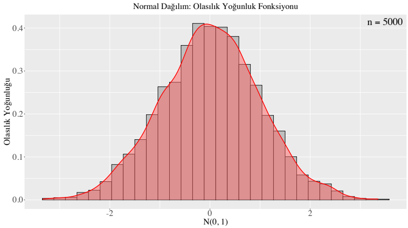
burada λ genellikle -5 ve 5 arasında değerler alan optimal üstel parametre değeridir.

- Box–Cox transformasyonu olası tüm λ değerlerini dikkate alır ve optimal λ değerini düzey formundaki veriyi normal dağılıma en yakın yapacak şekilde seçer.
  - Negatif düzey formundaki değerleri de kabul eden modifiye edilmiş Box–Cox transformasyonu için Bickel & Doksum (1981) çalışmasına bakınız.
- Eğer $\lambda = 1$ ise, $w_t = y_t - 1$ olur. Bu durumda, veri ölçek olarak aşağı kayar ancak verinin şekli değişmez. Yani, veri zaten normal dağılım yapmıştır ve Box–Cox transformasyonu gereksizdir. Diğer tüm λ değerleri için verinin şekli değişir.
- Şekil 32'de beta dağılımından rassal olarak çekilmiş verinin düzey formunda, logaritmik formda ve Box–Cox transformasyonu sonucunda yaptığı dağılım verilmiştir. Ayrıca, karşılaştırma amacıyla normal dağılım da gösterilmiştir.

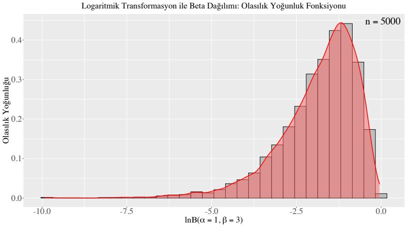
Beta Dağılımı: Matematiksel Transformasyonlar



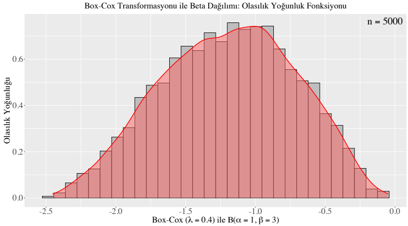
(a) Beta Dağılımı



(b) Normal Dağılımı



(c) Beta Dağılımı: Logaritmik Transformasyon

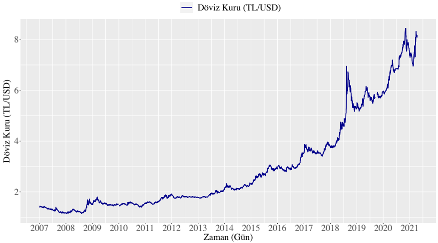


(d) Beta Dağılımı: Box–Cox Transformasyonu

Şekil 32: Beta Dağılımı: Matematiksel Transformasyonlar

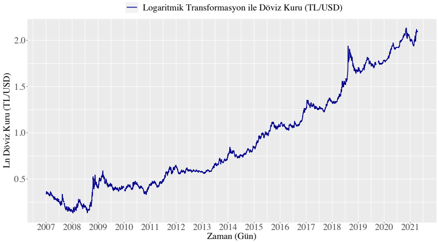
Döviz Kuru (TL/USD): Matematiksel Transformasyonlar

- Logaritmik formadaki veride 2017 sonrasındaki dalgalanmaların kısmen kaybolmuş ve 2007-2010 arasındaki dalgalanmalar ön plana çıkmış.
- Box–Cox transformasyonulu veride, 2017 sonrasındaki dalgalanmalar kaybolmuş ve 2007-2010 arasındaki dalgalanmalar çok daha belirgin.

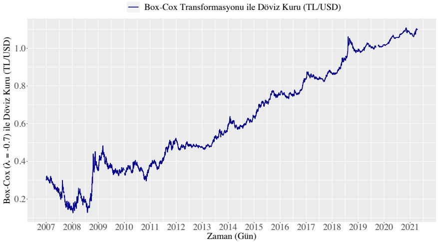


Döviz Kuru (TL/USD)

Kaynak: Yahoo Finance



Döviz Kuru (TL/USD): Logaritmik Transformasyon



Döviz Kuru (TL/USD): Box–Cox Transformasyonu

Bitcoin Fiyatları (BTC/USD) - Günlük Veri

Şekil 33: Bitcoin Fiyatları (BTC/USD)

Kaynak: Yahoo Finance

Zaman Serilerinin Dekompozizasyonu

- Trend, mevsimsellik ve döngüsellik zaman serisi örüntülerini daha önce Slayt 17'de incelemiştik.
- Zaman serileri verisi, çeşitli örüntüleri tekil olarak veya beraberce sergileyebilir.
- Bir zaman serisini, her biri temel bir örüntü kategorisini temsil eden bileşenlere bölmek zaman serisinin daha iyi anlaşılması açısından genellikle oldukça yararlıdır.
- Bir zaman serisi bileşenlerine ayırıldığında (dekompozizasyonu yapıldığında), genellikle trend ve dögüsellik tek bir **trend-döngüsellik** bileşeninde birleştirilir.
  - Basitleştirme amacıyla, bu bileşen genellikle **trend** olarak adlandırılır.
- Bu nedenle, bir zaman serisinin üç bileşenden oluştuğunu düşünebiliriz: **trend bileşeni**, **mevsimsellik bileşeni** ve **kalıntı bileşeni**.
  - Kalıntı bileşeni, zaman serisindeki trend ve mevsimsellik haricindeki herşeyi içerir.
  - Bazı zaman serilerinde (örneğin, en az günlük frekanstakilerde), farklı mevsim dönemlerine karşılık gelen birden fazla mevsimsellik bileşeni görülebilir.
- Bir zaman serisi analizinde dekompozizasyonu basitleştirmek amacıyla, öncelikle Slayt 15'de belirtilen ayarlamaların ve matematiksel transformasyonların yapılması yararlı olacaktır.

Zaman Serilerinin Dekompozizasyonu

- Uygulamada farklı dekompozizasyon methodları kullanılsa da, başlangıç noktası olarak her dekompozizasyon methodu iki farklı yöntem kullanır: toplamsal dekompozizasyon ve çarpımsal dekompozizasyon.

Toplamsal Dekompozizasyon (Additive Decomposition)

$$y_t = T_t + S_t + R_t$$

Çarpımsal Dekompozizasyon (Multiplicative Decomposition)

$$y_t = T_t \times S_t \times R_t$$

ya da

$$\ln y_t = \ln T_t + \ln S_t + \ln R_t$$

- Yukarıda, y_t zaman serisini, T_t trend bileşenini, S_t mevsimsellik bileşenini, R_t ise kalıntı bileşenini t döneminde ifade eder.

Zaman Serilerinin Dekompozizasyonu

- Trend bileşeni T_t, iş döngüleri dahil olmak üzere veri içindeki yavaş hareket eden orta ve uzun vadeli kısımları içerir.
- Mevsimsellik bileşeni S_t, genellikle her yıl veya her ay (her gün ve saat de olabilir) aynı zamanda tekrar eden dalgalanmaları içerir.
- Kalıntı bileşeni R_t, veriden trend ve mevsimsellik bileşenleri çıkarıldıktan sonra kalan kısmı ifade eder.
- Mevsimsel dalgalanmaların büyüklüğü veya trend etrafındaki dalgalanmalar, zaman serisinin düzeyine göre değişmiyorsa, toplamsal dekompozizasyon en uygun olan yöntemdir.
- Mevsimsellikteki varyasyon veya trend etrafındaki varyasyon, zaman serisinin seviyesiyle orantılı göründüğünde, çarpımsal dekompozizasyon daha uygundur.
    - Çarpımsal dekompozizasyon zaman serilerinde yaygın olarak kullanılanır.
    - Çarpımsal dekompozizasyon kullanmanın bir alternatifi, zaman serideki varyasyon zaman içinde sabit görünene kadar önce verileri dönüştürmek, ardından toplamsal dekompozizasyon kullanmaktır. Örnek: logaritmik transformasyon ve toplamsal dekompozizasyon kullanmak, çarpımsal dekompozizasyon kullanmaya eşdeğerdir.

Zaman Serilerinin Dekompozizasyonu

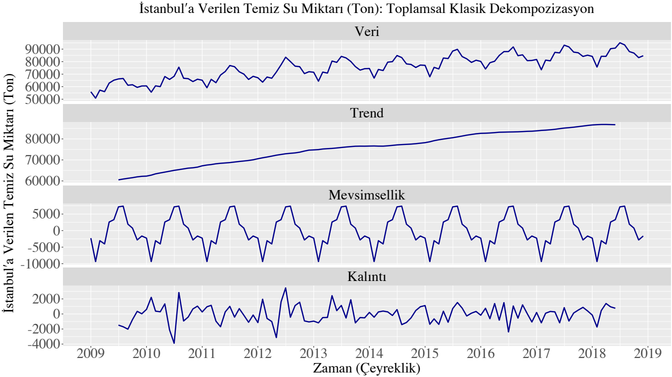
- Uygulamada, 4 farklı zaman serileri dekomompozizasyon methodu kullanılabilir.
- Klasik Dekompozizasyon
  - Nispeten basit bir yöntem olan klasik dekompozizasyon diğer zaman serisi dekompozizasyon methodlarının başlangıç noktasını oluşturur.
- X-11 Methodu
  - US Census Bureau tarafından geliştirilmiştir.
  - Detaylar için bakınız (Hyndman ve Athanasopoulos, 2018).
- SEATS Methodu
  - Bu method İspanya Merkez Bankası'nda geliştirilmiştir ve şu anda dünya çapında devlet kurumları tarafından yaygın olarak kullanılmaktadır.
  - "Seasonal Extraction in ARIMA Time Series" (SEATS) olarak da bilinir.
  - Detaylar için bakınız (Hyndman ve Athanasopoulos, 2018).
- STL Dekompozizasyonu
  - "Seasonal and Trend decomposition using Loess" (STL) olarak da bilinir.
  - X-11 ve SEATS methodları ile kıyaslandığında birçok avantajı vardır.
  - Detaylar için bakınız (Hyndman ve Athanasopoulos, 2018).

Klasik Dekompozizasyon

- Zaman serileri analizinde klasik dekompozizasyon hala yaygın olarak kullanılsa da, bazı problemleri nedeniyle genellikle diğer dekompozizasyon methodlarının kullanılması tavsiye edilmektedir.
- Klasik dekompozizasyon ile ilgili bazı problemler aşağıda özetlenmiştir.
  - Trend bileşeni T_t'nin tahmini, ilk birkaç ve son birkaç gözlem için mevcut değildir.
  - Trend bileşeni T_t'nin tahmini, verilerdeki hızlı artışları ve düşüşleri aşırı yumuşatma eğilimindedir.
  - Mevsimsellik bileşeni S_t'nin yıldan yıla sabit olduğu varsayılır. Pek çok zaman serisi için bu makul bir varsayımdır, fakat daha uzun seriler için değildir. Örneğin, klima kullanımı daha yaygın hale geldikçe elektrik talep modelleri zamanla değişmiştir.
  - Şoklardan kaynaklanan alışılmadık veri değerlerine karşı dirençli değildir.
- Diğer dekompozizasyon methodlarının klasik dekompozizasyonuna kıyasla birçok avantajı olsa da biz sadece sadece klasik dekompozizasyonu hesaplama detaylarına girmeden görsel olarak inceleyeceğiz.

Toplamsal Klasik Dekompozizasyon

- Daha önce Şekil 24'de gösterilen İstanbul'a Verilen Temiz Su Miktarı verisi, Şekil 34'de klasik dekompozizasyon ile bileşenlerine ayrılmış olarak gösterilmiştir.
- Mevsimsel dalgalanmaların büyüklüğü zaman serisinin düzeyine göre değişmediğinden, toplamsal dekompozizasyon kullanılmıştır.

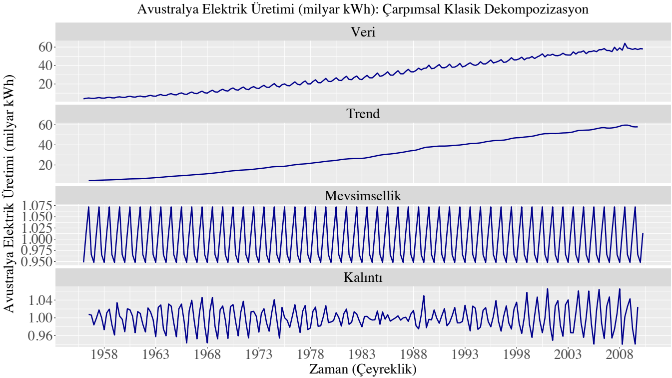


Şekil 34: İstanbul'a Verilen Temiz Su Miktarı: Toplamsal Klasik Dekompozizasyon

Kaynak: İstanbul Büyükşehir Belediyesi

Çarpımsal Klasik Dekompozizasyon

- Daha önce Şekil 20'de gösterilen Avustralya Elektrik Üretimi verisi, Şekil 35'de klasik dekompozizasyon ile bileşenlerine ayrılmış olarak gösterilmiştir.
- Mevsimsellikteki varyasyon zaman serisinin seviyesiyle orantılı göründüğünden, çarpımsal dekompozizasyon kullanılmıştır.



Şekil 35: Avustralya Elektrik Üretimi: Çarpımsal Klasik Dekompozizasyon

Kaynak: fpp2 R paketi

Hareketli Ortalama

- Şimdi, zaman serileri analizinde sıkça kullanılan ve klasik dekompozizasyonun ilk adımı olan hareketli ortalama (moving average) konusunu kısaca inceleyelim.

m. Dereceden Hareketli Ortalama (m-MA)

m. dereceden hareketli ortalama aşağıdaki gibi yazılabilir.

$$m\text{-MA}_t = \frac{1}{m}\sum_{j=-k}^{k} y_{t+j}, \quad m = 2k+1$$

burada hareketli ortalam merkezden alınmıştır. Yani, m-MA$_t$, t zamanındaki ve t zamanına k dönem kadar uzak zaman serilerinin ortalaması alınarak elde edilir.

- Örneğin, $k = 2$ ise $m = 5$'tir ve 5. dereceden hareketli ortalama (5-MA):

$$5\text{-MA}_t = \frac{1}{5}\sum_{j=-2}^{2} y_{t+j} = \frac{y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2}}{5}$$

- m-MA$_t$, klasik dekompozizasyondaki trend bileşeni T_t'yi ifade eder.
- Hareketli ortalama, verilerdeki rasgeleliğin bir kısmını ortadan kaldırarak düzgün bir trend bileşeni ortaya çıkarır.

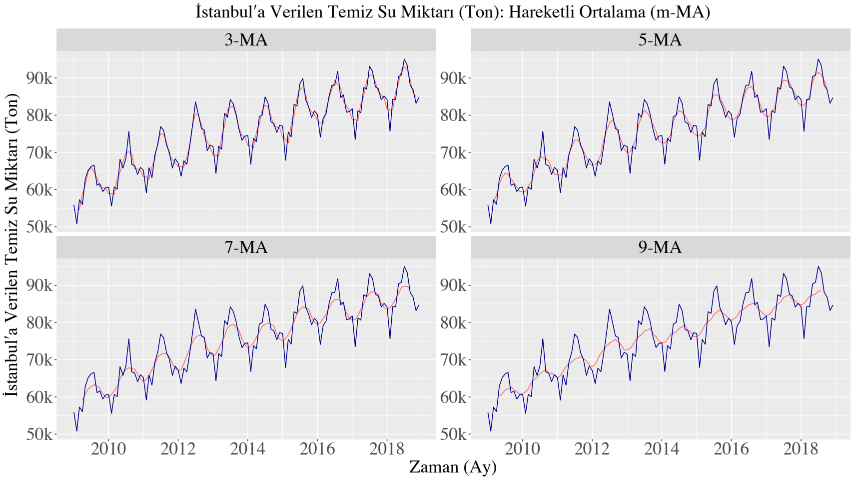
İstanbul'a Verilen Temiz Su Miktarı: Hareketli Ortalama

Tablo 4: İstanbul'a Verilen Temiz Su Miktarı: Hareketli Ortalama (m-MA)

Yıl	Ay	Temiz Su_t	3-MA	5-MA	7-MA	9-MA
2009	1	55926	NA	NA	NA	NA
2009	2	50838	54675	NA	NA	NA
2009	3	57261	54711	56587	NA	NA
2009	4	56034	58724	58444	59192	NA
2009	5	62878	61374	61516	60715	60231
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2018	8	93402	92182	90848	89689	88532
2018	9	88053	89440	89324	88883	NA
2018	10	86864	86043	87252	NA	NA
2018	11	83213	84935	NA	NA	NA
2018	12	84728	NA	NA	NA	NA

Notlar: İstanbul Büyükşehir Belediyesi datası kullanılmıştır.

İstanbul'a Verilen Temiz Su Miktarı: Hareketli Ortalama



Şekil 36: İstanbul'a Verilen Temiz Su Miktarı: Hareketli Ortalama (m-MA)

Kaynak: İstanbul Büyükşehir Belediyesi

Zaman Serilerinde Bağımlılığın Ölçülmesi

- Zaman serisi değişkenleri çoğunluklu kendi geçmiş değerlerine bağımlı olma eğilimindedir.
- Bu bağımlılığın özellikleri (yön, büyüklük, geçikme sayısı) zaman serisi verisini modellemede oldukça faydalıdır.
- Otokovaryans (autocovariance) ve otokorelasyon (autocorrelation) zaman serilerindeki bağımlılığı ölçmek için yaygın olarak kullanılmaktadır.
- Otokovaryans, otokorelasyon ve bağlantılı kavramları kısaca inceleyelim.

Örneklem Otokovaryansı

- Kovaryans ve korelasyon rassal X ve Y arasındaki doğrusal ilişkiyi ölçer.
- Zaman serileri analizinde, özellikle t dönemindeki değer ile önceki dönemlerdeki değerler (örneğin: $h = t - s$) arasındaki ilişkiyle ilgilenilir.
- Diğer bir deyişle otokovaryans ve otokorelasyonla ilgilenilir.

Örneklem Otokovaryansı

Otokovaryans, bir zaman serisi y_t'nin t dönemindeki değeri ve gecikmeli değerleri arasındaki doğrusal ilişkiyi ölçer. Örneklem otokovaryansı aşağıdaki gibi tanımlanır:

$$\widehat{Cov(y_t, y_{t-s})} = \hat{\gamma}_s = \frac{\sum_{t=s+1}^{n} (y_t - \bar{y}_{s+1:n})(y_{t-s} - \bar{y}_{1:n-s})}{n}$$

burada s gecikme değeridir; n ise gözlem sayısıdır; $\hat{\gamma}_s$ s. dereceden örneklem otokovaryansıdır; $\bar{y}_{s+1:n} = \frac{1}{n}\sum_{t=s+1}^{n} y_t$ ve $\bar{y}_{1:n-s} = \frac{1}{n}\sum_{t=1}^{n-s} y_t$ örneklem ortalamalarıdır.

Örneklem Otokovaryansı

- Örneğin, $s = 1$ ise birinci dereceden otokovaryans:

$$\widehat{Cov(y_t, y_{t-1})} = \hat{\gamma}_1 = \frac{\sum_{t=2}^{n} (y_t - \bar{y}_{2:n})(y_{t-1} - \bar{y}_{1:n-1})}{n}$$

- Örneğin, $s = 2$ ise birinci dereceden otokovaryans:

$$\widehat{Cov(y_t, y_{t-2})} = \hat{\gamma}_2 = \frac{\sum_{t=3}^{n} (y_t - \bar{y}_{3:n})(y_{t-2} - \bar{y}_{1:n-2})}{n}$$

- Örneğin, $s = 0$ ise otokovaryans formülü aslında varyansı verir:

$$\widehat{Cov(y_t, y_t)} = \hat{\gamma}_0 = \frac{\sum_{t=1}^{n} (y_t - \bar{y})(y_t - \bar{y})}{n} \quad \longrightarrow \quad \hat{\gamma}_0 = \widehat{Var(y_t)} = \frac{\sum_{t=1}^{n} (y_t - \bar{y})^2}{n}$$

Örneklem Otokorelasyonu

Örneklem Otokorelasyonu

Otokorelasyon, bir zaman serisi y_t'nin t dönemindeki değeri ve gecikmeli değerleri arasındaki doğrusal ilişkinin gücünü ölçer. Örneklem otokorelasyonu ise aşağıdaki gibi tanımlanır:

$$\hat{\rho}_s = \frac{\widehat{Cov(y_t, y_{t-s})}}{\sqrt{\widehat{Var(y_t)}}\sqrt{\widehat{Var(y_{t-s})}}} \longrightarrow \hat{\rho}_s = \frac{\hat{\gamma}_s}{\hat{\gamma}_0}$$

burada s gecikme değeridir; $\hat{\rho}_s$ s. dereceden örneklem otokorelasyonudur.

- Bu tanımın, istatistik derslerinde öğrendiğiniz korelasyon katsayısı tanımına benzer olduğunu unutmayın.
- Örneklem otokorelasyonu birbirinden s dönem kadar uzak olan y değerleri arasındaki doğrusal ilişkinin gücünü ölçer.
- $\hat{\rho}_1$, birinci dereceden örneklem otokorelasyonudur; $\hat{\rho}_2$, ikinci dereceden örneklem otokorelasyonudur; $\ldots$; $\hat{\rho}_s$, s. dereceden örneklem otokorelasyonudur.
- $\hat{\rho}_0$ her zaman 1'e eşittir.

Örneklem Otokorelasyon Fonksiyonu - Korelogram

- Önceden belirlenmiş bir maksimum gecikme değeri S'e kadar ($s = 1, 2, \ldots, S$) tüm örneklem otokorelasyonlarını hesaplayarak, bir zaman serisinin bağımlılık yapısını grafiğe dökebiliriz.
- Bu grafik zaman serisinin bağımlılık yapısının incelenmesinde kullanılır ve **örneklem otokorelasyon fonksiyonu** ya da **korelogram** olarak bilinir.
  - Örneklem otokorelasyon fonksiyonu İngilizce'de **sample autocorrelation function (SACF)** ya da **correlogram** olarak bilinir.
- Büyük örneklemde, Merkezi Limit Teoremi kullanılarak aşağıdaki durum gösterilebilir:

$$\hat{\rho}_s \sim N\left(0, \frac{1}{n}\right) \quad \longrightarrow \quad \sqrt{n}\hat{\rho}_s \sim N(0, 1)$$

- ρ_s için %95 güven aralığı şu şekilde bulunabilir:

$$\hat{\rho}_s - z_{\alpha/2} \cdot \frac{1}{\sqrt{n}} \leq \rho_s \leq \hat{\rho}_s + z_{\alpha/2} \cdot \frac{1}{\sqrt{n}}$$

$$\hat{\rho}_s - 1.96 \cdot \frac{1}{\sqrt{n}} \leq \rho_s \leq \hat{\rho}_s + 1.96 \cdot \frac{1}{\sqrt{n}}$$

Pür Rassal Süreç

- Şimdi, zaman serileri analizinde sıkça kullanılan ve stokastik bir süreç olan olarak pür rassal sürecin örneklem otokorelasyon fonksiyonunu inceleyelim.

Pür Rassal Süreç

Stokastik süreç $\{\epsilon_t : t = 1, 2, \ldots, n\}$ aşağıdaki koşulları sağlıyorsa ϵ_t'ye **pür rassal süreç** (white noise process) adı verilir.

$$E(\epsilon_t) = 0$$

$$\gamma_0 = Var(\epsilon_t) = \sigma_\epsilon^2$$

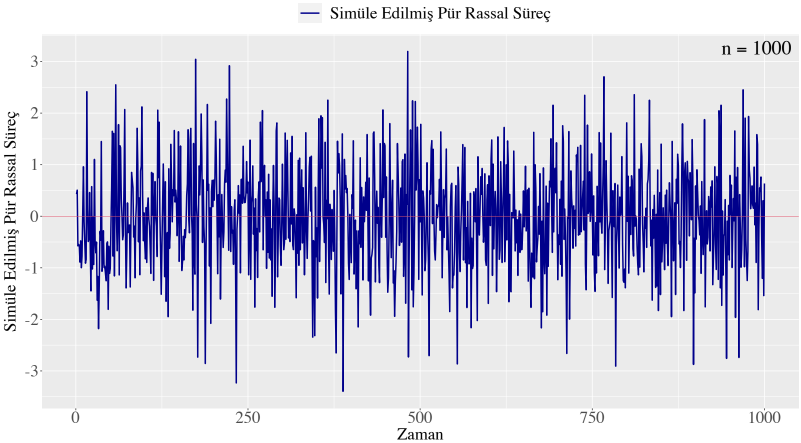
$$\gamma_{t-h} = \gamma_s = Cov(\epsilon_t, \epsilon_h) = 0, \quad t \neq h$$

burada s gecikme değeridir; ϵ_h ise pür rassal sürecin $h = t - s$ dönemindeki değeridir. Bu süreci kısaca $\epsilon_t \sim wn(0, \sigma_\epsilon^2)$ ile göstereceğiz.

- Pür rassal sürecin ortalaması sıfırdır ve varyansı sabit bir sayıdır.
- Pür rassal sürecin ortalaması ve varyansı zamana bağımlı değildir. Yani ortalama ve varyans t'ye göre değişmez.
- Ayrıca, herhangi bir zaman indeksi t ve h ($t \neq h$) için otokovaryans ve dolayısıyla otokorelasyon sıfıra eşit olduğundan, cari dönemdeki ϵ_t değeri geçmiş dönemlerdeki ϵ_h değerleriyle doğrusal olarak ilişkili değildir.

Pür Rassal Süreç

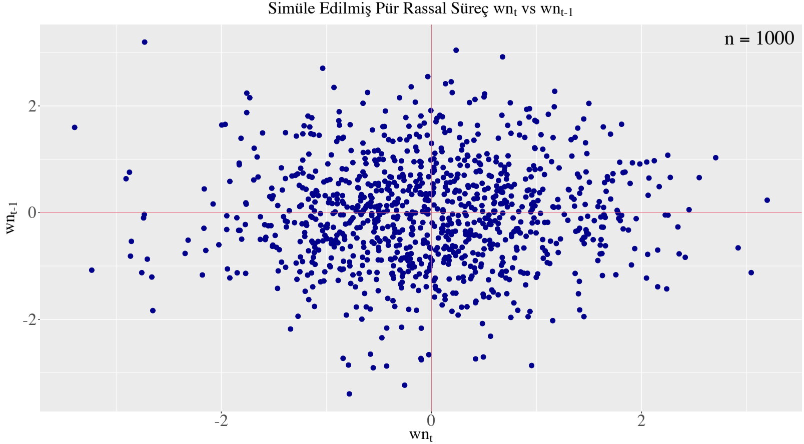
- Aşağıda simüle edilmiş bir pür rassal süreç örneği verilmiştir. Pür rassal süreçte trend, mevsimsellik ve döngüselliğin olmadığına dikkat edin!



Şekil 37: Simüle Edilmiş Pür Rassal Süreç

Pür Rassal Süreç

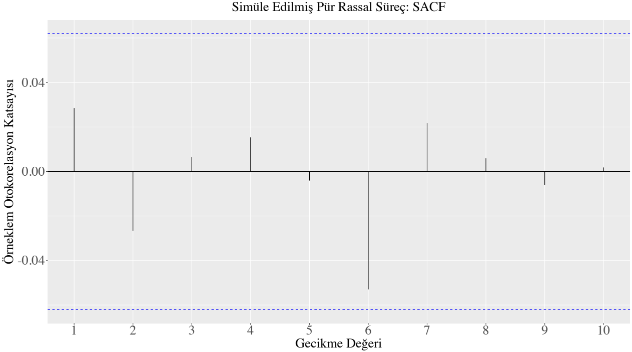
- Aşağıda simüle edilmiş bir pür rassal sürecin t dönemindeki değeri $\boldsymbol{wn_t}$ ve $t-1$ dönemindeki gecikmeli değeri $\boldsymbol{wn_{t-1}}$ arasındaki ilişki grafikle gösterilmiştir.



Şekil 38: Simüle Edilmiş Pür Rassal Süreç

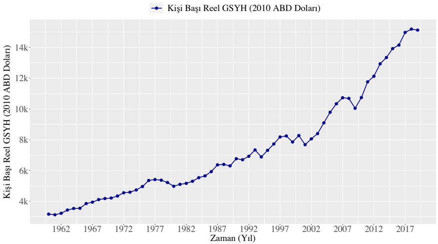
Pür Rassal Süreç: SACF

- Aşağıda simüle edilmiş bir pür rassal sürecin örneklem otokorelasyon fonksiyonu (SACF) %95 güven aralığı (mavi kesikli çizgi) kullanılarak grafikle gösterilmiştir.
- $\boldsymbol{wn_t}$ ve $\boldsymbol{wn_{t-s}}$ arasındaki SACF değerlerini gösteren dikey çizgiler güven aralığı sınırları içinde kaldığından bulunan SACF değerleri istatistiki olarak anlamsızdır.
- Yani, $\boldsymbol{wn}$ pür rassal sürecinde otokorelasyon yok diyebiliriz.

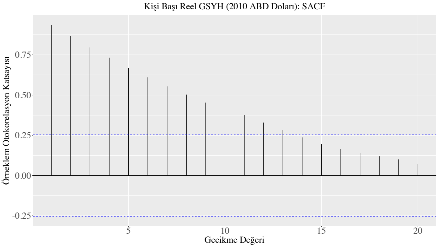


Şekil 39: Simüle Edilmiş Pür Rassal Süreç: SACF

Kişi Başı Reel GSYH (ABD Doları) - Türkiye: SACF



Kişi Başı Reel GSYH (ABD Doları) - Türkiye

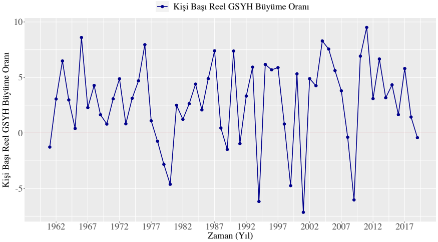


Kişi Başı Reel GSYH (ABD Doları) - Türkiye: SACF

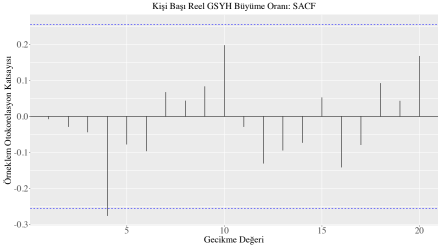
Kaynak: World Bank

- $\mathbf{GSYH}_t$ ve $\mathbf{GSYH}_{t-s}$ arasındaki SACF değerlerini gösteren dikey çizgiler güven aralığı sınırları dışında olduğundan (belli bir gecikmeye kadar istatistiki olarak anlamlı olduğundan), **GSYH** zaman serisinde otokorelasyon var diyebiliriz.
- Zaman serisi yukarıdaki gibi bir trende sahip olduğunda, düşük gecikmelerdeki SACF değerleri büyük ve pozitif olma eğilimindedir.
- Ancak, gecikme değeri (mesafe) arttıkça SACF değerleri azalır.

Kişi Başı Reel GSYH Büyüme Oranı - Türkiye: SACF



Kişi Başı Reel GSYH Büyüme Oranı - Türkiye



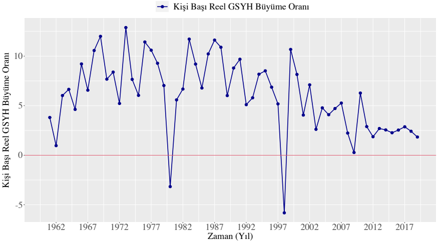
Kişi Başı Reel GSYH Büyüme Oranı - Türkiye: SACF

Kaynak: World Bank

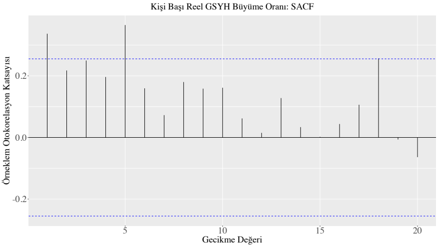
- $\mathbf{GR}_t$ ve $\mathbf{GR}_{t-s}$ arasındaki SACF değerlerini gösteren dikey çizgiler güven aralığı sınırları içinde kaldığından **GR** zaman serisinde otokorelasyon yok diyebiliriz.
- Bu durum kısacası **GSYH** zaman serisinin ilk farkları alındığında, yani **GR** zaman serisi hesaplandığında, otokorelasyonun ortadan kalktığını belirtir.
- Büyüme oranının nasıl hesaplandığını hatırlayın.

$$\mathrm{GR}_t = \frac{\mathrm{GSYH}_t - \mathrm{GSYH}_{t-1}}{\mathrm{GSYH}_{t-1}} * 100$$

Kişi Başı Reel GSYH Büyüme Oranı - Kore: SACF



Kişi Başı Reel GSYH Büyüme Oranı - Kore

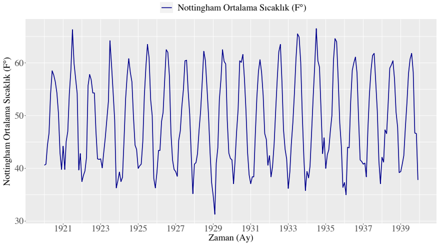


Kişi Başı Reel GSYH Büyüme Oranı - Kore: SACF

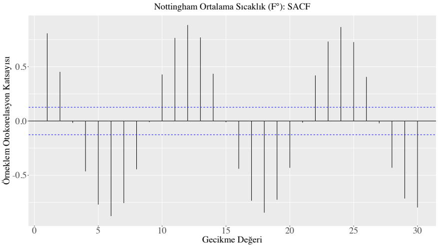
Kaynak: World Bank

- $\mathbf{GR}_t$ ve $\mathbf{GR}_{t-s}$ arasındaki SACF değerlerini gösteren dikey çizgiler güven aralığı sınırları dışında olduğundan (belli bir gecikmeye kadar istatistiki olarak anlamlı olduğundan), **GR** zaman serisinde otokorelasyon var diyebiliriz.
- Yukarıda Kore için verilen SACF grafiği daha önce gösterilen Türkiye grafiği ile karşılaştırıldığında, Kore için 1. ve 5. gecikme durumunda otokorelasyonun pozitif ve istatistiki olarak anlamlı olduğu görülüyor.

Nottingham Ortalama Sıcaklık - Aylık Veri: SACF



Nottingham Ortalama Sıcaklık

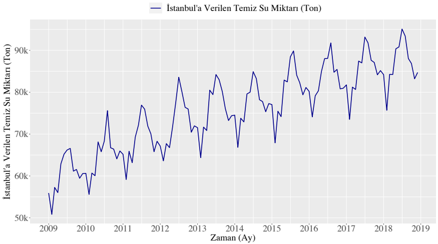


Nottingham Ortalama Sıcaklık: SACF

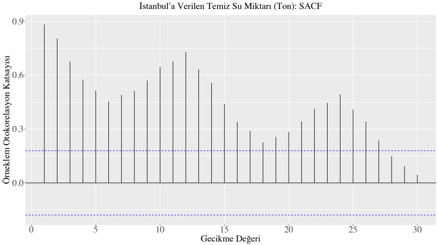
Kaynak: datasets R paketi

- $\mathbf{F}_t$ ve $\mathbf{F}_{t-s}$ arasındaki SACF değerlerini gösteren dikey çizgiler güven aralığı sınırları dışında olduğundan (belli gecikmelerde istatistiki olarak anlamlı olduğundan), **F** zaman serisinde otokorelasyon var diyebiliriz.
- Yukarıdaki gibi aylık zaman serilerinde yalnızca mevsimsellik olduğunda, fakat trend yokken:
  - SACF, 12. ve 24. mevsimsel gecikmelerde en yüksek değerleri alır.
  - SACF, 6. ve 18. mevsimsel gecikmelerde ise en düşük değerleri alır.

İstanbul'a Verilen Temiz Su Miktarı - Aylık Veri: SACF



İstanbul'a Verilen Temiz Su Miktarı



İstanbul'a Verilen Temiz Su Miktarı: SACF

Kaynak: İstanbul Büyükşehir Belediyesi

- $\mathbf{S}_t$ ve $\mathbf{S}_{t-s}$ arasındaki SACF değerlerini gösteren dikey çizgiler güven aralığı sınırları dışında olduğundan (belli gecikmelerde istatistiki olarak anlamlı olduğundan), **S** zaman serisinde otokorelasyon var diyebiliriz.
- Yukarıdaki gibi aylık zaman serilerinde mevsimsellik ve trend beraber olduğunda:
  - SACF, 12. ve 24. mevsimsel gecikmelerde en yüksek değerleri alır.
  - SACF, 6. ve 18. mevsimsel gecikmelerde ise en düşük değerleri alır.
  - SACF, artan trend nedeniyle gecikme değeri arttıkça küçülür.

Kaynaklar

Gujarati, D.N. (2009). *Basic Econometrics*. Tata McGraw-Hill Education.
Güriş, S. (2005). *Ekonometri: Temel Kavramlar*. Der Yayınevi.
Hyndman, R.J. ve G. Athanasopoulos (2018). *Forecasting: Principles and Practice*. OTexts.
Stock, J.H. ve M.W. Watson (2015). *Introduction to Econometrics*.
Wooldridge, J.M. (2016). *Introductory Econometrics: A Modern Approach*. Nelson Education.

