

Bulanık Kümeleme Analizi İle Türkiye ve AB Ülkelerinin Eğitim Göstergeleri Açısından Sınıflandırılması

Nesrin ALPTEKİN

Anadolu Üniversitesi İşletme Fakültesi, İşletme Bölümü Öğretim Üyesi,
8-8A

Özet: Küreselleşen dünyada, bilgiye erişebilen ve bilgiyi üretebilen ülkeler rekabet avantajı sağlamaktadır. Bilgiye erişmek ve bilgiyi üretme yeterliliğine sahip olmak eğitimle gerçekleşmektedir. Bu çalışma beşeri sermayenin bileşenlerinden biri olan eğitime ilişkin göstergelere göre AB üye ülkelerini ve Türkiye'yi k-ortalamlar ve bulanık c-ortalamlar kümeleme yöntemleri yardımıyla sınıflandırmaktadır. Bu çalışma ayrıca eğitim göstergeleri açısından Avrupa Birliği üye ülkelerine göre Türkiye'nin konumunu belirlemektedir. k-ortalamlar kümeleme analizinde dört kümeye ayrılırken, bulanık kümeleme analizinde ülkeler iki kümeye ayrılmıştır. AB üyesi ülkelere göre okul öncesi eğitim, yükseköğretimdeki öğrenci, yükseköğretimdeki kız öğrenci ve 18 yaşında eğitim almaya devam edenlere ait oranlar oldukça düşüktür. Elde edilen sonuçlara göre, eğitime katılımın artırılmasına yönelik iyileştirmelerin yapılması gerektiği sonucuna ulaşılmıştır.

Anahtar Kelimeler: Eğitim göstergeleri, Avrupa Birliği, Bulanık c- ortalamlar kümeleme analizi, k-ortalamlar kümeleme analizi

Classifying of Turkey and EU Countries by Using Fuzzy Clustering Analysis in Terms of Education Indicators

Abstract: In a globalizing world, countries that can access information and produce information provides a competitive advantage. To access to information and have sufficiency to produce knowledge are realized by education .This study classify the member states of EU and Turkey, according to the education indicators which is one of the components of human capital by using k-means clustering and fuzzy c-means clustering methods. This study also investigates the position of Turkey compared to the member states of European Union in terms of education indicators. In fuzzy c-means clustering analysis, countries are classified into two clusters, while they are classified in four clusters according to k-means clustering analysis. According to the EU member states, ratios of pre-school education, higher education students, students in higher education and the girls who continued to receive education in 18 -year-old in Turkey are very low. According to the obtained results, the reached conclusion is that it should be made improvements to increase participation in education.

Keywords: Education Indicators, European Union, Fuzzy c-means clustering analysis, k-means clustering analysis

e-mail:nesrinesen@anadolu.edu.tr

Giriş

Bir ülkenin ekonomik, siyasal ve sosyal gelişmişlik düzeyini belirleyen en önemli faktörlerden biri olan eğitimin amacı, bilgi seviyesi yüksek evrensel kültüre sahip sağlıklı bir toplum yetiştirmektir. Bir ülkenin refahı, o ülke insanların sürekli ve nitelikli bir eğitim alarak kazandıkları bilgi ve beceri ile ekonomik büyümeye yapabilecekleri katkıya bağlıdır. Bu nedenle sosyo-ekonomik gelişmenin ve verimlilik artışının en önemli unsuru, toplumun eğitim düzeyi olmaktadır (Çakmak, 2008: 35).

Günümüzde ülkelerin kalkınmışlık düzeyi milli gelirin yanı sıra, eğitim, sosyal, kültürel ve politik durumları ile de ölçülmektedir. İktisadi gelişme kişi başına düşen mal ve hizmet birimleriyle ifade edilebildiği gibi, kişi başına düşen eğitim ve sağlık harcamaları da gelişmişliğin önemli göstergeleri arasındadır. Buna paralel olarak bir ülkenin gelişmişliğinin temel göstergeleri olan okuryazarlık, okullaşma oranı ve ortalama yaşam süresi de kalkınmanın merkezine insanı yerleştirmektedir (Hoşgörür ve Gezgin, 2005: 5).

Gün geçtikçe artan bilgi artışı ve hızlı teknolojik gelişmeler nedeniyle Avrupa kıtasında bulunan ülkelerin diğer güçlü devletler karşısında gücünü muhafaza etmek ve dünya siyasetinde daha etkin olmak amacıyla bir araya gelmelerine neden olmuştur. İlk başlarda ekonomik bir

birliktelik olarak görülen Avrupa topluluğu daha sonraki yıllarda kültür, eğitim gibi sosyal konularda da ortak projeler geliştirmeye başlamıştır (Akbaş ve Özdemir, 2014).

AB'ye üye ülkeler, Türkiye'nin üye olma koşullarını yerine getirmesinde Türkiye'nin ekonomik ve sosyal gelişimini sağlamak ve AB kriterlerine uygun duruma gelmesi için çeşitli öneriler sunmaktadırlar. Türkiye'nin bu koşullara uygun hale gelmesinde eğitimin önemi oldukça büyüktür. AB'ye tam üyelik yolunda insan kaynağının eğitimi de önemli bir adım olacaktır (Taş ve Yenilmez, 2008: 162).

AB'ye aday ülke konumunda olan Türkiye, eğitimde kaliteyi arttırmak için çalışmalarını sürdürmektedir. Türkiye, Avrupa Birliği Eğitim ve Gençlik Programlarında 2004 yılından itibaren tam üye konumundadır.

Bu çalışmada, AB'ye uyum sürecindeki Türkiye'nin, AB ülkeleri ile temel eğitim göstergeleri açısından ne derece benzediği ya da birbirinden ayrıldığı, k-ortalamlar ve bulanık c-ortalamlar kümeleme analizi yöntemleriyle araştırılmaya çalışılmıştır.

1. Analiz Yöntemleri

Bu çalışmada, eğitim göstergeleri açısından Türkiye ve Avrupa Birliği Üye Ülkelerinin karşılaştırılmasında hiyerarşik olmayan kümeleme yöntemlerinden k-ortalamlar ve bulanık c-ortalamlar yöntemleri kullanılmıştır. Çok değişkenli bir istatistiksel

analiz yöntemi olan kümeleme analizinin öncelikli amacı birey ya da birimlerin temel özelliklerini dikkate alarak onları gruplandırmak ve böylece araştırmacıya özetleyici bilgiler sunmaktır (Kalaycı vd., 2005: 349).

Kümeleme analizi, X veri matrisinde yer alan ve doğal gruplamaları bilinmeyen birimleri ya da değişkenleri birbirleri ile benzer olan alt kümelere (grup, sınıf) ayırmaya yardımcı olan yöntemler topluluğudur. Kümeleme yöntemleri, uzaklık matrisinden yararlanarak, birimleri veya değişkenleri kendi içinde homojen ve kendi aralarında heterojen olacak şekilde gruplar oluşturmaya olanak sağlamaktadır (Koldere, 2008: 8). Kümeleme sonucunda, birimler küme içinde yüksek homojenlik, kümeler arasında yüksek heterojenliğin olması eğer kümeleme işlemi başarılı ise geometrik gösterimde küme içerisinde yer alan birimler birbirine yakın iken, farklı kümelerde yer alan birimlerin birbirinden uzakta yer almalarını sağlayacaktır (Akın ve Eren, 2012: 176).

Bağımlı ve bağımsız değişken ayırımı olmayan içsel bağımlılık içeren bir yöntem olan kümeleme analizinde, kümeler arasında yakınlıklar veya uzaklıkların hesaplanmasında hiyerarşik ve hiyerarşik olmayan olmak üzere iki farklı yöntem kullanılmaktadır. Hiyerarşik ve hiyerarşik olmayan yaklaşımların her ikisinde de amaç kümeler arasındaki farklılıkları ve kümeler içi

benzerlikleri en yüksek düzeye çıkarmaktır. Bir diğer deyişle, küme içi homojenliği artırırken, kümeler arası homojenliği azaltmaktır.

Kümeleme analizinde, hiyerarşik yöntemler birleştirici (toplamalı) yöntemler ve ayırıcı (bölünmeli) yöntemler olarak ikiye ayrılmaktadır. Uygulamada birleştirici yöntemler ayırıcı yöntemlere göre daha sık kullanılmaktadır. Hiyerarşik olmayan kümeleme başlığı altında pek çok yöntemden söz edilmektedir. Ancak bunlardan en çok kullanılanı MacQueen tarafından geliştirilmiş olan k -ortalamalar yöntemidir (Uzgören vd., 2013: 123).

2.1. k -ortalamalar Yöntemi

En çok bilinen kümeleme yöntemlerinden biri olan k -ortalamalar yöntemi, büyük sayıdaki birimlerden küçük sayıda kümeler oluşturmak için kullanılmaktadır. Bu yöntemin atama mekanizması, her bir verinin sadece bir kümeye ait olmasına izin vermektedir. Bu nedenle kesin bir kümeleme yöntemidir. Eşit büyüklükte küresel kümeleri bulmaya eğilimli olan bu yöntem, merkez noktanın kümeyi temsil etmesine dayalı bir kümeleme yöntemidir (Işık ve Çamurcu, 2007: 32). Bu yöntemde her iterasyonda yeni bir küme merkezi oluşturulur ve bir eleman yeniden hesaplanan yeni merkeze daha yakın ise o kümeye taşınır (Atalay ve Tortum, 2010: 337).

k-ortalamlar yönteminin aşamaları aşağıdaki gibidir (Ross, 2004: 100) :

- Birimler k adet kısma ayrılır.
- Her kısımdaki birimlerin dikkate alınan değişkenlerin her biri itibariyle ortalamaları alınır. Birimlerin, her grubun ortalama vektörüne uzaklıkları (genellikle Öklid uzaklığı) hesaplanır. Eğer birim, başlangıçta tanımlanan grubun merkezine en yakın ise grubunda kalır, aksi takdirde hangi grubun merkezine yakın ise o gruba katılır.
- Merkezler, yeni katılımlardan sonra tekrar hesaplanarak ikinci adım

tekrarlanır. Tekrar işlemi, değişiklik olmayana kadar devam eder.

- k-ortalamlar yönteminin amacı, kümeler içi kareler toplamını minimize ederek, yani küme içi değişkenliği en aza indirgeyerek p boyutlu değişkene sahip birimi k tane kümeye bölmektir (Koldere, 2008: 16-17).

x_1, x_2, \dots, x_n değişkenlerinin her biri, X uzayında birer noktayı ifade eden p değişkenli gözlem vektörleri olmak üzere, $a_{1n}, a_{2n}, \dots, a_{kn}$ aynı uzayda her gözlem grubu için küme merkezleri olarak belirlendiğinde

$$W_n = \frac{1}{n} \sum_{i=1}^n \min_{1 \leq j \leq k} \|x_i - a_{jn}\|^2 \quad (1)$$

formülü kullanılarak birimler en yakın kümeye atanmaktadır. k-ortalamlar yönteminin başarısı, uygun küme sayısının belirlenmesine bağlıdır. En pratik yöntem olarak küme sayısını $k \cong (n/2)^{1/2}$ olarak belirlemektir. Ancak örneklem sayısı büyüdükçe bu eşitlik iyi sonuçlar vermemektedir (Turanlı vd., 2006: 100).

2.2. Bulanık c-ortalamlar Yöntemi

Kümeleme analizinde, elemanların ya da birimlerin bir kümeye aitliği kesin, bulanık ve olasılıklı olmak üzere üç farklı durumda incelenebilir. Kesin kümelemede, birimler bir kümeye aittirler veya değildirler. Bulanık kümelemede, birimlerin aynı anda birden fazla kümeye aitliği söz konusu olabilir.

Olasılıklı kümeleme analizinde ise, bir birim bir kümeye aittir veya değildir, fakat bir birimin bir kümeye atanması bir olasılık dağılımına bağlıdır (Torra, 2005: 646).

Bulanık kümelemede Bezdek ve Hathaway (1987) tarafından ileri sürülen ve Kaufman ve Rousseeuw (1990) tarafından geliştirilen, bu çalışmada kullanılacak olan bulanık c-ortalamlar yöntemine ait algoritma

$$u_{ik} \geq 0, \sum_{j=1}^N u_{jk} = 1, i, j = 1, 2, \dots, n \text{ ve } k = 1, 2, \dots, K \quad (2)$$

kısıtları altında aşağıdaki amaç fonksiyonunu minimize etmeye çalışır.

$$C = \sum_{k=1}^K \frac{\sum_{i=1}^N \sum_{j=1}^N u_{ik}^2 u_{jk}^2 d(i,j)}{2 \sum_{j=1}^N u_{jk}^2} \quad (3)$$

Burada u_{ik} , i . birimin k . kümedeki bilinmeyen üyelik derecesini, u_{jk} , j . birimin k . kümedeki bilinmeyen üyelik olasılığını ve $d(i,j)$ ' de x_i ve x_j birimleri arasındaki farklılığı veya uzaklığı göstermektedir. x_i ve x_j birimleri arasındaki farklılığı veya uzaklığı gösteren $d(i,j)$, Öklid uzaklığı olarak tanımlıdır. Bulanık kümelemede her bir birimin tüm kümelere olan üyelik katsayıları toplamı daima bir olacak şekilde pozitifdir.

Bulanık kümelemede oluşan kümelere hangi küme sayısının uygun kümeleme olduğunu belirlemek için gölge istatistiği (silhouette statistics, SC) ve ortalama gölge istatistiğinden (average silhouette statistics, \overline{SC}) yararlanılır (Özdamar, 2004: 345).

Gölge istatistiği

$$s(i) = \frac{b(i) - a(i)}{\max[a(i), b(i)]}, \quad -1 \leq s(i) \leq 1 \quad (4)$$

formülü ile hesaplanır. Burada $a(i)$, i . birimin aynı küme içerisindeki diğer tüm birimlere olan uzaklıklarının ortalaması, $b(i)$ ise, i . birimin en yakın komşu olduğu ve elemanları arasındaki ortalama uzaklığın en küçük olduğu bir kümedeki elemanlar ile i . birimin uzaklıklarının ortalamasıdır (Tibshirani vd., 2001: 418).

Ortalama gölge istatistiği, bir kümedeki tüm birimlere ait gölge istatistiklerinin ortalaması olarak hesaplanır ve bulanık kümelemede uygun küme sayısını belirlemede maksimum değerli ortalama gölge istatistiğinden

yararlanılır (Özdamar, 2004:345). Bir veri setinde uygun kümeleme yapısı olması için ortalama gölge istatistiği değerinin, \overline{SC} , en az 0.50 olması beklenir.

Bulanık kümelemenin kesin kümeden ne kadar uzaklıkta olduğunun bir göstergesi Dunn ayrıştırma katsayısıdır. Bu katsayı elde edilen kümenin ne kadar bulanık olduğuna dair bir fikir vermektedir. Dunn ayrıştırma katsayısı, tüm üyelik katsayılarının, u_{ik} , kareler toplamının birim sayısına bölünmesiyle aşağıdaki gibi elde edilir (Kılıç vd., 2011: 33):

$$F(U) = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K u_{ik}^2 \quad (5)$$

$F(U)$ değeri daima $\left[\frac{1}{k}, k\right]$ aralığında yer alır.

Böylece birimlere ilişkin üyelik matrisi elde edilir. Küme sayısından bağımsız olarak Dunn ayrıştırma katsayısı

$$F_k(U) = \frac{K \cdot F(U) - 1}{K - 1} \quad (6)$$

formülü ile normalleştirilir. Normalleştirilmiş Dunn katsayısı, $F_k(U)$, 0 ile 1 arasında değerler alır. $F_k(U)$ değeri 1'e yaklaştıkça kesin kümelemeyi, 0' a yaklaştıkça da veri setindeki tam bulanıklığı göstermektedir. Bulanık kümelemede kullanılan bir diğer ayrıştırma katsayısı da, Kaufman ayrıştırma katsayısıdır:

$$D(U) = \frac{1}{N} \sum_{k=1}^K \sum_{i=1}^N (h_{ik} - m_{ik})^2 \quad (7)$$

ve aşağıdaki formül yardımıyla normalleştirilmiştir:

$$D_k(U) = \frac{K \cdot D(U) - 1}{K - 1} \quad (8)$$

Bulanık kümelemede optimal küme sayısı için, $F_k(U)$ ve $D_k(U)$ birlikte iyi birer göstergelerdir. $F_k(U)$ büyük ve $D_k(U)$ küçük olacak şekilde K küme sayısı belirlenmelidir.

2. Analiz ve Bulgular

Türkiye'nin temel eğitim göstergeleri açısından AB Ülkelerine göre konumunu

belirlemek ve benzerlik/farklılık gösteren ülkeleri saptamak için yapılan bu çalışmada, analizde kullanılan değişkenler Avrupa Birliği'nin resmi web sitesi olan Eurostat'tan elde edilmiştir.

Ülkelere ait temel eğitim göstergeleri, finansal ve finansal olmayan göstergeler olarak iki başlık altında toplanmıştır. En güncel göstergeler finansal olmayanlar göstergeler olarak 2011 yılına ait verilerdir. Bu nedenle analizde kullanılan veriler, eksik gözlem içermeyen ve 2011 yılına ait finansal olmayan 7 adet gösterge ile sınırlandırılmıştır:

Tablo 1. Değişkenler

| Değişkenler | Değişkenlerin Açıklaması |
|-------------|--|
| X_1 | Okul süresi beklentisi(yıl) |
| X_2 | 18 yaşında eğitim almaya devam edenlerin oranı (%) |
| X_3 | Okul öncesi eğitim alanların oranı (%) |
| X_4 | Öğrenci oranı (%) |
| X_5 | Yükseköğretimdeki kız öğrencilerin oranı (%) |
| X_6 | Yükseköğretimdeki öğrenci oranı (%) |
| X_7 | Meslek lisesinde eğitim gören erkek öğrencilerin oranı (%) |

Tablo 2. Temel Eğitim Göstergeleri

| Ülkeler | X1 | X2 | X3 | X4 | X5 | X6 | X7 |
|-------------|------|-------|-------|------|------|------|------|
| Almanya | 18,1 | 90,1 | 96,4 | 19,7 | 50,6 | 16,5 | 16,3 |
| Avusturya | 17,1 | 72,3 | 94,3 | 20,5 | 53,4 | 16,1 | 12,1 |
| Belçika | 19,5 | 89,9 | 98,1 | 26,4 | 55,3 | 12,6 | 12,4 |
| Bulgaristan | 16,4 | 81,7 | 86,6 | 17,7 | 55,1 | 12,4 | 17,5 |
| Çek Cumh. | 18,1 | 90,3 | 87,8 | 20,5 | 57,2 | 16,6 | 18,7 |
| Danimarka | 19,6 | 84,8 | 97,9 | 26,5 | 57,6 | 17,9 | 11,8 |
| Estonya | 17,8 | 87,3 | 89,1 | 21,9 | 59,7 | 11,9 | 16,3 |
| Finlandiya | 20,6 | 93,8 | 74,0 | 26,0 | 54,0 | 21,2 | 13,7 |
| Fransa | 16,4 | 76,6 | 100,0 | 23,0 | 54,8 | 22,1 | 18,4 |
| Hırvatistan | 15,8 | 66,9 | 70,6 | 18,3 | 57,3 | 12,3 | 14,3 |
| Hollanda | 19,0 | 89,1 | 99,6 | 23,1 | 51,8 | 9,4 | 20,6 |
| İngiltere | 16,9 | 61,3 | 97,0 | 22,8 | 56,4 | 19,5 | 19,9 |
| İrlanda | 17,2 | 94,9 | 96,1 | 25,9 | 51,7 | 21,1 | 15,7 |
| İspanya | 17,8 | 79,8 | 100,0 | 21,5 | 53,9 | 16,8 | 13,2 |
| İsveç | 19,8 | 95,8 | 95,3 | 26,2 | 59,1 | 15,6 | 11,3 |
| İtalya | 16,9 | 79,1 | 96,8 | 18,5 | 57,6 | 12,8 | 11,7 |
| Kıbrıs | 14,9 | 34,4 | 85,0 | 20,5 | 49,8 | 7,2 | 13,6 |
| Letonya | 18,1 | 95,6 | 92,7 | 21,4 | 61,1 | 12,8 | 11,4 |
| Litvanya | 19,1 | 100,1 | 84,2 | 24,2 | 59,0 | 22,6 | 9,9 |
| Lüksemburg | 15,0 | 70,4 | 95,6 | 19,9 | 51,9 | 3,0 | 9,9 |
| Macaristan | 17,7 | 89,2 | 94,5 | 21,2 | 55,9 | 8,5 | 10,7 |
| Malta | 16,0 | 58,2 | 100,0 | 19,8 | 56,0 | 6,3 | 12,9 |
| Polonya | 18,1 | 93,2 | 78,4 | 22,4 | 59,9 | 17,5 | 11,0 |
| Portekiz | 18,4 | 77,5 | 95,4 | 22,0 | 53,4 | 17,3 | 11,2 |
| Romanya | 16,1 | 73,3 | 82,0 | 19,7 | 56,0 | 16,0 | 17,8 |
| Slovakya | 16,5 | 84,5 | 76,9 | 20,8 | 59,6 | 18,0 | 16,9 |
| Slovenya | 18,4 | 92,1 | 89,8 | 20,5 | 60,6 | 17,4 | 16,0 |
| Yunanistan | 16,4 | 70,4 | 74,6 | 19,3 | 49,3 | 13,4 | 12,5 |
| Türkiye | 16,3 | 50,4 | 43,1 | 28,0 | 45,2 | 9,4 | 21,0 |

Kaynak: http://epp.eurostat.ec.europa.eu/portal/page/portal/education/data/main_tables,

Erişim Tarihi (24.03.2014).

Küme sayısı $k \cong (n/2)^{1/2}$ yaklaşımından kullanılarak yapılan kümeleme analizi

$k \cong (28/2)^{1/2} = 3,74 \cong 4$ olarak sonucunda oluşan kümeler Tablo 3'de

belirlenmiştir. SPSS 15.0 programı verilmiştir.

Tablo 3. k = 4 için k-Ortalamlar Yöntemine Göre Kümeleme

| Küme Numarası | Kümelerdeki Ülke Sayısı | Küme Elemanları |
|----------------------|--------------------------------|--|
| 1 | 12 | Almanya, Danimarka, Estonya, Fransa, İrlanda, İspanya, İsveç, Letonya, Litvanya, Macaristan, Polonya, Portekiz |
| 2 | 12 | Avusturya, Belçika, Bulgaristan, Çek Cumhuriyeti, Finlandiya, Hırvatistan, Hollanda, İtalya, Lüksemburg, |
| 3 | 3 | Kıbrıs, Malta, İngiltere |
| 4 | 2 | Yunanistan, Türkiye |

Normallik testine ait sonuçlar Tablo 4' de olduğundan küme sayısı 4 olarak verilmiştir. Kolmogorov-Smirnov normallik belirlendiğinde uzaklık değerlerinin normal testi sonuçlarına göre, $p = 0,636 > 0,05$ dağılım gösterdiği gözlenmiştir.

Tablo 4. Kolmogorov-Smirnov Normallik Testi Sonuçları

| | | Uzaklıklar |
|-----------------------------|-----------------------|-------------------|
| N | | 29 |
| Normal Parametreler | Ortalama | 14,7666 |
| | Standart Sapma | 4,37914 |
| Ekstrem Uzaklıklar | Mutlak | 0,138 |
| | Pozitif | 0,138 |
| | Negatif | -0,70 |
| Kolmogorov-Smirnov Z | | 0,745 |
| | | 0,636 |

Tablo 5'de küme merkezleri arasındaki uzaklıklar gösterilmektedir. Bu tabloya göre en çok birinci ve ikinci kümenin birbirine yakın, en çok da ikinci ve üçüncü kümenin birbirinden uzak olduğu söylenebilir.

Tablo 5. Küme Merkezleri Arasındaki Uzaklıklar

| Küme | 1 | 2 | 3 | 4 |
|------|--------|--------|--------|--------|
| 1 | | 27,269 | 39,340 | 45,996 |
| 2 | 27,269 | | 48,624 | 48,079 |
| 3 | 39,340 | 48,624 | | 37,800 |
| 4 | 45,996 | 48,079 | 37,800 | |

Tablo 6’da ise kümeleme analizinde kullanılan 7 değişkene ait varyans analizi sonuçları verilmiştir.

Tablo 6. Varyans Analizi

| | Küme | | Hata | | F | p |
|----------------------|--------------------|-----|--------------------|-----|--------|-------|
| | Kareler Ortalaması | Sd. | Kareler Ortalaması | Sd. | | |
| X₁ | 5,166 | 3 | 1,722 | 25 | 3,000 | ,050 |
| X₂ | 1400,570 | 3 | 90,414 | 25 | 15,491 | ,000* |
| X₃ | 8,164 | 3 | 7,864 | 25 | 1,038 | ,393 |
| X₄ | 712,176 | 3 | 86,263 | 25 | 8,256 | ,001* |
| X₅ | 10,750 | 3 | 7,332 | 25 | 1,466 | ,248 |
| X₆ | 50,555 | 3 | 10,249 | 25 | 4,933 | ,008* |
| X₇ | 1981,933 | 3 | 70,574 | 25 | 28,083 | ,000* |

Not: *İstatiksel açıdan anlamlı olan değişkenler belirtilmiştir($p < 0,05$).

Varyans analizi sonucunda kümeleri belirlemede en etkin olan değişkenler ortaya çıkarılır. Elde edilen bulgular, kümeleme analizinde X₃(öğrenci oranı) ve X₅(Yükseköğrenimdeki öğrenci oranı) değişkenlerine ait p değerleri 0,05’den büyük olduğundan bu değişkenler için kümeler arasında anlamlı bir farklılık yoktur. Kümelerin belirlenmesinde diğer 5 değişken etkili olmaktadır.

Bulanık c-ortalamlar kümeleme analizinde NCSS 9 programı kullanılmıştır. AB'ye üye 28 ülke ve Türkiye'nin temel eğitim göstergelerine göre bulanık kümeleme yöntemi ile sınıflandırılmasında farklı küme sayıları ($k = 2,3,4, 5$) için elde edilen ortalama gölge istatistiği, Dunn ayrıştırma katsayısı, normalleştirilmiş Dunn katsayısı ile Kaufman ayrıştırma katsayısı olan $D(U)$ ve normalleştirilmiş Kaufman ayrıştırma katsayısı olan $D_k(U)$ değerleri Tablo 7'de verilmiştir:

Tablo 7 Bulanık C-Ortalamlar Yöntemine Göre Kümeleme Analizi Sonuçları

| Küme Sayısı | \overline{SC} | $F(U)$ | $F_k(U)$ | $D(U)$ | $D_k(U)$ |
|-------------|-----------------|--------|----------|--------|----------|
| 2 | 0,5031 | 0,5613 | 0,1227 | 0,2258 | 0,4516 |
| 3 | 0,3460 | 0,3822 | 0,0733 | 0,4094 | 0,6141 |
| 4 | 0,3155 | 0,3067 | 0,0756 | 0,4951 | 0,6602 |
| 5 | 0,2040 | 0,2520 | 0,0650 | 0,6022 | 0,7527 |

Tablo 7 incelendiğinde, analizde kullanılan verilere en uygun küme sayısının $k = 2$ olduğu görülmektedir.

Bulanık kümeleme analizinde, uygun küme sayısının belirlenmesinde daha önce de belirtildiği gibi, \overline{SC} nin en az 0.50 olması beklenirken, $F_k(U)$ büyük ve $D_k(U)$ küçük olmalıdır. $k = 2$ için, $\overline{SC} = 0,5031$, $F_k(U) = 0,1227$ ve $D_k(U) = 0,4516$ olmaktadır.

Bulanık c-ortalamlar kümeleme analizine göre 1. Kümede yer alan ülkeler; Avusturya, Bulgaristan, Fransa, Hırvatistan, İngiltere, İrlanda, İtalya, Kıbrıs, Lüksemburg, Malta, Romanya, Slovakya, Yunanistan ve Türkiye'dir.

2. Kümede yer alan ülkeler ise Almanya, Belçika, Çek Cumhuriyeti, Danimarka, Estonya, Finlandiya, Hollanda, İspanya, İsveç, Letonya, Litvanya, Macaristan, Polonya, Portekiz ve Slovenya olmuştur. Ülkelerin hangi kümede yer aldıklarını gösteren küme üyelik değerleri Tablo 8' de verilmiştir.

Tablo 8. Ülkelerin Küme Üyelik Değerleri

| Ülkeler | Küme | 1. Küme Olasılığı | 2. Küme Olasılığı |
|-----------------|------|-------------------|-------------------|
| Almanya | 2 | 0,308 | 0,692 |
| Avusturya | 1 | 0,6128 | 0,3872 |
| Belçika | 2 | 0,3171 | 0,6829 |
| Bulgaristan | 1 | 0,7417 | 0,2583 |
| Çek Cumhuriyeti | 2 | 0,3194 | 0,6806 |
| Danimarka | 2 | 0,3217 | 0,6783 |
| Estonya | 2 | 0,3942 | 0,6058 |
| Finlandiya | 2 | 0,3839 | 0,6161 |
| Fransa | 1 | 0,6964 | 0,3036 |
| Hırvatistan | 1 | 0,7092 | 0,2908 |
| Hollanda | 2 | 0,3269 | 0,6731 |
| İngiltere | 1 | 0,6465 | 0,3535 |
| İrlanda | 1 | 0,5608 | 0,4392 |
| İspanya | 2 | 0,3808 | 0,6192 |
| İsveç | 2 | 0,3347 | 0,6653 |
| İtalya | 1 | 0,6605 | 0,3395 |
| Kıbrıs | 1 | 0,6246 | 0,3754 |
| Letonya | 2 | 0,3127 | 0,6873 |
| Litvanya | 2 | 0,309 | 0,691 |
| Lüksemburg | 1 | 0,6035 | 0,3965 |
| Macaristan | 2 | 0,4442 | 0,5558 |
| Malta | 1 | 0,6714 | 0,3286 |
| Polonya | 2 | 0,314 | 0,686 |
| Portekiz | 2 | 0,2767 | 0,7233 |
| Romanya | 1 | 0,7358 | 0,2642 |
| Slovakya | 1 | 0,7251 | 0,2749 |
| Slovenya | 2 | 0,2712 | 0,7288 |
| Yunanistan | 1 | 0,7397 | 0,2603 |
| Türkiye | 1 | 0,6511 | 0,3489 |

Bulanık c-ortalamlar yöntemine göre uygun küme sayısı 2 olurken, k-ortalamlar yöntemine göre seçilen küme sayısı 4 olduğu durumda elde edilen sonuçlara baktığımızda, k-ortalamlar yöntemi sonucunda elde edilen üçüncü küme elemanları İngiltere, Kıbrıs, Malta ile dördüncü küme elemanları olan Yunanistan ve Türkiye'nin bulanık c-ortalamlar yöntemine göre tek bir kümede yer aldıkları görülmektedir. Her iki yöntem için aynı küme sayısına göre, ($k = 2$ için), elde edilen sonuçları karşılaştırabilmek için bulanık c-ortalamlar ve k-ortalamlar yöntemleriyle elde edilen kümeler sırasıyla Tablo 9 ve Tablo 10'da verilmiştir.

Tablo 9. $k = 2$ için Bulanık c-ortalamlar Yöntemine Göre Kümeleme

| Küme Numarası | Kümelerdeki Ülke Sayısı | Küme Elemanları |
|----------------------|--------------------------------|---|
| 1 | 14 | Avusturya, Bulgaristan, Fransa, Hırvatistan, İngiltere, İrlanda, İtalya, Kıbrıs, Lüksemburg, Malta, Romanya, Slovakya, Yunanistan, Türkiye |
| 2 | 15 | Almanya, Belçika, Çek Cumhuriyeti, Danimarka, Estonya, Finlandiya, Hollanda, İspanya, İsveç, Letonya, Litvanya, Macaristan, Polonya, Portekiz, Slovenya |

Tablo 10. $k = 2$ için k-ortalamlar Yöntemine Göre Kümeleme

| Küme Numarası | Kümelerdeki Ülke Sayısı | Küme Elemanları |
|----------------------|--------------------------------|--|
| 1 | 8 | Hırvatistan, İngiltere, Kıbrıs, Lüksemburg, Malta, Romanya, Yunanistan, Türkiye |
| 2 | 21 | Almanya, Avusturya, Belçika, Bulgaristan, Çek Cumhuriyeti, Danimarka, Estonya, Finlandiya, Fransa, Hollanda, İrlanda, İtalya, İspanya, İsveç, Letonya, Litvanya, Macaristan, Polonya, Portekiz, Slovenya, Slovakya |

Küme sayısı $k = 2$ için, iki yönteme göre elde edilen kümelerdeki ülkeler farklılık göstermektedir. Bu farklılığın nedeni, bulanık c-ortalamlar yönteminin k-ortalamlar yöntemine göre başlangıç değerlerinden daha az etkilenmesidir.

Bulanık c-ortalamlar yöntemi genellikle k-ortalamlar yöntemine göre daha kararlı sonuçlar üretmektedir. Ayrıca, bulanık c-ortalamlar yöntemi k-ortalamlar yöntemine göre sıra dışı verilerden daha az etkilenmektedir.

Sonuç ve Tartışma

Bu çalışmada AB üye ülkeleri ve Türkiye, 2011 yılına ait finansal olmayan temel eğitim göstergeleri bakımından k-ortalamlar ve bulanık c-ortalamlar kümeleme yöntemlerine göre gruplandırılmıştır. k-ortalamlar ve bulanık c-ortalamlar kümeleme analizi sonucu

elde edilen sonuçlara bakıldığında, ülkeler k-ortalamlar yöntemine göre dört, bulanık c-ortalamlar yöntemine göre de iki kümede gruplanmışlardır. k-ortalamlar ($k = 4$ için) yönteminde üçüncü ve dördüncü kümelerde yer alan ülkeler olan İngiltere, Kıbrıs, Malta, Yunanistan ve Türkiye'nin bulanık c-ortalamlar yöntemine göre yapılan analizde aynı kümede toplandıkları görülmüştür. Küme sayısı $k = 2$ için, her iki yönteme göre elde edilen kümelerdeki ülkeler ise farklılık göstermektedir. Bu farklılığın nedeni, bulanık c-ortalamlar yönteminin k-ortalamlar yöntemine göre başlangıç değerlerinden daha az etkilenmesidir. Türkiye'de AB üyesi ülkelere göre "18 yaşında eğitim almaya devam edenlerin oranı", "Okul öncesi eğitim alanların oranı", "Yükseköğretimdeki öğrenci oranı" ve "Yükseköğretimdeki kız öğrenci oranı" düşüktür. Elde edilen analiz sonuçlarına göre eğitime katılım oranının ve eğitimli insan sayısının eğitim seviyesinin artışına paralel olarak arttırılmasına yönelik iyileştirmelerin yapılması gerekmektedir.

Kaynaklar

- Ayşen AA 2002.** *Kümeleme Analizi ile AB Ülkeleri ve Türkiye' nin Sürdürülebilir Kalkınma Açısından Değerlendirilmesi.* Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 29 319-332.
- Akın H, Eren B, Eren Ö 2012.** *OECD Ülkelerinin Eğitim Göstergelerinin Kümeleme Analizi ve Çok Boyutlu Ölçekleme Analizi ile Karşılaştırmalı Analiz.* ÖNERİ 10, 175-181.
- Atalay A, Tortum A 2010.** *Türkiye'deki İllerin 1997 -2006 Yılları Arası Trafik Kazalarına Göre Kümeleme Analizi".* Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 16, 335-343.
- Bezdek JC, Richard JH 1987.** *Clustering with Relational c-means Partitions From Pairwise Distance Data.* Mathematical Modelling 9, 435-439.
- Bolat BA 2011.** *Eğitim Göstergeleri Açısından Türkiye ve Avrupa Birliği'ne Üye Ülkelerin Karşılaştırılması.* İşletme İktisadi Enstitüsü Dergisi-Yönetim 22, 61-77.
- Çakmak Ö 2006.** *Eğitimin Ekonomiye ve Kalkınmaya Etkisi.* D.Ü. Ziya Gökalp Eğitim Fakültesi Dergisi, 11, 33-41.
- Doring C, Lesot MJ, Rudolf K 2006.** *Data Analysis with Fuzzy Clustering Method.* Computational Statistics & Data Analysis, 51, 192-214.
- Erkekoğlu H 2007.** *AB' ye Tam Üyelik Sürecinde Türkiye' nin Üye Ülkeler Karşısındaki Görelî Gelişme Düzeyi.* Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 14, 28-50.
- Girginer N 2013.** *A. Comparison of the Healthcare Indicators of Turkey and The European Union Members Countries Using Multidimensional Scaling Analysis and Cluster Analysis.* İktisat İşletme ve Finans 28, 55-72.
- Hoşgörür V, Gezgin G 2005.** *Ekonomik ve Sosyal Kalkınmada Eğitim.* Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi Dergisi 2 ,1-11.

Işık M, Çamurcu AY 2007. *K-Means, K-Medoids ve Bulanık C-Means Algoritmalarının Uygulamalı Olarak Performanslarının Tespiti.* İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 6, 31-45.

Kaufman L, Rousseeuw P.J 1990. *Findings Groups in Data: An Introduction to Cluster Analysis.* John Wiley, USA.

Kılıç İ 2011 *Bulanık Kümeleme Analizi ile Ülkelerin Turizm İstatistikleri Bakımından Sınıflandırılması.* İstatistikçiler Dergisi 4, 31-38.

Macqueen J 1967. *Some Methods for Classification and Analysis of Multivariate Observations.* Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability 281-297. *NCSS User's Guide-IV. Multivariate Analysis, Clustering, Meta-Analysis, Forecasting / Time Series, Operations Research, and Mass Appraisal,* ABDi 2006.

Öz B 2009. *Kümeleme Analizi ile Türkiye ve AB Ülkelerinin Beşeri Sermaye Göstergeleri Açısından Karşılaştırılması.* Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi 10, 1-30.

Özdamar K 2004. *Paket Programlar İle İstatiksel Veri Analizi (Çok Değişkenli Analizler).* Eskişehir: Kaan Kitabevi.

Taş U, Yenilmez F 2008. *Türkiye'de Eğitimin Kalkınma Üzerindeki Rolü ve Eğitim Yatırımlarının Geri Dönüş Oranı.* Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi 9, 155-186.

Tatlıdil H 1996. *Uygulamalı Çok Değişkenli İstatiksel Analiz.* Ankara: Akademi Matbaası.

Tibshirani R, Walther G, Trevor H 2001. *Estimating the Number of Clusters in a Data Set via the Gap Statistic.* Journal of Royal Statistical Society: Series B 63, 411-423.

Torra V 2005. *Fuzzy c-Means for Fuzzy Hierarchical Clustering.* Proceedings of the 14th IEEE Conference on Fuzzy Systems, 646-651.

Turanlı M 2006. *Avrupa Birliği'ne Aday ve Üye Ülkelerin Ekonomik Benzerliklerinin Kümeleme Analiziyle İncelenmesi.* İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi 5, 95-108.

Uzgören N 2013. *Türkiye’de İllerin Beşeri Sermayenin Unsuru Olan Temel Eğitim Göstergeleri Bakımından Sınıflandırılması: Temel Bileşenler ve Kümeleme Analizi Uygulaması.* Tisk Akademi 8, 118-133.

Yilanci V 2010. *Bulanık Kümeleme Analizi ile Türkiye’deki İllerin Sosyoekonomik Açından Sınıflandırılması.* Süleyman Demirel Üniversitesi İ.İ.B.F Dergisi, 15, 453-470, 2010.

Yılmaz, Ö, Kaya V 2005. *Genişleme Sürecindeki Avrupa Birliği: Ekonomik Performansa Dayalı Kümeleme Analizi.* Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 5, 361-376.