

-ARAŞTIRMA MAKALESİ-

## SAĞLAM KANONİK KORELASYON ANALİZİ İLE İNSANİ GELİŞİMİŞLİK VE DİJİTAL REKABETÇİLİK ENDEKSLERİ ARASINDAKİ İLİŞKİNİN DEĞERLENDİRİLMESİ\*

**Arzu ALTIN YAVUZ**

Doç. Dr.

Eskişehir Osmangazi Üniversitesi, Fen-Edebiyat Fakültesi İstatistik Bölümü

E-posta: aaltin@ogu.edu.tr

ORCID ID: <https://orcid.org/0000-0002-3277-740X>

**Barış ERGÜL**

Araş. Gör.

Eskişehir Osmangazi Üniversitesi, Fen-Edebiyat Fakültesi İstatistik Bölümü

E-posta: bergul@ogu.edu.tr

ORCID ID: <https://orcid.org/0000-0002-1811-5143>

### ÖZET

*İnsani Gelişme Endeksi (İGE), ülkeler arasındaki gelişmişlik farklılıklarını ortaya koymak amacı ile oluşturulmuş ve ülkeleri gelişmişlik sıraları ile sıralayan bir endekstir. İGE'nin hesaplanmasında son yıllarda, yaşam standardı, sağlık standardı ve okullaşma standardı olmak üzere üç endeks bileşeni kullanılmaktadır. Dijital Rekabetçilik Endeksi ise (DRE), hükümet uygulamalarının, iş modellerinin ve genel olarak toplumun dönüşümüne yol açan dijital teknolojileri ülkeleri konu alarak sıralayan bir endekstir. Bu unsurlar üç endeks bileşeni ile açıklanmaya çalışılmıştır: Bilgi, Teknoloji ve Geleceğe Hazır Olma. Bu çalışmada, 2016 yılı için açıklanan insani gelişmişlik ve dijital rekabetçilik endeks raporlarından yararlanılmıştır. Verilerine ulaşılan ülkeler için, İnsani Gelişmişlik ve Dijital Rekabetçilik veri setlerine ile ilk aşamada klasik kanonik korelasyon analizi uygulanmıştır. Sonraki aşamada, değişken setlerinin normallik analizleri yapılmıştır. Son aşamada ise, değişken setlerinin çok*

\* Bu makalede bilimsel araştırma ve yayın etiği ilkelerine uyulmuştur.

**Atıf (APA):** Altın Yavuz, A. & Ergül, B., (2020), Sağlam Kanonik Korelasyon Analizi ile İnsani Gelişmişlik ve Dijital Rekabetçilik Endeksleri Arasındaki İlişkinin Değerlendirilmesi, Sosyal ve Beşeri Bilimler Dergisi, 12 (1): 1-18.

*değişkenli normal dağılıma uymaması ve veri setinde aykırı değer bulunması sebebi ile sağlam kanonik korelasyon analizi uygulanmıştır. Sağlam kanonik korelasyon analizi ile elde edilen sonuçların daha güvenilir ve gerçekçi olduğu tespit edilmiştir. İnsani olarak gelişmiş toplumların, dijital dünyada da oldukça güçlü olacakları görülmektedir. Benzer şekilde, dijital gelişmişliği takip eden toplumlarında insanlığın gelişimine katkı sağlayacağı istatistiksel olarak ortaya çıkarılmıştır. Ek olarak, kanonik korelasyon analizinin varsayımlarının sağlanmaması durumunda önerilmiş olan sağlam kanonik korelasyon analizinin kullanılması, uygulamacılara daha güvenilir ve gerçekçi bilgi elde etmeleri konusunda yardımcı olacaktır.*

**Anahtar Kelimeler:** IGE, DRE, kanonik, sağlam, korelasyon

## **EVALUATION OF THE RELATIONSHIP BETWEEN HUMAN DEVELOPMENT AND DIGITAL COMPETITIVENESS INDICES VIA ROBUST CANONICAL CORRELATION ANALYSIS**

### **ABSTRACT**

*The Human Development Index (HDI) is an index that is developed with the aim of revealing differences in development between countries and ranking countries in terms of development. In recent years, three indices have been used in the calculation of HDI, including life standard, health standard, and schooling standard. The Digital Competitiveness Index (DCI) is an index ranking the subject of government practices, business models and the digital technologies that lead to the transformation of society in general. These elements were tried to be explained by three index components: Knowledge, Technology and Future, Readiness. In this study, human development and digital competitiveness index reports for 2016 were used. For the countries whose data has been reached, classical canonical correlation analysis has been applied to Human Development and Digital Competitiveness datasets in the first stage. In the next step, normality analysis of variable sets was performed. In the last stage, robust canonical correlation analysis was performed because the variable sets did not conform to the multivariate normal distribution and there were outliers in the data sets. The results obtained by robust canonical correlation analysis were found to be more reliable and realistic. It is seen that humanly developed societies will be quite powerful in the digital world. Similarly, it has been revealed that it will contribute to the development of humanity in societies that follow digital development. In addition, if the assumptions of the canonical correlation analysis are not met, the*

*use of the robust canonical correlation analysis proposed will help practitioners obtain more reliable and realistic information.*

**Keywords:** *HDI, DCI, canonical, robust, correlation*

**JEL Codes:** *015, R15*

## **Giriş**

Gelişme ve gelişmişlik günümüzde birçok ülkenin en önemli hedeflerinden biridir. Gelişmişlikle ilgili çalışmalar 1900'lü yılların başında hız kazanmış ve birçok çalışmaya konu olmuştur. Günümüzde ise sadece ülkelerin gelişmişliği değil, bölgesel gelişmişlik de önem kazanmıştır. Ülkeler veya bölgelerin büyümeleri ve gelişmeleri arasındaki ilişkiler yapılan çalışmalar sayesinde farklı bakış açıları doğrultusunda değerlendirilmiştir. Bu konuda yapılan ilk çalışmalarda, gelişmişliğin en büyük göstergesinin kişi başına milli gelir olduğu ortaya konmuştur.

1970'lere geldiğinde ise kalkınma kavramı yeniden tanımlanmış ve buna bağlı olarak gelişme kavramı da güncellenmiştir. Böylelikle, gelişmeyi kalkınmanın bir şartı olarak gören yeni bir yaklaşımla ülkelerin yada bölgelerin gelişmişliği değerlendirilmeye başlanmıştır. Bu yaklaşımda, ekonomik büyüme kavramı ile birlikte, yoksulluk, işsizlik, gelir dağılımı ve bölgeler arasındaki dengesizlikler de gelişme tanımının içinde değerlendirilmiştir. Tüm bu değişimler beraberinde insani gelişmişlik kavramının ortaya çıkmasına neden olmuştur (Ünal, 2008).

Birleşmiş Milletler Kalkınma Teşkilatı, ülkelerin kalkınma düzeylerini ölçmek için Human Development Index adıyla adlandırılan, İnsani Kalkınma Raporlarını yayınlamaya başlamıştır. Bu raporların içerisinde yer alan insani gelişme endeksi (İGE), ülkeler arasındaki gelişmişlik farklılıklarını ortaya koymak amacı ile oluşturulmakta ve her yıl yayınlanmaktadır. İGE temel ölçütleri, insani yaşam standardı, uzun ve sağlıklı yaşam, bilgiye erişim olarak kabul edilmektedir. Endeks değeri 0-1 aralığı içinde değer almaktadır. Eğer endeks değeri 0 değerine yakınsa bu durum minimum gelişmişliği, 1 değerine yakınsa maksimum gelişme düzeyini ifade etmektedir. İGE'nin hesaplanmasında bir süreç içerisinde gelişmiş ve son yıllarda İGE'nin hesaplanmasında 3 temel endeks değerinin ortalaması kullanılmaya başlanmıştır. Bunlar; yaşan standardı, sağlık standardı ve okullaşma standardı endeksidir (Firat ve Aydın, 2015).

Öte yandan, günümüz dünyasında gelişimin en hızlı olduğu alanlardan birisi dijital dünyadır. Ülkelerin ya da bölgelerin rekabetinde dijital dünyadaki

görünürlük büyük önem arz etmektedir. Bu süreçte dijital rekabetçilik kavramı ortaya atılmıştır. Dijital rekabetçilik kavramı, teknolojiye yatırım yapan ülkelerin ekonomi göstergeleri ile ilişkisini ortaya koyması nedeni ile ekonomi biriminin kapasitesi olarak tanımlanır. Hükümet uygulamalarının, iş modellerinin ve genel olarak toplumun dönüşümüne yol açan dijital teknolojileri benimsemek ve araştırmak olarak tanımlanmıştır. Bu sayede, firmalar veya ülkeler gelecekteki değer yaratımlarını güçlendirme fırsatını bulmaktadırlar.

Dijital rekabet edebilirlik çerçevesinin, örgütsel, kurumsal ve yapısal unsurları kapsayan faktörler üzerine inşa edilmesi gereği ortaya konulmuştur. Dahası, bu unsurların bilginin uygulanması, dönüşümde araştırmanın rolü, ilgili yönetmeliğin etkinliği, yeni teknolojilerin benimsenmesi ve ortaya çıkan değişimlere karşı açıklık ve esneklik içermesi de gerekmektedir (IMD, 2017). Tüm bu unsurlar altında dijital rekabetçilik üç faktör ile açıklanmaya çalışılmıştır. Bunlar; bilgi, teknoloji ve geleceğe hazır olma.

Bilgi Faktörü, yeni teknolojilerin keşfedilmesi, anlaşılması ve öğrenilmesi yoluyla dijital dönüşüm sürecinin altını çizen gerekli altyapıya atıfta bulunmaktadır. Bilgi faktörü de kendi içerisinde üç alt faktörü kapsamaktadır: Yetenek, Eğitim ve Bilimsel yatırım. Yetenek, belirli bir ekonomide mevcut beceri ve yetenek havuzudur. Yetenek havuzunun gücü ve gelişim düzeyi, işgücünün eğitim ve öğretime verdiği önceliğe bağlıdır. Bilimsel yatırım ise, bir ekonominin dijital dönüşümü için gerekli olan bilginin üretimini vurgulamaktadır.

Teknoloji Faktörü, dijital teknolojilerin geliştirilmesinin genel kapsamını değerlendirmektedir. Bu kapsamda, öncelikle iş faaliyetlerinin etkin bir şekilde yerine getirilmesine olanak veren destekleyici bir Düzenleyici Çerçeve alt faktörünü hesaba katmaktadır. Teknoloji faktörünü oluşturan ikinci bir alt faktör sermayedir. Sermaye, teknoloji ile ilgili gelişmede kullanılabilirliği ve mevcut yatırımı değerlendirmektedir. Aynı zamanda, belirli bir ekonomideki yatırım riski düzeyini de dikkate almaktadır. Son olarak Teknoloji faktörünü oluşturan üçüncü alt faktör Teknolojik Çerçeve alt faktörüdür. Bu faktör, bir ülkede mevcut fiziksel teknolojik altyapıyı ve kalitesini değerlendirmektedir (Bayraktar ve Kaya, 2016).

Geleceğe Hazır Olma faktörü, bir ekonominin hazırlığının dijital dönüşüme verdiği desteği ölçmektedir. Bu anlamda üç alt faktör içermektedir. Bunlar; Adaptif Tutumlar, İş Çevikliği ve Bilgi teknolojileri entegrasyonudur. Dijital Rekabetçilik, mevcut dijital teknolojilerin toplum tarafından emilmesini gerektirir. Dijital teknolojilerin emilmesi, bir toplumun dijital süreçlere katılma isteğini içeren belirli Adaptif Tutumlara ihtiyaç duyar. Geleceğe Hazır Olma, yeni teknolojilerin benimsenmesi açısından iş esnekliğini gerektirir. Bu anlamda, İş

Çevikliği, belirli bir ekonomideki firmaların yeni fırsatlardan yararlanmak için iş modellerini dönüştürebileceğini ima etmektedir. Bununla birlikte inovasyon ile iç içe bir süreçtir. Geleceğe Hazır Olma, son olarak, bilgi teknolojilerinin uygulamalarının ne kadar iyi olduğunu değerlendiren Bilgi Teknolojileri Entegrasyonuna ihtiyaç duymaktadır (IMD, 2017). Dijital Rekabetçilik Endeksi, belirtilen üç endeks bileşenini ifade eden endekslerin ağırlıklı ortalamaları yardımıyla hesaplanan bir endekstir (IMD, 2017).

İnsani gelişmişlik ve dijital rekabetçiliğin iç içe olduğu bu süreçte bir çok çalışma tek bir endekse odaklanmaktadır (Günel vd., 2017; Ünal, 2008).

Ülkemizde insani gelişmişlik adına yapılan çalışmaların sayısı sınırlıdır (Akder,1992; Yumuşak ve Tuna, 2002; Dikmen, 2004; Demir, 2006). Çalışmalarda, genel olarak mevcut olan durum ortaya konmaya çalışılmıştır.

Oysaki her iki endeks değeri birbirinden etkilenmektedir. Bu nedenle aralarındaki bu ilişkinin araştırılması gerekir. Dünyada ve Türkiye’de dijital rekabetçilik konusunda yapılan istatistik çalışma sayısının az olması nedeni ile literatürdeki bu boşluğu dolduracaktır.

Bu çalışmanın amacı, insani gelişmişlik ile dijital rekabetçilik arasındaki ilişkinin ortaya konmasıdır. İnsani gelişmişlik üzerinde etkili olan alt boyutları ile dijital rekabetçilik üzerinde etkili olan alt boyutlarının tespiti amaçlanmıştır. Böylelikle her iki endeks üzerinde etkili olan faktörlerin ortaya çıkartılmasını sağlamaktır. Bu amaçla, İnsani Gelişme Endeksi (İGE) ile Dijital Rekabetçilik Endeksi (DRE) alt boyutları değişken olarak alınmış ve ilgili alt boyutlar yardımı ile iki endeks değerlerinden oluşan veri setine kanonik korelasyon ve robust kanonik korelasyon analizleri uygulanmıştır. Ayrıca elde edilen sonuçlar her iki analiz yönteminin performansı bakımından da karşılaştırılmıştır.

Çalışmanın 2. bölümünde, Kanonik Korelasyon analizi tanıtılmıştır. Tüm istatistiksel tekniklerde olduğu gibi kanonik korelasyon analizinde de elde edilen sonuçların geçerliliği ve güvenilirliği, analizin dayandığı varsayımların sağlanmasına bağlıdır. Çalışmada bu varsayımların bazılarının sağlanmadığı tespit edilmiştir. Bu nedenle, kanonik korelasyon analizinin dayandığı varsayımların sağlanamaması durumunda uygulanan Sağlam Kanonik Korelasyon analizi açıklanmıştır. 3. bölümde, insani gelişmişlik ve dijital rekabetçilik endeksi kullanılarak kanonik ve sağlam kanonik korelasyon analizleri uygulanmıştır. Sonuç ve öneriler bölümünde, bulunan sonuçlar tartışılmış ve önerilerde bulunulmuştur.

## 1. Materyal ve Metot

İstatistik biliminde, tanımlanan en basit ilişki X ve Y ile gösterilen iki tesadüfi değişken arasındaki ilişkidir. Bu ilişkiye basit korelasyon adı verilmektedir. Değişken sayısının ikiden fazla sayıda olması durumunda, değişkenlerden biri ile geriye kalan diğer değişkenler arasında ilişki de hesaplanabilir. Bu hesaplanacak korelasyon katsayısı da çoklu korelasyon katsayısı olarak bilinmektedir.

En genel ve en karmaşık ilişki analizi olan kanonik korelasyon analizi kümeler arası ilişki olarak bilinmekte ve bir değişken kümesindeki değişkenlerin doğrusal bileşimi ile diğer kümedeki değişken bileşenleri arasındaki ilişkiyi araştırmada kullanılmaktadır (Işığışok, 1999). Bu analiz yönteminde, her iki kümede yer alan değişken sayısının eşit olması gerekmektedir. İki değişken kümesinden biri bağımlı diğeri bağımsız olarak düşünülebilir. Ancak birçok durumda bu nedensellik ilişkisi aranmamaktadır. Bir başka ifade ile her iki kümede bağımlı veya bağımsız olarak alınarak analiz gerçekleştirilebilir (Alpar, 2011). Kanonik korelasyon analizinin bağlantılı olduğu bir çok istatistiksel teknik bulunmaktadır.

Eğer bir bağımlı ve bir bağımsız değişken söz konusu ise basit korelasyona, bir bağımlı daha fazla bağımsız değişken söz konusu ise çoklu regresyona, yada bir bağımsız daha fazla bağımlı değişken söz konusu ise çok değişkenli regresyona dönüşmektedir (Karagöz, 2016).

Kanonik korelasyon analizi, öncelikle iki değişken seti arasındaki ilişkiyi maksimum yapacak doğrusal bileşenler elde etmeye yönelik bir analizdir. Başka bir deyişle, kanonik korelasyon analizinde, değişken setleri arasındaki doğrusal ilişkiler araştırılır. Bu doğrusal ilişkiden hareketle doğrusal bileşenler elde edilir. Bu doğrusal bileşenlere kanonik değişkenler adı verilir. Kanonik değişkenlerin benzer çiftleri arasındaki ilişki de kanonik korelasyon adını alır.

Kanonik korelasyon analizinde; değişken setlerine ait değişkenler arasındaki korelasyonu maksimum yapan doğrusal birleşimlerin belirlenmesi, iki değişken setinin istatistiksel olarak birbirinden bağımsız olup olmadığının test edilmesi, değişken setleri arası korelasyona en fazla katkıda bulunan her iki değişken setindeki değişkenlerin belirlenmesi ve bir değişken setinin diğer bir değişken seti tarafından ne ölçüde açıklanabildiğinin belirlenmesi amaçlanmaktadır (İlhan vd., 2013).

Kanonik Korelasyon Analizinin amacı, p boyutlu X rassal değişkeni ile q boyutlu Y rassal değişkeni arasında ilişkiyi tanımlamak ve ölçmektir. Orijinal

değişkenlerin  $a'X$  ve  $b'Y$  biçimindeki doğrusal kombinasyonlarının maksimum korelasyona sahip olması istenir. Kanonik Korelasyon aşağıdaki gibi ifade edilir:

$\alpha \in \mathbb{R}^p$  ve  $\beta \in \mathbb{R}^q$  vektörleri ifade etmek üzere,

$$(\alpha, \beta) = \operatorname{argmax}_{\alpha, \beta} |Corr(a'X, b'Y)| \quad (1)$$

biçimindedir.

$U=\alpha'X$  ve  $V=\beta'Y$  olarak ifade edilen değişkenler kanonik değişken setleri olarak ifade edilir. Bunu yapmak için, hem  $U$  hem de  $V$ 'nin birim varyanslı olmasını gerektiren bir normleştirme kısıtlaması eklenir. İlk kanonik korelasyon, (1) 'de elde edilen maksimum değere eşit olan iki kanonik değişken arasındaki korelasyonun mutlak değeri olarak tanımlanmaktadır. Daha yüksek sıralı kanonik değişkenler ve korelasyonlar (1) 'de olduğu gibi tanımlanmıştır, fakat ek kısıtlama altında,  $k$ 'nin kanonik bir varyasyonu,  $1 < k \leq \min(p, q)$  şartını sağlayan tüm kanonik varyasyonlarla ilişkisiz olmalıdır. (1) ile verilen denklemi çözmek için, rassal değişken  $Z=(X', Y')$  için kovaryans matrisi aşağıdaki gibi tanımlanır:

$$\Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}.$$

$\alpha \in \mathbb{R}^p$  ve  $\beta \in \mathbb{R}^q$  vektörleri, aşağıdaki matrislerin en büyük özdeğerlerine karşılık gelen özvektörlerdir.

$$\Sigma_{xx}^{-1} \Sigma_{xy} \Sigma_{yy}^{-1} \quad \text{ve} \quad \Sigma_{yy}^{-1} \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy}. \quad (2)$$

(2) ile verilen matrisler, aynı pozitif özdeğere sahiptir ve en büyük olanı ilk kanonik korelasyona eşittir. Bilinmeyen  $\alpha \in \mathbb{R}^p$  ve  $\beta \in \mathbb{R}^q$  kestirimi için örneklem kovaryans matrisinden yararlanılır. (2) eşitliği kullanılarak hesaplanan özdeğerler ve özvektörler sayesinde kanonik korelasyonlar elde edilir (Dehon vd., 2000).

Veri setinin içinde aykırı değerlerin bulunması durumunda, örneklem ortalaması ve varyansı bu değerlerden olumsuz etkilenmektedir. Bir başka ifadeyle, örneklem ortalaması ve varyansı sağlam tahmin ediciler değildir. Aykırı değerler ayrıca normallik varsayımının bozulmasına neden olduğundan dolayı, sağlam tahmin yöntemlerine ihtiyaç duyulmaktadır. Bunun için ortalama ve kovaryans matrislerinin sağlam tahmin edicilerinin kullanılması gerekmektedir. Karnaal (1991), çok değişkenli konum-ölçek parametrelerini  $M$  tahmin edicisini kullanarak vermişlerdir.

İki boyutlu rassal değişken  $z = (u, v)'$ , M-konum parametresi  $\mu(z)$ 'nin M tahmin edicisi ve M-yayıma matrisi  $C(z)$ , aşağıdaki denklemlerin çözülmesi ile elde edilir.

$$\mu = E \left[ w_1((z - \mu)'C^{-1}(z - \mu))z / E[w_1((z - \mu)'C^{-1}(z - \mu))] \right] \quad (3)$$

$$C = [w_2((z - \mu)'C^{-1}(z - \mu))(z - \mu)(z - \mu)'] \quad (4)$$

Burada,  $\mu$ ; tek değişkenli vektör ve  $C$ ; 2x2 boyutlu pozitif tanımlı simetrik matrisi göstermektedir.  $w_1$  ve  $w_2$ ; Huber M tahmin edicisinden bulunan ağırlık fonksiyonunu temsil etmektedir.

Bu konuda diğer yaklaşım, Croux ve Dehon (2002), Minimum Kovaryans Determinant tahmin edicisini kullanarak Sağlam Kanonik Korelasyon Analizini önermişlerdir. Kovaryans matrisinin kestirimi konusundaki son yaklaşım, verilerin işaretlerini kullanmaktır. Robust kovaryans matrisi ile çalışmak yerine, aykırı gözlemlerin etkisini doğrudan gözlemlerin “işaretleri” veya “sıraları” ile değiştirerek ve bu işaretlerin ve sıraların dışında sıradan bir kovaryans matrisini hesaplamak esasına dayalıdır. Bu konudaki diğer bir yaklaşım ise, izdüşüm yaklaşımıdır. Bu yaklaşımda, (1) eşitliğinde verilen  $|\text{Corr}(a'X, b'Y)|$  'nin maksimizasyonu yerine, korelasyon indislerinin maksimizasyonu yapılır. İndeks değerleri, tek değişkene sahip değişkenler arasında basit olarak Spearman korelasyon katsayılarının hesaplanmasına ihtiyaç duyar (Dehon vd., 2000).

Kanonik korelasyon analizinin uygulanabilmesi için değişken setlerinin ve bu değişken setlerinde yer alan değişkenlerin bazı varsayımlara uyması gerekmektedir. Bu varsayımlardan ilki değişken setlerinin çok değişkenli normal dağılıma sahip olmasıdır. İkinci varsayım ise, değişken setlerinden elde edilen kovaryans ve korelasyon matrislerinin parçalanabilir ve tersi alınabilir olması varsayımdır (Kalkan ve Özden, 2017; Keskin ve Özsoy, 2004).

Değişken setlerinin çok değişkenli normal dağılımdan gelip gelmediğinin test edilmesi için literatürde belirlenen farklı yaklaşımlar bulunmaktadır. Bu testler; Mardia'nın çok değişkenli normallik testleri, Henze-Zirkler çok değişkenli normallik testleri ve Royston'un çok değişkenli normallik testleridir (Korkmaz ve Gökşülük, 2014).

Mardia (1970) tarafından önerilen çok değişkenli normallik testi, çok değişkenli çarpıklık ve basıklık ölçülerine dayanmaktadır.



$m_{ij} = (x_i - \bar{x})'S^{-1}(x_j - \bar{x})$  kare mahalnobis uzaklığını göstermek üzere,

$$\alpha_{3,p} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n m_{ij}^3,$$

$$\alpha_{4,p} = \frac{1}{n} \sum_{i=1}^n m_{ii}^2$$

(5)

ile hesaplanmaktadır.

Çarpıklık için test istatistiği, Ki-kare dağılır ve aşağıdaki gibi hesaplanır:

$$\frac{n}{6} \alpha_{3,p} \sim \chi_{p(p+1)(p+2)/6}^2$$

(6)

Basıklık için test istatistiği, Normal dağılır ve aşağıdaki gibi hesaplanır:

$$\alpha_{4,p} \sim N(p(p+2), 8p(p+2)/n)$$

(7)

Henze ve Zirkler (1990) tarafından önerilen çok değişkenli normallik testi,

$$p = \text{değişken sayısı}, \beta = \frac{1}{\sqrt{2}} \left( \frac{n(2p+1)}{4} \right)^{1/p+4}, D_i = (x_i - \bar{x})'S^{-1}(x_i - \bar{x}),$$

$$D_{ij} = (x_i - x_j)'S^{-1}(x_i - x_j) \text{ olmak üzere,}$$

$$HZ = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \exp\left(-\frac{\beta^2}{2} D_{ij}\right) - 2(1 + \beta^2)^{-p/2} \sum_{i=1}^n \exp\left(-\frac{\beta^2 D_i}{2(1+\beta^2)}\right) + n(1 + 2\beta^2)^{-p/2} \quad (8)$$

ile bulunmaktadır.

Eğer veri seti çok değişkenli normal dağılıma uygun ise, HZ istatistiği log-normal dağılmaktadır.

Royston (1992) tarafından önerilen çok değişkenli normallik testi için  $W_j$  ( $j = 1, 2, \dots, p$ ) Shapiro-Wilks test istatistiği olsun.  $w_j = \log_{10}^{\square}(1 - W_j)$  ve  $Z_j = w_j - \mu\sigma$  normallik dönüşümünden elde edilen değerler olmak üzere,

$$R = \frac{e \sum_{i=1}^n K_j}{p} \sim \chi_e^2$$

(9)

ile hesaplanır. Burada,  $e = \frac{p}{1+(p-1)\bar{c}}$  ve  $K_j = (\Phi^{-1}(\Phi(-Z_j)/2))^2$  olarak hesaplanır. Burada  $c$ ,  $K_j$ 'ler arasındaki korelasyonu ve  $\Phi$  ise birikimli normal dağılımı göstermektedir.

## 2. Uygulama

Bu çalışmada, 2016 yılı için açıklanan insani gelişmişlik ve dijital rekabetçilik endeks raporlarından yararlanılmıştır. Verilerine ulaşılan 53 ülke için İnsani Gelişmişlik ve Dijital Rekabetçilik veri setlerine ile ilk aşamada klasik kanonik korelasyon analizi sonrasında değişken setlerinin normallik analizleri ve son aşamada ise değişken setlerinin çok değişkenli normal dağılıma uymaması ve veri setinde aykırı değer bulunması sebebi ile sağlam kanonik korelasyon analizi uygulanmıştır. Kanonik korelasyon analizi için sonuçların güvenilir olmasının sağlanması açısından örnekleme alınacak birim sayısının değişken sayısının 5 katı olmasının yeterli olacağı belirtmektedir (Keskin vd., 2005).

İnsani gelişmişlik veri setinde analize alınan değişkenler:

**X1:** Sağlık Standardı

**X2:** Okullaşma Standardı

**X3:** GSMH

Dijital rekabetçilik veri setinde analize alınan değişkenler:

**Y1:** Bilgi

**Y2:** Teknoloji

**Y3:** Geleceğe hazır olma

olarak belirlenmiştir.

Tablo 1'de Kanonik Korelasyon analizi sonucunda iki adet kanonik değişken çifti elde edilmiştir. Hesaplanan kanonik korelasyon katsayıları, Wilk's  $\lambda$  değerleri, hesaplanan ki-kare değerleri ile anlamlılık testleri sonuçları da Tablo 1'de gösterilmiştir. Tablo 1'den görüleceği gibi 3 kanonik değişken için hesaplanan korelasyon katsayılarının hepsinin anlamlı olduğu söylenebilir ( $p < 0.05$ ). Tablo 1'deki sonuçlara göre, ilk kanonik değişken çiftleri arasındaki kanonik korelasyon dikkate alınarak, insani gelişmişlik ve dijital rekabetçilik arasındaki ilişkinin yapısı ortaya konulmak istendiği zaman % 73.6 oranında gerçeğe yakın değerlendirme yapılabileceği ortaya konulmuş olmaktadır.

İlk üç kanonik fonksiyonların birikimli değerlerinden oluşan ilişkiye ait Wilk's  $\lambda$  değeri dikkate alınarak, insani gelişmişlik ve dijital rekabetçilik değişken setleri arasında paylaşılan ortak varyans 1-Wilk's  $\lambda=0.742$ 'dir. Başka bir ifade ile, ilk üç kanonik fonksiyon arasında paylaşılan ortak varyans %74'tür. İkinci ve üçüncü kanonik fonksiyonların birikimli değerlerinden oluşan ilişkiye ait Wilk's  $\lambda$  değerine göre, insani gelişmişlik ve dijital rekabetçilik veri setleri arasında paylaşılan ortak varyans 1-Wilk's  $\lambda=0.074$ 'tür. Kanonik değişkenler arasındaki ilişkinin en düşük olduğu üçüncü kanonik fonksiyonda insani gelişmişlik ve dijital rekabetçilik değişken setleri setleri yalnızca %.06'lik bir ortak varyans paylaşmaktadır (1-Wilk's  $\lambda=0.006$ ).

**Tablo 1.** İnsani Gelişmişlik ve Dijital Rekabetçilik değişken setleri için Kanonik değişkenlere ait korelasyon katsayıları, Kanonik  $R^2$ , Wilk's  $\lambda$  ve anlamlılık testleri

Kanonik korelasyon	Kanonik $R^2$	Wilk's $\lambda$	Ki-kare	sd	p
0.858	0.736	0.258	65.674	9	.000
0.546	0.298	0.926	32.575	4	.001
0.377	0.142	0.994	10.537	1	.048

İnsani gelişmişlik ve dijital rekabetçilik değişkenleri için, her bir değişkenin doğrusal kombinasyonunu oluşturan kanonik değişkenler ile gerçek değişkenler arasındaki standardize edilmiş korelasyonların yer aldığı kanonik katsayılar, Tablo 2'de verilmiştir. Standardize edilmiş kanonik katsayılar, bir değişken setindeki kanonik değişkenin oluşmasında, ilgili değişken setinde yer alan orjinal değişkenlerin katkısını göstermektedir.

Tablo 2'de insani gelişmişlik değişken setinde yer alan değişkenlerin kanonik değişkenleriyle olan ilişkisi;

$$U1 = -.004 * \text{Sağlık} - .003 * \text{Okullaşma} - .001 * \text{GSMH}$$

$$U2 = -.037 * \text{Sağlık} + .042 * \text{Okullaşma} + .001 * \text{GSMH}$$

$$U3 = -.012 * \text{Sağlık} + .011 * \text{Okullaşma} - .019 * \text{GSMH}$$

ile verilmiştir.

Tablo 2’de dijital rekabetçilik değişken setinde yer alan değişkenlerin kanonik değişkenleriyle olan ilişkisi;

$$V1 = .002 * \text{Bilgi} - .003 * \text{Teknoloji} - .006 * \text{Geleceğe hazır olma}$$

$$V2 = -.016 * \text{Bilgi} - .004 * \text{Teknoloji} + .015 * \text{Geleceğe hazır olma}$$

$$V3 = .011 * \text{Bilgi} - .018 * \text{Teknoloji} + .008 * \text{Geleceğe hazır olma}$$

olarak belirlenmiştir.

Kanonik değişkenlerin işaretleri dikkate alınmaksızın, insani gelişmişlik değişken setinde, sağlık standardı değişkeninin birinci kanonik değişkene en fazla katkı sağlayan değişken olduğu; okullaşma standardı değişkeninin ikinci kanonik değişkene; GSMH değişkeninin üçüncü kanonik değişkene en fazla katkı sağladığı görülmektedir. Dijital rekabetçilik değişken setinde, birinci kanonik değişkene en fazla katkı sağlayan değişkenin geleceğe hazır olma; bilgi değişkeni ise ikinci kanonik değişkene en fazla katkıyı; üçüncü kanonik değişkene de en fazla katkıyı sağlayan değişkenin teknoloji değişkeni olduğu görülmektedir.

**Tablo 2.** İnsani Gelişmişlik ve Dijital Rekabetçilik değişken setleri için Standardize Edilmiş Korelasyon Katsayıları

	İnsani Gelişmişlik				Dijital Rekabetçilik		
	1	2	3		1	2	3
<b>Sağlık Standardı</b>	-.004	-.037	-.012	<b>Bilgi</b>	.002	-.016	.011
<b>Okullaşma Standardı</b>	-.003	.042	.011	<b>Teknoloji</b>	-.003	-.004	-.018
<b>GSMH</b>	-.001	.001	-.019	<b>Geleceğe Hazır Olma</b>	-.006	.015	.008

İnsani gelişmişlik ve dijital rekabetçilik veri setleri arasındaki korelasyon 0.858 olarak bulunmuştur.

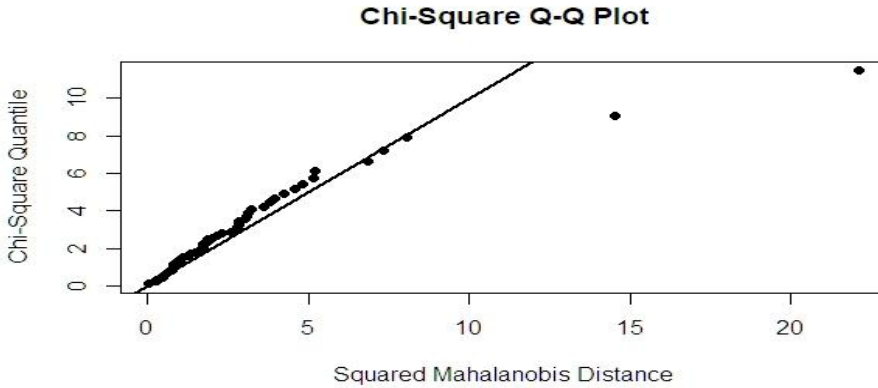
Sonraki aşamada elde edilen bu sonuçların güvenilir olup olmadığının saptanması için kanonik korelasyon analizinin dayandığı temel varsayım olan çok değişkenli normallik varsayımının kontrolü yapılmıştır. Bu amaçla, insani gelişmişlik ve dijital rekabetçilik değişken setlerindeki değişkenlerin yer aldığı veri seti için çok değişkenli normallik testi gerçekleştirilmiş ve elde edilen sonuçlar Tablo 3’te

verilmiştir. Tablo 3'e göre, 0,05 anlamlılık düzeyinde veri setlerinin çok değişkenli normal dağılımdan gelmediği görülmektedir ( $p < 0.05$ ).

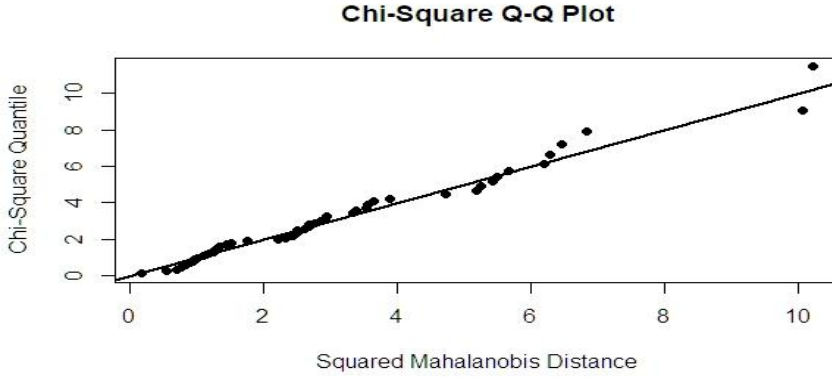
Ayrıca normallik varsayımının kontrolünde ve aykırı değerlerin tespitinde kullanılan Q-Q grafikleri her iki veri seti için çizilmiş ve Şekil 1-2'de verilmiştir. Şekil 1 ve 2 incelendiğinde, her iki veri setinde de aykırı değerlerin olduğu görülmektedir. Aykırı değerler, verinin kalan kısmından oldukça farklı yapıya sahip olan gözlemler olarak adlandırılmakta ve normallik varsayımını bozmaktadır.

**Tablo 3.** İnsani gelişmişlik ve Dijital Rekabetçilik değişken setleri için Çok değişkenli Normallik Testi Sonuçları

	Test	R	P	Çok Değişkenli Normal
<b>İnsani Gelişmişlik</b>	Royston	19.359	0.000	Uygun Değil
<b>Dijital Rekabetçilik</b>	Royston	15.539	0.043	Uygun Değil



**Şekil 1.** İnsani gelişmişlik değişkenleri için Q-Q Grafiği



**Şekil 2.** Dijital rekabetçilik değişkenleri için Q-Q Grafiği

Hem normallik testi sonucunda hem de Q-Q grafiği sonucunda veri setlerinin çok değişkenli normallik varsayımını sağlamadığı tespit edilmiştir. Dolayısıyla klasik kanonik korelasyon analizi sonuçları gerçeği yansıtmayacaktır. Bu nedenle aykırı değerlerin etkisini azaltan ve varsayımsal bozulmalara karşı dirençli olan sağlam Kanonik Korelasyon Analizi uygulanmıştır. Sağlam Kanonik Korelasyon analizi ile elde edilen sonuçlar Tablo 4’te verilmiştir.

**Tablo 4.** İnsani gelişmişlik ve Dijital rekabetçilik değişken setleri için Standardize Edilmiş Sağlam Kanonik Korelasyonlar

	İnsani Gelişmişlik				Dijital Rekabetçilik		
	1	2	3		1	2	3
<b>Sağlık Standardı</b>	.163	.168	-.928	<b>Bilgi</b>	-.226	.384	.861
<b>Okullaşma Standardı</b>	.986	.001	-.372	<b>Teknoloji</b>	.333	-.711	-.461
<b>GSMH</b>	.002	-.985	.003	<b>Geleceğe Hazır Olma</b>	.915	.589	-.217

Tablo 4’te insani gelişmişlik değişken setinde yer alan değişkenlerin sağlam kanonik değişkenleriyle olan ilişkisi incelendiğinde; 3 sağlam kanonik değişkene ilişkin oluşturulan denklemin şu şekilde kurulduğu görülmektedir.

$$U1 = .163 * \text{Sağlık} + .986 * \text{Okullaşma} + .002 * \text{GSMH}$$

$$U2= .168*Sağlık + .001*Okullaşma - .985*GSMH$$

$$U3= -.928*Sağlık - .372*Okullaşma + .003*GSMH$$

Tablo 4’de dijital rekabetçilik değişken setinde yer alan değişkenlerin sağlam kanonik değişkenleriyle olan ilişkisi incelendiğinde; 3 sağlam kanonik değişkene ilişkin oluşturulan denklem aşağıdaki gibidir.

$$V1= -.226*Bilgi+ .333*Teknoloji + .915*Geleceğe hazır olma$$

$$V2= .384*Bilgi - .711*Teknoloji + .589*Geleceğe hazır olma$$

$$V3= .861*Bilgi - .461*Teknoloji - .217*Geleceğe hazır olma$$

İnsani gelişmişlik değişken setinde, birinci kanonik değişkene en fazla katkı sağlayan değişkenin okullaşma; ikinci kanonik değişkene en çok katkıyı sağlayan değişkenin GSMH; üçüncü kanonik değişkene en fazla katkıyı sağlayan değişkenin de sağlık standardı olduğu görülmektedir. Dijital rekabetçilik değişken setinde, birinci kanonik değişkene en fazla katkı sağlayan değişkenin geleceğe hazır olma; teknoloji değişkeni ise ikinci kanonik değişkene en fazla katkıyı; üçüncü kanonik değişkene de en fazla katkıyı sağlayan değişkenin bilgi değişkeni olduğu görülmektedir.

İnsani gelişmişlik ve dijital rekabetçilik veri setleri arasındaki sağlam korelasyon 0.870 olarak bulunmuştur.

### **3. Sonuç ve Öneriler**

İnsanların sağlıklı bir yaşam, eğitim ve gelir imkanlarının olup olmadığının ölçülmesi Birleşmiş Milletlerin hazırladığı İnsani Gelişme Raporları ile ortaya konmaktadır. Bağımsız kuruluşlar tarafından hazırlanan Dijital Rekabetçilik Raporları da ülkelerin gelişmişlik ve dijital dünyadaki durumlarına ait sonuçlar hakkında bilgi vermektedir. Bu çalışmada, ülkeler için insani gelişmişlik ve dijital rekabetçilik veri setlerine Kanonik Korelasyon Analizi ve Sağlam Kanonik Korelasyon Analizleri uygulanmıştır.

Kanonik değişkenler arasındaki korelasyonun en yüksek olduğu birinci kanonik fonksiyon için, insani gelişmişlik ve dijital rekabetçilik veri setlerine ait ortak varyans %74.2 olarak bulunmuştur. Birinci kanonik korelasyonun çıkarılmasından sonra kalan iki kanonik fonksiyon için insani gelişmişlik ve dijital rekabetçilik veri setleri arasında paylaşılan ortak varyans %7.4’tür. İlk iki kanonik fonksiyonun çıkarılmasından sonra kalan üçüncü kanonik fonksiyon için insani gelişmişlik ve dijital rekabetçilik veri setleri %.06’lik bir ortak varyans paylaşmaktadır.

Kanonik Korelasyon Analizi sonucunda, insani gelişmişlik değişken setinde birinci kanonik değişkene en fazla katkı sağlayan değişkenin sağlık standardı; ikinci kanonik değişkene en çok katkısını sağlayan değişkenin okullaşma standardı; üçüncü kanonik değişkene en fazla katkısını sağlayan değişkenin GSMH olduğu görülmektedir. Sağlam Kanonik Korelasyon analizi sonucunda ise, birinci kanonik değişkene en fazla katkı sağlayan değişkenin okullaşma; ikinci kanonik değişkene en fazla katkısını GSMH; üçüncü kanonik değişkene en çok katkısını sağlık standardı değişkeninin olduğu belirlenmiştir. Veri setinde aykırı değer bulunduğundan sağlam kanonik korelasyon analizi sonuçları daha güvenilirdir. Ayrıca, klasik kanonik korelasyon analizindeki standardize edilmiş katsayılar zayıf bir ilişkiyi ifade ederken, sağlam kanonik korelasyon analizinde bulunan standardize edilmiş katsayılar ilişkinin daha güçlü olduğunu ifade etmektedir. Aykırı değerlerin etkisinin giderilmesi ilişkinin daha doğru bir şekilde ortaya konmasına neden olmuştur.

Benzer şekilde, Kanonik Korelasyon Analizi sonucunda, dijital rekabetçilik değişken setinde birinci kanonik değişkene en fazla katkı sağlayan değişkenin geleceğe hazır olmadır. Bilgi değişkeni ise ikinci kanonik değişkene en fazla katkısını sağlamaktadır. Üçüncü kanonik değişkene en fazla katkısını sağlayan değişkenin teknoloji değişkeni olduğu görülmüştür. Sağlam Kanonik Korelasyon analizi sonucunda ise, birinci kanonik değişkene en fazla katkı sağlayan değişkenin benzer şekilde geleceğe hazır olma olduğu tespit edilmiştir. Teknoloji değişkeninin ikinci kanonik değişkene en fazla katkısını sağladığı görülmüştür. Üçüncü kanonik değişkene de en fazla katkısını sağlayan değişken ise bilgi değişkenidir. Dijital rekabetçilik değişken seti için de, sağlam kanonik korelasyon analizinde bulunan standardize edilmiş katsayılar klasik kanonik korelasyon analizinde elde edilen katsayılara göre daha büyüktür. Bu durum değişkenler arasındaki ilişkinin aslında klasik yöntemle belirlenenden daha güçlü olduğunu göstermektedir.

Kanonik Korelasyon Analizi için insani gelişmişlik ve dijital rekabetçilik setleri arası korelasyon değeri 0.858, Sağlam Kanonik Korelasyon Analizi için aynı setler arasındaki korelasyon değeri ise 0.870 olarak bulunmuştur.

Yukarıdaki tüm açıklamalar ışığında, sağlam kanonik korelasyon analizi ile elde edilen sonuçların daha güvenilir ve gerçekçi olduğu görülmektedir. İnsani olarak gelişmiş toplumların, dijital dünyada da oldukça güçlü olacakları ve benzer şekilde dijital gelişmişliği takip eden toplumlarında insanlığın gelişimine katkı sağlayacağı istatistiksel olarak ortaya çıkarılmıştır. Ayrıca, kanonik korelasyon analizinin varsayımlarının sağlanmaması durumunda önerilmiş olan sağlam



kanonik korelasyon analizinin kullanılması, uygulamacılara daha güvenilir ve gerçekçi bilgi elde etmeleri konusunda ışık tutacaktır.

### **Kaynaklar**

- Akder, H. (1992). “İnsanca Gelişme Raporu ve Türkiye Profili”, İnsanca Gelişme I. Türkiye Konferansı, UNDP.
- Alpar, R. (2011). Uygulamalı Çok Değişkenli İstatistiksel Yöntemler, 3. Baskı, Ankara: Detay Yayıncılık.
- Bayraktar, Y. & Kaya, H. İ. (2016). “Yeni Ekonomi ve Değişen Rekabet Anlayışı: Karşılaştırmalı Bir Analiz”, Bilgi Ekonomisi ve Yönetimi Dergisi, Cilt: 11, Sayı:1.
- Croux, C. & Dehon, C. (2012). “Analyse Canonique Base Sur Des Estimateurs Robustes De La Matrice De Covariance”, La Revue de Statistique Applique, Vol: 2.
- Dehon, C., Filzmoser, P. & Croux, C. (2000). Robust Methods for Canonical Correlation Analysis. Data Analysis, Classification, and Related Methods.
- Demir, S. (2006). “Birleşmiş Milletler Kalkınma Programı İnsani Gelişme Endeksi ve Türkiye Açısından Değerlendirme”, DPT Sosyal Sektörler ve Koordinasyon Genel Müdürlüğü.
- Dikmen, M. K. (2004): “Büyüme ve İnsani Gelişme Arasındaki Bağlantılar: Uluslararası Ampirik Bir Çalışma”, İktisat İşletme ve Finans Dergisi, Cilt: 19, Sayı: 222.
- Fırat, E. & Aydın, A. (2015). “İnsani Kalkınma Endeksine Göre Türkiye'nin Eğitim Endeks Göstergelerinin OECD Ülkeleri ile Karşılaştırılması”, Selçuk Üniversitesi İ.İ.B.F. Sosyal Ekonomik Araştırmalar Dergisi, Sayı: 29.
- Günel, F.E., Çağlar, A., Uyar, S. G., Karadeniz, O. & Yeşilyurt, M. E. (2017). “Türkiye’de İllere Göre İnsani Gelişme Endeksi”, Pamukkale Üniversitesi SBE Dergisi, Sayı: 27.
- Henze, N. & Zirkler, B. (1990). “A Class of Invariant Consistent Tests for Multivariate Normality”, Communications in Statistics-Theory and Methods, 1990, 19(10).
- IMD World Digital Competitiveness Ranking (2017).

- Işığışık, E. (1999). “Kanonik Korelasyon Çözümlemesi: Bursa'daki 500 Büyük Firmanın Girdi ve Çıktı Değişkenleri Üzerine Bir Uygulama”, Marmara Üniversitesi İİBF Ekonometri Bölümü, IV. Ulusal Ekonometri ve İstatistik Sempozyumu Bildirileri.
- İlhan, M., Çetin, B., Sünkür, M. & Yılmaz, F. (2013). “Ders Çalışma Becerileri ile Akademik Risk Alma Arasındaki İlişkinin Kanonik Korelasyon ile İncelenmesi”, Eğitim Bilimleri Araştırmaları Dergisi, Cilt: 3, Sayı:2.
- Karagöz, Y. (2016). SPSS ve AMOS 23 Uygulamalı İstatistiksel Analizler, Nobel Yayınevi, Ankara.
- Karnel, G. (1991). Robust Canonical Correlation and Correspondence Analysis, In: The Frontiers of Statistical Scientific and Industrial Applications. American Sciences Press.
- Kalkan, S. & Özden, Ü. H. (2017). “Dünya Üniversitelerinin İtibarını Etkileyen Değişkenlerin Kanonik Korelasyon Analizi ile Belirlenmesi”, Social Sciences Research Journal, Volume:6, Issue:2.
- Keskin, S., Kor, A. & Başpınar, E. (2005). “Akkeçi Oğlaklarında Kesim Öncesi ve Kesim Sonrası Ölçülen Bazı Özellikler Arasındaki İlişki Yapısının Kanonik Korelasyon Analizi ile İrdelenmesi”, Tarım Bilimleri Dergisi, 11(2).
- Korkmaz, S. & Göksülük, D. R. (2014). Programming Language MVN Package. R Programming.
- Mardia, K.V. (1970). “Measures of Multivariate Skewness and Kurtosis with Applications”, Biometrika, 57(3).
- Royston, P. (1992). “Approximating the Shapiro-Wilk W-Test for Non-Normality”, Statistics and Computing, 2(3).
- Thompson, B. (1985). Canonical Correlation Analysis. Sage Publication Ltd., London.
- Ünal, Ç. (2008). “İnsani Gelişmişlik Endeksine Göre Türkiye'nin Bölgesel Farklılıkları”, Coğrafi Bilimler Dergisi, Cilt: 6, Sayı: 2.
- Yumuşak, İ. G. & Tuna, Y. (2002). “Kalkınmışlık Göstergesi Olarak Beşeri Kalkınma İndeksi ve Türkiye Üzerine Bir Değerlendirme”, İstanbul Üniv. İktisat Fakültesi Mecmuası, Cilt: 52, Sayı: 1.