



## Bulanık kümeleme analizi ile ülkelerin turizm istatistikleri bakımından sınıflandırılması

İbrahim Kılıç

Afyon Kocatepe Üniversitesi,  
Veteriner Fakültesi, Biyoistatistik  
Anabilim Dalı, 03200-Afyon, Türkiye

[kilicibrahim@hotmail.com](mailto:kilicibrahim@hotmail.com)

Oktaç Emir

Anadolu Üniversitesi, İşletme  
Fakültesi, Konaklama İşletmeciliği  
Bölümü, 26470-Eskişehir, Türkiye

[oktayemir@anadolu.edu.tr](mailto:oktayemir@anadolu.edu.tr)

Gonca Kılıç

Afyon Kocatepe Üniversitesi,  
Turizm İşletmeciliği ve Otelcilik  
Yüksekokulu, 03200-Afyon, Türkiye

[goncakilic@aku.edu.tr](mailto:goncakilic@aku.edu.tr)

### Özet

*Bu çalışmada, dünyanın hemen her bölgesinden veya kıtasından örneklerin yer aldığı 30 ülkenin, 2007 yılına ait dokuz değişken ile verilen turizm istatistikleri kullanılarak bulanık kümeleme analizi ile sınıflandırılması amaçlanmıştır. Çalışmada, her bir küme sayısı ( $k=2,3,4, \dots$ ) için ortalama gölge istatistiği, Dunn katsayısı ve normalleştirilmiş Dunn katsayısı olarak bilinen Nonfuzziness Index hesaplanmıştır. Bununla birlikte, her bir küme sayısı için belirlenen küme üyelik fonksiyonları kullanılarak ülkelere ait veri setine Diskriminant Analizi uygulanmış ve doğru sınıflandırılma oranları saptanmıştır. Elde edilen bu parametreler vasıtasıyla en uygun küme yapısı veya sayısı belirlenmeye çalışılmıştır.*

**Anahtar Kelimeler:** Çok değişkenli analiz; Bulanık kümeleme analizi; Turizm istatistikleri.

### Abstract

#### Classification of countries in terms of tourism statistics using fuzzy clustering analysis

*In this study, it was aimed to classify 30 countries, which include samples from almost all the regions or continents of the world, with fuzzy clustering analysis by using the tourism statistics given with nine variables belonging to the year 2007. In the study, average Silhouette Coefficient, Dunn Coefficient and Nonfuzziness Index known as normalized Dunn Coefficient were computed for each cluster number ( $k=2,3,4, \dots$ ). In addition to this, by using cluster membership functions determined for each cluster number, discriminant analysis was applied to the data set belonging to the countries and correct classification percentages were determined. With the help of these obtained parameters, the most appropriate cluster structure and number was tried to be determined.*

**Keywords:** Multivariate analysis; Fuzzy clustering analysis; Tourism statistics.

## 1. Giriş

Turizm istatistikleri ile ilgili yapılan çalışmalar kapsamında, Dünya Turizm Örgütü (World Tourism Organization, WTO) ve Dünya Turizm ve Seyahat Konseyi (World Travel and Tourism Council, WTTC) gibi kuruluşların değerlendirmelerinin hemen hemen tamamında, sadece turizm geliri veya turist sayısı gibi değişkenler tek başına dikkate alınmakta ve ülkelere ilişkin yapılan karşılaştırmalarda konunun bütünlüğü doğru bir şekilde yansıtılmamaktadır. Oysaki yapılacak karşılaştırmalarda ülkelerin turizm potansiyellerini ilgilendiren tüm değişkenlerin eşzamanlı olarak ele alınması gerekir. Diğer taraftan, yapılan değerlendirmelerde ülkelerin büyüklüğü gibi önemli bir faktör göz ardı edilerek sadece turizm gelirleri ve turist sayıları dikkate alınmaktadır. Oysa, bir ülkeyi turizm faaliyetleri veya göstergeleri açısından değerlendirmek ve diğer ülkelerle karşılaştırma yapabilmek için gerek nüfus, gerekse ekonomik

bakımından büyüklüğünün mutlaka hesaba katılması gerekir. Zira nüfus ve ekonomik hacim bakımından büyük bir ülke için yılda sözgelimi 50 milyon turist çok şey ifade etmeyebilirken, başka bir küçük ülke için ise belki de 5 milyon turist oldukça önemlidir. Göreli bu sayılara bakarak büyük olan ülkenin turizm istatistikleri bakımından önde olduğu veya küçük olan ülkenin turizm açısından geride olduğu şeklinde değerlendirme yapılması hatalı olabilmektedir. Bu durumda, turizm gelirleri ve turist sayılarının yanı sıra turizm göstergelerinin önemli parametreleri olan oda sayısı, yatak kapasitesi, doluluk oranı, ortalama kalış süresi, turizm gelirlerinin GSYH içindeki payı, turizme ilişkin mal ihracatının toplam ihracattaki payı ve turizme ilişkin hizmet ihracatının toplam ihracattaki payı gibi ortalama ve oranların da yansıtıldığı pek çok değişkenin de dikkate alınması gerekir. Bununla birlikte, farklı ölçü birimlerinden (sayı, zaman, oran vb) oluşan bu değişkenlerin standartlaştırılarak ve eşzamanlı olarak çok değişkenli istatistiklerle değerlendirilmesinin daha doğru bir yaklaşım olduğu kabul edilebilir. Çünkü standartlaştırma işlemi ile farklı ölçü birimlerinde ve çok farklı büyüklüklerde olan değişkenlerin, bu yüzden daha fazla veya daha az ağırlıklandırılmaları önlenerek ve böylece, birimlere ait gerçek sınıflama işlemi yapılabilecektir.

Kümeleme analizleri kullanılarak gerek şehirlerin gerekse ülkelerin farklı özelliklerine göre sınıflandırılmasına yönelik literatürde pek çok çalışma yapılmıştır. Şahin ve Hamarat [15] tarafından yapılan çalışmada G-10 Avrupa Birliği ve OECD ülkelerinin sosyo-ekonomik benzerliklerinin bulanık (fuzzy) kümeleme analizi ile belirlenmesi amaçlanmıştır. Terlemez [16] ise kümeleme analizi ile Avrupa Birliği'ne aday ülkelerin ekonomik durumlarını incelemiştir. Atbaş [2] tarafından yapılan çalışmada ise Türkiye'deki illerin suç istatistikleri bakımından farklı kümeleme yöntemleri ile kümelendirilmesi yapılmış uygun küme sayısının tespit edilmesine çalışılmıştır. Geler [8], Koç [11] ve Yıllancı [19] tarafından yapılan çalışmalarda da Sosyo-ekonomik özelliklerine göre Türkiye'deki illerin kümeleme analizi ile sınıflandırılması yapılmıştır.

Bu çalışmada, dünyanın hemen her bölgesinden veya her kıtasından örneklerin yer aldığı 30 ülkenin, 2007 yılına ait dokuz değişken ile verilen turizm istatistikleri kullanılarak bulanık kümeleme analizi (Fuzzy Clustering Method) ile sınıflandırılması amaçlanmıştır. Dünya Turizm Örgütü (WTO) [17,18] 2007 yılı turizm istatistiklerinde 50'ye yakın ülkenin istatistikleri yer almasına rağmen oda sayısı, yatak kapasitesi ve doluluk oranı gibi bazı değişkenler için 20 ülkenin istatistiklerinde eksiklikler bulunduğu için bu çalışmada dokuz değişken açısından tam olarak sunulan 30 ülkenin istatistikleri kullanılmıştır.

## 2. Yöntem

Bu çalışmada kullanılan veri grubunu, 30 ülkenin 2007 yılına ait turizm istatistikleri [17,18] oluşturmaktadır. 30 ülkeye ait turizm istatistikleri; *ziyaretçi sayısı*, *turizm geliri*, *oda sayısı*, *yatak kapasitesi*, *doluluk oranı*, *ortalama kalış süresi*, *turizm gelirlerinin GSYH içindeki payı*, *turizme ilişkin mal ihracatının toplam ihracattaki payı* ve *turizme ilişkin hizmet ihracatının toplam ihracattaki payı* olmak üzere dokuz değişkenden oluşmaktadır.

Araştırmada kullanılan değişkenler farklı ölçü birimlerinde (sayı, \$, gün, yüzde) olup rakamsal olarak da oldukça farklı büyüklükte dirler. Bu nedenle, farklı ölçü birimlerinde ve çok farklı büyüklüklerde olan değişkenlerin diğer değişkenlere göre daha fazla ya da daha az ağırlıklandırılmalarını önlenmek amacıyla ham verilerin yanı sıra standartlaştırılmış veriler de kullanılmıştır. Verilerin standartlaştırılmasında, bir değişkene ait her bir değerden ortalama değerlerin çıkarılarak sonucun standart sapma değerine bölünmesiyle  $[(x-\mu)/\sigma]$  elde edilen değerler kullanılmıştır. Bu çerçevede, çalışmada gerek standartlaştırılmamış ve gerekse standartlaştırılmış veriler için elde edilen bulgular karşılaştırma yapılabilmesi için birlikte sunulmuştur.

Çalışmada, ülkelerin turizm istatistikleri bakımından sınıflandırılması amacı ile bulanık kümeleme yöntemi (Fuzzy Clustering Method) uygulanmıştır. Klasik kümeleme yöntemleri her bir birim için kesin karar alırlar ve bir kümeye atarlar. Sonuçları itibarıyla yaklaşık aynı sonuçları veren kümeleme algoritmalarında bazı birimlerin farklı kümelerde yer aldığı gözlenebilmektedir. Bu tip durumlarda birimlerin küme üyeliklerinde bir bulanıklık söz konusu olmakta ve birimlerin küme üyeliklerinde bir kararsızlık ortaya çıkmaktadır. Bulanık kümeleme yöntemi bu tip durumları tanımlamak amacı ile

geliştirilen önemli bir yöntemdir [15]. Daha açık bir ifade ile bulanık kümeleme yöntemi, birimler kümelerine göre birbirinden belirgin bir şekilde ayrılmıyorsa veya küme üyeliklerinde bazı birimler kararsız bir durum yaratıyorsa uygun bir yöntem olarak ortaya çıkmaktadır.

Bulanık kümeleme yönteminde çoğunlukla Kaufman ve Rousseeuw [9] tarafından geliştirilen bulanık C algoritması kullanılmaktadır. Burada uzaklıklar ve küme üyeliklerinden aşağıda verilen formül (Eşitlik 1) ile hesaplanan C amaç fonksiyonunu minimize etmek hedeflenir [13].

$$C = \sum_{v=1}^k \frac{\sum_{i,j=1}^n u_{iv}^2 u_{jv}^2 d(ij)}{2 \sum_{j=1}^n u_{jv}^2} \quad i, j=1, \dots, n \text{ ve } v=1, \dots, k \quad (1)$$

Burada,  $d(ij)$ ,  $i$  ve  $j$ . birimler arasındaki uzaklık (benzerlik);  $u_{iv}$ ,  $i$ . birimin  $v$ . kümeye bilinmeyen üyeliğini ve  $u_{jv}$ ,  $j$ . birimin  $v$ . kümeye bilinmeyen üyeliğini tanımlar. Bulanık kümelemede her bir birimin tüm kümelerine olan üyelik katsayıları toplamı daima bir olacak şekilde pozitifdir. Bulanık kümelemenin, kesin kümelemeden ne kadar uzakta olduğu Dunn ayrıştırma katsayısıyla belirlenir. Bu katsayı elde edilen kümenin ne kadar bulanık olduğuna ilişkin bir fikir verir. Dunn Ayrıştırma Katsayısı, tüm üyelik katsayılarının ( $u_{iv}$ ) kareler toplamının birim sayısına bölünmesiyle hesaplanır (Eşitlik 2):

$$F(u) = \frac{1}{n} \sum_{i=1}^n \sum_{v=1}^k u_{iv}^2 \quad (2)$$

$F(u)$  her zaman  $[1/k, 1]$  aralığında bulunur. Böylelikle birimlere ilişkin üyelik matrisi elde edilir. Küme sayısından bağımsız olarak bu katsayı aşağıdaki formül (Eşitlik 3) ile normalleştirilir:

$$F_k(u) = \frac{kF(u) - 1}{k - 1} \quad (3)$$

Bulanıksızlık Endeksi (Nonfuzziness Index) olarak isimlendirilen Normalleştirilmiş Dunn Katsayısı ( $F_k(u)$ ) 0 ile 1 aralığında bir değer alır. Ayrıca, uygun kümelemede çekirdek sayısı ve bu çekirdek noktalarına göre belirlenen kümelerin uygunluğu veya kümelerin kararlılık yapısı için gölge istatistiği (Silhouette Coefficient (SC)) kullanılır ve bu değerlerin ortalaması Ortalama Gölge İstatistiği ( $\overline{SC}$ ) olarak bilinir. Sonuç olarak, bulanık kümeleme yönteminde Ortalama Gölge İstatistiği ( $\overline{SC}$ ) ve Dunn ayrıştırma katsayısı ( $F(u)$ ) uygun küme yapısını veren parametrelerdir ve bir veri setinde uygun kümeleme yapısı olması için  $\overline{SC}$  değerinin en az 0.50 olması beklenir [3,14-15].

Bu çalışmada, birimler arasındaki uzaklık ölçütü Öklid uzaklığı ile belirlenmiş ve her bir küme sayısı ( $k=2,3,4,\dots$ ) için ortalama gölge istatistiği, Dunn ayrıştırma katsayısı ve normalleştirilmiş Dunn ayrıştırma katsayısı hesaplanmıştır.

Kümeleme analizlerinde en uygun küme sayısının tespit edilmesine yönelik olarak literatürde [1,4-7,12] pek çok yaklaşım vardır. Bulanık kümeleme yöntemi gibi aşamalı olmayan kümeleme yöntemlerinde küme sayısının belirlenmesi için en çok tercih edilen yöntem;  $k=2, 3, 4, \dots$  biçiminde ardışık olarak küme sayısını bir artırarak oluşan kümelemede birimlerin hangi kümeye ait olduğunu belirten küme üyelik fonksiyonlarını kullanarak yeni veri yapısına ayırma (diskriminant) analizi uygulamaktır [13]. Bilindiği gibi kümeleme analizlerinde birimler belirli sayıda kümeler ayrıldıklarında hangi kümeye ait olduklarını veren değerler küme kodu (küme üyelik fonksiyonu) olarak adlandırılır. Ayırma analizinde ise hangi gruba ya da kümeye ait olduğu bilinen  $n$  birimin  $p$  tane özelliği kullanılarak gerçek gruplarına veya sınıflarına optimal düzeyde atanmalarını sağlayacak fonksiyonlar türetilir ve birimlerin gerçek veya doğru sınıflama oranları elde edilir. Bu çerçevede kümeleme analizi ile bulunan küme üyelik fonksiyonları bir

değişken olarak ele alınarak birimlerin p özelliğine göre ayırma analizi uygulandığında doğru sınıflama oranları elde edilebilir. Böylelikle en yüksek doğru sınıflama oranına sahip küme sayısı en iyi parçalama olarak kabul edilebilir.

Bu çalışmada da ayırma analizinin önemli bir varsayımı olan çok değişkenli normallik testine yönelik olarak grup kovaryans matrislerinin homojenliği testi olarak bilinen Box'S M testi uygulanmıştır. Bu çerçevede, her bir küme sayısı ( $k=2,3,4,\dots$ ) için belirlenen küme kodları veya üyelik fonksiyonları kullanılarak ülkelere ait veri setine ayırma analizi uygulanmış ve doğru sınıflandırılma oranları saptanmıştır. Elde edilen bu parametreler vasıtası ile en uygun küme yapısı veya sayısı belirlenmeye çalışılmıştır. Araştırmada kullanılan verilerin analizinde NCSS 2007 ve SPSS 14.01 programlarından yararlanılmıştır.

### 3. Bulgular

30 Ülkenin turizm istatistiklerine göre bulanık kümeleme yöntemi ile sınıflandırılmasına yönelik farklı küme sayıları ( $k=2, 3, 4, \dots$ ) için elde edilen ortalama gölge istatistiği, Dunn ayrıştırma katsayısı, normalleştirilmiş Dunn ayrıştırma katsayısı ve küme üyelik fonksiyonlarına göre uygulanan Ayırma Analizi doğru sınıflandırılma oranları Çizelge 1'de verilmiştir.

**Çizelge 1.** Ülkelerin turizm istatistiklerine ilişkin bulanık kümeleme analizi sonuçları

Küme Sayısı (k)	Standartlaştırılmış veriler için				Ham veriler için			
	D.S.O.	$\overline{SC}$	F(u)	$F_k(U)$	D.S.O.	$\overline{SC}$	F(u)	$F_k(U)$
2	%80.0	0.418	0.742	0.484	%73.3	0.444	0.529	0.058
3	<b>%86.7</b>	<b>0.543</b>	<b>0.783</b>	<b>0.675</b>	<b>%80.0</b>	<b>0.519</b>	<b>0.537</b>	<b>0.306</b>
4	%73.3	0.418	0.682	0.575	%70.0	0.386	0.517	0.356
5	%70.0	0.418	0.671	0.589	%63.3	0.230	0.449	0.311
6	%63.3	0.235	0.634	0.560	%60.0	0.215	0.416	0.299
7	%60.0	0.342	0.621	0.558	%56.7	0.208	0.425	0.329
8	%56.7	0.313	0.584	0.524	%56.7	0.195	0.465	0.389

D.S.O.: Ayırma (diskriminant) analizi sonucunda elde edilen doğru sınıflandırılma oranı,  $\overline{SC}$ : Ortalama gölge istatistiği, F(u): Dunn ayrıştırma katsayısı,  $F_k(U)$ : Normalleştirilmiş Dunn ayrıştırma katsayısı.

Çizelge 1'deki bulgular, gerek standartlaştırılmış gerekse ham veriler için ortalama gölge istatistiği ve Dunn ayrıştırma katsayılarının, en iyi kümeleme yapısının küme sayısı 3 olduğunda ortaya çıktığını göstermektedir. 3 küme için ortalama gölge istatistiği  $\overline{SC}=0.543$  olarak bulunmuş olup bu değer 3 küme için uygun küme yapısının oluştuğunu göstermektedir. Normalleştirilmiş Dunn katsayısı ise ( $F_k(U)=0.675$ ) minimum bulanıklığın veya maksimum bulanıksızlığın 3 küme ( $k=3$ ) için olarak hesaplandığını göstermektedir. Diğer taraftan, her iki veri seti için Ayırma Analizi doğru sınıflandırılma oranı en yüksek olan küme sayısı da 3'tür.

Ülkelerin hangi kümede yer aldıklarını belirten küme üyelik değerleri Çizelge 2 ve Çizelge 3'te sunulmuştur. Buna göre, birinci kümede 5 ülke, ikinci kümede 9 ülke ve üçüncü kümede ise 16 ülke yer almıştır. Her bir kümeye ait gölge istatistikleri (SC), her iki veri seti için en kararlı veya homojen kümenin ikinci küme ( $SC=0.688$ ;  $SC=0.645$ ) olduğunu göstermektedir. En kararsız veya bulanık küme ise üçüncü küme olmuştur.

Çalışmada, 3 küme için elde edilen Ayırma Analizi doğru sınıflandırılma istatistikleri Çizelge 4'te sunulmuştur. Analizde grup kovaryans matrislerinin homojenliğini test eden Box'S M testi için elde edilen sonuçlar, çok değişkenli normal dağılım varsayımı altında uygulanabilecek olan doğrusal ayırma analizinin uygulanabilirliğini ortaya koymuştur (Box'S M=3.629;  $p>0.05$ ). Buna göre, genel doğru sınıflandırılma oranı standartlaştırılmış veriler için %86.7, ham veriler için ise %80 olarak hesaplanmıştır. Standartlaştırılmış verilerde ilk kümenin tamamı doğru sınıflanırken, Çizelge 2'ye göre 2. kümede yer alan Portekiz üçüncü kümede; Çizelge 2'ye göre 3. kümede yer alan Türkiye, İspanya ve Slovenya'dan iki ülke (Türkiye ve İspanya) 2. kümede ve bir ülke (Slovenya) de 1. kümede yer almıştır.

Ham verilerde yapılan sınıflandırmada ise Çizelge 3'e göre 1. kümede yer alan Hollanda, Slovenya ve Macaristan'dan Hollanda 3. kümede, Slovenya ve Macaristan ise 2. kümede yer almıştır. Bununla birlikte, Çizelge 3'te 2. kümede yer alan Filipinler 1. kümede, 3. kümede yer alan Avusturya 2. kümede, Polonya ise 1. kümede yer almıştır.

**Çizelge 2.** Ülkelerin küme üyelik değerleri (standartlaştırılmış veriler için)

<i>Ülkeler</i>	<i>Üye Küme</i>	<i>1.Küme Değerleri</i>	<i>2.Küme Değerleri</i>	<i>3.Küme Değerleri</i>
Paraguay	1	<b>0.729</b>	0.134	0.137
Norveç	1	<b>0.727</b>	0.135	0.138
Macaristan	1	<b>0.724</b>	0.139	0.137
Şili	1	<b>0.722</b>	0.138	0.140
Slovakya	1	<b>0.716</b>	0.145	0.139
Almanya	2	0.000	<b>1.000</b>	0.000
İtalya	2	0.000	<b>1.000</b>	0.000
Avusturya	2	0.001	<b>0.995</b>	0.004
Çin	2	0.001	<b>0.987</b>	0.012
Fransa	2	0.001	<b>0.987</b>	0.012
Polonya	2	0.023	<b>0.966</b>	0.011
Hollanda	2	0.061	<b>0.906</b>	0.033
ABD	2	0.001	<b>0.842</b>	0.157
Portekiz*	2	0.015	<b>0.650</b>	0.335
Suudi Arabistan	3	0.001	0.001	<b>0.998</b>
Filipinler	3	0.001	0.001	<b>0.998</b>
Panama	3	0.009	0.014	<b>0.977</b>
Tunus	3	0.009	0.014	<b>0.977</b>
Jamaika	3	0.009	0.014	<b>0.977</b>
Mısır	3	0.009	0.014	<b>0.977</b>
Avustralya	3	0.001	0.045	<b>0.954</b>
Ürdün	3	0.046	0.025	<b>0.929</b>
Hırvatistan	3	0.042	0.030	<b>0.928</b>
Fas	3	0.042	0.030	<b>0.928</b>
Belarus	3	0.104	0.098	<b>0.798</b>
İsrail	3	0.105	0.098	<b>0.797</b>
Finlandiya	3	0.106	0.099	<b>0.796</b>
Türkiye*	3	0.001	0.424	<b>0.575</b>
İspanya*	3	0.001	0.429	<b>0.571</b>
Slovenya*	3	0.314	0.145	<b>0.541</b>
<b>SC</b>		<b>0.540</b>	<b>0.688</b>	<b>0.463</b>

Çizelge 3. Ülkelerin küme üyelik değerleri (ham veriler için)

Ülkeler	Üye Küme	1.Küme Değerleri	2.Küme Değerleri	3.Küme Değerleri
Finlandiya	1	<b>0.841</b>	0.072	0.087
İsrail	1	<b>0.805</b>	0.105	0.090
Şili	1	<b>0.774</b>	0.109	0.117
Norveç	1	<b>0.766</b>	0.101	0.133
Belarus	1	<b>0.752</b>	0.140	0.108
Paraguay	1	<b>0.745</b>	0.145	0.110
Azerbaycan	1	<b>0.741</b>	0.145	0.113
Hollanda*	1	<b>0.536</b>	0.158	0.306
Slovenya*	1	<b>0.496</b>	0.379	0.125
Macaristan*	1	<b>0.481</b>	0.282	0.237
Fas	2	0.053	<b>0.863</b>	0.084
Tunus	2	0.066	<b>0.825</b>	0.109
Hırvatistan	2	0.119	<b>0.713</b>	0.168
Jamaika	2	0.146	<b>0.709</b>	0.145
Ürdün	2	0.166	<b>0.698</b>	0.137
Mısır	2	0.113	<b>0.643</b>	0.244
Portekiz	2	0.112	<b>0.621</b>	0.268
Filipinler*	2	0.297	<b>0.444</b>	0.259
Fransa	3	0.074	0.115	<b>0.811</b>
İtalya	3	0.088	0.132	<b>0.781</b>
Çin	3	0.112	0.116	<b>0.772</b>
Meksika	3	0.088	0.144	<b>0.769</b>
İspanya	3	0.064	0.169	<b>0.766</b>
Türkiye	3	0.071	0.230	<b>0.699</b>
Almanya	3	0.244	0.170	<b>0.586</b>
Avustralya	3	0.134	0.284	<b>0.582</b>
ABD	3	0.179	0.264	<b>0.557</b>
Suudi Arabistan	3	0.253	0.206	<b>0.541</b>
Avusturya*	3	0.144	0.341	<b>0.516</b>
Polonya*	3	0.330	0.251	<b>0.419</b>
<b>SC</b>		<b>0.516</b>	<b>0.645</b>	<b>0.438</b>

Çizelge 4. Üç küme için ayırma analizi doğru sınıflandırılma oranları

İstatistikler		Tahmin Edilen Grup Üyeliği						
		Standartlaştırılmış veriler için			Ham veriler için			
		1.Küme	2.Küme	3.Küme	1.Küm e	2.Küme	3.Küme	
Gerçek Grup Üyeliği	Sayı	1.Küme	5	0	0	7	2	1
		2.Küme	0	8	1	1	7	0
		3.Küme	1	2	13	1	1	10
	Oran	1.Küme	%100.0	%0.0	%0.0	%70.0	%20.0	%10.0
		2.Küme	%0.0	%88.9	%11.1	%12.5	%87.5	%0.0
		3.Küme	%6.2	%12.5	%81.3	%8.3	%8.3	%83.4
Box's M= 3.629; p=0.063		Standartlaştırılmış veriler için genel doğru sınıflandırılma oranı = %86.7 Ham veriler için genel doğru sınıflandırılma oranı = %80.0						

#### 4. Sonuç

Turizm istatistikleri bakımından 30 ülkenin bulanık kümeleme analizi ile sınıflandırılması amacı taşıyan bu çalışmada elde edilen sonuçlara göre, standartlaştırılmış ve ham veri setlerinin her ikisi için gerek ortalama gölge istatistiği gerek Dunn ayrıştırma katsayıları ve gerekse Ayırma Analizi doğru sınıflandırılma oranı en yüksek olan küme sayısı 3 olarak belirlenmiştir. Uygulanan bulanık kümeleme analizine göre, birinci kümede Paraguay, Norveç, Macaristan, Şili, Slovakya; ikinci kümede Almanya, İtalya, Avusturya, Çin, Fransa, Polonya, Hollanda, ABD, Portekiz ve üçüncü kümede ise Arabistan, Filipinler, Panama, Tunus, Jamaika, Mısır, Avustralya, Ürdün, Hırvatistan, Fas, Belarus, İsrail, Finlandiya, Türkiye, İspanya, Slovenya ülkeleri yer almıştır.

Ayırma Analizi doğru sınıflandırılma oranlarına göre, ülkeler 3 gruba ayrıldığında standartlaştırılmış verilerde 30 ülkenin %86.7'si, ham verilerde ise %80'i doğru kümede yer almışlardır. Küme üyelik değerleri düşük olan bazı ülkeler (standartlaştırılmış veriler için: Portekiz, Türkiye, İspanya, Slovenya; ham veriler için: Hollanda, Slovenya, Macaristan, Filipinler, Avusturya, Polonya) kararsız bir yapı göstermişler ve sınıflandırma açısından bulanıklık sergilemişlerdir.

Klasik kümeleme yöntemlerinde olduğu gibi bulanık kümeleme yönteminde de n birim belirli sayıdaki kümelere ayrılır, ancak bu birimlerden bazıları herhangi bir kümeye girmeye zorlanmaz. Bulanık kümeler kümedeki birimin üyeliği olarak tanımlanan 0 ile 1 arasındaki her bir birimi belirleyen fonksiyonlardır. Birbirine çok benzeyen birimler aynı kümede yüksek üyelik ilişkisine göre yer alırlar. Bundan dolayı bulanık kümeleme yöntemi, birimlerin kümeye ya da kümelere ait olabilmeye katsayılarını hesaplar. Üyelik katsayılarının toplamı daima 1'e eşittir. Böylelikle birim en yüksek üyelik katsayısına sahip olduğu kümeye atanır. Üyelik fonksiyonları, kümedeki elemanlar sürekli veya süreksiz olsun bir bulanık kümedeki bulanıklığı karakterize eden fonksiyonlardır. Klasik kümeleme yöntemlerinde ise her bir birim sıfır olmayan sadece bir üyelik katsayısına sahiptir ve bu değer daima 1'dir [3,9,13,15]. Bu durum, diğer klasik kümeleme yöntemlerinin, bulanık çözümlemenin sınırlı bir durumu olduğunu göstermektedir [10]. Bu bilgiler doğrultusunda klasik kümeleme yöntemlerinden farklı olarak bulanık kümeleme yöntemi ile bireylerin küme üyeliklerindeki kararsızlık veya bulanıklık ortaya konulduğu için gerek turizm gerekse farklı alanlarda, birimlerin farklı özelliklerine göre sınıflandırılmasına yönelik yapılacak çalışmalarda daha açıklayıcı bilgiler verebilecektir.

#### Kaynaklar

- [1] M.R. Anderberg, 1973, *Cluster Analysis for applications*, Academic Press, New York.
- [2] A.C.G. Atbaş, 2008, *Kümeleme analizinde küme sayısının belirlenmesi üzerine bir çalışma*, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara.
- [3] J.C. Bezdek, S.K. Pal, 1992, *Fuzzy Models For Pattern Recognition: Methods That Search For Structures in Data*, IEEE Press, New York.
- [4] R.B. Calinski, J. Harabasz, 1974, A dendrite method for cluster analysis, *Communications in Statistics*, 3, 1-27.
- [5] K.S. Dinçer, K. Özdamar, 1992, Kümeleme çözümlemesinde uygun kümeleme ölçütlerinin karşılaştırılması, *Hacettepe Fen ve Mühendislik Bilimleri Dergisi*, 14, 17-33.
- [6] N.A. Erilli, U. Yolcu, E. Eğrioğlu, C.H. Aladağ, Y. Öner, 2011, Determining the most proper number of cluster in fuzzy clustering by using artificial neural networks, *Expert Systems with Applications*, 38/3, 2248-2252.
- [7] B. Everitt, 1974, *Cluster Analysis*, Heinmann, London.
- [8] D. Geler, 2005, *Sosyo-ekonomik değişkenliklerine göre illerin kümelenmesi*, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara.
- [9] L. Kaufman, P.J. Rousseeuw, 1990, *Finding Groups Data: An Introduction to Cluster Analysis*, John Wiley and Sons Inc, New York.
- [10] İ. Kılıç, 2008, *Canlı ağırlık ve bazı vücut ölçüleri kullanılarak Karayaka ve Bafra (Sakız x Karayaka G<sub>1</sub>) koyunlarının çok değişkenli istatistiksel yöntemler ile incelenmesi*. Ankara Üniversitesi Sağlık Bilimleri Enstitüsü, Doktora Tezi, Ankara.

- [11] S. Koç, 2001, *İllerin sosyo-ekonomik özelliklere göre sınıflandırılması*, Çukurova Üniversitesi 5. Ulusal Ekonometri ve İstatistik Sempozyumu, Çukurova Üniversitesi, Adana.
- [12] F.H.C. Marriot, 1971, Practical Problems in a method of Cluster Analysis, *Biometrics*, 27, 501-514.
- [13] K. Özdamar, 2004, *Paket Programlar ile İstatistiksel Veri Analizi (Çok Değişkenli Analizler)*, Kaan Kitabevi, Eskişehir.
- [14] S. Sharma, 1996, *Applied Multivariate Techniques*, John Wiley and Sons Inc, New York.
- [15] M. Şahin, B. Hamarat, 2002, *G-10 Avrupa Birliği ve OECD Ülkelerinin Sosyo-Ekonomik Benzerliklerinin Fuzzy Kümeleme Analizi İle Belirlenmesi*, ODTÜ VI. International Conference in Economics, 11-14.
- [16] L. Terlemez, 2001, *Kümeleme analizi ile Avrupa Birliğine aday ülkelerin ekonomik durumlarının incelenmesi*, Anadolu Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Eskişehir.
- [17] WTO, 2009, *Compendium of Tourism Statistics-Data 2003-2007*, World Tourism Organization (UNWTO), Madrid.
- [18] WTO, 2010, *UNWTO World Tourism Barometer. Interim Update April 2010*, <http://www.unwto.org>.
- [19] V. Yılandı, 2010, Bulanık kümeleme analizi ile Türkiye'deki illerin sosyoekonomik açıdan sınıflandırılması, *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 15/3, 453-470.