

## 2005 GENÇ BAYANLAR DÜNYA VOLEYBOL ŞAMPİYONASINA KATILAN SPORCULARIN KÜMELEME ANALİZİ İLE İNCELENMESİ

Gürol ZIRHLIOĞLU<sup>1</sup>, Süha KARACA<sup>2</sup>

<sup>1</sup>Yüzüncü Yıl Üniversitesi Van Meslek Yüksekokulu

<sup>2</sup>Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Beden Eğitimi Bölümü

### ÖZET

Kümeleme analizi, birimleri veya değişkenleri benzerlik ölçülerine göre homojen gruplara bölmeye yarayan çok değişkenli bir istatistiksel analiz yöntemidir. Birim ya da değişkenlere ait uygun gruplar oluşturmak için aşamalı veya aşamalı olmayan kümeleme yöntemleri kullanılır. Yapılan çalışmada, 2005 yılında Dünya Genç Bayanlar Voleybol karşılaşmalarında final ve yarı final oynayan 4 takımın sporcularının analizi yapılmıştır. Sporcuların fiziksel ve teknik özelliklerine göre kümeleme yapıldıktan sonra kümeler arasında farklılık meydana getiren faktörler belirlenmiştir. Elde edilen varyans analizi sonuçlarına göre smaç yüksekliği, blok yüksekliği, smaç sayısı ve toplam sayı değişkenleri için kümeler arasında anlamlı farklılıkların olabileceği sonucu elde edilmiştir.

**Anahtar Sözcükler:** Kümeleme analizi, Aşamalı olmayan kümeleme yöntemi, k-Ortalamalar yöntemi.

### ASSESSMENT OF 2005 JUNIOR WOMEN'S WORLD CHAMPIONSHIP PLAYERS WITH CLUSTER ANALYSIS

#### ABSTRACT

Cluster analysis is a multivariate statistical analysis method which divides units or the variables according to similarity criterion. Hierarchical or nonhierarchical cluster methods are used in order to constitute suitable groups of units or variables. In this study, the players of four teams which played final and semi-final in 2005 Junior Women's World Championship were analysed. According to physical and technical characteristics of the players, the factors that created differences between clusters were obtained after cluster had been done. According to the results of variance analysis it is concluded that significant differences between clusters might be obtained for pike height, block height, spike scores and total scores.

**Key Words:** Cluster analysis, Nonhierarchical cluster method, k-Mean method.

Geliş tarihi : 01.12.2006

Yayına kabul tarihi : 14.03.2007

## GİRİŞ

Kümeleme analizi,  $n \times p$  boyutlu  $X$  veri matrisinde yer alan ve grupları belli olmayan gözlemleri birbirleri ile benzer olan alt kümelere ayırmaya yardımcı olan yöntemler topluluğudur. Bu yöntem; birimleri,  $p$  değişkene göre hesaplanan ve benzerlik ölçülerine göre homojen olan gruplara bölmeye yarar. Kümeleme analizi, aşamalı kümeleme yöntemi kullanılırken kümelerin sayısına ve küme yapılarına ilişkin herhangi bir varsayımda bulunmaz. Ancak, aşamalı olmayan kümeleme yöntemleri ile yapılan değerlendirmelerde küme sayısının belirlenmesine yönelik bazı yaklaşımlar söz konusudur. Kümeleme analizi araştırmanın amacına bağlı olarak değişkenleri veya bireyleri gruplama imkanına sahiptir (Özdamar, 1999).

Kümelemede pek çok yöntem bulunmakta ve bunlar farklı başlıklar altında toplanabilmektedir (Dođan, 2002). Genel olarak birim ya da değişkenleri uygun gruplara ayırırken grupları belirlemede izlenen yaklaşımlara göre;

1. Aşamalı Kümeleme Yöntemleri (Hierarchical Clustering Methods)

2. Aşamalı Olmayan Kümeleme Yöntemleri (Nonhierarchical Clustering Methods) olarak iki gruba ayrılmaktadır.

Aşamalı kümeleme yöntemleri, değişkenleri değişik aşamalarda bir araya getirerek kümeleri belirlemeyi ve bu kümelere girecek olan elemanların hangi uzaklık düzeyinde küme elemanı olduğunu belirleyen bir yöntemdir. Kümeleme; iki gözlemin uzaklıkları veya benzer-

likleri dikkate alınarak yapılır (Özdamar, 1999).

Aşamalı olmayan kümeleme yöntemleri; diğer kümeleme yöntemleri gibi birimlerin kendi içinde homojen ve kendi aralarında heterojen kümelere ayrılmasını hedefler ve prototip kümeler aracılığı ile alt popülasyonların parametre tahminlerini yapmayı amaçlar. Bu yöntemler ile birimlerin uygun oldukları kümelere toplanmaları ve  $n$  birimin  $k$  sayıda kümeye parçalanması hedeflenir. Aşamalı olmayan kümeleme yöntemlerine ait yaklaşımlar içinde en fazla kullanılanları  $k$ -Ortalamalar Yöntemi ve Yığıma Kümeleme Yöntemidir (Özdamar, 1999; Rencher, 2002).

Günümüzde istatistiksel yöntemler spor alanında özellikle takım sporlarında sıklıkla kullanılmaktadır. Takımların gerek maç esnasında gerekse maç sonrası değerlendirmelerle, yapılan hatalar ve eksiklikler ile olumlu veya olumsuz hareketlerin belirlenmesi doğru sporcuların seçiminde yönlendirici olmaktadır. Hızla gelişen ve yaygınlaşan modern voleybolda da, oyuncunun ve takımın performanslarının amaçlarının ölçüsü, değerlendirmesi, analizi ve yorumu üst seviye takımların bilimsel eğitiminin ve gelişiminin ayrılmaz bir parçasıdır (Hançer, 2002).

Voleybol, oyun kuralları hızlı değişim gösterdiği için daha az hata yapılması gereken bir konuma ulaşmıştır. Oyun içerisinde yapılan hataları giderme, hızlı taktik geliştirme ve dolayısı ile oyunu lehine çevirme giderek zorlaşmıştır. Voley-

bolda iyi bir hücum organizasyonu kadar savunma organizasyonu da oldukça önemlidir. Karşılaşma esnasında ve sonrasında yapılan değerlendirmeler genellikle servis atışları, servis karşıma vuruşları, blok dokunuşları ve hücum vuruşları üzerinedir. Genellikle sporcuların ve takımların performans değerlendirmeleri de bu faktörler üzerinden yapılmaktadır (Korkmaz ve Gültekin, 2000).

Yapılan bu çalışmanın amacı, aşamalı olmayan kümeleme analizi yöntemi ile 2005 Dünya Genç Bayanlar Voleybol karşılaşmalarında final ve yarı final oynayan takım sporcularının fiziksel ve teknik özelliklerine göre oluşturdukları kümeler arasındaki farklılıkların hangi faktörlerden kaynaklandığını belirlemektir.

## YÖNTEM

**Verilerin Toplanması:** Yapılan çalışmada, 2005 yılında Türkiye’de yapılan Genç Bayanlar U20 Dünya Şampiyonasına katılan 12 ülke takımlarından final ve yarıfinal oynayan Brezilya, İtalya, Sırbistan Karadağ ve Çin takımlarının sürekli oynayan oyuncularını değerlendirilmeye alınmıştır. Bu amaçla, Uluslararası Voleybol Federasyonu (FIVB)’nin internet sitesinde yer alan verilerden yararlanılmıştır. Sporculara ait kişisel bilgilere “Team&Coaches” seçeneğinden ulaşılabilirken, sporcuların maç performans bilgilerine ise “Best Players” seçeneğinden ulaşılmıştır. Müsabakalarda yer almayan yedek oyunculara ait herhangi bir veriye rastlanmamıştır. Çalışmada, pa-

sör ve libero oyuncularını ele alınan faktörler bakımından takım içerisinde etkin olmadıklarından dolayı değerlendirme dışında tutulmuştur.

Internet sitesinde yer alan sporculara ait kişisel bilgilerden sadece isim, boy, ağırlık, smaç yüksekliği, blok yüksekliği, smaç sayısı, blok sayısı, servis sayısı ve toplam sayı bilgilerine ait veriler kullanılmıştır. Değerlendirmeye alınan sporcular bütün müsabakalarda en fazla süreyle oyunda kalan oyunculardır.

**Verilerin Analizi:** Yapılan çalışmada sporculara ait çeşitli özellikler dikkate alınarak benzerlik ve farklılıkları ortaya koyabilmek amacıyla aşamalı olmayan kümeleme analizi yöntemi uygulanmıştır. Bu yöntemde birimlerin uygun olan kümelerde toplanmaları ve  $n$  birimin  $k$  sayıda kümeye parçalanması hedeflenir (Özdamar, 1999). Veriler üzerinde herhangi bir standartlaştırma yapılmadan orijinal veriler dikkate alınmıştır. Kümeleme analizinde, küme sayısının belirlenmesi konusunda son yıllarda yoğun çalışmalar yapılmaktadır. Çalışmada, küme sayısının belirlenmesinde kullanılan pratik yollardan biri olarak kabul edilen,

$$k = \sqrt{\frac{n}{2}}$$

eşitliği kullanılmıştır (Doğan, 2002). Burada,  $n$  kümelenecek birey sayısıdır.

Verilerin aşamalı olmayan kümeleme yöntemine göre değerlendirilmesinde “ $k$ -Ortalamlar” yöntemi kullanılmıştır.

$k$ -Ortalamlar yöntemi birimleri kümelere parçalamakta ve oluşan kümele-

re ait parametre tahminlerini vermektedir. Bu yöntem  $n$  birimin  $k$  kümeye ayrılmasında birimin  $p$  boyutlu uzayda gösterimi olan en yaygın ortalamaya sahip çekirdek noktalı bir kümeye atanmasını içerir. Bireylerin kümelere atanmasında uzaklık ölçüsü olarak Mahalanobis Uzaklığı kullanılmaktadır (Doğan, 2002; Prakash ve Hoff, 2002). Bu uzaklık ölçüsü, öklid uzaklığının genelleştirilmiş bir şeklidir.  $n \times p$  boyutlu bir veri matrisinde birimler arasındaki öklit uzaklığı;

$$d(i, j) = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

şeklinde belirtilir (Özdamar, 1999). Burada;

$$i = 1, 2, \dots, n$$

$$j = 1, 2, \dots, n$$

$n$  = birey / birim sayısı

$p$  = değişken sayısı

$x_{ik}$  ve  $x_{jk}$  ise sırasıyla  $i$ . ve  $j$ . bireyle-

rin  $k$ . özelliğine ait değerlerdir. Mahalanobis Uzaklığı ise,

$$d^2(i, j) = (\bar{x}_i + \bar{x}_j)' S^{-1} (\bar{x}_i - \bar{x}_j)$$

eşitliği ile hesaplanır (Johnson ve Wichern, 1988). Burada,  $\bar{X}_i$ . bireye ait gözlem vektörü,  $\bar{X}_j$ . bireye ait gözlem vektörü,  $t$  vektör transpozitesi ve  $S^{-1}$  ise benzerlik matrisinin tersidir. Kullanılan  $k$ -ortalama yönteminin gerçekten birbirinden farklı kümeler oluşturup oluşturmadığını anlamak için her bir değişken ayrı ayrı dikkate alınarak tek yönlü varyans analizi yapılmıştır.

#### BULGULAR

Aşamalı olmayan kümeleme yöntemi ile yapılan değerlendirmelerde sporcuların gruplara göre dağılımları Tablo 1'de belirtildiği gibidir.

Sporcuların fiziksel ve teknik özellikleri dikkate alınarak uygulanan  $k$ -ortalama yöntemini ile elde edilen kümelere

**Tablo 1.** Aşamalı olmayan kümeleme yöntemi ile elde edilen gruplar.

1. Grup	2. Grup	3. Grup
S.O. (Brezilya)	A.S. (Brezilya)	R.B. (Brezilya)
V.A. (İtalya)	I.G. (İtalya)	S.V. (İtalya)
G.D. (İtalya)	L.C. (İtalya)	J.V. (Sırb.Karadağ)
D.K. (Sırb. Karadağ)	M.Y. (Çin)	F.R. (Brezilya)
A.J. (Sırb. Karadağ)	N.Y. (Çin)	S.H. (Sırb. Karadağ)
M.M. (Sırb. Karadağ)	T.M. (Brezilya)	
	N.Y. (Çin)	
	Y.Z. (Çin)	
	C.W. (Çin)	

**Not:** Tabloda grupları oluşturan oyuncuların isimlerinin baş harfleri verilmiştir.

## Voleybolda Kümeleme Analizi

**Tablo 2.** Sporculara ait kümeler, tanımlayıcı istatistikler ve varyans analiz sonuçları.

	Boy	Kilo	Smaç Yüks.	Blok Yüks.	Smaç Sayısı	Blok Sayısı	Servis Sayısı	Toplam Sayı
	$\bar{x} \pm s_{\bar{x}}$	$\bar{x} \pm s_{\bar{x}}$	$\bar{x} \pm s_{\bar{x}}$	$\bar{x} \pm s_{\bar{x}}$	$\bar{x} \pm s_{\bar{x}}$	$\bar{x} \pm s_{\bar{x}}$	$\bar{x} \pm s_{\bar{x}}$	$\bar{x} \pm s_{\bar{x}}$
1.Küme (n=6)	185.0±2.28	67.83±3.31	294.7±6.62	280.3±6.02	48.5±10.37	12.5±4.72	4.5±2.35	65.5±6.89
2.Küme (n=9)	188.8±4.09	73.22±7.55	313.6±3.00	299.1±6.62	43.4±7.29	19.6±9.77	6.8±1.99	69.8±14.21
3.Küme (n=5)	184.2±4.02	71.2±4.76	299.2±9.78	282.8±8.41	89.4±12.74	14.6±4.93	4.6±2.88	108.6±8.47
Genel (n=20)	186.5±4.05	71.1±6.09	304.3±10.59	289.4±11.18	56.45±21.7	16.2±7.84	5.6±2.48	78.2±20.98
F	3.282	1.479	18.376	16.267	38.629	1.715	2.275	24.730
P	0.062	0.256	0.00006	0.00011	0.0000	0.21	0.133	0.00001

ait tanımlayıcı istatistikler ve varyans analizi sonuçları Tablo 2’de verilmiştir.

Varyans Analizi değerlerine göre dik-kate alınan değişkenlerden smaç yüksekliği, blok yüksekliği, smaç sayısı ve toplam sayı değişkenleri için kümeler arasında istatistiksel olarak anlamlı farklılıklar bulunmuştur.

### TARTIŞMA

Spor karşılaşmalarının istatistik analizleri oyuncuların ve takımın performansı ile ilgili verileri elde etmede kullanılan en önemli yöntemdir. Bu yöntem maç analizinin neredeyse her aşamasında kullanılabilir. İstatistiksel analiz yöntemiyle veri elde etmenin, veriyi kullanmanın ve yorumlamanın oyun hareketlerinin sınıflandırılması, hareketlerin değerlendirilmesi, elde edilen bilginin kaydedilmesi, veri işleme, verinin analizi ve yorumu gibi bir çok yolu vardır (Hançer, 2002; Korkmaz ve Gültekin, 2000).

Yapılan çalışmada, Dünya Genç Bayanlar Voleybol müsabakalarına katılıp yarı final ve final oynayan takımların sporcularına ait fiziksel ve teknik özellikler dikkate alınarak aşamalı olmayan kü-

meleme yöntemleri ile sınıflandırmalar yapılmıştır.

Yapılan varyans analizi sonucunda smaç yüksekliği, blok yüksekliği, smaç sayısı ve toplam sayı değişkenleri için kümeler arasında anlamlı farklılıkların olabileceği sonucu dikkate alınarak Tablo 2’deki tanımlayıcı istatistikler incelendiğinde bu değişkenler dışında kalan boy, kilo, blok sayısı ve servis sayısı değişkenlerine ait ortalamaların genel ortalamalara yakın olduğu sonucu elde edilmiştir. Bir voleybol karşılaşmasının kazanılmasında önemli etkenlerden olan smaç yüksekliği, blok yüksekliği, smaç sayısı ve toplam sayı değişkenlerine ait ortalamaların ise, genel ortalamalardan farklı olduğu görülebilmektedir. Üçüncü kümenin smaç yüksekliği genel ortalamadan düşük olmasına rağmen bu küme sporcularının smaç ile elde ettikleri sayı diğer kümelerden daha fazladır. İkinci kümeye ait elde edilen sonuçlar incelendiğinde, boy ve smaç yüksekliği değerlerinin yüksek olmasına karşın smaç ile elde edilen sayının daha düşük olduğu görülmüştür. Bu durum smaç ile elde edilen sayıda boy ve smaç yüksekliğinden ziyade vuruş tekniğinin, gücün,

taktiklerdeki hız ve deđişkenliđin daha önemli olabileceđi sonucu elde edilmiştir (Hançer, 2002). Boy ve smaç yüksekliđi birlikte incelendiđinde smaç ile elde edilen sayıları etkilemediđi sonucu elde edilmiştir ( $p>0.05$ ). İkinci kümenin blok yüksekliđinin diđer kümelere oranla daha fazla olması bu küme elemanlarının daha fazla blok sayısı elde etmelerine neden olabilecek bir durum olarak gözlenmiştir.

Voleybolda yapılan hücumların veya savunmaların sayı ile sonuçlanması temel hedeftir. Bu durum düşünöldüğünde üçüncü kümeyi oluşturan sporcuların ön plana çıktığı görölmüştür. Bu kümeyi oluşturan sporculardan ikisi turnuvayı birinci olarak bitiren Brezilya takımının sporcuları, iki tanesi ikinci olan Sırbistan Karadađ takımının sporcuları ve bir tanesi ise dördüncü takım olan İtalya takımının sporcusudur. Bu sporcular turnuvadaki en skorler oyuncular olup, oluşan en ideal küme olarak nitelendirilebilir.

Kümeleme analizinde birbirlerine benzer gözlemlerin alt kümeleri oluşturulduğu dikkate alındığında, özellikle takım sporlarında kadro oluşumu esnasında, oyun içinde deđişim yaparken, takımların transfer dönemlerinde oyuncu seçiminde maliyet planlamalarında benzer özelliđe sahip olan sporcuların deđerlendirilmesinde bu yöntemin etkili sonuçlar verebileceđi düşünölmektedir.

**Yazışma Adresi (Corresponding Address)**

Dr. Gürol ZIRHLIOĐLU  
Yüzüncü Yıl Üniversitesi  
Van Meslek Yüksekokulu  
VAN  
e-posta: gurol@yyu.edu.tr

**KAYNAKLAR**

- Dođan, İ. (2002). Kümeleme Analizi ile Seçim. **Türk J Vet Anim Sci**, 26, 47-53.
- Hançer, A. (2002). 1987 Çekoslovakya Voleybol Antrenör Kursu Notları. (www.voleybolum.com).
- Johnson, R.A. & Wichern, D.W. (1988). **Applied Multivariate Statistical Analysis**. New Jersey: Prentice Hall, Englewood Cliffs.
- Korkmaz, F. & Gültekin, O. (2000). 1999 Yılı Avrupa Kupa Galipleri Kupası Bayanlar Voleybol Final Karşılaşmalarının Analizi. **Voleybol Bilim ve Teknoloji Dergisi**, 2, 25-31.
- Özdamar, K. (1999). **Paket Programlar ile İstatistiksel Veri Analizi (Çok Deđişkenli Analizler)**. Eskişehir: ETAM A.Ş. Matbaa Tesisleri.
- Prakash, P. & Hoff, B. (2002) Microarray Gene Expression Data Mining with Cluster Analysis Using GeneSight™. Application Note GS10. BioDiscovery, Inc.
- Rencher, A.C. (2002). **Methods of Multivariate Analysis**. New York: Wiley.