

# Sorgulayıcı Veri Çözümlemesine Bir Örnek: İllere Göre İşsizlik Oranları, 2008-2013

Ümit ŞENESEN\*

Bir şeyi ne kadar İYİ YAPTIĞINI ölçmeye kalkışmadan  
önce ne YAPABİLECEĞİNİ anlamak önemlidir.  
John. W. Tukey (1977: s.v)

**Öz:** Bu yazıda sorgulayıcı veri çözümleme (SVC) yöntemlerinin kısa bir tanıtımından sonra, her biri çok yalın, anlayıp anlatması çok kolay ama fazla hassas olmayan bazı araçlar yardımıyla illerimizin işsizlik oranları ele alınarak incelenecektir. Burada elde edilmeye çalışılacak birtakım ipuçları, daha sonra başka araştırmacılar tarafından çıkarsama yöntemleriyle daha ayrıntılı, daha hassas biçimde incelenerek konu aydınlığa kavuşturulabilecektir. Sorgulayıcı veri çözümlemesinin ilk çözümleme adımı olduğu, çıkarsamalı çözümlemenin de bunun tamamlayıcısı olduğu belirtilmelidir. Bu yazıda önce SVC yöntemlerinin nitelikleri, önemli öğeleri açıklanacak, sonra *işlenmiş çizelge, dal-yaprak, beşli özet, kutu-çizim* gibi birkaç sorgulayıcı veri çözümleme aracının kullanılışı gösterilecektir<sup>1</sup>. Bu araçlar yardımıyla illere göre işsizlik oranları incelenerek aydınlatılması gerekli noktalar ortaya konacaktır.

**Anahtar kelimeler:** sorgulayıcı veri çözümlemesi (SVC), illere göre işsizlik, Türkiye

**An example of exploratory data analysis: Provincial unemployment rates of Turkey, 2008-2013**

**Abstract:** This paper starts with a brief introduction to the exploratory data analysis methods. Then it examines the provincial unemployment rates of Turkey for 2008-2013 period by using these methods on official TURKSTAT data. A short explanation of a few very simple methods, such as coded-table, stem-and-leaf, five-point-summary and box-plot, is given and they are implemented on the data. It was found that in general the provincial unemployment rates

---

\* Prof, Dr.

<sup>1</sup> Tukey (1977) ile Velleman, Hoaglin (1980) bu konuda temel kaynak olarak gösterilebilir. Türkçe kısa bir tanıtım için bkz. Şenesen (2015); daha ayrıntılı bilgi için bkz. Şenesen (2004).

have increased in 2009 and decreased later on until 2013. But in 2012 and 2013 a contrary movement has been developed for some provinces. They are mostly Southeast Anatolia provinces and their unemployment rates strikingly high for the last two years of the study period. This finding is exactly what was expected from an exploratory data analysis: to get a few clues about some phenomenon which should be investigated further by making use of inferential statistical methods. It is leaved to the readers (or other investigators) to think and bring out the reasons that caused this unusual and unexpected increase of unemployment rates for those provinces named in the study.

**Keywords:** exploratory data analysis (EDA), provincial unemployment, Turkey

## Sorgulayıcı Veri Çözümlemesi

Bilindiği gibi herhangi bir değişkenin, örneğimizde Türkiye illerinin belli bir yıldaki işsizlik oranlarının, küçükten büyüğe doğru dizilmesi bir dağılım oluşturur. Bu dağılım incelenirken *düzye*, *yayılım*, *biçim*, eğer varsa *dışadüşen* (*outlier*) gözlemler gibi birkaç niteliği dikkate alınır. Düzye, dağılımı oluşturan gözlem değerlerinin genelde toplu olarak bulunduğu orta bölgeyi temsil eden bir göstergedir. En yaygın kullanılan düzye göstergesi *aritmetik ortalamadır* ama dağılım çarpıklaştıkça işlevini yerine getirmesi zorlaşır, hatta olanaksız bile olabilir. Bu bakımdan *ortanca* daha berk, daha dayanıklı bir göstergedir; bu nedenle SVC yöntemlerinde sıkça kullanılır. Yayılım için de en sık kullanılan gösterge *varyans* ya da *standart sapmadır* ama aritmetik ortalamaya bağlı olduklarından çarpık dağılımlarda bunlar da gözden düşer. Onların yerine, daha berk olan üst ve alt *dördebölenlerin* farkı ( $D_{ii} - D_a$ ) olan *dördebölenler aralığı* ( $DA$ )<sup>2</sup> ya da en üst ve en alt değerlerin farkını ( $E_{ii} - E_a$ ) gösteren *Yayılım aralığı* ( $YA$ ) geçer. Biçim konusu SVC’de birkaç bakımdan ele alınır: dağılımın çarpıklık durumu, tepe sayısı, sivrilik-basıklık. Son olarak, eğer varsa Dışadüşen(ler), değerleri öbür gözlemlerden açık ara büyük ya da küçük gözlemlerdir. Bunların incelenmesi bize dağılıma ilişkin epeyce bir fikir verir. Eğer birden çok veri kümesi, örneğimizde illere göre işsizlik oranlarının çeşitli yıllardaki dağılımları, karşılaştırılacaksa bu nitelikleri göz önüne almak gerçekten işe yarayabilir. Bu niteliklerine göre incelenen veri kümesi ya da kümeleri, eğer belirgin biçimde çarpık ise bunların çarpıklıklarının düzeltilmesi gerekebilir. Bu da verilerin dönüştürülmesiyle yapılır. Kare, karekök, logaritma ya da tersini alma ( $1/\text{gözlem değeri}$ ) gibi işlemlerden biri bu işi görebilir. Veri kümesini çarpıklıktan kurtarıp

<sup>2</sup> Dördebölen, İngilizce *quartile*, dördebölenler aralığı ( $DA$ ) *intraquartile range* ( $IQR$ ) karşılığıdır.

normal dağılıma yaklaştırma,  $t$ ,  $F$  sınamaları gibi pek çok çıkarsama yöntemi için gerekli normal dağılmış anakütle varsayımını sağlamaya yöneliktir. Eğer bu işlemler birden çok veri kümesi için yapılıyorsa genellikle bir başka şaşırtıcı sonuç daha elde edilebilir: Yayılımları hayli farklı kümeler, dönüştürme sonucunda yaklaşık eşit yayılımlara sahip olur. Söylemeye bile gerek yok ki bu son özellik de aynı çıkarsama yöntemlerince kullanılan varyans eşitliği varsayımını sağlamaktadır.

## İşlenmiş Çizelge

Şimdi TÜİK'ten alınmış illere göre işsizlik oranlarının bu yöntemlerle incelenmesine geçelim. Bu veriler 2008-2013 arasındaki durumu yansıtmaktadır. 2013 sonrası her il göre veri yayımlanmamaktadır. TÜİK bu verileri, il kodları (plaka numaraları) sırasına göre vermektedir. Ama böyle bir gösterim, olup biteni anlamayı kolaylaştırmamaktadır. Verilerde saklı olan yapıyı anlamak için önce bu verileri işlenmiş çizelgeye çevirelim. Verilerde saklı yapı derken, her yıldaki dağılımın niteliklerini anlamaya, sonra bunları birbirleriyle karşılaştırmaya yarayan ipuçlarından söz edeceğiz. Bu amaçla her bir il için altı yıllık işsizlik oranı dizisinin ortanca düzeyini bularak illeri bu ortanca değerine göre en yüksekte en düşüğe doğru sıralayalım. Sonra da her yılı için birer ortanca değeri hesaplayıp bunu da çizelgenin altına ekleyelim. Ayrıca her işsizlik oranını da yüksekliğine göre farklı boyda sayılarla gösterelim. Her sayının büyüklüğünü çizelgenin yanına eklenen açıklama kutularında gösterelim. Sözelimi %25 ya da daha yüksek işsizlikleri 17, %22,5 ile 25 arasındakileri 16, ..., %5 ile 7,5 arasındakileri 9, %5'in altındakileri 8 boyutunda belirtelim. Böylece Çizelge 1 (a ve devamı b) elde edilmiş olur.

Çizelge 1b'nin en alt satırındaki ortancalara bakarak işsizliğin 2008-2013 arasındaki seyrini hemen görebiliriz. Bu verilere göre 2009'da işsizlik 2 puandan fazla artarak %13 düzeyine yaklaşmış, sonraki yıllarda hep azalarak 2013'te %7,5'in altına inmiştir.<sup>3</sup> Bunu Türkiye'nin genel işsizlik seyri olarak yorumlayabiliriz. Demek ki işsizlik yurdumuzda 2009'dan 2013'e kadar düşüş eğilimine girmiştir.

İllerin ortancalarına bakıldığında en yüksek oranlı dokuz ilden yedisi Güneydoğu Anadolu illeridir (Şırnak, Siirt, Adıyaman, Hakkâri Mardin, Diyarbakır, Batman); üçüncü sıradaki Adana ile altıncı sıradaki İzmir de bu bölgeden önemli oranda göç alan illerdir. Bu saptama gerek iktisatçıların gerek toplumbilimcilerin ilgisini çekebilecek bir ipucu olarak görülebilir.

<sup>3</sup> Burada bütün illerin işsizlik oranlarının toplam il sayılarına bölünmesiyle bulunacak aritmetik ortalamanın bir anlam taşımadığını okuyucular görmüş olmalıdır. Çünkü her il için ayrı ayrı hesaplanan işsizlik oranlarında bölme işleminin paydasında yer alan istihdam sayıları aynı değildir. Bu sayılar bilinirse tartılı aritmetik ortalamayla Türkiye için genel işsizlik oranı bulunabilir. TÜİK'in yayımladığı oranlar bunlardır.

Çizelgenin geri kalan bölümünde yer alan her ilin her yıldaki işsizlik oranının zaman içinde nasıl bir seyir izlediği de farklı gösterim büyüklükleriyle iyice anlaşılır kılınmaya çalışılmıştır. Sözgelimi ilk sıradaki 73 kodlu Şırnak, Türkiye'nin genel seyrine uyararak 200'den sonra peş peşe iki yıl 5-6 puanlık düşüşler göstermiş, ama sonraki yıllarda ülke genelinin tersine bir yön çizerek 2012'de %19'u, 2013'te %20'yi aşmıştır. Benzer yönde gelişmeler Siirt, Hakkari, Mardin, Diyarbakır, Batman için de söz konusudur. Adana ve İzmir de de 2013'te bir miktar işsizlik artışı dikkat çekmektedir. Benzer duruma öbür illerimizde pek rastlanmamaktadır. İncelenmesi gereken bir ipucu da budur. Ele alınan dönemin son iki yılında ne olmuştur da bu ilerimizde böylesi bir işsizlik artışı ortaya çıkmıştır?

## Dal-Yaprak, Beşli Özet, Kutu-Çizim

Her yıl için ayrı bir dal-yaprak düzenlenerek bunlar hem ayrı ayrı incelenebilir, hem de karşılaştırılabilir. Bu sayede gözlem değerlerini yoğunlaşma bölgeleri, yayılma aralıkları kolayca görülebilir; çarpıklıkları, sivrilikleri, tepe sayıları hakkında fikir edinilebilir; bazı gözlemler öbürlerinden önemli ölçüde küçük ya da büyük ise bunlar hemen fark edilebilir.

Önce bir dal-yaprak çiziminin nasıl düzenlendiğine bakalım. Bu gösterimde her işsizlik oranının %10'luk basamağını, %1'lik basamağından ayıracağız. İlkine *dal*, ikincisine *yaprak* diyeceğiz. % 10'dan düşük işsizlik oranlarının dal değeri 0 (sıfır) olacaktır. Her satıra sadece iki yaprak değeri koyma kararını vermiş olalım.<sup>4</sup> Öyleyse bir satırdaki yaprak değerleri 0 ile 1, sonraki satırda 2, 3, sonrakinde 4,5, ..., birinde de 8,9 yer alacaktır. İlk yılı örnek alalım. Çizelge 1'in 2003 yılı sütunundaki işsizlik oranlarına bakmamız gerekir. Çizelge 1b'nin en alt satırındaki Ardahan ve Bayburt illerinin %3.7 ve %4.7 olan işsizlik oranlarındaki 3 ve 4'ü yaprak olarak alacak, 7'yi atacağız. Son basamak kaç olursa olsun, (9 bile olsa) yine de bunu yukarı doğru yuvarlamada kullanmayacağız. SVCÇ'nin hassaslık kaybı özelliği işte buradan kaynaklanmaktadır. Bu iki il için dal 0 olacaktır. Çizelge 2'de 2008 yılı dal-yaprak gösteriminde bu iki il en üsttedir. %3 küsur işsizlik oranı ilk satırda, %4 küsur olanı ikinci satırdadır. Çizelge 1a'daki ilk ilin işsizlik oranı %22,1'dir Bunu %20 ve %2 diye ayırıştırıp dal 2, yaprak 2 diye gösterelim, son basamağı da kesip atalım. Yani %22,1'i, %22 küsur diye okuyacağız. Bu ilin işsizlik değeri Çizelge 2'de 2008 gösteriminde en alttadır.

Dal ve yaprak satırlarının her birinde en soldaki sütunda birikimli sıklıklar yer almaktadır. 2008 yılında ilk satırda %3 küsur işsizlik oranıyla bir il olduğundan en solda 1 görülmekte, bir alttaki satırda %4 ve %5 küsurluk işsizliği olan 9 il bulunduğundan en sola (1+ 9 =) 10 sayısı konmuştur. Sonraki satırlarda görülen 28

<sup>4</sup> Bu kararın nasıl alınabileceğini görmek için bkz. Tukey (1977:s.11-16) ya da Şenesen (2004: s.271-280).

ve 40 sayıları da böyle elde edilmiştir. Bir alttaki satırda durum farklıdır. Burada da aynı kural kullanılmış olsaydı en solda 57 sayısını görmemiz gerekirdi ama bu sayı, ortanca ilin sıralamadaki yerinden yani  $(81+1)/2 = 41$ 'den büyük olduğundan birikimli il sayısı bu satıra konmamış, onun yerine sadece %10 ve %11 küsur işsizlik oranına sahip 17 il olduğunu gösteren sayı ayrıç içinde (17) yer almıştır. 2008 yılı gösteriminin en alt satırına gidip %22 küsurluk işsizlik oranına sahip ilin hizasına en sola 1 konmuş, onun üstündeki satırlarda da aynı birikimli il sayısı kuralına uyarak uygun sayılar yerleştirilmiştir.

Yan yana iki yılın gösteriminde aynı dal-yaprak değerlerinin aynı hizaya getirilmesi karşılaştırmayı kolaylaştırır. 2008 ile 2009 gösterimlerinde buna uyulmuştur.

Önce sadece 2008 yılının dal-yaprak gösterimini inceleyelim, sonra öbür yıllarla karşılaştırmalar yapalım. Hemen görebildiğimiz gibi %10'dan düşük 40 il olduğuna göre tam ortaya düşen 42. ilde işsizlik %10 küsurdur. Yani ortanca işsizlik oranı bu düzeydedir. En üst oran %22 ile en alt oran %3 arasındaki yayılım aralığı da % 19 dolayındadır. 41. il ortanca olduğuna göre alt ve üst dördebölenler de alttan ve üstten  $(41+1)/2 = 21$ . illerdir. Demek ki alt dördebölen %7, üst dördebölen %12 küsur işsizlik oranlarıdır. Yani illerimizin yarısında işsizlik oranları %7 ile %12 arasındadır. Başka deyişle 2008'de illerimizin dörtte biri %7 ya da daha az, dörtte biri de %12 ya da daha çok işsizlik oranı sergilemektedir. Bu iki dördebölen arasındaki fark ( $DA$ ) %5 kadardır. Dağılım tam normal olmasa da çarpıklık görece azdır. Bu yılda işsizlik oranı %10'dan düşük iller ile %10 ve daha yüksek iller hafifçe birbirinden ayrı iki küme izlenimi vermektedir ama bu kesin bir gözlem değildir.

Önceden belirttiğimiz gibi ondalık işaretinin sağındaki değerleri attığımızdan bu saptamalar biraz kabadır ama ilk bakışta görülemeyen yapıyı epeyce aydınlatmaktadır. 2008 yılının dal-yaprak gösterimini öbür yıllarla karşılaştırdığımızda 2009'da işsizlik oranlarının biraz yükseldiğini ( $Ortc \sim \%12$ ), sonraki yıllarda düştüğünü görürüz. Ortanca değeri 2010'da %10, 2011'de %8, 2012 ve 2013'te %7 dolayına inmiştir. Bu değerlerin Çizelge 1b'nin son satırındaki biraz daha hassas ortancalarla uyumlu olduğu açıktır.

En alttaki işsizlik oranı ( $E_a$ ) 2009'da artık %3 değil % 4 düzeyindedir ama en üstteki oran ( $E_{\#}$ ) %22'den %26 küsura çıkmıştır. Böylece 2009'da yayılım aralığının ( $YA = 26 - 4 = 22$ ) genişleyip izleyen iki yılda  $(19 - 4 = 15$  ve  $16 - 4 = 12$  daraldığını ama son iki yılda  $(25 - 4 = 21$  ve  $23 - 4 = 19)$  yeniden genişlediği anlaşılmaktadır. Bu olurken en düşük işsizlik oranının %4 küsurda kaldığı, en yüksekini ise yıldan yıla değiştiği görülmektedir.

Benzer durum dördebölenler aralığında da yaşanmıştır.  $DA$  değerleri 2009'da  $(17 - 10 =)$  %7'ye yükselmiş, sonra sırasıyla  $(12 - 6 =)$  %6'ya,  $(10 - 6 =)$  %4'e,  $(9 - 5 =)$  %4'e düşmüş ama 2013'te yeniden  $(11 - 6 =)$  %5'e çıkmıştır.

Kısaca belirtmek gerekirse illerimiz arasında işsizlik oranları hem ülke bütününde hem de orta düzeydeki iller arasında birbirinden uzaklaşmaya başlamıştır.

İllerdeki işsizlik oranlarının dağılımlarının çarpıklığı 2009'dan sonraki iki yılda epeyce azalmışken son iki yılda çok belirgin biçimde artmış, büyük değerlere doğru uzayan kuyruklar iyice ortaya çıkmıştır. İllerin, %10'dan az ve %10 ya da daha yüksek işsizlik oranlarına sahip iki farklı öbeğe ayrılma durumları, pek göze batmayan biçimde de olsa, 2008-2011 arasında da sürmüş, çok sayıda ilimizde işsizlik oranları az çok sabit kalırken bazı illerimizde olağandışı artışlar belirmiştir. Son iki yılda kimi illerin işsizlik oranlarının ötekilerle arası iyice açılarak adeta kopuşlar yaşanmıştır.

Beşli özet, yukarıda sözünü ettiğimiz  $E_a$ ,  $D_a$ ,  $DA$ ,  $D_{ii}$ ,  $E_{ii}$  değerlerinden oluşur. Bunların hassas hesaplanmış değerleri Çizelge 3'te gösterilmiştir. Bu değerler kullanılarak her yılın dışadüşen illeri (varsa) ortaya çıkarılabilir. Tukey (1977:44) dışadüşen değerleri bulmak için şu terimleri tanımlar:

$$\text{Adım} = 1,5 * DA \quad \text{Alt Eşik} = D_a - \text{Adım} \quad \text{Üst Eşik} = D_{ii} + \text{Adım}$$

Bu iki eşik değerinin dışında kalan gözlemler dışadüşendir. Sözgelimi 2008 yılında  $DA = 5,85$  olduğuna göre  $\text{Adım} = 1,5 * 5,85 = 8,775$  olur.  $\text{Alt Eşik} = 6,95 - 8,75 = -1,825$ 'tir. İşsizlik oranı ekisi olamayacağına göre altd dördebölen yoktur.  $\text{Üst Eşik} = 12,8 + 8,75 = 21,55$  çıkar; bundan büyük gözlemler dışadüşen sayılmalıdır. Şırnak ilinin %22.1 işsizlik oranı dışadüşendir. Yani 2008 yılında Şırnak öbür illerimizden ciddi ölçüde yüksek bir işsizlik oranına sahiptir. Benzer hesaplamalar öbür yıllar için de yapıldığında 2009-2011 döneminde hiçbir dışadüşen gözleme rastlanmaz ama 2012'de 5 ve 2013'te 7 dışadüşen vardır:

2012: Batman %25,0, Mardin %20,9, Siirt %20,0, Şırnak %16,4, Adıyaman %15,8

2013: Batman %23,4, Mardin %20,6, Siirt %20,5, Şırnak %20,1, Diyarbakır %18,7, Şanlıurfa %16,3, İzmir %15,4.

Anlaşılabacağı üzere 2012 ve 2013'te adı geçen illerde adeta işsizlik depremi yaşanmıştır. İlk çizelgede ipuçlarını gördüğümüz bu çarpıcı değişim burada bütün açıklığıyla gözler önüne serilmiştir. Hepsi Güneydoğu Anadolu Bölgesinde yer alan bu olağandışı yüksek işsizlik sahibi illere, oralardan yoğun göç alan İzmir de eklenmiştir. Nedenlerinin araştırılması gereken bu durum yadsınamaz biçimde açığa çıkmıştır.

Beşli-özet değerleri Çizim 1'de bu dışadüşenlerle birlikte yer almaktadır. Kutu-çizimler, beşli özet değerleri ile dışa düşenlerin çizime dönüştürülmesidir. Kutunun ortasındaki yatay çizgi  $Ort$ , kutunun alt ve üst sınırları da  $D_a$  ve  $D_{ii}$  değerlerini, kutulardan iki yöne uzanan çizgilerin uçları, (dışadüşen yoksa)  $E_a$  ve  $E_{ii}$  değerlerini gösterir. Dışadüşen varsa bu değerler de ilgili il kodlarıyla birlikte ayrıca belirtilir. Bu çizimler yukarıda sözü edilen bulguların bir kez daha gösterimidir. Burada beşli özet değerlerinin zaman içindeki seyri gözle kolayca izlenebilir. Ayrıca

kutunun hem içinde hem dışındaki çarpıklıklar daha açık biçimde sergilenir.  $Ortc$  bize 2009'daki yükselişle sonraki düşüşü karşılaştırmalı olarak açıkça gösterir. Kutunun boyu, işsizlik oranı bakımından ortalarda yer alan %50 ilimizin her yıl hangi sınırlar içinde olduğunu yansıtır. Bu illerin işsizlik oranları 2009'da birbirinden hayli uzaklaşmakta ama sonraki yıllarda birbirine iyice yaklaşmaktadır. Kutunun alt sınırı ( $D_a$ ) da 2009'da yukarı kaymasına kaşın daha sonra düşmüş; son yılda biraz yükselmiştir. Ayrıca ortaldaki %50 ilin  $Ortc$  değere göre çarpıklıklarının derecesi de bu çizimde belirginleşir. 2009'daki üst yarıda sıkışma, alt yarıda açılma varken 2013'te bunun tam tersi bir durum göze çarpmaktadır. Öbür yıllarda kutu içinde pek çarpıklık yoktur. Ama bu çizimin belki de en önemli yanı dışadüşen illerin işsizlik oranlarının ötekilerden ne kadar yukarıda olduğunu sergilemesidir. 2009 yılında Adana ilindeki %26,5 işsizlik oranı her ne kadar öbür illerden hayli büyükse de genel işsizlik oranlarından hatırı sayılır derecede uzak olmadığından dışadüşen değildir.

### Dönüştürmeler

2008-2013 arasında illerin işsizlik oranlarının dağılımı her yılda büyük değerlere doğru çarpıktır. Çizelge 4'teki özet göstergeler de hem bu çarpıklığı, hem de son iki yıldaki büyük sıvrilikleri yansıtır. Bu dağılımları çarpıklıktan biraz olsun kurtarmak için gözlemler, doğal logaritmaları alınarak dönüştürülmüştür. Elde edilen sonuçlar Çizim 2'de sergilenmiştir. Bu yeni çizimi yorumlamak için şunu iyice vurgulamak gerekir. Çizim 1'de herhangi iki ilin işsizlik oranlarını gösteren iki nokta arasındaki uzaklık o illerin işsizlik oranları arasındaki farkı gösterirken Çizim 2'de aynı illerin logaritması alınmış değerlerini gösteren iki nokta arasındaki fark, artık işsizlik oranlarının farkını değil, birinin öbüründen yüzde kaç daha büyük ya da küçük olduğunu temsil eder. Buradaki uzaklıklar karşılaştırılırken bunun akılda tutulması gerekir. Çizim 1'de, herhangi bir yılda kutunun üst yarısının uzunluğu,  $D_{\bar{u}} - Ortc$  yüzde puan ile gösterilirken, alt yarısının uzunluğu  $Ortc - D_a$  yüzde puan kadardır. Ama Çizim 2'de aynı noktalar arasındaki uzaklıklar, üst yarıda  $\ln D_{\bar{u}} - \ln Ortc = \ln (D_{\bar{u}} / Ortc)$ ; alt yarıda  $\ln Ortc - \ln D_a = \ln (Ortc / D_a) =$  demektir. Yeni çizimde son iki yılda dışadüşen illerin sayısı azalarak 4'e inmiştir. Yani 2012 ve 2013'te Batman, Mardin, Siirt, Şırnak illerindeki işsizlik oranları, ilgili yılın diyelim  $Ortc$  değerine oranlanırsa elde edilen oranlar, diğer bütün illerin  $Ortc$  değere oranlanmasından bulunacak değerlere göre olağandışı büyük olduğunu söyleyebiliriz. Demek ki bu dört ilin son iki yıldaki işsizlik oranlarında gerçekten önemli bozulmalar ortaya çıkmıştır.

## Sonuç

Özetlersek, bu sorgulayıcı veri çözümlemesi bize izi sürülecek bazı ipuçları vermiştir. Ülke genelinde 2009'dan 2013'e kadar işsizlik oranlarında genel bir düşme eğilimi vardır. Bu düşüş Türkiye'nin büyük bölümünde 2012 ve 2013 yıllarında da sürerken, her ne olursa, Güneydoğu Anadolu illerimiz ile bu bölgeden yoğun göç alan kimi illerimizde bu son iki yılda tersine bir gelişim gözlenmiştir. Bu bölgenin çoğu ilinde işsizlik oranları bu iki yıl içinde öbür bütün illere göre hem mutlak hem oran olarak ciddi ölçüde artmıştır. Bunun nedenlerinin araştırılması gerekir.



**KAYNAKÇA:**

Şenesen, Ü. (2004) *İstatistik: Sayıların Arkasını Anlamak, Literatür (son baskı 2017)*.

Şenesen, Ü. (2015) "İktisat arařtırmalarında sorgulayıcı veri çözümlemesi" *Ekonomi-tek, Cilt: 4, No: 2, Mayıs, s. 1-23*.

Tukey, J. W. (1977) *Exploratory Data Analysis*, Addison Wesley.

Velleman, P. F., D. C. Hoaglin (1980) *Applications, Basics, and Computing of Exploratory Data Analysis (ABC of EDA)*, Duxbury. The Internet-First University Press 2004'te bu kitabın tıpkıbasımını yayımlamıřtır. Kitap Cornell Üniversitesi kütüphanesinin řu adresinden ücretsiz indirilebilir: <http://dspace.library.cornell.edu/handle/1813/62>.

Daha geniş bir kaynakça için bkz: Şenesen (2004 ve 2015).

**Çizelge 1a. İllerdeki işsizlik oranları (%) 2008-2013 (TÜİK verilerinden düzenlenmiştir)**

İl kodu	İl adı	2008-13	İllerin İşsizlik Oranı (%)						Simgeler	
		Ortanca	2008	2009	2010	2011	2012	2013	İşsizlik %	font
73	Şırnak	18,20	22,1	17,0	11,2	12,0	19,4	20,1	25 +	17
56	Siirt	16,35	17,9	14,8	12,7	11,8	20,0	20,5	22,5 <	25 16
1	Adana	16,15	20,5	26,5	19,1	11,9	11,4	13,2	20 <	22,5 15
2	Adıyaman	15,95	16,5	17,9	10,2	16,1	15,8	9,1	17,5 <	20 14
30	Hakkari	15,15	18,3	19,7	17,4	12,9	9,9	11,7	15 <	17,5 13
35	İzmir	14,95	11,8	16,2	15,1	14,7	14,8	15,4	12,5 <	15 12
47	Mardin	14,90	17,0	12,8	9,1	12,3	20,9	20,6	10 <	12,5 11
21	Diyarbakır	14,60	15,7	20,6	13,5	8,6	7,3	18,7	7,5 <	10 10
72	Batman	14,20	14,3	13,5	11,7	14,1	25,0	23,4	5 <	7,5 9
80	Osmaniye	14,20	16,3	19,4	14,4	12,4	10,0	14,0	<	5 8
27	Gaziantep	13,90	16,8	17,4	13,4	14,4	11,2	6,9		
31	Hatay	13,30	17,7	19,0	13,9	12,7	11,4	12,2		
77	Yalova	12,90	12,1	17,8	14,4	13,7	12,1	11,0		
65	Van	12,80	13,0	15,6	17,2	12,6	7,8	10,3		
33	Mersin	12,75	13,1	17,6	14,1	9,2	9,5	12,4		
13	Bitlis	12,65	14,2	14,3	16,3	11,1	9,1	10,6		
63	Şanlıurfa	12,60	12,8	17,0	12,4	8,0	6,2	16,3		
41	Kocaeli	12,35	11,6	17,0	13,8	13,1	11,1	10,1		
49	Muş	12,35	13,2	16,2	16,4	11,5	8,9	10,4		
46	Kahramanmaraş	12,20	12,8	16,0	12,8	11,1	9,7	11,6		
38	Kayseri	11,70	11,1	14,1	14,1	12,3	9,3	9,9		
34	İstanbul	11,55	11,2	16,8	14,3	11,8	11,3	11,2		
23	Elazığ	11,45	15,5	18,1	12,6	10,3	9,0	7,8		
12	Bingöl	11,05	14,8	17,8	12,5	9,6	8,3	7,0		
6	Ankara	11,00	11,8	13,6	12,1	9,4	9,5	10,2		
9	Aydın	10,80	12,4	16,4	13,5	9,2	7,9	6,9		
79	Kilis	10,65	10,9	14,9	10,1	12,6	10,4	7,7		
71	Kırıkkale	10,65	11,1	19,3	12,0	10,2	8,9	8,0		
14	Bolu	10,60	11,5	11,5	10,6	10,6	9,5	9,6		
62	Tunceli	10,40	17,9	18,6	11,5	9,3	7,6	8,1		
58	Sivas	10,35	10,7	13,2	13,6	9,9	7,8	10,0		
54	Sakarya	10,25	8,5	11,0	11,4	10,9	9,6	9,4		
81	Düzce	9,90	10,2	14,6	13,4	9,6	8,5	8,7		
39	Kırklareli	9,85	11,5	13,6	10,8	8,9	7,9	8,0		
66	Yozgat	9,80	12,2	11,0	12,7	8,5	6,1	8,6		
40	Kırşehir	9,65	11,1	15,6	10,8	8,5	7,4	7,3		
48	Muğla	9,55	10,1	12,1	10,5	9,0	8,0	7,3		
7	Antalya	9,50	9,7	12,7	11,1	9,3	8,5	7,9		
26	Eskişehir	9,40	9,9	15,2	10,5	8,9	8,4	8,5		
44	Malatya	9,35	11,5	14,6	9,5	9,2	6,8	7,8		
32	Isparta	9,30	8,1	10,5	12,0	9,7	8,9	8,7		
22	Edirne	9,10	14,3	13,4	9,2	9,0	7,5	7,8		

Çizelge 1b. İllerdeki işsizlik oranları (%) 2008-2013 (devam)

İl kodu	İl adı	2008-13 Ortanca	İllerin İşsizlik Oranı (%) (devam)					Simgeler		
			2008	2009	2010	2011	2012	2013	İşsizlik %	font
76	İğdir	9,10	6,7	12,6	11,1	10,4	7,8	6,9	17,5 < 20	14
4	Ağrı	9,05	6,5	10,1	11,7	10,9	8,0	6,8	15 < 17,5	13
16	Bursa	9,00	10,8	14,7	10,5	7,5	7,1	6,6	12,5 < 15	12
59	Tekirdağ	8,95	9,1	11,9	9,6	8,8	7,1	7,2	10 < 12,5	11
51	Niğde	8,50	9,3	13,8	9,4	7,7	6,1	6,1	7,5 < 10	10
20	Denizli	8,30	9,4	13,2	11,2	7,2	6,9	6,5	5 < 7,5	9
78	Karabük	8,00	7,5	8,4	11,5	8,0	7,8	8,0	< 5	8
50	Nevşehir	7,95	9,0	9,9	8,4	7,5	6,1	6,0		
68	Aksaray	7,85	10,0	13,1	9,2	6,5	5,8	5,8		
42	Konya	7,55	10,7	10,8	8,2	6,9	6,2	4,7		
67	Zonguldak	7,55	6,9	7,3	10,7	8,1	7,5	7,6		
18	Çankırı	7,25	7,7	9,5	9,1	6,3	6,0	6,8		
11	Bilecik	7,10	8,2	8,5	7,7	6,2	6,4	6,5		
36	Kars	7,00	4,1	7,4	7,4	9,0	6,4	6,6		
10	Balıkesir	6,95	7,9	9,1	8,0	5,5	5,5	6,0		
55	Samsun	6,95	7,8	7,3	7,8	5,5	5,9	6,6		
61	Trabzon	6,75	6,1	6,6	6,3	7,0	6,9	7,4		
5	Amasya	6,65	8,2	6,9	6,7	5,2	5,6	6,6		
3	Afyonkarahisar	6,60	7,7	10,8	7,6	4,8	4,6	5,6		
45	Manisa	6,60	9,8	11,7	8,1	4,7	4,4	5,1		
64	Uşak	6,60	9,2	12,0	7,8	4,7	4,2	5,4		
70	Karaman	6,60	7,7	7,5	8,5	5,7	4,9	4,2		
57	Sinop	6,55	6,9	10,4	8,7	5,5	5,3	6,2		
17	Çanakkale	6,55	7,0	7,3	7,0	4,8	5,0	6,1		
25	Erzurum	6,50	6,2	7,9	6,5	6,5	6,5	6,6		
15	Burdur	6,35	6,4	5,1	6,3	6,5	5,8	6,9		
43	Kütahya	6,25	6,5	9,1	6,5	4,6	4,2	6,0		
28	Giresun	6,20	6,3	6,5	6,0	6,1	6,1	6,5		
19	Çorum	6,10	7,5	6,3	6,5	4,7	5,1	5,9		
24	Erzincan	6,10	6,1	7,6	6,1	5,6	5,7	6,7		
74	Bartın	6,10	5,2	6,4	10,2	6,0	5,7	6,2		
8	Artvin	6,00	5,7	5,6	5,7	6,3	6,3	7,1		
52	Ordu	5,90	4,9	5,9	5,9	5,9	5,6	6,1		
60	Tokat	5,90	5,9	5,9	6,7	5,4	5,7	6,6		
53	Rize	5,85	5,2	4,6	6,0	5,9	5,8	6,7		
29	Gümüşhane	5,80	5,0	4,4	5,8	6,1	5,8	7,2		
37	Kastamonu	5,75	5,3	7,3	6,4	5,3	5,3	6,2		
75	Ardahan	5,10	3,7	4,2	7,5	7,4	4,4	5,8		
69	Bayburt	4,75	4,7	4,5	4,7	4,9	4,8	6,2		
	Ortanca	9,3	10,1	12,8	10,6	8,9	7,5	7,4		



**Çizelge 3. 2008-2013 yılları için beşli özetler**

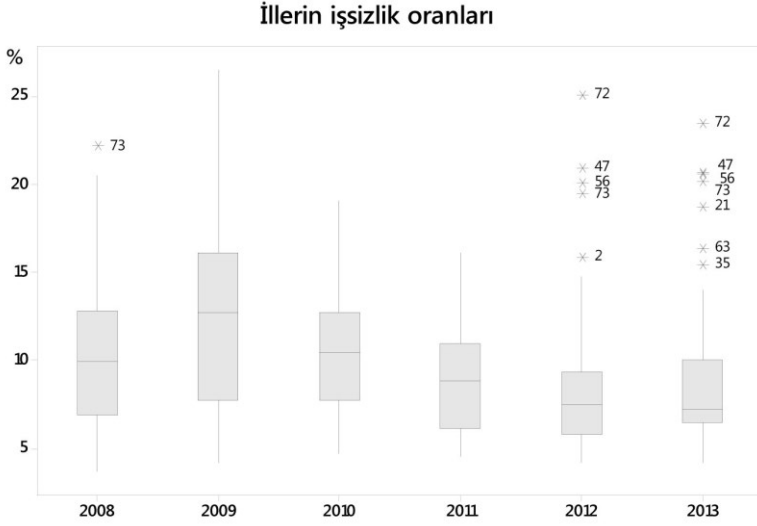
Yıl	$E_a$	$D_a$	$Ortc$	$D_{\bar{u}}$	$E_{\bar{u}}$	$YA$	$DA$
2008	3,7	6,95	10	12,8	22,1	18,4	5,85
2009	4,2	7,75	12,7	16,1	26,5	22,3	8,35
2010	4,7	7,75	10,5	12,7	19,1	14,4	4,95
2011	4,6	6,15	8,9	11	16,1	11,5	4,85
2012	4,2	5,8	7,5	9,4	25	20,8	3,6
2013	4,2	6,5	7,3	10,05	23,4	19,2	3,55

$E_a$ : En alt değer,  $D_a$ : Alt dördebölen,  $Ortc$ : Ortanca,  $D_{\bar{u}}$ : Üst dördebölen,  $E_{\bar{u}}$  En üst değer,  $YA$ : Yayılım aralığı,  $DA$ : Dördebölenler aralığı

**Çizelge 4. 2008-2013 yılları için özet göstergeler**

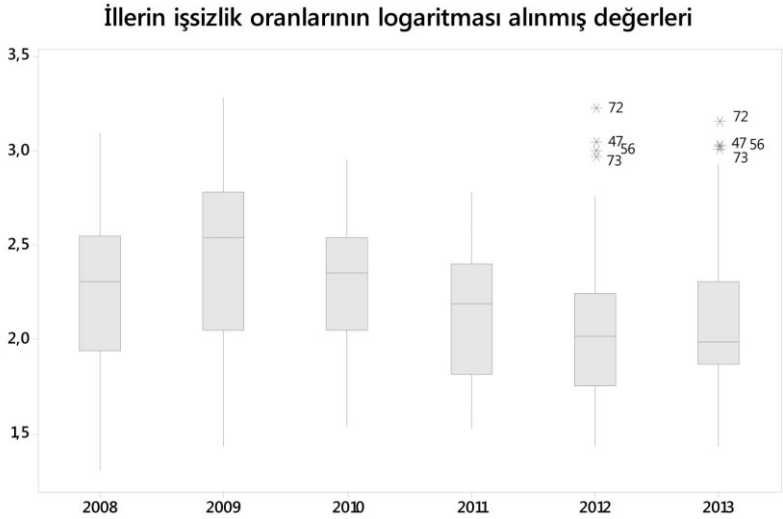
Yıl	Ar. Ort	Kırpık Ort	St.Sap.	Çarpıklık	Basıklık
2008	10,401	10,223	4,097	0,65	-0,07
2009	12,309	12,234	4,717	0,19	-0,40
2010	10,456	10,338	3,212	0,34	-0,44
2011	8,796	8,692	2,882	0,38	-0,74
2012	8,258	7,759	3,795	2,31	6,41
2013	8,801	8,341	3,878	2,00	3,94

Çizim 1. İllerin işsizlik oranları (%)



73 Şırnak, 72 Batman, 47 Mardin, 56 Siirt, 2 Adıyaman, 21 Diyarbakır, 63 Şanlıurfa, 35 İzmir

Çizim 2. İllerin işsizlik oranlarının doğal logaritmaları alınarak dönüştürülmüş değerleri



72 Batman, 47 Mardin, 56 Siirt, 73 Şırnak