

T.C
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

ÇOK BOYUTLU ÖLÇEKLEME VE KÜMELEME ANALİZİ YARDIMIYLA
SAĞLIK GÖSTERGELERİNİN İNCELENMESİ

EKİN İŞILDAR

2501130157

TEZ DANIŞMANI

DOÇ.DR. HANDAN YOLSAL

İSTANBUL, 2017



YÜKSEK LİSANS
TEZ ONAYI

ÖĞRENCİNİN

Adı ve Soyadı : EKİN İŞILDAR

Numarası : 2501130157

Anabilim/Bilim Dalı : EKONOMETRİ

Tez Savunma Tarihi : 22.06.2017

Danışman : DOÇ.DR.HANDAN YOLSAL

Tez Savunma Saati : 11:00

Tez Başlığı : ÇOK BOYUTLU ÖLÇEKLEME VE KÜMELEME ANALİZİ YARDIMIYLA SAĞLIK
GÖSTERGELERİNİN İNCELENMESİ

TEZ SAVUNMA SINAVI, Lisansüstü Öğretim Yönetmeliği'nin 36. Maddesi uyarınca yapılmış, sorulan sorulara alınan cevaplar sonunda adayın tezinin KABULÜ'NE OYBİRLİĞİ / OYÇOKLUĞUYLA karar verilmiştir.

JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
1- PROF.DR.A.KARUN NEMLIOĞLU		Kabul
2- PROF.DR.DİLEK ALTAŞ		Kabul
3- DOÇ.DR.HANDAN YOLSAL		Kabul

YEDEK JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
1- PROF.DR.ESEN YILDIRIM		
2- DOÇ.DR.MELDA AKIN		

ÖZ

ÇOK BOYUTLU ÖLÇEKLEME VE KÜMELEME ANALİZİ YARDIMIYLA SAĞLIK GÖSTERGELERİNİN İNCELENMESİ

Ekin IŞILDAR

Bu çalışmada, İstatistiki Bölge Birimleri Sınıflamasına (İBBS) göre belirlenen Düzey-2 bölgeleri seçilen sağlık göstergeleri açısından karşılaştırılmıştır. Karşılaştırma çok değişkenli istatistiksel analiz yöntemlerinden çok boyutlu ölçekleme (ÇBÖ) ve kümeleme analizleri kullanılarak yapılmıştır. Böylece Düzey-2 bölgeleri arasında var olan benzerlik ve benzemezliklerin ortaya konulması amaçlanmıştır.

Çalışmada ÇBÖ analizi olarak ALSCAL algoritmasından, kümeleme analizi olarak ise temelde hiyerarşik ve hiyerarşik olmayan yöntemler olmak üzere dört farklı kümeleme yönteminden yararlanılmıştır. Bölgelerin sağlık göstergeleri açısından incelenmesinden önce, her iki yöntemin teorik altyapıları ayrıntılı bir biçimde ele alınmıştır. Ayrıca bu iki yöntemde de kullanılan yakınlıklar ve standartlaştırma gibi temel kavramlara da çalışmanın birinci bölümünde yer verilmiştir. Sonuç olarak, bölgelerin her iki yöntemle sağlık göstergelerine göre çözümlemelerinin birbirini desteklediği ve bölgelerin yaklaşık olarak üç farklı grupta konumlandığı görülmüştür. İstanbul alt bölgesi özellikle sağlık personeli ve ekipman sayısı bakımından diğer bölgelerden ayrılmış ve tek başına bir grup oluşturmuştur. Aynı şekilde Güneydoğu Anadolu bölgesi, Van ve Ağrı alt bölgelerinin birbirine yakın konumlandığı ve özellikle BÖH, TDH ve KÖH gibi demografik değişkenler açısından diğer bölgelerden farklılaştığı sonucuna ulaşılmıştır.

Anahtar Kelimeler: Çok Boyutlu Ölçekleme Analizi, Kümeleme Analizi, ALSCAL Algoritması, İstatistiki Bölge Birimleri Sınıflaması

ABSTRACT

EXAMINATION OF HEALTH INDICATORS WITH MULTIDIMENSIONAL SCALING AND CLUSTERING ANALYSIS

Ekin ISILDAR

In this study, the level-2 regions determined by the Statistical Region Units Classification were compared in terms of selected health indicators. The comparison was made using multidimensional scaling and clustering analyzes from Multivariate statistical analysis methods. Thus, it is aimed to reveal similarities and dissimilarities between the level-2 regions. In addition, basic concepts such as proximity and standardization used in these two methods are given in the first chapter.

In the study, ALSCAL algorithm was used as a multidimensional scaling analysis and four different clustering methods were used as clustering analysis, mainly hierarchical and non-hierarchical methods. Before the regions are examined in terms of health indicators, theoretical backgrounds of both methods have been discussed in a detailed manner. It has been seen that the solutions of the regions according to the health indicators in both methods support each other and the regions are located in approximately three different groups. In particular by the number of health personnel and equipment, the sub-region of Istanbul has been separated from other regions and has formed a group on its own. Similarly, the Southeastern Anatolia region, Van and Agri sub-regions are located close to each other and differentiated from other regions in terms of demographic variables, especially infant mortality rate, total fertility rate and crude death rate.

Key Words: Multidimensional Scaling Analysis, Clustering Analysis, ALSCAL Algorithm, Statistical Region Units Classification

ÖNSÖZ

Sağlık göstergeleri ülkelerin gelişmişlik düzeylerinin en önemli belirleyicilerindendir. Bu göstergelerden ülke içinde bölgelerarası gelişmişlik düzeyini belirlemekte ve sosyal refahın dağılımını görmek amacıyla yararlanılır.

Bu çalışmada Türkiye’de 2015 yılına ait sağlık göstergeleri yardımıyla bölgelerarası kıyaslama yapılmıştır. Çalışmada İstatistiki Bölge Birimleri Sınıflamasına (İBBS) göre belirlenen Düzey-2 bölgeleri çok değişkenli istatistiksel analiz yöntemlerinden çok boyutlu ölçekleme ve kümeleme analizi kullanılarak karşılaştırılmıştır. Buna göre bölgeler sağlık göstergeleri açısından çok boyutlu ölçekleme ile üç boyutta temsil edilmiştir. Sağlık göstergelerinin kümeleme analizine göre genellikle üç kümeye ayrıldığı görülmüştür. Her iki yöntemde de en belirgin şekilde farklılaşan bölgenin, beklendiği gibi, İstanbul olduğu belirlenmiştir. Ayrıca Doğu ve Güneydoğu bölgelerindeki Van ve Ağrı alt bölgelerinin birbirine yakın kümelandiği görülmüştür.

Tez çalışmam esnasında, araştırmalarımın her aşamasında beni yönlendiren; bilgi, öneri ve yardımlarını esirgemeyen ve yoğun çalışma temposunda bana zaman ayıran tez danışman hocam sayın Doç. Dr. Handan YOLSAL’a ve eğitim hayatım boyunca benden desteklerinin esirgemeyip her zaman yanımda olan annem Emine IŞILDAR ve babam Birol IŞILDAR’a en içten duygularıyla teşekkür ederim.

İÇİNDEKİLER

ÖZ.....	iii
ABSTRACT.....	iv
ÖNSÖZ.....	v
TABLolar LİSTESİ.....	ix
ŞEKİLLER LİSTESİ.....	xi
GİRİŞ	1

BİRİNCİ BÖLÜM

ÇOK DEĞİŞKENLİ İSTATİSTİKSEL ANALİZ YÖNTEMLERİ İLE İLGİLİ TEMEL KAVRAMLAR

1.1.Çok Değişkenli İstatistiksel Analizde Kullanılan Farklı Yaklaşımlar.....	3
1.1.1. Çok Değişkenli Bağımlı Yöntemler	4
1.1.2. Çok Değişkenli İç Bağımlı Yöntemler	4
1.2. Veri Yapısı	5
1.3. Ölçek Türleri	5
1.4. Yakınlık Ölçüleri	6
1.4.1. Benzemezlik Ölçüleri.....	7
1.4.2. Benzerlik Ölçüleri	11
1.4.2.1. Nicel Değişkenler İçin Benzerlik Ölçüleri	11
1.4.2.2. İki Sınıflı Veriler İçin Benzerlik Ölçüleri	12
1.5. Standardizasyon	14

İKİNCİ BÖLÜM

ÇOK BOYUTLU ÖLÇEKLEME

2.1.Çok Boyutlu Ölçeklemenin Tarihi Gelişimi	16
2.2. Tanım ve Kullanım Alanları	16

2.3. Çok Boyutlu Ölçekleme Analizi Türleri	19
2.3.1. Klasik Çok Boyutlu Ölçekleme	21
2.3.1.1.Öklid Uzaklıklarından Koordinat Değerlerinin Elde Edilmesi...21	
2.3.1.2. Boyut Sayısının Belirlenmesi	25
2.3.2. Metrik Olmayan Çok Boyutlu Ölçekleme.....	26
2.4. Uyum İyiliğinin Değerlendirmesi.....	28

ÜÇÜNCÜ BÖLÜM

KÜMELEME ANALİZİ

3.1. Tanım ve Genel Bilgiler.....	31
3.2. Kümeleme Analizi İçin Varsayımlar	33
3.3. Kümeleme Analizi Yöntemleri	33
3.4. Hiyerarşik Kümeleme Yöntemi.....	34
3.4.1. Birleştirici Hiyerarşik Kümeleme Analizi Algoritması	35
3.4.2. Tek Bağlantı Yöntemi	36
3.4.3. Tam Bağlantı Yöntemi	37
3.4.4. Ortalama Bağlantı Yöntemi	38
3.4.5. Merkezi Yöntem	39
3.4.6. WARD Yöntemi	40
3.5. Hiyerarşik Olmayan Kümeleme Analizi	41

DÖRDÜNCÜ BÖLÜM

SAĞLIK GÖSTERGELERİNİN İNCELENMESİ

4.1 Türkiye’de Sağlık Politikalarının Tarihsel Gelişimi	46
4.2. Sağlık Göstergelerinin Yıllara Göre İncelenmesi	48
4.3. İstatistiki Bölge Birimleri Sınıflandırması.....	54
4.4. İBBS Düzey-2 Bölgelerinin Sağlık Göstergeleri Açısından İncelenmesi	56
4.5. Literatür İncelemesi.....	61

4.6. Sağlık Göstergelerinin İstatistiksel Analizi	63
4.6.1. Çok Boyutlu Ölçekleme Sonuçları.....	64
4.6.2. Kümeleme Analizi Sonuçları	71
SONUÇ	80
KAYNAKÇA	.83
EKLER.....	88

TABLÖLAR LİSTESİ

Tablo 1.1. 2×2 Boyutlu Gözlem Tablosu	13
Tablo 1.2. İki Sınıflı Değişkenler İçin Benzerlik Katsayıları	14
Tablo 1.3. Standardizasyon Yöntemleri	15
Tablo 2.4. Stress Değerlerinin Uyum Bakımından Değerlendirilmesi	29
Tablo 3.1. Klasik Birleştirici Hiyerarşik Kümeleme Analizi Yöntemleri	42
Tablo 4.1. 1923 Yılına Ait Sağlık Göstergeleri	46
Tablo 4.2. 1950 ve 1960 Yıllarına Ait Sağlık Göstergeleri	47
Tablo 4.3. 2005-2015 Yılları Arası Sağlık Personeli Sayısı	49
Tablo 4.4. Yıllara Göre Yataklı Tedavi Kurumlarındaki Cihaz Sayısı, Türkiye	49
Tablo 4.5. Türkiye’de İBBS’ye Göre Sınıflandırılmış Bölgeler	55
Tablo 4.6. Düzey-2 Bölgeleri İçin Analizde Kullanılacak Değişkenler	63
Tablo 4.7. Üç Boyut İçin Elde Edilen İterasyonlar ve Kruskal Stress Değeri	64
Tablo 4.8. ÇBÖ Sonuçlarına Göre Üç Boyutlu Uzay İçin Uyarıcı Koordinatlar	65
Tablo 4.9. Üç Boyutlu Uzayda Değişkenler İçin Elde Edilen İterasyonlar ve Kruskal Stress Değeri	69
Tablo 4.10. En Uzak Komşuluk Yöntemi ile Elde Edilen Birleştirici Kümeleme Tablosu	72
Tablo 4.11. Ortalama Bağlantı Yöntemi ile Elde Edilen Birleştirici Kümeleme Tablosu	74
Tablo 4.12. WARD Yöntemi ile Elde Edilen Birleştirici Kümeleme Tablosu.....	76
Tablo 4.13. k-Ortalamlar (k=3) Yöntemi ile Elde Edilen Kümeler.....	77
Tablo 4.14. k- Ortalamalar Yönteminde Son Küme Merkezleri.....	78
Tablo 4.15. k- Ortalamalar Varyans Analizi Tablosu	79

ŞEKİLLER LİSTESİ

Şekil 3.1. Tek Bağlantı Yöntemi ile Oluşturulan Dendogram Örneği	37
Şekil 3.2. Tam Bağlantı Yöntemi ile Oluşturulan Dendogram Örneği	38
Şekil 3.3. Gruplariçi Ortalama Bağlantı Yöntemi ile Oluşturulan Dendogram Örneği	39
Şekil 3.4. Merkezi Yöntem ile Oluşturulan Dendogram Örneği	39
Şekil 3.5. Ward Yöntemi ile Oluşturulan Dendogram Örneği	40
Şekil 4.1. Yıllara Göre Toplam Sağlık Harcamaları (Milyon TL).....	50
Şekil 4.2. Toplam Sağlık Harcamasının Gayrisafi Yurtiçi Hasılaya Oranı (%)	50
Şekil 4.3. Yıllara Göre Kaba Ölüm Hızı (KÖH) (%)	51
Şekil 4.4. Yıllara Göre Bebek Ölüm Hızı (BÖH) Oranları (%)	52
Şekil 4.5. Yıllara Göre Tıbbi Ürün Pazarı (Hastane Satışları Hariç, Üretici Fiyatlarıyla)	53
Şekil 4.6. Yıllara Göre Sağlık Hizmetlerinden Genel Memnuniyet Oranı (%)	53
Şekil 4.7. İBBS Düzey-2 Bölgelerinde Bebek Ölüm Hızı Oranları (2015)	57
Şekil 4.8. İBBS Düzey-2 Bölgelerinde Beş Yaş Altı Bebek Ölüm Hızı Oranları (2015)	57
Şekil 4.9. İBBS Düzey-2 Bölgelerinde Diş Hekimi Sayısı (2015)	58
Şekil 4.10. İBBS Düzey-2 Bölgelerinde Eczacı Sayısı (2015).....	58
Şekil 4.11. İBBS Düzey-2 Bölgelerinde Toplam Hekim Sayısı (2015)	59
Şekil 4.12. İBBS Düzey-2 Bölgelerinde Kaba Ölüm Hızı Oranları (2015)	60
Şekil 4.13. İBBS Düzey-2 Bölgelerinde SGK İş Göremez Gün Sayısı (2015)	60
Şekil 4.14. İBBS Düzey-2 Bölgelerinde Toplam Doğurganlık Hızı (TDH) (2015) ..	61
Şekil 4.15. Üç Boyutlu Uzay İçin Türetilmiş Uyarıcılar Gösterimi	67
Şekil 4.16. Üç Boyutlu Uzay İçin Serpilme Diyagramı	67
Şekil 4.17. İki Boyutlu Uzay İçin Türetilmiş Uyarıcılar Gösterimi	68

Şekil 4.18. Üç Boyutlu Uzayda Değişkenler İçin Türetilmiş Uyarıcılar Gösterimi ..	70
Şekil 4.19. Üç Boyutlu Uzayda Değişkenler İçin Serpilme Diyagramı	70
Şekil 4.20. En Uzak Komşuluk Yöntemi ile Oluşan Dendogram	73
Şekil 4.21. Ortalama Bağlantı Yöntemi ile Oluşan Dendogram	75
Şekil 4.22. WARD Yöntemi ile Oluşan Dendogram	76

GİRİŞ

Bir olaya etki eden faktörlerin deneysel olarak kontrol altında tutulduğu klasik istatistiksel yöntemlerde her seferinde tek bir faktörün etkisi incelenmektedir. Gözlem ya da ölçüm yoluyla elde edilen ve değişken olarak adlandırılan özelliklerin çok sayıda olması durumunda ise araştırmalarda klasik istatistiksel yöntemlerin kullanılması olanaksız olacaktır. Ayrıca gerçek hayatı konu alan bilimsel çalışmalarda, incelenen birim ve değişken sayısı birden fazla olup birbirleriyle ilişki halinde olmaktadır. Bundan dolayı incelenen olayların bütün yönleriyle değerlendirilmesi, yapılan araştırmanın güvenilir ve anlamlı sonuçlar verebilmesi için bir zorunluluk oluşturmaktadır. Bu şekilde birden fazla faktörün etkisi altında olan olaylarda çok değişkenli veri ve bunların analizinde kullanılmak üzere çok değişkenli istatistiksel analiz yöntemlerine başvurulmalıdır.

Çok değişkenli istatistiksel analiz yöntemleri veri setlerinde yer alan değişkenlere bağlı olarak bağımlı ve iç bağımlı olmak üzere iki alt grupta incelenebilir. Bu durumda öncelikle bağımlı değişken sayısı ve kullanılan değişkenlerin ölçek yapıları belirlenmelidir.

Bu çalışmada, çok değişkenli iç bağımlı istatistiksel yöntemlerden çok boyutlu ölçekleme ve kümeleme analizleri yardımı ile İBBS Düzey-2 26 Bölgenin sağlık göstergeleri açısından incelenmesi ve elde edilen sonuçların iki yöntem için karşılaştırılması amaçlanmıştır. Bu nedenle ilk bölümde çok değişkenli istatistiksel analiz yöntemlerinden kısaca bahsedilecek ve çalışmanın devamında kullanılacak olan benzerlik ve uzaklık ölçülerine yer verilecektir. Çalışmanın ikinci ve üçüncü bölümlerinde çok boyutlu ölçekleme ve kümeleme analizlerinin; aşamaları, varsayımları ortaya konularak her iki yöntem de ayrıntılı bir şekilde ele alınacaktır. Son bölümde ise Türkiye’de sağlık ve sağlık politikalarının gelişim süreci incelenecek ve 26 bölge için sağlık göstergesi verilerine çok boyutlu ölçekleme ve kümeleme analizleri uygulanarak elde edilen sonuçların karşılaştırılması ve bulguların yorumlanması ile çalışma sonlandırılacaktır.

BİRİNCİ BÖLÜM

ÇOK DEĞİŞKENLİ İSTATİSTİKSEL ANALİZ YÖNTEMLERİ İLE İLGİLİ TEMEL KAVRAMLAR

Bir olayın, temel özelliklerini ve gelişimini görmek amacıyla tek değişkenli analiz yöntemlerinden yararlanmak yeterli olurken; olaylar arasındaki ilişkiyi ve etkileşimi ortaya çıkartmak, olayların gelişimini takip etmek amacıyla iki veya daha fazla değişkenli analizlerin kullanılması tercih edilmektedir¹. Bu amaçla geliştirilen çok değişkenli istatistiksel analiz (ÇDİA) yöntemlerinin oldukça uzun bir geçmişi vardır. 19 yy başlarında olasılık ve matematikte yaşanan gelişmeler ile birlikte özellikle Laplace (1811) ve Gauss (1823) tarafından iki değişkenli normal olasılık yoğunluk fonksiyonu üzerine yapılan çalışmalar, F.Galton (1889)'un korelasyon ve regresyon analizi ile ilgili çalışmaları, çok değişkenli normal dağılım üzerine geliştirdiği teori, K.Pearson'nın farklı korelasyon katsayıları ile ilgili teorisi, Fisher'in normal dağılıma dayanan örnekleme teorisi ve faktör analizi de dahil olmak üzere yaptığı çalışmalar öncelikle istatistiğin ve istatistiğin bir dalı olan ÇDİA yöntemlerinin gelişmesine ve yaygınlaşmasına yol açmıştır².

Önceleri istatistikte tek değişken için tek değişkenli normal dağılım ve merkezi limit teoremi ortaya konulmuş ve ardından çok değişkenli normal dağılım üzerine çalışılmıştır. Bilgisayar teknolojisinin gelişmesiyle birlikte, özellikle konuyla ilgili paket programlarının yaygınlaşması sonucunda, ÇDİA yöntemleri istatistik ve ekonomi başta olmak üzere neredeyse bütün disiplinlerde araştırmacılar tarafından tercih edilen bir yöntemler bütünü olarak kullanılmaya başlanmıştır.

ÇDİA yöntemleri, diğer istatistiksel yöntemlerde olduğu gibi bilimsel araştırmalarda sayısal olarak belirtilen sonuçlarının özetlenmesinde, yorumlanmasında ve bu sonuçların karar verme mekanizmasında kullanılmasını

¹Karun Nemlioğlu, **Benzerlik Analizleri: Varsayımlara Uymayan Kategorik Verilerde Çok Değişkenli Analiz**, İstanbul, Beşir Kitapevi, Temmuz 2005, s. 4.

²Theodore .W.Anderson, **An Introduction to Multivariate Statistical Analysis**, Third Edition, John Wiley and Sons, New York, 2003, p. 3-4.

hedeflemektedir³. Buradan ÇDİA yöntemlerinin veriyi betimleme ve keşfetme yöntemlerini ve aynı zamanda veri hakkında istatistiksel çıkarım yapabilme yöntemlerini kapsadığı sonucuna varılabilir. Bu tür yöntemlerin çoğunda kullanılan değişkenlerin tümü verinin özelliklerini ve izlediği patikayı ortaya çıkartmak için bir bütün olarak sınanmalıdır. Bu yöntemlerin büyük bir çoğunluğu betimsel olup, bazıları çıkarsamalı yöntemlerdir. Genel anlamda burada kullanılan yöntemlerin tümünde amaç; veride var olan bilgiye ait bir ipucu çıkartmak ve verinin ne söylediğini keşfetmektir⁴.

Kullanılacak ÇDİA yönteminin belirlenmesi analizde yer alan veri türüne, karşılaşılan problem türlerine ve ulaşılmak istenilen amaca bağlıdır. Örneğin; anakütle özelliklerinin bilinmediği yapılar hakkında gruplar oluşturulması ya da var olan gruplara yeni birimlerin atanmasının sağlanması (kümeleme), veri kümesinin değişkenliğinin daha az sayıda değişken tarafından açıklanması (veri indirgeme), çok değişkenli ortalamaların birlikte sınanması gibi amaçlar için yararlanılabilir⁵.

İncelenen olayların daha az sayıda parametre ile ifade edilmesi olarak da tanımlanan ÇDİA, çok daha farklı amaçlar için uygulanabilmektedir. Bu amaçlardan en önemlisi basitleştirme amacıdır. Böylece büyük veri kümeleri nispeten daha az parametre ile özetlenebilir. ÇDİA yöntemleri aynı zamanda keşfedici yöntemler olup, veri kümesini sınamaktan ziyade hipotez oluşturmayı amaçlar⁶.

1.1. Çok Değişkenli İstatistiksel Analizde Kullanılan Farklı Yaklaşımlar

ÇDİA yöntemleri veri kümelerinde yer alan değişkenlere göre bağımlı ve iç bağımlı olmak üzere iki alt gruba ayrılabilir.

³ Selim Gündüz, “Uzaklık Fonksiyonlarının Çok Boyutlu Ölçekleme Algoritmalarındaki Etkinliğinin İncelenmesi ve Uygulamalar”, Çukurova Üniversitesi, İstatistik Ana Bilim Dalı, Yüksek Lisans Tezi, s. 1.

⁴ B. Everitt, **An R and S-Plus Companion to Multivariate Analysis**, Springer-Verlag London Limited, 2005, p. 1.

⁵ Kazım Özdamar, **Paket Programlar ile İstatistiksel Veri Analizi: Minitab 16- IBM SPSS 21**, Cilt 2, Yenilenmiş 9. bsk, Nisan Kitapevi, 2013, s. 2.

⁶ Christopher Chatfield; Alexander J. Collins, **Introduction To Multivariate Analysis**, Springer-Science+Business Media, 1980.

1.1.1. Çok Değişkenli Bağımlı Yöntemler

Bağımlı değişkenlerin bağımsız değişkenler tarafından tahmin edildiği ve değişkenler arasındaki ilişkinin matematiksel bir model yardımıyla ortaya konulduğu istatistiksel analiz yöntemleridir⁷. Bu yöntemlerde değişkenlerde bağımlı bağımsız sınıflandırması yapılmaktadır. Değişkenler arasındaki yapısal ilişki ve fonksiyonel form ile ilgili olan bağımlı yöntemlere; Hotelling T^2 , çok değişkenli varyans analizi (MANOVA), ayırma analizi, çok değişkenli kovaryans analizi (MANCOVA), regresyon, lojistik regresyon analizi gibi yöntemler örnek olarak verilebilir⁸.

1.1.2. Çok Değişkenli İç Bağımlı Yöntemler

İç bağımlı yöntemlerde amaç veri kümesinin temel boyutlarını ortaya çıkarmak olduğundan tüm değişkenler eşanlı olarak analiz edilmekte ve değişkenler arasında bağımlı-bağımsız ayrımı yapılmasına gerek duyulmamaktadır⁹. Değişkenler arasındaki ilişkinin doğrusal olarak tanımlandığı bu analiz türlerinde iki uç nokta ile karşılaşılır: kullanılan değişkenler birbirlerinden tamamen bağımsız olabildiği gibi birbirleriyle çoklu doğrusal bağlantı içinde de bulunabilir. İç bağımlı yöntemlere temel bileşenler analizi, faktör analizi, kümeleme ve çok boyutlu ölçekleme analizleri örnek olarak verilebilir.

Bağımlı ve iç bağımlı ayırımından farklı yaklaşımlarda bulunmaktadır. ÇDİA yöntemlerinin bir kısmı değişkenler arasındaki ilişkiyi ele alırken (R analizi) bir kısmı ise bireyler arasındaki ilişkiyi ele almaktadır (Q analizi). Ancak bazı durumlarda bu yöntemler arasındaki ayrım çok net değildir¹⁰. Örneğin, birimler arası ilişkiler incelenirken birimlerin hangi kümeden veya gruptan geldiği araştırılmak istenildiğinde kümeleme analizi ve ayırma analizi gibi yöntemlere başvurulur. Bu gibi yöntemleri kapsayan yaklaşım sınıflandırma yaklaşımı olarak anılır. Diğer yaklaşımlar birimler arasındaki ilişkileri kümelemekten ziyade sıralamaya ve

⁷ Yüksel Ünlükaplan, “Çok Değişkenli İstatistiksel Yöntemlerin Peyzaj Ekolojisi Araştırmalarında Kullanımı”, Çukurova Üniversitesi, Peyzaj Mimarlığı Anabilim Dalı, Doktora Tezi, Adana, 2008, s. 14.

⁸ Chatfield, Collins, **a.g.e.**, p. 8.

⁹ Ünlükaplan, **a.g.e.**, s. 15.

¹⁰ Chatfield, Collins, **a.g.e.**, p. 8-9.

ölçeklemeye dayanan yaklaşımlar olmaktadır. Bu açıdan bakıldığında iç bağımlı yöntemlerden olan kümeleme analizinin aynı zamanda bir sınıflandırma yöntemi, aynı şekilde iç bağımlı yöntemlerden olan ÇBÖ yönteminin de aynı zamanda bir ölçekleme yöntemi olduğu söylenebilir¹¹.

ÇDİA yöntemlerini sınıflandırırken analizde kullanılan değişkenlerin bağımlı-bağımsız yapısının yanında ölçek yapılarının da bilinmesi gerekmektedir. Ancak bu noktada çalışmada kullanılan notasyonları göstermekte yarar vardır.

1.2. Veri Yapısı

" y_{ij} " r'inci gözlemde Y_j rasgele değişkeninin değeri olmak üzere $n \times p$ Y gözlemler matrisi ($i = 1, 2, \dots, n; j = 1, 2, \dots, p$);

$$Y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1p} \\ y_{21} & y_{22} & \dots & y_{2p} \\ \vdots & \vdots & \dots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{np} \end{bmatrix} = [Y_{(1)} \quad Y_{(2)} \quad \dots \quad Y_{(p)}] = \begin{bmatrix} Y'_1 \\ Y'_2 \\ \vdots \\ Y'_p \end{bmatrix} \quad (1.1)$$

Burada $Y_{(j)}$, $p \times 1$ boyutlu Y_j değişken vektörünü ve y_i de bu değişkenin i 'nci ($i = 1, 2, \dots, n$) birimde aldığı değeri göstermektedir. Y matrisi "Veri Matrisi" olarak adlandırılır ve genellikle $n > p$ olmalıdır. Geometrik olarak, y_i gözlemi p -boyutlu bir uzayda bir noktayı temsil eder. Dolayısıyla Y 'nin satırları p -boyutlu uzayda n noktayı yansıtacaktır¹².

1.3. Ölçek Türleri

"Ölçek", verilerin özelliklerinin tespit edilmesinde en önemli kavramlardan biridir ve veri analizlerinde dikkat edilmesi gereken noktaların başında verilerin

¹¹ A.e.

¹² Parimal Mukhopadhyay, **Multivariate Statistical Analysis**, World Scientific, (Çevrimiçi), https://books.google.com.tr/books?id=XrQYtClGJ3UC&printsec=frontcover&dq=parimal+mukhopadhyay&hl=tr&sa=X&redir_esc=y#v=onepage&q=parimal%20mukhopadhyay&f=false , p. 9-10, 22.04.2014.

ölçek türlerinin belirlenmesi gelmektedir. Amaca uygun istatistik analiz tekniklerinin uygulanması buna bağlıdır¹³.

Steven (1946) tüm ölçek türlerinin sınıflama, sıralama, aralık ve oran olmak üzere dört grupta sınıflandırılabilceğini ileri sürmüştür¹⁴. r ve s birimlerinin Y değişkeni üzerindeki değerleri sırasıyla y_r ve y_s olsun. Böyle bir tasarı üzerinde ölçüm düzeyleri aşağıdaki şekilde tanımlanacaktır.¹⁵

1. Sınıflama (Nominal) ölçeği sadece sınıflar arasında ayırım yapmaktadır. Bunun anlamı; r ve s açısından sadece $y_r = y_s$ ve $y_r \neq y_s$ olduğu söylenebilir.
2. Sıralama (Ordinal) ölçeği nesnelerin sıralanmasını teşvik etmektedir. $y_r = y_s$ ve $y_r \neq y_s$ arasındaki ayırımın yanı sıra; eşitsizlik durumunda $y_r > y_s$ ve $y_r < y_s$ arası ayırım daha da rafine edilmiştir.
3. Aralık (Interval) ölçeği, iki nesne arasındaki farka anlamlı bir ölçü tayin etmiştir. Sadece $y_r > y_s$ olduğunu değil aynı zamanda r 'nin s 'den $(y_r - y_s)$ birim farklı olduğunu da belirtmektedir.
4. Oran (Ratio) ölçeği, mutlak sıfır noktasına sahip aralık ölçektir. Eğer $y_r > y_s$ ise, sonrasında r 'nin s 'den y_r/y_s kez üstün olduğu söylenebilmektedir.

Bu ölçek tanımları sınıflamadan oran ölçeğe kadar hiyerarşik olarak sıralanmaktadır. Her bir ölçek sıralama olarak altındaki tüm ölçeklerin tüm özelliklerini içermektedir.

Nitel ya da kategorik olan değişkenler çoğu zaman sınıflama ve sıralama ölçüm düzeyine sahipken buna karşılık nicel değişkenler ise aralık ya da oran ölçüm düzeyine sahip değişkenler olmaktadır¹⁶.

¹³ Nemlioğlu, **a.g.e.**, s. 1.

¹⁴ Subhash Sharma, **Applied Multivariate Techniques**, 1996, John Wiley and Sons, p. 1.

¹⁵ Michael R. Anderberg, **Cluster Analysis for Applications: Probability and Mathematical Statistics**, Air Force Systems Command United States Air Force, Academic Press New York San Francisco London, p. 27.

¹⁶ **A.e.**

1.4. Yakınlık Ölçüleri

Genel bir terim olan yakınlık (proximity), araştırma altındaki birimlerin tüm benzerlik ve benzemezliklerinin göstergesidir. Yakınlıklar, doğrudan veya dolaylı olarak iki farklı şekilde elde edilebilir. Çoğu uygulamada dolaylı yakınlıkların kullanımı daha yaygın olmasına rağmen; psikoloji ve pazar araştırması gibi alanlarda doğrudan belirlenen yakınlıklar kullanılmaktadır.

Benzerlik ve benzemezlik ölçüleri ters bir şekilde yakından ilişkilidir. Benzerliği ölçtüğü bilinen bir s fonksiyonu tanımlandığında, uygun bir dönüşümle karşılık gelen bir benzemezlik ölçüsü elde etmek çok kolay olacaktır.¹⁷

Kümeleme analizi ve çok boyutlu ölçekleme gibi çok değişkenli istatistik yöntemlerin birçoğunda başlangıç noktası Y veri matrisi yerine, n gözlem arasındaki benzemezliklerden ya da benzerliklerden elde edilen nxn boyutundaki matrislerdir¹⁸. Çoğu durumda bu benzerlik ya da benzemezlik ölçüleri doğrudan gözlenmez ancak veri matrislerinden elde edilebilirler. Her bir birim çifti ile ilgili benzerlik ya da benzemezliğini göstermek amacıyla nxn boyutlu matris oluşturma'nın birçok farklı yolu bulunmaktadır¹⁹.

1.4.1. Benzemezlik Ölçüleri

Tüm kaydedilen değişkenler sürekli olduğunda birimler arasındaki yakınlıklar genellikle benzemezlik ölçüleri ya da uzaklık ölçüleri biçiminde niceliği belirtir. Burada benzemezlik ölçüsü, δ_{rs} , metrik eşitsizliğin yerine getirildiği bir uzaklık ölçü birimidir.

$$\delta_{rk} + \delta_{ks} \geq \delta_{rs} \quad (1.2)$$

rk, ks ve rs birey çiftleri için. Bu çalışmada benzemezlik ölçüsü olarak alınan ölçüler, d_{rs} elemanları ile D uzaklık ölçüleridir.

¹⁷ Chatfield, Collins, **a.g.e.**, p. 191.

¹⁸ Reha Alpar, **Uygulamalı Çok Değişkenli İstatistiksel Yöntemler**, 4. bsk, Ankara, Detay Yayıncılık, 2013 s. 167.

¹⁹ Chatfield, Collins, **a.g.e.**, p. 191.

Diğer bir ifadeyle, benzemezlik ölçüleri iki nesne arasındaki uzaklıkları göstermektedir ve iki şehir arasındaki uzaklık gibi doğrudan ölçülebilirdir²⁰. İki gözlem birbirinden uzak olduğunda büyük benzemezlik değeri, bu gözlemler birbirlerine yakın olduğunda ise küçük benzemezlik değeri elde edilecektir²¹.

p-boyutlu uzayda, (y_r, y_s) gözlem çiftleri arasındaki uzaklıkları (distance) göstermek üzere D uzaklık matrisi aşağıdaki özelliklere sahip metrik tanıma dayalıdır.

Y veri matrisi ve tüm y_r, y_s ve $y_k \in Y$ için;

1. r ve s gibi iki gözlem için d_{rs} uzaklığı 0'dan büyük olmalı ve r ve s gözlemleri arasındaki uzaklık; s ve r gözlemleri arasındaki uzaklığa eşit olmalıdır. Bu özellik simetri olarak adlandırılır.

$$d(y_r, y_s) \geq 0 \quad (1.3)$$

$$d(y_r, y_s) = d(y_s, y_r) \quad (1.4)$$

2. r, s ve k gözlemleri için üçgen eşitsizliği olarak adlandırılan bu özellik, basit şekilde bir üçgenin herhangi bir kenarının uzunluğunun diğer iki kenarın uzunlukları toplamından küçük olduğunu göstermektedir.

$$d(y_r, y_s) \leq d(y_r, y_k) + d(y_s, y_k) \quad (1.5)$$

3. r ve s gözlemleri için $d_{rs} = 0$ ise bu durumda r ve s aynı gözlemdir. $d_{rs} \neq 0$ olması durumunda ise r ve s'nin farklı gözlemler oldukları söylenebilir. Bu özellik belirlilik (ing) olarak adlandırılır.

$$d(y_r, y_s) = 0 \Leftrightarrow y_r = y_s \quad (1.6)$$

d_{rs} , (y_r, y_s) gözlem çiftleri arasındaki uzaklıkları göstermek üzere n birimli veri setinden elde edilen D uzaklıklar matrisi;

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} & \dots & d_{1n} \\ d_{21} & d_{22} & d_{23} & \dots & d_{2n} \\ d_{31} & d_{32} & d_{33} & \dots & d_{3n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ d_{n1} & d_{n2} & d_{n3} & \dots & d_{nn} \end{bmatrix} \quad (1.7)$$

²⁰ Chapter 435, "Multidimensional Scaling", NCSS Statistical Software, (Çevrimiçi), https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Multidimensional_Scaling.pdf , p. 435-4, 14.08.2015.

²¹ Alphar, **a.g.e.**, s. 168.

şeklinde oluşturulur. Uzaklık tanımlarındaki koşullar gereğince $r = s$ olduğunda $d_{rr} = 0$ ve $d_{rs} = d_{sr}$ şeklinde simetri olmasından dolayı D uzaklık matrisi:

$$D = \begin{bmatrix} 0 & d_{12} & d_{13} & \dots & d_{1n} \\ d_{12} & 0 & d_{23} & \dots & d_{2n} \\ d_{13} & d_{23} & 0 & \dots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{1n} & d_{2n} & d_{3n} & \dots & 0 \end{bmatrix} \quad (1.8)$$

biçiminde de yazılabilir²²

En yaygın olarak kullanılan yakınlık ölçüsü, Öklid uzayında olağan uzaklıkların genellemesi olan, en azından oran ölçeği ile ölçülmüş değişkenler arasındaki Minkowski uzaklığıdır. Minkowski ölçüsü;

$$d_{rs} = \left(\sum_{k=1}^d |y_{rk} - y_{sk}|^r \right)^{1/r} \quad (1.9)$$

olarak tanımlanır. Burada; r parametre, d veri nesnelerinin boyut sayısı, y_{rk} ve y_{sk} sırasıyla r'inci ve s'inci nesnelerin k'ıncı bileşenini göstermektedir.

Öklid ve city-blok uzaklıklarının her ikisi de genel Minkowski uzaklığının ya da L_r -normunun (uzaklık) özel bir durumudur. r'nin belirli değerleri için bilinen Minkowski uzaklıkları şöyledir:

1. $r = 1$ City block (Manhattan, L_1) uzaklığı
2. $r = 2$ Öklid uzaklığı (L_2)

İki nokta arasındaki uzaklığın en yaygın ölçüsüdür.

3. $r \rightarrow \infty$ “supremum” ($L_{\max}$, L_{∞}) uzaklığı²³

Öklid Uzaklığı:

r ve s iki nokta arasındaki gerçek uzaklık Öklid uzaklığı kullanılarak sayısal olarak hesaplanabilir.

$$d_{rs} = \left[\sum_{k=1}^p (y_{rk} - y_{sk})^2 \right]^{1/2} \quad (1.10)$$

²² Sharma, **a.g.e.**, p. 44-45; Gündüz, **a.g.e.**, s. 25-26.

²³ “An Introduction to Cluster Analysis For Data Mining”, (Çevrimiçi), http://www-users.cs.umn.edu/~han/dmclass/cluster_survey_10_02_00.pdf, p. 11, 24.01.2016.

Burada p boyut sayısını, d_{rs} uzaklığı ve y_{rk} k'inci sütün ve r'inci satırdaki değeri göstermektedir. Öklid uzaklığı ünlü Pisagor teoreminin bir uzantısıdır²⁴.

Karesel Öklid uzaklığı:

Karesel Öklid uzaklığı,

$$d_{rs} = \sum_{k=1}^p (y_{rk} - y_{sk})^2 \quad (1.11)$$

formülü ile hesaplanır ve burada d_{rs} r ve s birimleri arasındaki karesel uzaklığı göstermektedir²⁵.

Ağırlıklandırılmış Öklid Uzaklığı:

Analizde yer alan değişkenler aynı ağırlık ile ölçeklendirilmemiş olduğu durumda kullanılır.

$$d_{rs} = (\sum_{k=1}^p w_k^2 (y_{rk} - y_{sk})^2)^{1/2} \quad (1.12)$$

olarak elde edilir. w_k k'inci değişkenin değişim aralığı (range) ya da standart sapmasını (s_k) çarpma işlemine göre tersidir. Aynı zamanda w_k değeri olarak s_k 'nin tersi kullanılması durumunda Karl-Pearson uzaklığı elde edilecektir²⁶.

City Block/ Manhattan Uzaklığı:

Manhattan uzaklığı da denilen City-Block uzaklığı, birimler arasındaki var olan mutlak farkların toplamı alınarak hesaplanır. City-Block uzaklık ölçüsü,

$$d_{rs} = \sum_{k=1}^p |y_{rk} - y_{sk}| \quad (1.13)$$

şeklinde ifade edilir²⁷.

²⁴ Chapter 435, **a.g.e.**, p. 435-2.

²⁵ Sharma, **a.g.e.**, p. 187.

²⁶ Hüseyin Tatlıdil, **Uygulamalı Çok Değişkenli İstatistiksel Analiz**, Hacettepe Üniversitesi, Fen Fakültesi İstatistik Bölümü, s. 254.

²⁷ Azize Celile Günay ATBAŞ, “Kümeleme Analizinde Küme Sayısının Belirlenmesi Üzerine Bir Çalışma”, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara, 2008, s. 14.

Chebychev Uzaklıklığı:

Birimler arasındaki mutlak farkın en büyüğü olarak tanımlanan Chebychev uzaklığı,

$$d_{rs} = \max_k^p |(y_{rk} - y_{sk})| \quad (1.14)$$

şeklindedir²⁸.

Mahalanobis Uzaklığı:

Mahalanobis uzaklığı, değişkenler arasındaki korelasyon ya da kovaryansı göz önüne alan, iki nokta arasındaki istatistiksel uzaklık olarak tanımlanır. Eğer değişkenlerin varyansı bire eşit ve değişkenler birbirleri ile ilişkisiz ise; Mahalanobis uzaklığı Öklid uzaklığına dönüşecektir. Anlaşılacağı üzere Öklid uzaklığı Mahalanobis uzaklığının özel bir durumudur.

p değişkenli durumda; r ve s gözlemleri arasındaki Mahalanobis uzaklık ölçüsü:

$$Md_{rs} = (y_r - y_s)' S^{-1} (y_r - y_s) \quad (1.15)$$

şeklindedir. Burada y, px1 boyutlu koordinatlar vektörünü, S ise pxp boyutundaki kovaryans matrisini göstermektedir²⁹.

1.4.2. Benzerlik Ölçüleri

Benzerlik ölçüleri iki gözlem arasındaki yakınlık değeri ile ifade edilmektedir. Büyük benzerlik değeri bu gözlemlerin birbirine yakın ya da benzer olduğunu, küçük benzerlik değeri ise gözlemlerin uzak ya da benzemez olduğunu göstermektedir³⁰. Değişken türlerine göre benzerlik ölçüleri de farklılaşmaktadır.

²⁸ Alpar, **a.g.e.**, s. 172.

²⁹ Sharma, **a.g.e.**, p. 42.

³⁰ Alpar, **a.g.e.**, s. 168.

1.4.2.1. Nicel Değişkenler İçin Benzerlik Ölçüleri

p boyutlu uzayda verilen iki nesne y_r ve y_s arasındaki benzerlik ölçüsü aşağıdaki koşulları sağlamalıdır:

$$1. \quad 0 \leq s_{rs} \leq 1 \text{ tüm } y_r \text{ ve } y_s \text{ nesneleri için,} \quad (1.16)$$

$$2. \quad s_{rs} = 1 \text{ ancak ve ancak } y_r = y_s \quad (1.17)$$

$$3. \quad s_{rs} = s_{sr} \quad (1.18)$$

Bu koşulları sağlayan benzerlik ölçüsü verildiğinde, $d_{rs} = 1 - s_{rs}$ ilişkisi ya da diğer bazı azalan fonksiyonlar kullanılarak farklılık ölçüsü oluşturulabilir. Ancak elde edilen bu ölçü her zaman metrik formda olmayacaktır³¹.

En yaygın benzerlik ölçüsü olarak, y_r ve y_s nesneleri arasındaki Pearson momentler çarpımı korelasyonu kullanılmaktadır. $r, s = 1, 2, \dots, n$ olmak üzere şöyle tanımlanır;

$$q_{rs} = \frac{\sum_{j=1}^p (y_{rj} - \bar{y}_r)(y_{sj} - \bar{y}_s)}{\left[\sum_{j=1}^p (y_{rj} - \bar{y}_r)^2 \sum_{j=1}^p (y_{sj} - \bar{y}_s)^2 \right]^{1/2}} \quad (1.19)$$

burada $\bar{y}_r = \sum_j y_{rj}/p$ ve $\bar{y}_s = \sum_j y_{sj}/p$ olacaktır. Ancak $-1 < q_{rs} < 1$ olduğundan pozitiflik koşulu sağlanmamaktadır. Bu durumu düzeltmek için, q_{rs} yerine $(1 - q_{rs}^2)$ ya da $|q_{rs}|$ nicelikleri kullanılabilir³².

1.4.2.2. İki Sınıflı Veriler İçin Benzerlik Ölçüleri

Öge çiftleri genellikle belli özelliklerin varlığı ya da yokluğu temeline göre karşılaştırılır. Özellik varsa 1 değerini yoksa 0 değerini aldığı varsayılan iki sınıflı (binary) değişken tanımlanarak, özelliklerin varlığı ya da yokluğu matematiksel olarak ifade edilebilir³³.

³¹ Neil H. Timm, **Applied Multivariate Analysis**, Springer, 2002, p. 519.

³² A.e.

³³ Richard A. Johnson, Dean W. Wichern, **Applied Multivariate Statistical Analysis**, Sixth Edition, Pearson Prentice Hall, 2007, p. 674.

İki sınıflı veriler için benzerlik ölçüleri hesaplanırken uyuşum ve uyuşmazlığın nasıl ağırlıklandırıldığı dikkate alınmaktadır³⁴. Çünkü bazı durumlarda, (1-1) uyuşumu (0-0) uyuşumuna göre benzerliğin daha güçlü göstergesi olabilmektedir. Bu gibi durumlarda (0-0) eşleşmesini önemsememek ya da tamamen göz ardı etmek mantıklı olabilir. Benzerlik katsayılarını tanımlamak için çok çeşitli uygulamalar önerilmiştir ve bu uygulamalar (1-1) ve (0-0) uyuşumu ile ilgili farklılaştırılmış yaklaşımlara imkan vermektedir.

Bu düzenleri tanımlamak için r ve s öğelerinin uyuşma ve uyuşmazlık sıklıkları kontenjans tablosu biçiminde düzenlenirse:

Tablo 1.1: 2x2 Boyutlu Gözlem Tablosu

		Gözlem s		Toplam
		1	0	
Gözlem r	1	a	b	a+b
	0	c	d	c+d
Toplam		a+c	b+d	p=a+b+c+d

olacaktır³⁵. y_r ve y_s gibi iki öğe karşılaştırıldığında c ve d uyuşum göstermeyen çiftler için sıklığı gösterirken a ve b uyuşan çiftlerin sıklığını gösterecektir.

³⁴ Timm, **a.g.e.**, p. 520.

³⁵ Timm, **a.g.e.**, p. 518; Johnson, Wichern, **a.g.e.**, p. 675.

Tablo 1.2: İki Sınıflı Değişkenler İçin Benzerlik Katsayıları*

	Benzerlik Katsayısı	Açıklama
Russel ve Rao Benzerlik Ölçüsü	$s_{rs} = \frac{a}{p}$	Sadece 1-1 uyuşumuna önem verir.
Basit Benzerlik Ölçüsü	$s_{rs} = \frac{a + d}{p}$	1-1 uyuşumu ve 0-0 uyuşumuna eşit ağırlık verir.
Jaccard Benzerlik Ölçüsü	$s_{rs} = \frac{a}{a + b + c}$	0-0 uyuşumu önemsizdir. Hem paydada hem de payda dikkate alınmaz.
Parçalı Benzerlik Ölçüsü	$s_{rs} = \frac{2a}{2a + b + c}$	0-0 uyuşumu önemsizken; 1-1 uyuşumuna iki kat ağırlık verir.
Rogers ve Tanimoto Benzerlik Ölçüsü	$s_{rs} = \frac{a + d}{a + d + 2(b + c)}$	Uyuşma göstermeyen çiftlere iki kat ağırlık verir.
Sokal ve Sneath Benzerlik Ölçüsü -1	$s_{rs} = \frac{2(a + d)}{2(a + d) + b + c}$	1-1 uyuşması ve 0-0 uyuşmasına iki kat ağırlık verir.
Sokal ve Sneath Benzerlik Ölçüsü -2	$s_{rs} = \frac{a}{a + 2(b + c)}$	0-0 uyuşması önemsizdir ve uyuşmayan çiftler iki kat ağırlıklandırılır.
Kulezynski Benzerlik Ölçüsü	$s_{rs} = \frac{a}{b + c}$	0-0 uyuşması hariç tutularak uyuşanların uyuşmayan çiftlere oranıdır.

***Kaynak:** Neil Timm, Applied Multivariate Analysis, p. 521; Richard A.Johnson, Dean W.Wichern, Applied Multivariate Statistical Analysis, p. 675.

1.5. Standardizasyon

ÇDİA’de yer alan değişkenler farklı ölçüm düzeylerine sahip olduklarında (sıklık, iki sınıflı, oran vb. gibi), değişkenlerdeki bu ölçek farklılıkları yanlılığa neden olmaktadır. Diğer yandan veri matrisinde ortalama ve varyanslar açısından birbirlerinden farklı değişkenler yer aldığında, ortalaması ve varyansı büyük olan değişkenler diğer değişkenleri belli oranda baskılamakta ve analizdeki etkinlik ve rollerini nispi olarak azaltmaktadır. Böyle durumlarda verilerin standartlaştırılması

daha doğru sonuçlara ulaşılmasını sağlayacaktır³⁶. Ayrıca dağılım şekli değişkenlerin standartlaştırılmasından etkilenmemektedir³⁷.

Değişkenlerin standartlaştırılması için kullanılan yöntemler şunlardır³⁸:

y_{ij} = j 'inci değişkene ait i 'inci birimin gözlem değeri

$\bar{y}_j$ = j 'inci değişkene ait örneklem ortalaması

S_j = j 'inci değişkene ait örneklem standart sapması

R_j = j 'inci değişkene ait değişim aralığı

$$R_j = \text{en büyük } y_{ji} - \text{en küçük } y_{ji} \quad j = 1, 2, \dots, p \quad (1.20)$$

Tablo 1.3: Standardizasyon Yöntemleri*

Yöntemler	Formüller
Standart değişken dönüşümü	$z_{ij} = \frac{y_{ji} - \bar{y}_j}{S_j}$
Değişim aralığını 1 birim haline getirmek	$x_{ji} = \frac{y_{ji}}{R_j}$
Değişim aralığını (0-1) haline getirmek	$x_{ji} = \frac{y_{ji} - \text{en küçük } y_{ji}}{R_j}$
En büyük değer 1 olacak şekilde dönüştürme	$x_{ji} = \frac{y_{ji}}{\text{en büyük } y_{ji}}$
Ortalama 1 olacak şekilde dönüştürme	$x_{ji} = \frac{y_{ji}}{\bar{y}_j}$
Standart Sapması 1 olacak şekilde dönüştürme	$x_{ji} = \frac{y_{ji}}{S_j}$

***Kaynak:** Reha Alphar, Uygulamalı Çok Değişkenli İstatistiksel Yöntemler s.96-99; Kazım Özdamar, Paket Programlar ile İstatistiksel Veri Analizi: Minitab 16-IBM SPSS 21, s.267-268.

³⁶ Özdamar, **a.g.e.**, s. 266.

³⁷ Alphar, **a.g.e.**, s. 99.

³⁸ Alphar, **a.g.e.**, s. 96-99 ; Özdamar, **a.g.e.**, s. 267-268.

İKİNCİ BÖLÜM

ÇOK BOYUTLU ÖLÇEKLEME

Kullanımı son yıllarda giderek yaygınlaşan bir ÇDİA yöntemi olan çok boyutlu ölçekleme (ÇBÖ) analizinin temel amacı, veri kümesinin yapısını nispeten eğitimsiz insan gözü tarafından kolayca anlaşılabilir bir biçimde yansıtmaktır¹.

2.1. Çok Boyutlu Ölçeklemenin Tarihi Gelişimi

ÇBÖ tarihi, her biri kabaca on yıla karşılık gelen ve son derece yenilikçi çalışmalar ile ayrılan dört evreye bölünebilir².

- İlk evre, ÇBÖ sorunun tanımlandığı ve ilk metrik çözümün sağlandığı, Torgerson (1952)'nin çığır açan çalışmasıdır.
- İkinci on yılda ise metrik olmayan ÇBÖ üzerine Kruskal (1964) ve Shepard (1962)'nin öncülük ettiği çalışmaları ve Coombs (1964)'in veri teorisi üzerine çalışmalarını sürdürdüğü görülmüştür.
- Üçüncü on yıl Carrol ve Chang (1970)'in bireysel farklılıklar çok boyutlu ölçekleme üzerine eğilim belirleyici çalışması ile başlamış ve Takane, Young, de Leeuw (1977) ile de Leeuw ve Heiser (1980)'in katkıları ile sürmüştür.
- 1980'li yıllarda ise maksimum olasılıklı ÇBÖ ile ilgili gelişmeler görülmüştür. Ramsay (1982) ve Takane (1980a, 1980b) çalışmaları bu konuda örnek olarak verilebilir.

¹ Forrest W.Young, **Multidimensional Scaling: History, Theory, and Applications**, ed. Robert M.Hamer, New Jersey, Lawrence Erlbaum Associates, 1987, p. 67.

² Young, **a.g.e.**, p. 15.

2.2. Tanım ve Kullanım Alanları

ÇBÖ, veri kümesinin “gizli yapısını” ortaya çıkartmak için araştırmacılara olanak tanıyan matematiksel yöntemler kümesidir. Çok boyutlu ölçeklemede temel çıktı, noktaların geometrik yapılandırılmasıyla oluşturulan, uzaysal (spatial) gösterimdir. Bu uzaysal gösterimde her nokta bir nesneye (subject/object) karşılık gelir ve buradaki yapılandırma ilgili nesneler arasındaki “gizli yapıyı” ifade eder³.

Nesneler arasındaki ilişkileri tanımlamak için kullanılan ÇBÖ analizi, geometrik uzayda p değişken üzerinde ölçülen n nesne ya da birim arasındaki uzaklığa bağlı olarak k boyutlu uzayda ($k < p$) nesneleri göstermeyi amaçlar ve yöntem herhangi bir dağılım varsayımı gerektirmez⁴.

ÇBÖ yöntemi her türlü nesne arasındaki yakınlıkların girdi olarak kullanıldığı yöntemler sınıfındandır. Burada girilen ham veri genellikle araştırma altındaki nesneler ya da uyarıcıların (stimuli) küresel benzerlik ya da benzemezlikleri olmaktadır⁵. Bir diğer ifadeyle n nesne için; farklı nesne çiftleri arasında $m = n(n - 1)/2$ adet benzerlik/benzemezlik değerleri vardır ve bu benzemezlikler temel veriyi oluşturmaktadır. Ancak benzerliklerin ya da benzemezliklerin sayısal olarak elde edilemediği durumlarda, diğer bir ifadeyle sıralama ölçeği ile ölçülmüş değişkenler için temel veri benzerliklerin sıra sayısı olacaktır⁶. Ayrıca ÇBÖ; nesneler, bireyler, denekler ya da başka bir uyarıcıya ilişkin veriler üzerine de uygulanmaktadır. Burada; nesneler cansız şeyleri, bireyler ya da denekler insanlar veya hayvanları, uyarıcılar ise maddi olmayan varlıkları ifade etmelerine rağmen, bu dört terim analizde genellikle birbirlerinin yerine kullanılmaktadır⁷.

³ Joseph B.Kruskal, Myron Wish, **Multidimensional Scaling**, ed. Eric M.Uslaner, Sage University Paper, 1978, series: Quantitative Applications in the Social Sciences, p. 7.

⁴ Nuray Girginer, “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, 323, 2013, p. 60.

⁵ Kruskal, **a.g.e.**, p. 7. ; Florian Wickelmaier, An Introduction to MDS, Sound Quality Research Unit, Aalborg University, Denmark, May 4, 2003, p. 4.

⁶ Johnson, Wichern, **a.g.e.**, p. 708.

⁷ Trevor F.Cox, Michael A.A.Cox, **Multidimensional Scaling**, Second Edition, Chapman & Hall/CRC, 2001, p. 15.

ÇBÖ daha dar bir tanım ile uzaydaki noktaların nesneleri temsil ettiği, daha düşük boyutlu bir uzaya yönelik araştırmadır. Uzaydaki noktalar arası uzaklıklar $\{d_{rs}\}$ orijinal benzemezliklere $\{\delta_{rs}\}$ mümkün olduğunca iyi eşleştirilebilmelidir⁸. Bu eşleştirme $f = \delta_{rs} \rightarrow d_{rs}(Y)$ fonksiyonu ile gösterilebilir. Burada amaç; f fonksiyonu yardımıyla gerçek değerlere en yakın uzaklık değerlerini elde etmektir⁹.

ÇBÖ verileri, ölçekleme algoritmalarından herhangi biri kullanılarak çeşitli şekillerde analiz edilebilir. ALSCAL, PROXSCAL ve INDSCAL algoritmaları bu örneklemelerden bir kaçıdır¹⁰. Analizde, nesneler arasında elde edilen uzaklıklar kümesine ait yapının ortaya konulması amaçlandığından ALSCAL (alternating least squares scaling) algoritmasından yararlanılacaktır¹¹.

ALSCAL algoritması; bireysel farklılıklar modelleri, Torgerson (1952), Shepard (1962) ve Kruskal (1964) tarafından tartışılan bireysel olmayan farklılıklar modelleri ve çok boyutlu açıklayıcı modeller için kullanılır. ALSCAL algoritması kullanılan herhangi bir analizde; bir ya da daha fazla farklılıklar matrisine sahip olunabilir, bu farklılıklar sınıflama, sıralama, aralık ve oran ölçme düzeylerinde tanımlanmış olabilir ve matrisler asimetrik, simetrik veya dikdörtgen olabilir¹². Ayrıca ALSCAL algoritmasında nokta sayısı, denek sayısı ya da gözlemler ile ilgili bir sınırlama yoktur. Ancak indirgendiği boyut sayısı altı ile sınırlandırılmıştır¹³.

Herhangi bir ÇBÖ yönteminde çözüm yorumu; yansıtma, dönüştürme ve döndürme altında değişmez. ÇBÖ yöntemlerinin yorumlanmasında üç önemli nokta vardır. Bunlar¹⁴;

⁸ Cox and Cox, **a.g.e.**, p. 12.

⁹ Fikri Ağgün, “Çok Boyutlu Ölçekleme Analizinin İncelenmesi ve Bir Uygulama”, Zootekni Anabilim Dalı, Yüzüncü Yıl Üniversitesi, Yüksek Lisans Tezi, Van, 2011, s. 20.

¹⁰ Michael C.Hout, Megan H.Papesh, Stephen, D.Goldinger, “Multidimensional Scaling”, **Wiley Interdiscip Rev. Cogn. Sci.**, January 1, 4(1), 2013, p. 93-103.

¹¹ Neyran Orhunbilge, **Çok Değişkenli İstatistik Yöntemler**, İstanbul Üniversitesi, Yayın No:4942, İşletme Fakültesi, Yayın no:286, İstanbul, 2010, s. 535.

¹² Young, **a.g.e.**, p.

¹³ Forrest W.Young, Yoshio Takane and Rostyslaw Lewyckij, “ALSCAL: A Nonmetric Multidimensional Scaling Program With Several Individual- Differences Options”, University of North Carolina, North Carolina, Chapel Hill, 1978, p. 451.

¹⁴ David J.Bartholomew, Fiona Steele, İrini Moustaki, Jane I.Galbraith, **Analysis of Multivariate Social Science Data**, Second Edition, Taylor&Francis Group, LLC, 2008, p. 58.

- Elde edilecek yapılandırma, noktalar arası uzaklıklarda değiştirme olmadan, doğrudan yansıtılabilir.
- Satır veya sütun koordinatlarından bir sabit çıkartarak ya da ekleyerek orijin değiştirilirse; noktalar arası uzaklıklar bu değişimden etkilenmez.
- Noktalar arası uzaklıklar etkilenmeden, noktalar kümesi döndürülebilir. Bu eksenlerin döndürülmesi anlamına gelir.

Ölçme düzeylerinin tümüne ait veri türleri için uygun olan ÇBÖ analizinde veri türleri önemli bir yer tutar. Çünkü uygulamada hangi yaklaşımın kullanılacağı bu veri türlerine göre belirlenmektedir¹⁵. Ayrıca; farklı ölçme birimli değişkenler kullanılıyorsa hesaplanan uzaklık ölçülerinde aynı etkiye sahip olabilmeleri için bu değişkenler standart değerlere dönüştürülmelidir. Ancak böylelikle gerçek durum yansıtılabilir¹⁶.

2.3. Çok Boyutlu Ölçkleme Analizi Türleri

Ölçkleme yöntemi klasik ölçkleme ve metrik olmayan çok boyutlu ölçkleme olarak iki ana türe ayrılır. Bu yöntemlerden ilki Torgerson tarafından önerilmiştir. Yöntemde orijinal metrikler kullanıldığından dolayı metrik ya da klasik ÇBÖ olarak adlandırılmıştır. Ancak veriler sınıflama ya da sıralama ölçeği ile ölçülmüş olduğunda, bu yöntemin kullanılması uygun değildir. Bu nedenle J.B. Kruskal ve R.N.Shepard tarafından sadece farklılıkların sıra özelliklerini kullanan bir yöntem geliştirilmiştir¹⁷.

Hemen hemen bütün ÇBÖ modelleri matematiksel açıdan metriktir, oysa ÇBÖ algoritmalarının –psikometrik anlamda- bazıları metrik, bazıları ise metrik değildir¹⁸.

Klasik olarak, metrik ÇBÖ problemlerinde, r ve s iki nesne arasındaki gerçek uzaklık y_r^G ve y_s^G ile gösterilirken, $\delta_{rs} = \|y_r^G - y_s^G\|$ olacak şekilde bir veri matrisi

¹⁵ Sığırlı ve ark., “Türkiye ve Avrupa Birliği’ne Üye Ülkelerin Sağlık Düzeyi Ölçütlerinin Çok Boyutlu Ölçkleme Analizi ile İncelenmesi”, **İnönü Üniversitesi Tıp Fakültesi Dergisi**, 13, 2, 2006, s. 81-85.

¹⁶ Esra Yiğit, “Çok Boyutlu Ölçkleme Yöntemlerinin İncelenmesi ve Bir Uygulama”, Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, s. , İstatistik Anabilim Dalı; Alpar, **a.g.e.**, s. 6.

¹⁷ Chapter 435, **a.g.e.**, p. ; Chatfield, **a.g.e.**, p. 190.

¹⁸ Young, **a.g.e.**.

$Y_{n \times p}$ olduğu varsayılır. Y veri matrisi genellikle gözlenmez ve bunun yerine y_i 'ler k boyutlu uzayda olmak üzere $Y = [y_1, y_2, \dots, y_n]'$ gözlem kümesini elde etmeye çalışılır. Böylece uzaklıklar $d_{rs} = \|y_r - y_s\|$ ve $d_{rs} \approx \delta_{rs}$ olacaktır. Uzaklıklar ile ilgili bir model;

$$d_{rs} = \alpha + \beta \delta_{rs} + e_{rs} \quad \text{tüm } r, s \text{ için} \quad (2.1)$$

burada e_{rs} rastlantısal ölçüm hatasını gösterir. (d_{rs}, δ_{rs}) çiftleri ile ilgili bir serpilme diyagramı üretildiğinde, aşağıdaki eşitlik kullanılarak tahmini uzaklıklar kümesi elde edilebilir;

$$\hat{d}_{rs} = \hat{\alpha} + \hat{\beta} \delta_{rs} \quad (2.2)$$

Burada uygun değerler, $\hat{d}_{rs}$, uzaklık değildir ancak basitçe uzaklıklara bağlı sayılar oldukları söylenebilir ve farklılıklar (disparities) olarak adlandırılır.

Yakınlıklar δ_{rs} düzenlenirse;

$$\delta_{r_1 s_1} < \delta_{r_2 s_2} < \dots < \delta_{r_m s_m} \quad (2.3)$$

Burada $m = n(n-2)/2$ ve $r_i < s_i$ farklı tamsayılar olarak (2.2)'deki denkleme göre;

$$\delta_{rs} < \delta_{uv} \Rightarrow \hat{d}_{rs} \leq \hat{d}_{uv} \quad \text{tüm } r < s, u < v \quad (2.4)$$

Bu önerme zayıf monotonluk kısıtlaması olarak adlandırılır.

Ancak ÇBÖ'nin birçok uygulamasında gerçek yakınlık ölçüleri kullanışlı değildir ve yakınlık ölçüleri ancak (2.3)'teki gibi sıralanabilir. Burada sıralanan nicelikler n nesneye ilişkin algılanan kararlardır. Bununla birlikte, δ_{rs} ve d_{rs} arasındaki ilişkiyi (1)'deki gibi belirlemek mümkün olmayabilir, onun yerine;

$$d_{rs} = f(\delta_{rs}) + e_{rs} \quad \text{tüm } r, s \text{ için} \quad (2.5)$$

f fonksiyonu, zayıf monotonluk kısıtlaması ile, bilinmeyen monoton artan bir fonksiyondur. Metrik olmayan olarak adlandırılan bu ölçekleme işleminde δ_{rs} uzaklıklarını oluşturmak için sadece uzaklıkların sıra sayıları kullanılmaktadır¹⁹.

2.3.1. Klasik Çok Boyutlu Ölçekleme

Klasik çok boyutlu ölçekleme; noktalar arası benzemezliklerden, ki bu benzemezlikler genellikle öklid uzaklıklarıdır, noktaların yapılandırmasını elde etmek amacıyla cebirsel bir yeniden yapılandırma yöntemidir²⁰. Burada yakınlık matrisi verilerinin, bir haritadan ölçülen mesafeler gibi, metrik özellikler sergilediği varsayılır. Bundan dolayı Klasik ÇBÖ uzayında uzaklıklar; gerçek yakınlıklar arasındaki oranları ve aralıkları olabildiğince iyi muhafaza eder.

Klasik ÇBÖ'yi diğer yöntemlerden ayıran en önemli nokta; farklı başlangıç noktalarına gerek duymadığından, yinelemeli işlemleri gerektirmeyen analitik bir çözüm sağlamasıdır²¹.

2.3.1.1. Öklid Uzaklıklarından Koordinat Değerlerinin Elde Edilmesi

p boyutlu Öklid uzayında n noktanın koordinatları verildiği varsayıldığında r'inci ve s'inci noktaları arasındaki Öklid uzaklığı aşağıdaki formül yardımıyla hesaplanır;

$$y_r = (y_{r1}, \dots, y_{rp})' \quad r = 1, \dots, n \quad (2.6)$$

$$d_{rs}^2 = (y_r - y_s)'(y_r - y_s) \quad (2.7)$$

p boyutlu Öklid uzayında n noktanın gerçek koordinatları biliniyorsa eğer; her nokta çifti arasındaki Öklid uzaklıkları kolayca hesaplanabilir. Ancak ÇBÖ'de istenilen; gerçek koordinatlar ile uzaklık değerlerinin elde edilmesinden ziyade

¹⁹ Timm, **a.g.e.**, p. 542.

²⁰ Chatfield, Collins, **a.g.e.**, p. 198.

²¹ Wickelmaier, **a.g.e.**, p. 10.

uzaklıklar verildiğinde koordinatların ve dolayısıyla Y matrisinin doğru bir biçimde şekillendirilmesidir²².

Bir diğer ifadeyle klasik çok boyutlu ölçekleme algoritması, koordinat matrisi Y'nin iç-çarpım matrisi $B = YY'$ den özdeğer ayrıştırması yoluyla elde edilebilmesi gerçeğine dayanmaktadır²³. Bu nedenle B iç-çarpım matrisi, bilinen karesel uzaklıklar $\{d_{rs}^2\}$ yardımıyla elde edilir ve sonrasında $B = YY'$ formunda çarpanlarına ayrılarak bilinmeyen koordinatlar bulunur²⁴. Bu aşamada öncelikle iç-çarpım matrisi B elde edilir.

Öklid uzaklıkları kümesi verildiğinde, bu uzaklıklardan elde edilen noktaların tek bir gösterimi olmayacaktır. Bunun anlamı yapılandırmanın konum ve yönünün tayin edilemeyeceğidir. Genellikle konum problemi yapılandırmanın ağırlık merkezi orijine yerleştirilerek üstesinden gelinir²⁵.

$$\sum_{r=1}^n y_{ri} = 0 \quad (i = 1, \dots, p) \quad (2.8)$$

Öklid uzaklıkları yardımıyla B iç-çarpım matrisini elde etmek için;

$$d_{rs}^2 = y_r' y_r + y_s' y_s - 2y_r' y_s \quad (2.9)$$

eşitliği kullanılabilir ve buradan;

$$\frac{1}{n} \sum_{r=1}^n d_{rs}^2 = \frac{1}{n} \sum_{r=1}^n y_r' y_r + y_s' y_s \quad (2.10)$$

$$\frac{1}{n} \sum_{s=1}^n d_{rs}^2 = y_r' y_r + \frac{1}{n} \sum_{s=1}^n y_s' y_s$$

$$\frac{1}{n^2} \sum_{r=1}^n \sum_{s=1}^n d_{rs}^2 = \frac{2}{n} \sum_{r=1}^n y_r' y_s$$

Olacaktır. Eşitlikte yerine konulursa;

²² Tatlıdil, **a.g.e.**, s. 271.

²³ Wickelmaier, **a.g.e.**, p. 10.

²⁴ Cox, Cox, **a.g.e.**, p. 43 ; Chatfield, Collins, **a.g.e.**, p. 198.

²⁵ Chatfield, Collins, **a.g.e.**, p. 198.

$$b_{rs} = y'_r y_s \quad (2.11)$$

$$b_{rs} = -\frac{1}{2}(d_{rs}^2 - \frac{1}{n} \sum_{r=1}^n d_{rs}^2 - \sum_{s=1}^n d_{rs}^2 + \frac{1}{n^2} \sum_{r=1}^n \sum_{s=1}^n d_{rs}^2)$$

$$a = a_{rs} - a_{r.} - a_{.s} + a_{..} \quad (2.12)$$

$$\text{Burada } a_{rs} = -1/2 d_{rs}^2 \text{ ve} \quad (2.13)$$

$$a_{r.} = n^{-1} \sum_s a_{rs}, \quad a_{.s} = n^{-1} \sum_r a_{rs}, \quad a_{..} = n^{-2} \sum_r \sum_s a_{rs}$$

Elde edilir. Böylelikle; uzaklıklar matrisine “çifte ortalama” işlemi uygulanarak pozitif yarı tanımlı B matrisi elde edilecektir.

Bir sonraki aşama elde edilen bu iç-çarpım matrisi yardımıyla gerçek koordinatların elde edilmesi sürecidir. $Y = [y_1, \dots, y_n]'$ (nxp) koordinatlar matrisini göstermek üzere iç-çarpım matrisi B;

$$B = YY' \quad (2.14)$$

Olarak ifade edilebilir. Burada B'nin rankı;

$$r(B) = r(YY') = r(y) = p \quad (2.15)$$

olacaktır. Elde edilen matris simetrik, pozitif yarı tanımlı ve p ranklıdır. Bundan dolayı p negatif-olmayan ve n-p adet sıfır özdeğere sahip olacaktır.

Koordinatları yeniden kazanmak amacıyla B matrisini spektral ayrışma açısından tekrar yazarsak;

$$B = \Lambda \Lambda' \quad (2.16)$$

Burada $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, B matrisinin özdeğerlerinin $\{\lambda_i\}$ köşegen matrisi ve $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_n]$ karşılık gelen özvektörler matrisini göstermektedir²⁶. Kareler toplamlarının λ_i 'ye eşit olması amacıyla bu özvektörler ölçeklendirilirler. $f_i = \sqrt{(\lambda_i)} v_i$ düzenlemesi yapılabilir. Bu durumda, mümkün koordinat matrisi Y;

²⁶ Cox, Cox, **a.g.e.**, p. 42-44.

$$Y = [f_1, f_2, \dots, f_k] \quad (2.17)$$

Olacaktır. Y nxk boyutlarında olacağından k boyutlarda bir yapılandırma elde edilecektir. Başka bir ifadeyle, r'inci nokta ya da bireye ait koordinat $\{f_i\}$ 'nin r'inci bileşeni tarafından verilecektir²⁷.

Aşağıdaki adımlarla klasik ÇBÖ algoritması özetlenebilir²⁸.

- **Adım1:** Nesneler arasındaki uzaklıklar $\{\delta_{rs}\}$ elde edilir.
- **Adım 2:** İç-çarpım matrisi B'nin elde edilebilmesi amacıyla $A = \left[-\frac{1}{2} \delta_{rs}^2\right]$ matrisi düzenlenir.
- **Adım 3:** $B = [a_{rs} - a_r. - a_{.s} + a_{..}]$ eşitliği ile iç-çarpım matrisi elde edilir.
- **Adım 4:** Özdeğerler $(\lambda_1, \dots, \lambda_{n-1})$ ve ilgili özvektörler $(v_1, \dots, v_{n-1})$ bulunur ve $v_i^T v_i = \lambda_i$ olacak şekilde özvektörler normalize edilir. Eğer B matrisi pozitif-yarı tanımlı değil ise; (özdeğerlerin birkaçı negatif olacağından);
 - i- Negatif değerler gözardı edilerek devam edilir.
 - ii- Uzaklık değerlerine uygun bir c sabiti eklenerek, $\delta'_{rs} = \delta_{rs} + c(1 - \delta_{rs})$, ikinci adıma geri dönülür.
- **Adım 5:** p uygun boyut sayısı seçilir. Mümkün olduğunca bunun için pozitif özdeğerlerin sayısı kullanılır.
- **Adım 6:** p boyutlu Öklid uzayında n noktanın koordinatları $x_{ri} = v_{ir}$ ($r = 1, \dots, n ; i = 1, \dots, p$) tarafından elde edilir.

Çözüm elde edildikten sonra uyumun ne derecede sağlandığının ölçüsü elde edilmek istenilebilir. Bu özellik; yeterince iyi bir uyum elde etmek için gerekli boyut sayısını yargılamak için de kullanılır. Bu amaçla $\sum_{r < s} (d_{rs} - \delta_{rs})^2$ kareler toplamına bakılabilir. (En küçük kareler durumunda en iyi uyum elde edildiğinden matematiksel olarak uygundur.) Bununla birlikte basit kareler toplamı uzaklıkların ölçüldüğü ölçeğe bağlı olacağından; uzaklıkların aynı birime indirgenmesi amacıyla kare toplamalarının normalize edilmesi tercih edilir. Uyum iyiliği ölçüsü olarak;

²⁷ Chatfield, Collins, **a.g.e.**, p. 199.

²⁸ Cox, Cox, **a.g.e.**, p. 48-49. ; Ağgün, **a.g.e.**, s. 29.

$$\sqrt{\frac{\sum_{i < j} (d_{ij} - \delta_{ij})^2}{\sum_{i < j} d_{ij}^2}} \quad (2.18)$$

kullanılabilir. Stres ya da normalleştirilmiş stres olarak adlandırılan bu ölçüyü hesaplamadan başka yolları da vardır. Alternatif bir stres ölçüsü eşitliğin bölen kısmında d_{rs} ile δ_{rs} değiştirilmesiyle elde edilir. Sıfıra yakın stress değerleri; ÇBÖ çözümünün orijinal δ_{rs} uzaklıkları için iyi bir uyuma sahip olduğunu gösterecektir²⁹.

2.3.1.2. Boyut Sayısının Belirlenmesi

Belirlenen boyut sayısı koordinat eksenlerinin sayısını vereceğinden, ÇBÖ algoritmalarının birçoğunda olduğu gibi klasik ÇBÖ’de de boyut sayısının belirlenmesi önemlidir. Önemli olan bir diğer nokta ise doğru boyut sayısından çok yorumlanabilir düzeyde boyut sayısının elde edilmesidir³⁰.

Çözüm süresi boyunca ortaya çıkan özdeğerlerin göreceli büyüklüğü dikkate alınır ve sonrasında bu büyüklükler boyut sayısını belirlemek için kullanılır³¹. Eğer iç-çarpım matrisi B pozitif yarı tanımlı ise, sıfır olmayan özdeğerlerin sayısı gerekli boyut sayısını verecektir. Ters durumunda ise, pozitif tanımlı bir matris olmadığında, pozitif özdeğerlerin sayısı uygun boyut sayısını ifade eder. Buradan da anlaşılacağı üzere özdeğerler, benzemezliklerin temsil edilmesi için ihtiyaç duyulan boyut sayısının bir göstergesi olmaktadır. Ancak bunlar uzay için istenilen maksimum boyutlardır³².

Uygun boyutun elde edilmesi için geliştirilmiş; özdeğerlerin büyüklükleri ile elde edilen uyum katsayıları da mevcuttur. k boyut sayısı için uyum katsayıları;

$$\frac{\sum_{i=1}^k |\lambda_i|}{\sum_{i=1}^n |\lambda_i|} \quad yada \quad \frac{\sum_{i=1}^k \lambda^2}{\sum_{i=1}^n \lambda^2} \quad (2.19)$$

²⁹ Bartholomew et. al, **a.g.e.**, p. 61.

³⁰ Ulaş Akkücü: “Çok Boyutlu Ölçekleme Tekniği olarak Torgersen Ölçekleme Yöntemi ve Temel Bileşenler Analizi ile Karşılaştırılması”, **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi**, Aralık 2009, sayı:25, s. 315., Yiğit, **a.g.e.**, s. 11.

³¹ Chapter 435, **a.g.e.**, p. 435-3.

³² Cox, Cox, **a.g.e.**, p. 48.

şeklinde elde edilebilir. 0.80 oranı genellikle iyi uyumun göstergesidir. Gerçek farklılık matrisi Öklid uzaklıklarından oluşuyor ise, ki bu mükemmel uyumun göstergesi olmaktadır, bu durumda boyut sayısı nesne sayısına eşit olacak ve bu oranlar 1 değerini alacaktır. Ancak bu oranlardan çok elde edilen haritanın yorumlanabilir olması önemli olduğundan uyum katsayıları ile basitlik arasında bir ilişki kurulmalıdır³³.

2.3.2. Metrik Olmayan Çok Boyutlu Ölçekleme

n öge; büyüklüklerinden ziyade $n(n-1)/2$ adet benzerliklerinin sıralanması kullanılarak daha düşük boyutlu bir koordinat sisteminde düzenlenebilir. Shepard (1962) yakınlıklar ve uzaklıklar arasında sadece sıralı bir ilişki varsayarak metrik MDS çözümü elde etmenin mümkün olduğunu göstermiştir. Sadece bu sıra bilgileri kullanılarak bir gösterim elde edildiğinden; süreç metrik olmayan çok boyutlu ölçekleme olarak adlandırılır³⁴.

Klasik ölçeklemede olduğu gibi burada da temel amaç; n nesne arasında verilen farklılıklar ile noktalar arası uzaklıkların uyum içinde olduğu, n nokta bulmak ve mümkünse sadece iki ya da üç boyutta nesnelerin temsil edilmesini sağlamaktır. Klasik ölçeklemeden temel farkı, analizde sadece farklılıkların sıra sayılarının kullanılmasıdır³⁵.

Algısal haritalama (perceptual mapping) olarak da bilinen bu yaklaşımda yer alan benzerlik matrisleri, nesneler kümesi ile ilgili yargılara ya da algılara dayalı olarak ölçülen yakınlıklardan oluşmaktadır ve dolayısıyla sıralı yapıdadır. Bu durumda, bir nesnenin diğerinden daha iyi olarak algılandığı genellikle bilinir. Ancak ne kadar iyi olduğu ile ilgili bir yorum yapılamaz. Burada sadece yakınlık ölçüleri sıralanabilir³⁶.

Son olarak metrik olmayan ÇBÖ’de önemli bir diğer nokta; analiz sonucunda elde edilen yapılandırmanın farklılık değerleri ile aynı ilişki düzenine sahip

³³ Akkücüük, **a.g.e.**, s. 315.

³⁴ Johnson, Wichern, **a.g.e.**, p. 707 ; Kruskal, Wish, **a.g.e.**, p. 23.

³⁵ Chatfield, Collins, **a.g.e.**, p. 204.

³⁶ Alpar, **a.g.e.**, s. 405 ; Timm, **a.g.e.**, p. 544.

olmasıdır³⁷. d_{rs} 'lerden elde edilen uyumlu uzaklıklar, $\hat{d}_{rs}$, δ_{rs} 'ler ile aynı sıra ilişkisine ya da ters sıra ilişkisine sahip olacaktır³⁸.

Metrik olmayan ÇBÖ'de veri matrisi nitel yapıda olduğundan analitik çözüm mümkün değildir ve bunun yerine yinelemeli bir yaklaşım kullanılmalıdır. p^* boyutlarda; tahmini bir başlangıç yapılandırması ile başlanır, tahmini yapılandırmada her nokta çifti arasındaki Öklid uzaklıkları hesaplanır ve gözlenen uzaklıklar $\{\delta_{rs}\}$ ile türetilmiş uzaklıklar $\{d_{rs}\}$ karşılaştırılır. $\{d_{rs}\}$ 'ler $\{\delta_{rs}\}$ 'ler ile aynı sıralama düzenine sahip ise, tatmin edici bir yapılandırma bulunduğu söylenebilir. Ancak uygulamada, d 'ler ile δ 'lar arasında mükemmel bir uyumun sağlandığı yapılandırma bulmak genellikle mümkün olmayacaktır. Bunun yerine, türetilmiş uzaklıkların gözlenen farklılıklara “mümkün olduğunca yakın” düzenlemeye sahip olduğu bir yapılandırma bulmak amaçlanmıştır.

Herhangi bir tahmini yapılandırmanın uyum iyiliği, d 'ler ve δ 'lar arasında monoton en küçük kareler regresyon ilişkisi oluşturularak bulunur. $\hat{d}_{rs}$, regresyon eğrisi üzerinde δ_{rs} 'ye karşılık gelen noktayı simgeler ve monotonluk koşulu olarak; $\delta_{rs} < \delta_{tv}$ ise $\hat{d}_{rs} \leq \hat{d}_{tv}$ koşulunu sağlamalıdır³⁹.

Türetilmiş yapılandırma ile verilen farklılıkların ne denli örtüştüğünü değerlendirmek amacıyla, “stress” olarak adlandırılan bir istatistik hesaplanır. Bu stress değeri;

$$S = \sqrt{\frac{S^*}{T^*}} = \sqrt{\frac{\sum (d_{rs} - \hat{d}_{rs})^2}{\sum d_{rs}^2}} \quad (2.20)$$

şeklinde tanımlanır. Burada $\hat{d}_{rs}$ değerleri, $\hat{d}_{rs}$ 'lerin d_{rs} değerleri ile aynı sıra düzeninde olması ile S 'i minimize eden değerlerdir⁴⁰. Her zaman 0 ile 1 arasında

³⁷ Ağgün, **a.g.e.**, s. 30.

³⁸ Bartholomew, **a.g.e.**, s. 62.

³⁹ Chatfield, Collins, **a.g.e.**, p. 205-206.

⁴⁰ Kruskal, “Nonmetric MDS: A Numerical Method”, **Psychometrika**, v.29, (2), 1964, p. 115-129.

değer alan stress 0.1'den küçük herhangi bir değer aldığı anda, türetilen yapılandırma tarafından nesnelerin iyi bir gösterimi olduğu şeklinde yorumlanmaktadır⁴¹.

Kruskal'ın algoritması verilen başlangıç yapılandırmasından yerel minimumu bulmak için en dik inişi kullanır. Başlangıç yapılandırmasının seçimi, küresel yerine yerel bir minimum bulmak için önemlidir. Çoğu kez başlangıç yapılandırması için metrik ÇBÖ çözümü kullanılması önerilmektedir⁴².

Metrik olmayan ÇBÖ'de temel adımlar şöyledir⁴³;

- **Adım 1:** Noktaların rasgele bir yapılandırması bulunur.
- **Adım 2:** Noktalar arası d uzaklıkları hesaplanır.
- **Adım 3:** En uygun biçimde ölçeklenmiş veri $f(p)$ elde etmek amacıyla yakınlıkların en uygun/optimal monoton dönüşümleri bulunur.
- **Adım 4:** Optimal biçimde ölçeklenmiş veri ve noktaların yeni yapılandırmasından elde edilen uzaklıklar arasındaki stress minimize edilir.
- **Adım 5:** Birkaç kriter ile stress değerleri karşılaştırılır. Stress değeri yeterince küçük ise algoritma sonlandırılır, değilse ikinci adıma geri dönlür.

2.4. Uyum İyiliği Değerlendirmesi

Bir veri analiz problemi olarak, analizin uygulandığı model tarafından belirli bir veri kümesinin ne derecede iyi temsil edildiğini belirlemek amacıyla bir ifade gereklidir. Uyum iyiliği istatistiği için en uygun seçim elde edilen stress değerleridir⁴⁴.

Stress değerlerinin yorumlanması nesnelerin sayısına (I) ve boyut sayısına (R) bağlı olmaktadır. Kural olarak, $I > 4R$ olduğu durumda stress değerinin yorumu I ve R'ye karşı duyarlı değildir. Bununla birlikte I ve R birbirlerine yakın olduğunda büyük değişimler meydana gelir. Stress ölçüsünün yorumu ancak, I R'ye göre yeterince büyük ise geçerlidir. Ayrıca küçük stress değeri iyi uyumu belirtirken

⁴¹ Johnson, Wichern, **a.g.e.**, p. 709.

⁴² Chapter 435, **a.g.e.**, p. 435-5.

⁴³ Wickelmaier, **a.g.e.**, p. 15.

⁴⁴ Chapter 435, **a.g.e.**, p. 435-3.

yüksek değerler kötü bir uyumun göstergesi olmaktadır. Kruskal (1964a) uyum iyiliği açısından stress değerlerinin yorumlanması amacıyla bazı kurallar sağlamıştır⁴⁵.

Tablo 2.4: Stress Değerlerinin Uyum Bakımından Değerlendirilmesi

Stress	Uyum İyiliği
>0.20	Kötü Uyum
0.10	Vasat Uyum
0.05	İyi Uyum
0.025	Mükemmel Uyum
0.00	Kusursuz Uyum

Kruskal stress-1 ve stress-2 olarak ifade ettiği ve sadece payda olarak farklı bir ölçek faktörü kullanıldığı iki farklı stress formülü elde etmiştir⁴⁶.

$$Stress - 1 = \left[\frac{\sum_{r < s} (d_{rs} - \hat{d}_{rs})^2}{\sum_{r < s} d_{rs}^2} \right]^{1/2} \quad Stress - 2 = \left[\frac{\sum_{r < s} (d_{rs} - \bar{d}_{rs})^2}{\sum_{r < s} (d_{rs} - \bar{d}_{rs})^2} \right]^{1/2} \quad (2.21)$$

ÇBÖ çözümünün uygunluğunu değerlendirmek için stress değerlerine ek olarak yamaç diyagramı (scree plot) ve Shepard diyagramları da kullanılmaktadır⁴⁷. Yamaç diyagramı boyut sayılarına karşı çizilen stress değerlerini göstermektedir. Boyut sayısının artmasıyla stress değeri monoton bir şekilde azalacağından, kabul edilebilir en düşük boyut sayısı aranmalıdır. Bunun için yamaç diyagramında; boyut sayısının artmasıyla stress üzerinde küçük bir artışın meydana geldiği nokta, “dirsek” incelenmelidir⁴⁸. Gözlenen uzaklıklar ile türetilmiş uzaklıklar arasındaki ilişki *Shepard Diyagramı* ile değerlendirilebilir. Doğrusal ya da doğrusal olmayan formlar yardımıyla elde edilen uzaklıklar ile hangi modelin verilere uyum sağladığı

⁴⁵ J. B. Kruskal, “Multidimensional Scalling By Optimizing Goodness of Fit To A Nonmetric Hypothesis”, **Psychometrika**, vol.29, (1), March, 1964.

⁴⁶ Kruskal, Wish, **a.g.e.**, p. 26; Ağgün, **a.g.e.**, s. 32.

⁴⁷ Wickelmaier, **a.g.e.**, p. 13.

⁴⁸ Bartholomew et. al., **a.g.e.**, p. 65; Wickelmaier, **a.g.e.**, p. 13.

gösterilir⁴⁹. Shepard diyagramının daha az yayılması (saçılımı) iyi uyumu işaret etmektedir⁵⁰.

Son olarak, girdi olarak kullanılan uzaklık verisinin analiz sonucunda elde edilen farklılıklar ile ne oranda açıklandığının göstergesi olan korelasyon endeksinin karesi R^2 'de uyum iyiliğinin bir ölçüsü olarak kullanılabilir. Herhangi bir korelasyon yorumunda olduğu gibi 1 mükemmel uyumun sıfır (0) ise hiç uyum olmadığının göstergesidir. Bunun yanında $R^2 \geq 0.60$ olması uyum için “kabul edilebilir sınır” olarak değerlendirilir⁵¹.

⁴⁹ Özdamar, **a.g.e.**, s. 455.

⁵⁰ Wickelmaier, **a.g.e.**, p. 14.

⁵¹ Tuğba Gürçaylılar Yenidoğan, “Pazarlama Araştırmalarında Çok Boyutlu Ölçekleme Analizi: Üniversite öğrencilerinin marka algısı üzerine bir araştırma”, **Akdeniz İ.İ.B.F Dergisi**, (15), 2008, s. 154; Alphar, **a.g.e.**, s. 413.

ÜÇÜNCÜ BÖLÜM

KÜMELEME ANALİZİ

Kümeleme analizinin temel amacı, bireyler (nesneler/birimler vs.) kümesinin, eğer varsa, “doğal gruplaşmalarını” bulmaktır. Genel olarak kümeleme analizi, bireyler kümesini birbirini dışlayan, kapsamlı gruplar kümesi şeklinde parçalara ayırmayı amaçlamaktadır. Burada, grup içindeki bireylerin birbirlerine benzer; farklı gruplar içindeki bireylerin ise birbirlerinden farklı olması beklenir.

3.1. Tanım ve Genel Bilgiler

Kümeleme analizi, gruplanmamış Y veri matrisinde yer alan gözlemleri ya da değişkenleri sahip oldukları özelliklere bağlı olarak homojen alt gruplar halinde ayırmak amacıyla geliştirilen yöntemler topluluğudur¹.

Eğer veriler nxn boyutlu $Y=[y_{ij}]$ matrisi olarak temsil ediliyorsa;

$$Y = \begin{bmatrix} y'_1 \\ \vdots \\ y'_i \\ \vdots \\ y'_n \end{bmatrix} \quad (3.1)$$

kümeleme analizi; Y'nin satırlarının, ki bu satırlar genellikle öğeleri veya nesneleri temsil eder, k belli küme içine bölüneceği bir sınıflandırma planı geliştirmeyi amaçlamaktadır².

Özellikle karmaşık veri setleri içerisindeki sınıflandırmaları keşfetmek için tasarlanan istatistiksel işlemleri tanımlamada kullanılan bir terim olan kümeleme analizi, bağımlı ve bağımsız değişken ayrımı içermeyen iç bağımlı yöntemlerden

¹ Özdamar, **a.g.e.**, s. 255.

² Timm, **a.g.e.**, p. 515.

biridir³. Kümeleme analizinde yer alan veriler arasındaki gruplaşmayı ortaya çıkartmak için, yakınlık ölçüsü tanımlanmalıdır. Bir diğer ifadeyle kümeleme analizi, ölçülen nesneleri birbirleri ile olan benzerliklerine göre sınıflandırmayı amaçladığından “benzerlik” ve “benzemezlik” kavramları analizde önem taşımaktadır. Yakınlık ölçüsünün seçimi, konu, ölçüm ölçeği ve analiz edilen değişkenlerin türleri gibi bazı kriterlere bağlıdır⁴ ve birinci bölümde ayrıntılı olarak ele alınmıştır.

Kümeleme analizi, bir grup içindeki tüm gözlemler benzer iken farklı gruplardaki gözlemler benzer olmayacak şekilde gözlem kümesini belli (sınırlı) sayıda bilinmeyen gruplar ya da kümeler halinde böler⁵.

Kümeleme analizinde temel amaç birimlerin (veya değişkenlerin) doğal gruplanmasını keşfetmek ve nispeten homojen bir grup halinde birimleri düzenlemektir⁶. Bunun yanında kümeleme analizi için; (a) veri keşfi, (b) veri indirgeme, (c) hipotez üretme ve (d) gruplara dayalı tahmin gibi çeşitli amaçlardan da söz edilebilir⁷. Ayrıca kümeleme analizi nesneleri gruplar halinde sınıflandırmak için kullanılan bir araçtır ve daha düşük boyutlu bir uzayda, nesnelerin geometrik gösterimi ile ilgili değildir. Uzayın boyutluluğunu araştırmak için çok boyutlu ölçekleme yöntemi kullanılabilir⁸.

Günümüzde biyoloji, zooloji, kimya, yerbilimleri, tıp, mühendislik, işletme ve ekonomi bilimi ve sosyal bilimlerin literatürleri kümeleme analizi çalışmaları ile doludur⁹. Son olarak veri madenciliği alanında, büyük veri gruplarından kurtulup işe yarar özet veriler elde etmek amacıyla yaygın olarak kullanılır¹⁰.

³ Paul A.Gore jr., Cluster Analysis, Department of Psychology, Southern Illinois University, (Çevrimiçi), [https://www.hse.ru/data/2012/01/04/1262163878/Gore%20Paul%20A.%20\(2000\)%20Cluster%20Analysis.pdf](https://www.hse.ru/data/2012/01/04/1262163878/Gore%20Paul%20A.%20(2000)%20Cluster%20Analysis.pdf), 26.01.2016, p. 298; Orhunbilge, **a.g.e.**, s. 473.

⁴ Girginer, **a.g.e.**, p. 61 ;Timm, **a.g.e.**, p. 515.

⁵ Timm, **a.g.e.**, p. 515.

⁶ Johnson, Wichern, **a.g.e.**, p. 673 ; Gore, **a.g.e.**, p.

⁷ Chatfield, Collins, **a.g.e.**, p. 214.

⁸ Timm, **a.g.e.**, p. 515.

⁹ Gore, **a.g.e.**, p. 298.

¹⁰ Günay Atbaş, **a.g.e.**, s. 11.

Grup içerisindeki verilerin birbirine benzerlikleri ve farklı gruplarda yer alan verilerin ne ölçüde ayrı olduklarına bakılarak; yapılan kümeleme işleminin ne kadar başarılı olduğuna karar verilir¹¹.

Aynı veri kümesine uygulanacak farklı sınıflandırma yöntemleri farklı sonuçlar meydana getirebilir. Bu nedenle tek bir kümeleme yöntemine dayanan analiz sonuçları değerlendirileceğine birkaç farklı yöntemin denenmesi daha güvenilir sonuçlar ortaya çıkaracaktır¹².

3.2. Kümeleme Analizi İçin Varsayımlar

Araştırmada yer alan gözlemlerin/birimlerin yapısal özelliklerini ölçmek amacıyla geliştirilmiş objektif bir yöntem olan kümeleme analizi, küme yapıları ya da kümelerin sayısına ilişkin herhangi bir varsayımda bulunmaz. İstatistik özelliklerden daha çok matematiksel özelliklere sahip olduğundan burada normallik varsayımı, diğer çok değişkenli istatistik analiz yöntemlerinde önemli kabul edilen, nispeten önemsiz kalır ve uzaklık değerlerinin normalliği yeterli kabul edilir.¹³ Aynı şekilde doğrusallık ve eşit varyanslılık gibi varsayımlar da kümeleme analizi için fazla önem taşımaz¹⁴.

Kümeleme analizinde başarılı sonuçlar alınması, güvenilir bir örnek seçimine ve seçilen bu örneğin ana kütle yapısını doğru bir şekilde temsil etmesine diğer bir deyişle örneklemin temsil gücüne bağlıdır. Dikkate alınması gereken bir diğer konu değişkenler arasındaki çoklu bağlantı durumudur. Diğer çok değişkenli yöntemler için önemli olan çoklu bağlantı kümeleme analizi için de önemlidir ancak etkisi biraz farklı olmaktadır. Kümeleme analizinde, aralarında çoklu bağlantı bulunan değişkenler üstü kapalı olarak diğerlerine göre daha fazla ağırlıklandırılmış olur¹⁵.

¹¹ Kumru Didem Atalay, “Kümeleme Analizi”, (Çevrimiçi), <http://tip.baskent.edu.tr/kw/upload/600/dosyalar/cg/sempozyum/ogrsmpzsnm13/13.P9.pdf> , 18.09.2016.

¹² Günay Atbaş, **a.g.e.**, s. 1.

¹³ Şeref Kalaycı, **SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri**, 6. Baskı, Asil Yayın, Ankara, 2014, s. 358; Girginer, **a.g.e.**, p. 61.

¹⁴ Orhunbilge, **a.g.e.**, s. 477.

¹⁵ Ünlükaplan, **a.g.e.**, s. ;Kalaycı, **a.g.e.**, s. 358.

3.3. Kümeleme Analizi Yöntemleri

Kümeleme analizini uygulamak için tasarlanan algoritmalar genellikle, hiyerarşik ve hiyerarşik olmayan kümeleme yöntemleri olarak iki ana sınıfa ayrılır. Hiyerarşik yöntemlerde, sırayla en benzer nesne çiftleri ile başlayarak kümeler oluşturulur ve adım adım daha yukarı kümelenme şekillenir. Hiyerarşik olmayan yöntemlerde ise nesne çiftlerinin tüm dağılımları değerlendirilir ve sonrasında bu nesne çiftleri belirli sayıda gruplar içine sınıflandırılır¹⁶.

Kümeleme analizinde ilk aşama benzerlik ölçüsünün seçilmesidir. Daha sonra kullanılacak kümeleme tekniğinin türü (hiyerarşik yada hiyerarşik olmayan gibi) ile ilgili karar verilir. Üçüncü adımda, seçilen teknik için kümeleme yönteminin türü (hiyerarşik kümeleme tekniğinde Centroid yöntemi gibi) belirlenir. Dördüncü adımda ise küme sayısı hakkında bir karar verilir ve son olarak kümeleme sonucu yorumlanmaktadır¹⁷.

3.4. Hiyerarşik Kümeleme Analiz Yöntemleri

Aşamalı kümeleme yöntemleri olarak da bilinen bu yöntemde amaç; benzerlikleri göz önüne alınan birimleri ya da değişkenleri belirli düzeylerde bir araya getirmektir. Bu süreçte, kümeler hiyerarşik olarak birleştirilir ve bir küme diğer bir küme ile ilk kez birleştiriliyorsa diğer adımlarda bu kümeden ayrılamaz. Birimlerin (ya da değişkenlerin) ilk aşamada kaç küme oluşturduğuna ve küme elemanlarının belirlenmesinde hangi kriterlerin göz önüne alındığına bakılarak iki gruba ayırmak mümkündür¹⁸.

Birleştirici (agglomerative) hiyerarşik kümeleme analizi yöntemlerinde her bir nesne bir küme olarak düşünülür. Böylelikle, başlangıçta nesneler kadar çok kümeler var olacaktır. Sonrasında en az uzaklığa sahip iki küme yeni bir kümede birleştirilir ve her adımda küme sayısının bir eksiği küme oluşur¹⁹.

¹⁶ Timm, **a.g.e.**, p. 523 ;Girginer, **a.g.e.**, p. 62.

¹⁷ Sharma, **a.g.e.**, p. 187.

¹⁸ Özdamar, **a.g.e.**, s. 269.

¹⁹ Günay Atbaş, **a.g.e.**, s. 15 ; Johnson, Wichern, **a.g.e.**, p. 680.

Ayırıcı (divisive) hiyerarşik kümeleme yöntemlerinde ise birleştirici hiyerarşik yöntemlerin tam tersi yönünde çalışır. Başlangıçta tüm gözlemlerin büyük bir küme oluşturduğu düşünülür ve sonrasında birbirlerine benzemeyen gözlemler tek tek ayıklanarak daha küçük kümeler meydana getirilir. Gözlemlerin her biri tek başına bir küme oluşturduğunda süreç tamamlanmış olur²⁰.

Literatüre bakıldığında, ayırıcı hiyerarşik kümeleme yöntemlerinin sıklıkla tercih edilen bir yöntem olmadığı, çoğu zaman birleştirici hiyerarşik yöntemlerin kullanıldığı görülmektedir²¹.

Hiyerarşik yöntemlerde kümeleme sürecinin işleyişi ve aşamaların daha kolay anlaşılabilmesi için “dendrogram” (ağaç diyagramı) adı verilen grafiklerden yararlanılır. Her birimin kümeleme süreci başında oluşturduğu kümeler ağacın dalları ile gösterilirse; süreç sonunda tüm birimlerin toplandığı küme ağacın gövdesini temsil edecektir²².

3.4.1. Birleştirici Hiyerarşik Kümeleme Analizi Algoritması

Birleştirici kümeleme analizi için çoğu algoritma, nesneler arasındaki uzaklık ya da benzerlik çiftlerinin matrisi ile başlamaktadır²³ ve adımlar aşağıdaki gibidir²⁴:

Adım 1: Her biri tek bir varlık içeren n küme ile başlar. $D = \{d_{rs}\}$ $n \times n$ boyutlu simetrik uzaklıklar (ya da benzerlikler) matrisini göstermektedir.

Adım 2: Kümelerin en yakın (en benzer) çifti için uzaklık matrisi incelenir. Burada d_{rs} ; r ve s en benzer kümeleri arasındaki uzaklığı göstermektedir.

Adım 3: r ve s kümeleri birleştirilir. (a) r ve s kümelerine karşılık gelen satır ve sütunlar silinir ve (b) (rs) kümesi ile kalan kümeler arasındaki uzaklığı veren satır ve sütun eklenerek uzaklıklar matrisinde girişler güncellenir.

²⁰ Şenol Çelik, “Kümeleme Analizi ile Sağlık Göstergelerine Göre Türkiye’deki İllerin Sınıflandırılması”, *Doğuş Üniversitesi Dergisi*, 14, (2), 2013, s. 181.

²¹ Özdamar, *a.g.e.*, s. 269.

²² Tatlıdil, *a.g.e.*, s. 255.

²³ Gerry P. Quinn, Michael J. Keough, *Experimental Design and Data Analysis*, Cambridge University Press, New York, 2002, p. 490.

²⁴ Johnson, Wichern, *a.g.e.*, p. 681-682.

Adım 4: 2 ve 3. Adımlar n-1 kez tekrarlanır. Algoritma sonlandığında tüm nesneler tek bir kümede olacaktır.

Bağlantı, merkezi ve hata varyansı yöntemleri olmak üzere üç kategoriye ayrılan birleştirici hiyerarşik yöntemler arasında bir tek bağlantı algoritması nesneleri ya da değişkenleri kümelemek için kullanılabilir. Bunun dışında kalan iki yöntem sadece nesneleri kümelemektedir²⁵.

İki küme arasındaki uzaklığın hesaplanması amacıyla çok sayıda farklı kurallar ya da yöntemler ileri sürülmüştür²⁶. İlk adım (birinci kümenin oluşumu) tüm yaklaşımlar için aynıdır ancak ilk adımdan sonra, çeşitli yöntemler kümeler arasındaki uzaklığı hesaplamak için kullanılan prosedüre göre değişiklik göstermektedir²⁷.

Birleştirici kümeleme yöntemlerinde dendogram örnekleri için Ek-3'deki veriler kullanılmıştır.

3.4.2. Tek Bağlantı Yöntemi

Bu yöntemde, iki küme arasındaki uzaklık iki en yakın üye ya da komşu arasındaki mesafe olarak tanımlanır²⁸. Bu nedenle en yakın komşuluk yöntemi de denir ve farklı veri yapılarındaki kümelerin tanımlanmasına olanak verdiğinden sıklıkla tercih edilmektedir²⁹. Burada “en yakın komşu” kavramı en küçük uzaklığı ya da en büyük benzerliği ifade etmektedir³⁰.

İlk olarak uzaklıklar matrisinde en küçük uzaklığa sahip iki gözlem birleştirilerek birinci küme oluşturulur. Sonrasında bir sonraki en küçük uzaklık tespit edilir ve tüm gözlemler tek bir kümede gruplanana kadar bu süreç devam eder³¹.

²⁵ Timm, **a.g.e.**, p. 515.

²⁶ Sharma, **a.g.e.**, p. 191.

²⁷ **A.e.**

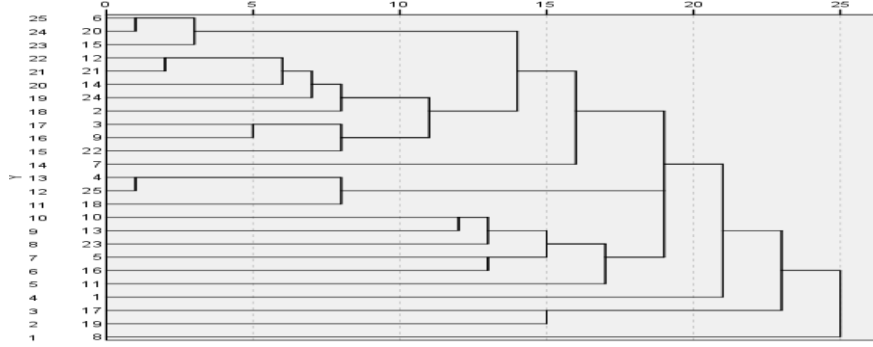
²⁸ Statistics: 3.1 Cluster Analysis, (Çevrimiçi) <http://www.statstutor.ac.uk/resources/uploaded/clusteranalysis.pdf>, p. 2, 15.09.2016.

²⁹ Alphar, **a.g.e.**, s. 324.

³⁰ Johnson, Wichern, **a.g.e.**, p. 682.

³¹ Alphar, **a.g.e.**, s. 324.

Şekil 3.1: Tek Bağlantı Yöntemi ile Oluşturulan Dendogram Örneği



Burada; k'ncü kümenin, r ve s'inci birimlerden oluşan birleştirilmiş küme ile ilişkisi uzaklık ölçütü olarak,

$$d_{k(r,s)} = \text{Min}(d_{kr}, d_{ks}) \quad (3.2)$$

eşitliği ile ifade edilmektedir. Bu eşitlikte; $d_{k(r,s)}$ oluşturulan r ve s kümeleri ile k. küme arasındaki uzaklığı, d_{kr} k'ncü küme ile r'inci küme arasındaki uzaklığı ve d_{ks} k'ncü küme ile s'inci küme arasındaki uzaklığı göstermektedir³².

Bu yöntem diğerlerine nazaran basittir ve doğal kümeler küresel ya da eliptik formda olmadığı zaman diğer yöntemlerden daha iyi sonuçlar vermektedir³³. Temel dezavantajı sözde “zincirleme” etkiye sahip olmasıdır³⁴.

3.4.3. Tam Bağlantı Yöntemi

Tam bağlantı yöntemi en yakın komşu yönteminin tam tersidir. İki küme arasındaki uzaklık, iki kümedeki tüm mümkün gözlem çiftleri arasındaki uzaklığın maksimumu olarak tanımlanır³⁵.

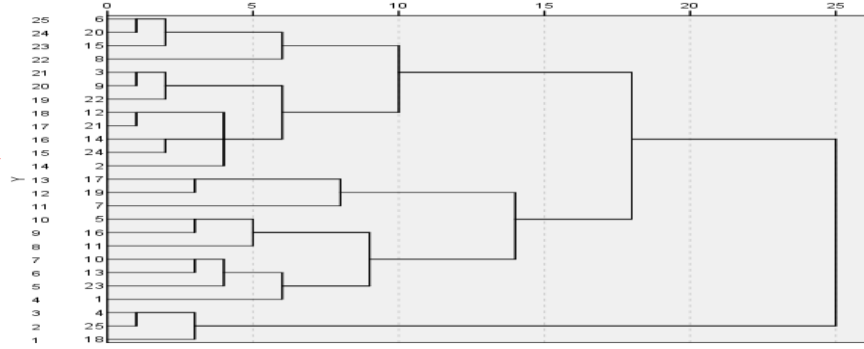
³² Tatlıdıl, **a.g.e.**, s. 256. ; Atbaş, **a.g.e.**, s. 16.

³³ Statistics: 3.1 Cluster Analysis, **a.g.e.**, p. 2.

³⁴ Chatfield, Collins, **a.g.e.**, p. 228.

³⁵ Sharma, **a.g.e.**, p. 192.

Şekil 3.2: Tam Bağlantı Yöntemi ile Oluşturulan Dendogram Örneği



En uzak komşuluk yöntemi olarak da adlandırılan bu yöntemde, ilk olarak uzaklıklar matrisindeki en yakın iki gözlem ile kümelemeye başlanır ve sonrasında küme gözlemleri arasındaki en büyük uzaklık ya da en az benzerlik dikkate alınarak yeni iki küme oluşturulur³⁶. Tam bağlantı yönteminde uzaklıklar,

$$d_{k(i,j)} = \text{Max} (d_{ki}, d_{kj}) \quad (3.3)$$

eşitliği ile elde edilmektedir.

3.4.4. Ortalama Bağlantı Yöntemi

İki küme arasındaki uzaklık, iki küme içerisindeki tüm nesne çiftleri arasındaki ortalama uzaklık olarak hesaplanır. Oldukça sağlam bir yöntem olarak kabul edilir³⁷. İlk kümenin oluşturulması tek bağlantı ve tam bağlantı yöntemlerinde olduğu gibidir³⁸. Sonrasında grup içi ve gruplararası varyans değeri küçük olan kümeler bir araya getirilir. Aşırı değerlerden etkilenmeyen bu yöntemde kümelerdeki tüm birimler işleme sokulmaktadır³⁹.

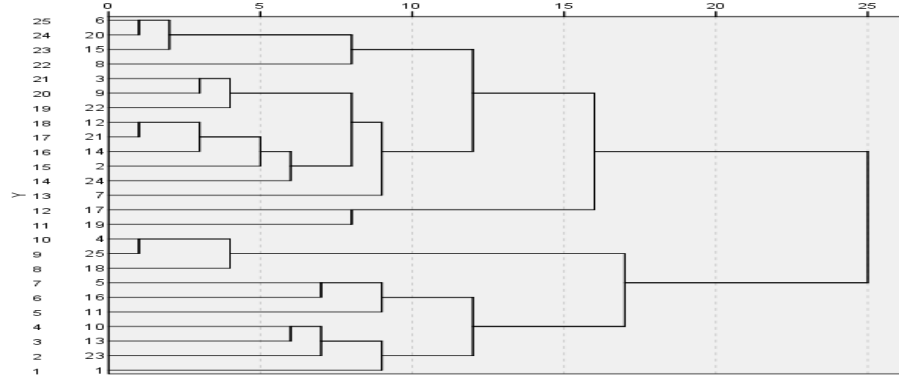
³⁶ Alpar, **a.g.e.**, s. 330.

³⁷ Statistics: 3.1 Cluster Analysis, **a.g.e.**, p. 3.

³⁸ Günay Atbaş, **a.g.e.**, s. 17.

³⁹ Orhunbilge, **a.g.e.**, s. 475.

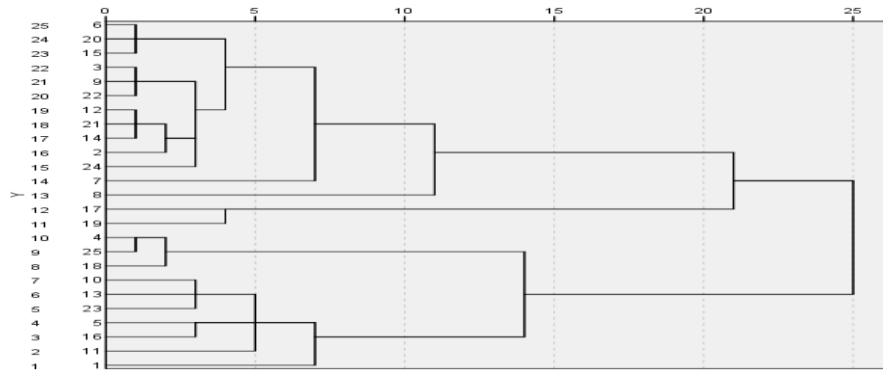
Şekil 3.3: Gruplariçi Ortalama Bağlantı Yöntemi ile Oluşturulan Dendogram Örneği



3.4.5. Merkezi Yöntem

Kümelerde yer alan birimlerin ortalama değeri ilgili kümenin merkezi olarak kabul edilen bu yöntemde; kümeler arasındaki uzaklık hesaplanırken, her bir küme kümenin merkezi tarafından temsil edilir⁴⁰. Diğer hiyerarşik kümeleme yöntemlerine göre aşırı değerlerden en az etkilenen yöntemdir ve genellikle karesel Öklid uzaklığı ile kullanılır⁴¹.

Şekil 3.4: Merkezi Yöntem ile Oluşturulan Dendogram Örneği



⁴⁰ Sharma, **a.g.e.**, p.190; Orhunbilge, **a.g.e.**, s. 474.

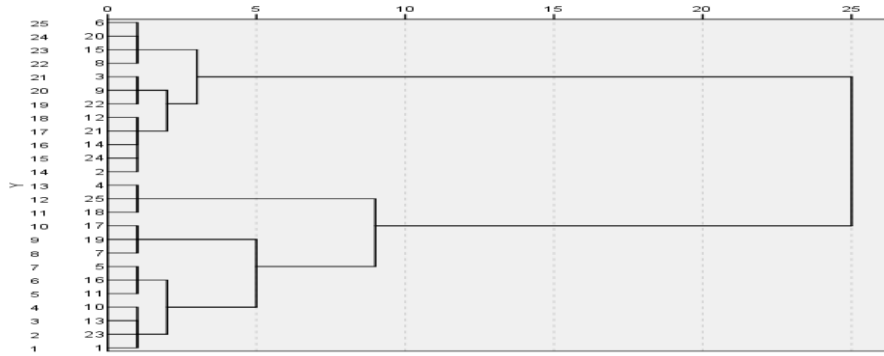
⁴¹ Alphar, **a.g.e.**, s. 333.

3.4.6. WARD Yöntemi

Ward yöntemi, kümeler içi homojenliği en üst düzeye çıkararak kümeleri oluşturur ve toplam grup içi ya da küme içi kareler toplamını en aza indirmeyi amaçlar⁴².

k küme için, ESS_k , kümedeki her birimin küme ortalamasından sapmaların kareleri toplamını verir. Mevcut durumda k küme varsa, ESS , ESS_k 'ların toplamı yada $ESS = ESS_1 + ESS_2 + \dots + ESS_k$ olarak tanımlanmaktadır. Analizin her adımında, tüm mümkün küme çiftlerinin birleşmesi dikkate alınır ve iki küme, ESS 'de en küçük bir artışa neden olan birleşme ile (minimum bilgi kaybı), birleştirilir.

Şekil 3.5: Ward Yöntemi ile Oluşturulan Dendogram Örneği



Başlangıçta, her küme tek bir birimden oluşacağından ve tüm ESS_k 'lar ve dolayısıyla ESS sıfır olmaktadır. Tüm kümeler, n birimli tek bir grup da birleştirildiğinde, ESS değeri,

$$\sum_{j=1}^n (y_j - \bar{y})' (y - \bar{y}) \quad (3.4)$$

eşitliği ile elde edilir. Burada, y_j j'inci birim ile ilgili çok değişkenli ölçümü ve $\bar{y}$ tüm birimlerin ortalamasını göstermektedir⁴³.

⁴² Sharma, **a.g.e.**, p. 193.

⁴³ Johnson, Wichern, **a.g.e.**, p. 693.

Ward yönteminin, kümeleri birbirlerine eşit sayıda birim içerecek biçimde oluşturma gibi yanlı eğilimleri ile birlikte sayıca az birime sahip kümeleri birleştirme eğiliminin de olduğu belirtilmektedir⁴⁴.

Ayrıntılı tanımlamalarını yaptığımız bu beş yöntem Tablo 3.1’de yardımıyla özetlenmiştir⁴⁵.

3.5. Hiyerarşik Olmayan Kümeleme Analizi Yöntemleri

Birimleri, p değişkenler açısından, küme içi homojen ve kümeler arası heterojen bir yapıda olacak şekilde kümelere ayırmayı amaçlayan yöntemlerdir⁴⁶. Araştırmacının küme sayısı ile ilgili bir ön bilgisi varsa ya da anlamlı olacak küme sayısına karar verebiliyorsa hiyerarşik olmayan kümeleme yöntemlerinin kullanılması tercih edilir⁴⁷.

Hiyerarşik olmayan kümeleme yöntemleri değişkenlerden ziyade birimleri k küme içine gruplamak için tasarlanmıştır. Kümelerin sayısı, k, önceden belirlenir ya da kümeleme işleminin bir parçası olarak tanımlanır⁴⁸. Bu çerçevede görev araştırmacınıdır. Küme başlangıç noktaları elde edilirken konu ile ilgili önceki çalışmalardan ya da bilgilerden yararlanabilir ya da kümeleme öncesinde hiyerarşik yöntemler uygulayarak yaklaşık bir sonuca ulaşılabilir⁴⁹. Ayrıca bu yöntemlerde süreç D uzaklıklar matrisi ile değil, ham veri matrisi Y kullanılarak başlar ve nesneler arasındaki ilişkileri hiyerarşik (sıra düzenli) bir biçimde göstermediğinden hiyerarşik kümelemenin aksine; kümeleme işlemi süresince nesneler kümelere yeniden atanırlar⁵⁰.

Bu tür yöntemlerde kaç küme olması muhtemel genellikle bilmek zordur ve bu nedenle analizin birçok kez tekrarlanması gerekebilir ve analiz, başlangıç küme merkezlerinin seçiminde çok hassas davranabilir. Bunlar hiyerarşik olmayan

⁴⁴ Orhunbilge, **a.g.e.**, s. 475.

⁴⁵ Brian S.Everitt, Sabine Landau, Morven Leese, Daniel Stahl, **Cluster Analysis**, 5th Edition, 2011, John Wiley and Sons, p. 79.

⁴⁶ Alphar, **a.g.e.**, s. 341.

⁴⁷ Tatlıdil, **a.g.e.**, s. 258.

⁴⁸ Johnson, Wichern, **a.g.e.**, p. 696.

⁴⁹ Alphar, **a.g.e.**, s. 342.

⁵⁰ Timm, **a.g.e.**, p. ; Quinn, Keough, **a.g.e.**, p. 491.

Tablo 3.4: Klasik Birleştirici Hiyerarşik Kümeleme Analizi Yöntemleri

Yöntem	Alternatif Adı	Genellikle Kullanılan	Kümeler Arası Uzaklık Tanımları	Açıklamalar
Tek bağlantı yöntemi Sneath (1957)	En yakın komşu	Benzerlik veya Uzaklık	Nesne çiftleri arasındaki minimum uzaklık	Özellikle büyük veri kümelerinde dengesiz ve dağınık kümeler üretme eğilimindedir. Küme yapısını dikkate almaz.
Tam bağlantı yöntemi Sorensen (1948)	En uzak komşu	Benzerlik veya uzaklık	Nesne çiftleri arasındaki maksimum uzaklık	Eşit çaplı, kompakt kümeler bulma eğilimindedir (Nesneler arası maksimum uzaklık). Küme yapısını dikkate almaz.
(Grup) Ortalama bağlantı yöntemi Sokal ve Michener (1958)	UPGMA	Benzerlik veya uzaklık	Nesne çiftleri arasındaki ortalama uzaklık	Küçük sapmalar ile kümeleri birleştirme eğilimindedir. Tek ve tam bağlantı yöntemlerinin arasında bulunur. Küme yapısını dikkate alır ve nispeten dayanıklıdır.
Merkezi (Centroid) Yöntem Sokal ve Michener (1958)	UPGMC	Uzaklık (ham veri gerektirir)	Ortalama vektörleri (centroid) arasındaki karesel Öklid uzaklığı	Varsayılan noktalar Öklid uzayında temsil edilebilir (geometrik yorumu için). Çok sayıda kümeleneşmiş iki gruptan daha çok birleştirilmiş küme hakimdir.
Ağırlıklandırılmış Ortalama Bağlantı Yöntemi McQuitty (1966)	WPGMA	Benzerlik veya uzaklık	Nesne çiftleri arasındaki ağırlıklandırılmış uzaklık	UPGMA’ da olduğu gibi, ancak küçük kümelerdeki noktalar büyük kümelerdeki noktalara göre daha çok ağırlıklandırılır. (Küme boyutları muhtemelen değişken olduğunda yararlıdır).
Medyan Bağlantı Yöntemi Gower (1967)	WPGMC	Uzaklık (ham veri gerekli)	Ağırlıklandırılmış centroidler arasındaki karesel Öklid uzaklıkları	Geometrik yorum için varsayılan noktalar Öklid uzayında temsil edilebilir. Yeni grup birleştirilmiş gruplar arasında konum olarak aradır.
Ward’s Yöntemi Ward (1963)	En küçük kareler toplamı	Uzaklık (ham veri gerekli)		Geometrik yorumlama için varsayılan noktalar Öklid uzayında temsil edilebilir. Aynı büyüklükte, küresel kümeler bulma eğilimindedir. Aykırı değerlere duyarlıdır.

yöntemlerin iki önemli dezavantajıdır⁵¹. Bunlara ek olarak, bu tür yöntemlerde olası küme sayısı önceden bilindiğinden ve buna bağlı olarak kümeler arası ilişkiler dikkate alınmadığından analiz sonucunda bir dendrogram oluşmaz².

Hiyerarşik olmayan kümeleme analizi, büyük veri setleri söz konusu olduğunda kullanılan olma eğilimindedir³. Hiyerarşik olmayan algoritmaların çoğu; (1) başlangıç küme ağırlık merkezi ya da çekirdeklerini elde etmede kullanılan yöntem ve (2) yeniden atanan gözlemler için kullanılan kural bakımından değişiklik gösterir⁴. Bu çalışmada en sık kullanılan hiyerarşik olmayan yöntemlerden k-ortalamlar yöntemi ele alınacaktır.

K-ortalamlar yöntemi, birimleri/nesneleri önceden tanımlanmış (k) kümenin içine bölerek çalışır ve sonrasında nesnelerin küme üyeliği, küme içi varyansa karşı kümeler arası varyansın oranını en üst düzeye çıkartmak gibi bazı kriterlere dayanarak iteratif olarak yeniden değerlendirilir⁵. Kümelemenin her aşamasında kontrol edilen “küme içi homojenlik” ve “kümeler arası heterojenlik” düzeyleri en üst seviyeye ulaştığında atama işlemleri son bulmaktadır⁶.

Mac Queens, k-ortalamlar terimini her bir birimin en yakın ağırlık merkezine (ortalamaya) sahip kümeye atandığı bir algoritma tanımlamak için önermiştir⁷.

K ortalamlar yöntemi, aşağıdaki adımları izleyerek birimleri kümeler içine böler⁸;

Adım 1: Verilerden elde edilen bilgilere dayanarak, ilk k nokta çekirdek nokta olarak alınır. Bu noktaların her birinde, p değişkenler küme ortalama vektörü olarak kabul edilir ve küme ortalama vektöründen her birimin uzaklıkları hesaplanır.

⁵¹ Statistics: 3.1 Cluster Analysis, a.g.e., p. 4.

⁵² Ünlükaplan, a.g.e., s. 22.

⁵³ Statistics: 3.1 Cluster Analysis, a.g.e., p.4.

⁵⁴ Sharma, a.g.e., p. 202.

⁵⁵ Quin, Keough, a.g.e., p. 491.

⁵⁶ Özdamar, a.g.e., s. 296.

⁵⁷ Johnson; Wichern, a.g.e. p.

⁵⁸ Girginer, a.g.e., s. 323.

Adım 2: Geriye kalan $n-k$ birim, en yakın ortalama vektör kümesine atanır. Her atama sonrası, kümenin ortalama vektörü yeniden değerlendirilir. Böylece, çekirdek noktaların ortalama vektörleri değiştirilmiş olur ve birimlerin yeni oluşan küme ortalama vektörüne uzaklıkları tekrar hesaplanır.

Adım 3: Kümeler arası varyansın maksimum, küme içi varyansın minimum olduğu duruma kadar tüm birimler k kümeye atanır. Bu süreçte kümeleme tekrarlayan bir yaklaşımla elde edilir. Küme içi kovaryans matrisinin minimum olduğu koşul sağlanıncaya ve varyans farklılığı kriterine eşit ya da daha küçük oluncaya kadar parçalama (bölme) işlemi devam eder.

DÖRDÜNCÜ BÖLÜM

SAĞLIK GÖSTERGELERİNİN İNCELENMESİ

Genel olarak sağlık kavramı, insanların hasta olmaması şeklinde algılanmasına rağmen; Dünya Sağlık Örgütü (DSÖ)'e göre sağlık, “sadece hastalık ve sakatlığın olmayışı değil, bedensel, ruhsal ve sosyal yönden tam bir iyilik hali” olarak tanımlanmıştır¹.

Sağlık hizmetlerinin amacı, toplumun belirli bir kesimine yönelik değil, tamamına ilişkin genel sağlık koşullarını geliştirmek ve iyileştirmektir². Diğer bir ifadeyle iyi bir sağlık hizmeti almak insanların; ırk, din, siyasi görüş, ekonomik ya da sosyal koşulları ile ilgili herhangi bir ayrım yapmaz ve her insanın temel hakkıdır. Buradan da anlaşılabacağı gibi toplumsal bir özelliğe sahiptir³.

Ekonomik gelişmişlik düzeyi ile halk sağlığı arasında yakın ve karşılıklı nedensel bir ilişki vardır. Yapılan çalışmalarda, ekonomik göstergelerdeki iyileşmelerin sağlık göstergeleri üzerinde olumlu bir etkiye sahip olduğu gözlenmiştir. Refah düzeyi yüksek olan gelişmiş ülkeler, insan gücüne yatırım yapmasından dolayı ve buna ek olarak ekonomik gelişmenin ana unsuru olan hastalık tedavisi ve insan sağlığının korunması amacıyla sağlık hizmetlerine her geçen yıl GSYH'nın çok daha fazlasını fon olarak tahsis etmektedir⁴.

Toplumun geleceğini etkileyen sağlık sektörü, 1960 yılından bu yana dünya çapında kapsamlı çalışmalara uğramıştır. Birçok ülkede sağlık sistemlerini geliştirmek amacıyla yeni uygulamalar ve sistem değişiklikleri yapılmıştır. Günümüzde kritik konularda tam ve zamanında veri elde ederek; sağlık hizmetlerine

¹ Constitution of The World Health Organization, (Çevrimiçi), Basic Documents, Forty-Fifth Edition, October, 2006, http://www.who.int/governance/eb/who_constitution_en.pdf , p. 1, 25.04.2017.

² Filiz Ersöz, “OECD’ye Üye Ülkelerin Seçilmiş Sağlık Göstergelerinin Kümeleme ve Ayırma Analizi ile Karşılaştırılması”, **Türkiye Klinikleri J Med Sci**, 29 (6), 2009, s. 1651.

³ WHO, Filiz Ersöz, **a.g.e.**, s. 1651

⁴ Filiz Ersöz, “Türkiye ile OECD Ülkelerinin Sağlık Düzeyleri ve Sağlık Harcamaları Analizi”, **İstatistikçiler Dergisi**, 2, 2008, s. 96., Girginer, **a.g.e.**, p. 56.

eriřim, eřitlik, saęlık politikaları, performans vb. gibi saęlık sistemleri ile ilgili gözetim ve deęerlendirme yapmak mümkündür. Bu amaçla binlerce gösterge kullanılabilir⁵. Ancak göstergelerin teknik çeřitlięi ve ülkelerin saęlık sistemlerinde ki farklılıklardan dolayı yaşam beklentisi, doğum ve ölüm oranları, saęlık harcamaları gibi temel saęlık göstergelerinin kullanılması yaygındır⁶.

4.1. Türkiye’de Saęlık Politikalarının Tarihsel Geliřimi

Osmanlı döneminde var olan tıp geleneęi, Cumhuriyetin kurulması ile birlikte geliştirilmiř; tüm kurumların ve hizmet politikalarının elde edilmesinde Avrupa ülkeleri örnek alınmıřtır⁷.

1920-1923 yılları arasındaki dönemde, Türkiye Büyük Millet Meclisi (TBMM)’ nin kurulması ile saęlık hizmetleri öncelikli bir devlet görevi olarak incelenmiř ve saęlık hizmetlerinin farklı bir koldan idaresi için Saęlık Bakanlığı kurulmuřtur⁸.

1923-1946 yılları arasında Cumhuriyet’in ilan edilmesi ile birlikte saęlık bakanlıęına atanan Dr. Refik Saydam, saęlık hizmetlerinde örgütlenme ve hizmetlerin yurt geneline yayılmasında önemli bir başarı elde etmiřtir. O yıllardan günümüze kadar ulařan veriler Tablo 4.1’de gösterilmektedir⁹.

⁵ Canser Boz, Haydar Sur, Selma Söyüç: “The Similarities and Differences Analysis of OECD Countries in Terms of Health System Indicators”, **ACU Saęlık Bil Derg**, 3, 2016, s. 155.

⁶ Girginer, **a.g.e.**, s. 56; Boz ve ark., **a.g.e.**, s.155.

⁷ Saęlık Bakanlığı Tarihçe, (Çevrimiçi), <https://www.saglik.gov.tr/TR,11492/tarihce.html> , 06.03.2017.

⁸ Recep Akdur, “Türkiye’de Saęlık Hizmetleri ve Avrupa Topluluęu Ülkeleri ile Kıyaslanması”, A.Ü.T.F. Halk Saęlığı Anabilim Dalı, Ankara, 1999, s. 11.

⁹ T.C. Saęlık Bakanlığı Stratejik Plan 2013-2017, (Çevrimiçi), <https://sgb.saglik.gov.tr/content/files/stratejikplan20132017/index.html> , s. 20, 06.03.2017.

Tablo 4.1: 1923 Yılına Ait Sağlık Göstergeleri*

	1923
Yataklı Tedavi Kurumu	86
Hasta Yatak Sayısı	6437
Hekim Sayısı	554
Eczacı Sayısı	69
Hemşire Sayısı	4
Sağlık Memuru Sayısı	560
Ebe Sayısı	136

*Kaynak: T.C. Sağlık Bakanlığı Stratejik Plan 2013-2017, Aralık 2012, s. 20.

Refik Saydam dönemi olarak adlandırılan bu dönemde, yerel idareler politikası izlenerek hastane açılması desteklenmiş ve büyük illerde, Ankara, İstanbul, Sivas gibi, bakanlığa bağlı numune hastaneleri açılmıştır. Merkezi yönetimin hizmet alanı koruyucu hekimlik, yerel yönetimlerin hizmet alanı ise tedavi edici hekimlik olarak ayrılmıştır. 1930 yılında sağlık hizmetlerinin anayasası olma özelliği gösteren “Umumi Hıfzıssıhha Kanunu” çıkarılmıştır. Bu kanunla sağlık uygulamalarının ve politikasının ana noktaları tanımlanmıştır¹⁰. Ayrıca yine bu dönemde cüzzam, sıtma, verem, frengi gibi bulaşıcı hastalıklarla mücadele programlarının başlatılmıştır¹¹.

1946-1960 yılları arasında “Birinci On Yıllık Milli Sağlık Planı” onaylanmıştır. Cumhuriyet döneminin ilk yazılı sağlık planı olma özelliğine sahip olan bu plan kanunlaşmamış olmasına rağmen içeriği ile ülkemizin sağlık yapılanmasında önemli bir yer elde etmiştir. Yine bu dönemde; yerel yönetimin elinde olan sağlık hizmetlerinin merkezden yönetilmesine kararlaştırılmış ve böylece koruyucu sağlık hizmetleri ile tedavi edici hekimlik bir arada verilmesi planlanmıştır. Buna bağlı olarak geçmişte sadece birkaç büyük kentin yararlandığı hizmetlerden kırsal bölgelerinde faydalanması amaçlanmıştır.

¹⁰ Akdur, a.g.e., s. 12.

¹¹ TC. Sağlık Bakanlığı Stratejik Plan 2013-2017, a.g.e., s. 21.

Tablo 4.2: 1950 ve 1960 Yıllarına Ait Sağlık Göstergeleri*

	1950	1960
Hekim Sayısı	3.020	8.214
Hemşire Sayısı	721	1.658
Ebe Sayısı	1.285	3.219
Kurum Sayısı	118	442
Yatak Sayısı	14.581	32.398

***Kaynak:** T.C. Sağlık Bakanlığı Stratejik Plan 2013-2017, Aralık 2012, s. 23.

1960-1980 yılları arasında yaygın, sürekli ve il içerisinde bütünleşmiş bir yapılandırmaya gidilmiştir. Artan nüfusa karşılık 1965 yılında çıkartılan “Nüfus Planlaması Hakkında Kanun” ile birlikte nüfus arttırıcı, doğumu teşvik eden, politikadan ziyade nüfus artışında sınırlama sağlayan, doğum kontrolü, politikalar uygulanmaya başlanmıştır. Yine bu dönemde, “Genel Sağlık Sigortası” için bir kanun taslağı oluşturulmuş ancak bu taslak kanunlaşamamıştır.

1982 yılı anayasasının ilgili maddeleri ile vatandaşların sosyal güvenlik haklarına sahip olmaları ve bu hakların kullanılması devletin güvencesi altına alınmıştır. 1990’lı yıllarda sosyal güvenlik kurumlarının tek çatı altında toplanması, aile hekimliği gibi konularda yapılan reform çalışmaları uygulama alanı bulamayıp teoride kalmıştır. Bu dönemin sağlıkta gerçekleşen en önemli olayı 1992 yılında çıkarılan yasa ile sosyal güvencesi olmayan dar gelirli vatandaşların yeşil kart uygulaması ile sağlık güvencesi kapsamına alınmasıdır¹².

2003 yılından günümüze kadar geçen sürede sağlıkta dönüşüm programı çerçevesinde köklü değişiklikler yapılmıştır. Bu dönemde; emekli sandığı, SSK ve Bağ-Kur SGK adı altında tek bir kurumda birleştirildi ve SSK hastaneleri başta olmak üzere diğer kamu hastaneleri sağlık bakanlığına devredildi. “112 Acil Sağlık” hizmetinin alanı genişletilerek köylere de hizmet götürülmeye başlandı. Yine bu dönemde, aile hekimliği uygulaması hayata geçirilmiş, ilaçlardan alınan vergiler düşürülerek ilaç fiyatlandırma sistemi değiştirilmiştir. Son olarak bulaşıcı hastalık göstergelerinde, ekonomik açıdan gelişmiş ülkelerin seviyesine gelinmiştir¹³.

¹² T.C. Sağlık Bakanlığı Stratejik Plan 2013-2017, **a.g.e.**, s. 25-27, T.C. Sağlık Bakanlığı Tarihçe, **a.g.e.**

¹³ **A.e.**

4.2. Sağlık Göstergelerinin Yıllara Göre İncelenmesi

Kamu sağlık politikasının temel amacı, bir ülkenin sağlık durumunu korumak ve geliştirmektir. Bu nedenle, insan sağlığına katkıda bulunan faktörleri tanımlamak önemli olmaktadır¹⁴. Bebek ölüm oranları, ölüm nedenleri ve hastalıklar, ülkedeki doktor, hemşire, eczacı sayısı gibi sağlık sistemine ait göstergeler ve ortalama yaşam süresi gibi birçok faktör, bir toplumun sağlık düzeyini etkilemektedir.

20.yy’da dünya geneline bakıldığında sağlık sektöründe köklü gelişmeler olduğu görülmüştür. Beslenme ve sağlık koşullarındaki iyileşmeler, tıbbın ve tıptaki teknolojinin gelişmesi, bulaşıcı olan hastalıklara karşı getirilen aşılar, sağlık sektöründe sosyal devlet politikalarında zamanla ortaya çıkan gelişmeler günümüzde insanın ortalama ömrünün uzamasına neden olmuştur.

Tablo 4.3: 2005-2015 Yılları Arası Sağlık Personeli Sayısı*

Yıl	Hekim	Diş Hekimi	Hemşire	Sağlık Memuru	Ebe	Eczacı
2005	100.853	18.149	78.182	62.035	43.541	22.756
2006	104.475	18.332	82.626	67.514	44.483	23.140
2007	108.402	19.278	94.661	79.441	47.175	23.977
2008	113.151	19.959	99.910	89.540	47.673	24.778
2009	118.641	20.589	105.176	93.550	49.357	25.201
2010	123.447	21.432	114.772	99.302	50.343	26.506
2011	126.029	21.099	124.982	110.862	51.905	20.089
2012	129.772	21.404	134.906	122.663	53.466	26.571
2013	133.775	22.295	139.544	134.488	53.427	27.012
2014	135.616	22.996	142.432	149.616	52.838	27.199
2015	141.259	24.834	152.803	145.943	53.086	27.530

*Kaynak: Yıllara Göre İstatistikler/TÜİK

Artan nüfus ile birlikte meydana gelen hastalıklar ve hasta sayısında da artış meydana gelmekte ve bunun sonucunda hastane sayısı, hekim ve hemşire sayıları, diş hekimi, eczacı sayıları gibi göstergeler de yıllara göre değişiklik gösterecektir. Ayrıca her yıl açılan yeni üniversiteler ile birlikte Tıp Fakültesi sayılarında artış olması şaşırtıcı değildir¹⁵.

¹⁴ Girginer, a.g.e., s. 56.

¹⁵ Çelik, a.g.e., s. 176.

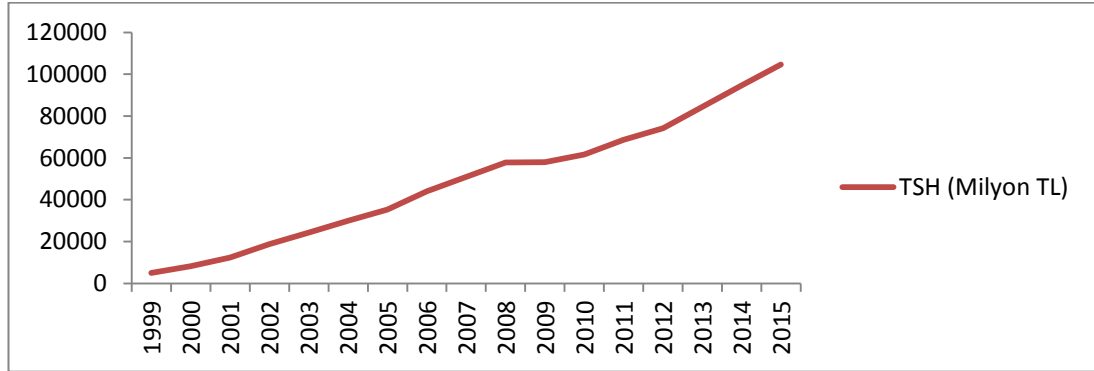
Tablo 4.4: Yıllara Göre Yataklı Tedavi Kurumlarındaki Cihaz Sayısı, Türkiye*

	2002	2011	2012	2013	2014	2015
MR	58	709	720	751	757	794
BT	323	974	1.017	1.058	1.071	1.119
Ultrason	1.005	3.775	4.282	4.756	5.286	5.518
Doppler Ultrason	681	2.091	2.480	2.793	3.151	4.015
EKO	259	1.181	1.379	1.542	1.793	1.897

***Kaynak:** T.C. Sağlık Bakanlığı, Sağlık İstatistikleri Yıllığı 2015, Sağlık Araştırmaları Genel Müdürlüğü, Bölüm 7, s. 108.

Tablo 4.4’de yataklı tedavi kurumlarındaki tıbbi cihaz sayılarının yıllara göre değişimi verilmiştir. Manyetik Rezonans (MR) Görüntüleme ünitesi, Bilgisayarlı Tomografi (BT) cihazı, Doppler Ultrason sayıları yıllara oranla incelendiğinde artan bir trende sahip olduğu görülmektedir. 2015 verilerine bakıldığında; yataklı tedavi kurumlarında 1.000 kişiye düşen MR sayısı 143, BT sayısı 174, Ultrason sayısı 348, Doppler Ultrason sayısı 138 ve EKO sayısının 88 olduğu görülmüştür¹⁶.

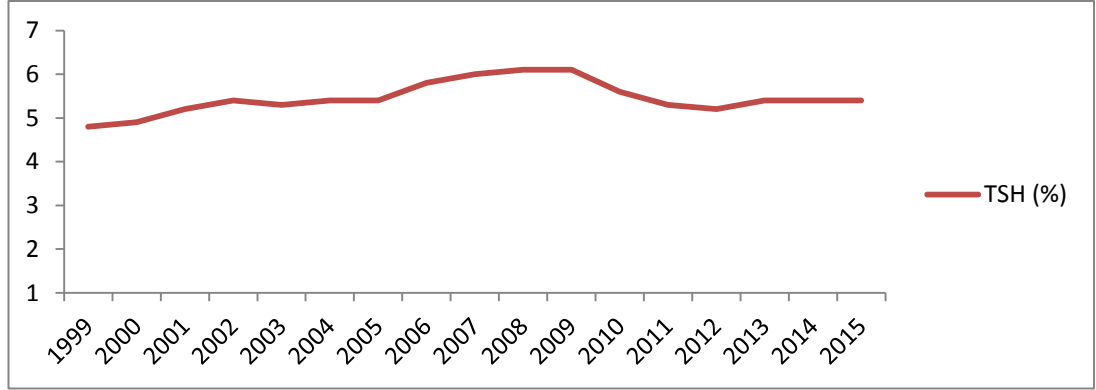
Şekil 4.1: Yıllara Göre Toplam Sağlık Harcamaları (Milyon TL)*



***Kaynak:** Yıllara Göre İstatistikler/TÜİK

¹⁶ TC. Sağlık Bakanlığı Sağlık İstatistikleri Yıllığı: 2015, Sağlık Araştırmaları Genel Müdürlüğü, Bölüm 7, s. 108.

Şekil 4.2: Toplam Sağlık Harcamasının Gayrisafi Yurtiçi Hasılaya Oranı (%)*



*Kaynak: Yıllara Göre İstatistikler/TÜİK

Sağlık harcamaları sağlık göstergesi olarak kullanılan önemli bir ölçüttür ve bu alana yapılan yatırımlar insana yarım olarak görüldüğünden ekonomik büyüme üzerine etkisi uzun döneme yayılarak görülmektedir¹⁷. Türkiye’de yıllara göre sağlık harcamaları incelendiğinde; 1999 yılında 4.985 TL olan harcamanın 2015 yılı itibariyle 104.568 TL’ye yükseldiği görülmektedir. Toplam sağlık harcamasının GSYH içindeki oranının yıllara oranla yatay bir seyir izlediği söylenebilir. %4,8 olan bu oran zaman zaman içerisinde çok fazla bir artış göstermeyerek %5,4’e çıkmıştır.

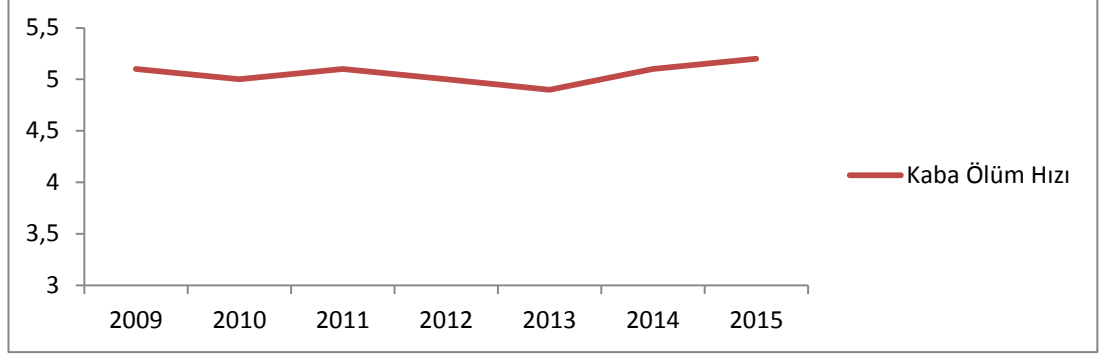
Toplam sağlık harcamalarının yanında bu harcamalar içerisindeki kamu ve özel sektörün payını bilmek; “sosyal devlet anlayışı ile birlikte devletin sahip olduğu sorumluluklarını ne ölçüde yerine getirdiğini göstermede önemli olmanın yanında ileriki dönemlerde kamu içerisindeki sağlık sektörünün yükünü öngörmekte de önemlidir¹⁸. Toplam sağlık harcaması içindeki cari sağlık harcaması payı incelendiğinde 2014 yılında %93,8 olan bu oranın 2015 yılında %92,6’ ya düştüğü; aynı şekilde genel devlet sağlık harcamasının payı incelendiğinde ise 2014 yılında %77,4 olan oranın 2015 yılında %78,5’ e yükseldiği görülmüştür. Bunlara ek olarak

¹⁷ İbrahim Güran Yumuşak, Durmuş Çağrı Yıldırım, “Sağlık Harcamaları İktisadi Büyüme İlişkisi Üzerine Ekonometrik Bir İnceleme”, **The Journal of Knowledge Economy & Knowledge Management**, Volume IV, Spring, 2009, s. 61.

¹⁸ Yusuf Çelik, “Türkiye Sağlık Harcamalarının Analizi ve Sağlık Harcama Düzeyinin Uygunluğunun Değerlendirilmesi”, **Sosyal Güvenlik Dergisi**, Cilt.1, Sayı.1, Haziran, 2011, s. 65.

2014 yılında 1.232TL olan kişi başı sağlık harcamasının 2015 yılında %9.2 oranında artarak 1.345TL ye yükselmiştir¹⁹.

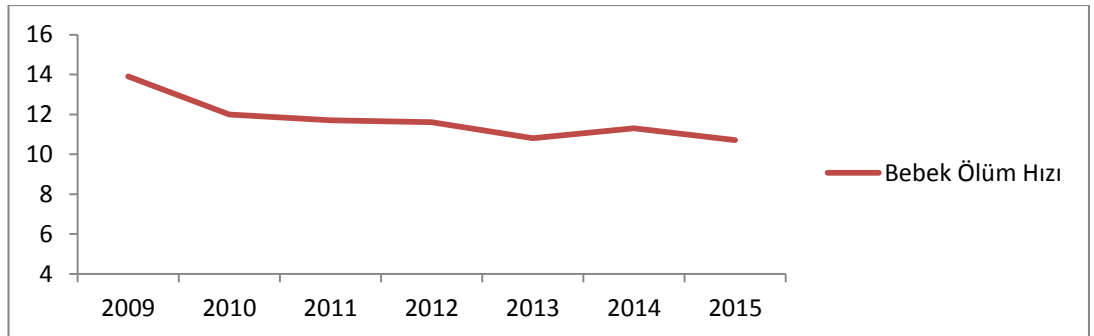
Şekil 4.3: Yıllara Göre Kaba Ölüm Hızı (KÖH) (%)*



*Kaynak: Yıllara Göre İstatistikler/TÜİK

“Bin kişi başına düşen ölüm sayısı” olarak ifade edilen kaba ölüm hızı²⁰ (KÖH), artan nüfus ile birlikte ölüm sayılarında da meydana gelen artmadan dolayı yıllar içerisinde çok fazla değişiklik göstermemiştir. 2009 yılında binde 5.1 olan bu oran 2015 yılında binde 5.2’ ye yükselmiştir.

Şekil 4.4: Yıllara Göre Bebek Ölüm Hızı (BÖH) Oranları (%)*



*Kaynak: Yıllara Göre İstatistikler/TÜİK

¹⁹ Sağlık Harcamaları İstatistikleri, TÜİK haber bülteni, (Çevrimiçi), <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=21527> , 10.05.2017.

²⁰ “Belli bir yıl içinde her bin nüfus başına düşen ölüm sayısı, $KÖH = \frac{\bar{O}}{N} \times 1000$ “, Ölüm İstatistikleri, TÜİK.

Toplumun sađlık düzeyini belirlemede kullanılan bir ölçüt olan bebek ölüm hızı²¹ oranları, ülkenin gelişmişlik ve refah düzeyini de gösterdiğinden önemlidir. Bu nedenle, sađlık politikalarının uygulandığı öncelikli alanlardan bir tanesidir. Türkiye uyguladığı somut politikalar ile Türkiye genelinde bebek ve anne ölüm oranlarını azaltmayı amaçlamıştır²².

Bebek ölüm hızı (BÖH) yıllara göre incelendiğinde, düşme eğiliminde olduğu açıkça görülmektedir. 2009 yılında binde 13.9 olan bu oran 2015 yılında bin canlı doğumda 10.7 bebeğe düşmüştür.

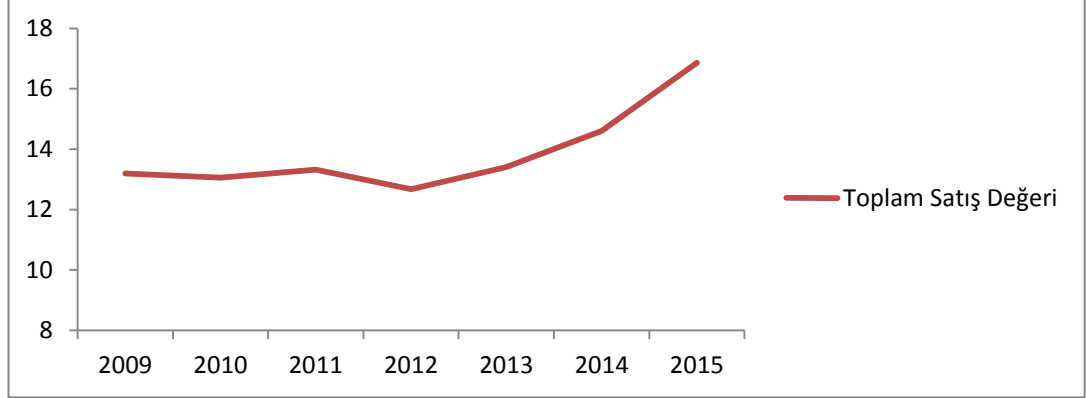
Tüm bu incelediğimiz sađlık göstergelerinin yanında ilaç sektörünün de güçlü olması ile etkin ve kaliteli sađlık hizmetlerine ulaşılabilir. Diğer bir ifadeyle ilaç sektöründe meydana gelen gelişmeler ülkelerin ekonomik kalkınmalarında önemli etkiye sahiptir. Bunun yanında salgın hastalıklar, savaş gibi durumlarda ülkelerin kendi ilaç ihtiyacını karşılayabilecek düzeyde olması önemlidir. Türkiye tıbbi ürün pazarı²³ gelişimi incelendiğinde; %15,5 büyüme oranı ile 2015 yılında 16,86 milyar TL'ye ulaştığı görülmüştür.

²¹ “Belli bir yıl içinde her bin canlı doğum başına düşen bebek ölüm sayısı, $BÖH = \frac{BÖ}{D} \times 1000$ ”, Ölüm İstatistikleri, TÜİK.

²² Esra Özpınar, “Bebek Ölüm Hızı Türkiye Genelinde Azalırken Neden Suriyeli Nüfusun Fazla Olduğu Yerlerde Artıyor?”, **Türkiye Ekonomi Politikaları Araştırma Vakfı**, No: 201619, Mayıs, 2016, s. 1.

²³ İEİS-İlaç Endüstrisi İşverenler Sendikası, Türkiye İlaç Pazarı, (Çevrimiçi), <http://www.ieis.org.tr/ieis/tr/indicators/33/turkiye-ilac-pazari> , 06.04.2017.

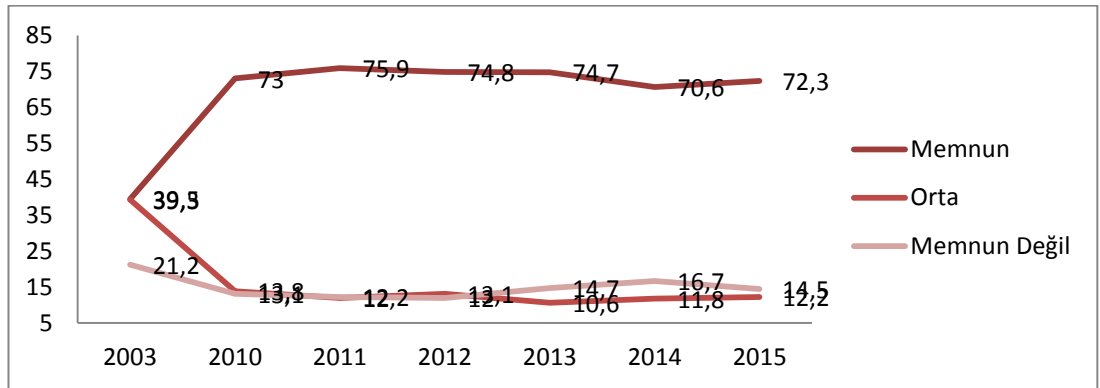
Şekil 4.5: Yıllara Göre Tıbbi Ürün Pazarı* (Hastane Satışları Hariç, Üretici Fiyatlarıyla)**



* Sağlık Bakanlığı'ndan ruhsatlı ilaçları ve ilaç dışı tıbbi ürünleri kapsar.

***Kaynak:** İEİS-İlaç Endüstrisi İşverenler Sendikası, Türkiye İlaç Pazarı.

Şekil 4.6: Yıllara Göre Sağlık Hizmetlerinden Genel Memnuniyet Oranı* (%)



***Kaynak:** Sağlık Hizmetlerinde Memnuniyet Oranı, (Çevrimiçi), <http://www.tkhk.gov.tr/Dosyalar/bd6db74478224854955cf55e208c3c28.pdf>, 11.05.2017.

Son olarak sağlık hizmetlerinden genel memnuniyet oranı yıllara göre incelendiğinde; 2003 yılında %39.5 olan memnun oranının 2015 yılına gelindiğinde %72.3'e çıktığı söylenebilir. Yine aynı şekilde %21.2 memnun değil oranının %14.5'e düştüğü görülmüştür.

4.3. İstatistiki Bölge Birimleri Sınıflandırması

Belli bir bölgenin toplumsal yapısının, iktisadi ve kültürel boyutlarda olumlu değişimi bölgesel kalkınma olarak tanımlanabilir. Ülke kalkınmasına göre daha karmaşık bir yapı içeren bölgesel kalkınma, yerel kalkınma ile bir paralellik göstermektedir. Diğer bir ifadeyle ulusal kalkınmanın temeli bölgesel kalkınmaya dayanmaktadır. Ancak ekonomik kalkınma gerçekleşmeden bölgesel ve yerel kalkınmadan söz etmemiz mümkün değildir²⁴.

Disiplinler arası bir ifade olan bölge kavramı coğrafi bir terim olarak, fiziki ve beşeri açıdan benzerlik gösteren yer anlamındadır. Türkiye 1941 tarihinde toplanan I. Coğrafya kongresinde 7 bölgeye ayrılmıştır²⁵.

Avrupa Birliği'ne adaylık ile birlikte Türkiye yeni bölge kavramları ile tanışmıştır. AB'ye tam üyelik ölçütü olarak, 1970'lerde Avrupa Topluluğu İstatistik Ofisi (EUROSTAT) tarafından geliştirilen²⁶ İstatistiki Bölge birimleri sınıflandırmasının (İBBS) (The Nomenclature of Territorial Units for Statistics (NUTS)) kullanılması; bölgelerin bu sınıflandırma ile belirlenmesi ve yerel, bölgesel ve ülke genelinde planlamaların bu ölçülere göre yapılması kararlaştırılmıştır.

1999 yılında Türkiye'nin AB'ye adaylığının kabul edilmesiyle, 2001 yılında Devlet Planlama Teşkilatı, İç İşleri Bakanlığı ve Devlet İstatistik Enstitüsünün görevlendirilmesiyle istatistiki bölgelerin belirlenmesine başlanmıştır. Tam üyelikle birlikte oluşturulan bu bölgeler değişime uğrayabilmektedir.

²⁴ Ümran Şengül; Seyedhadi Eslemian; Miraç Eren: "Türkiye'de İstatistiki Bölge Birimleri Sınıflamasına Göre Düzey-2 Bölgelerinin Ekonomik Etkinliklerinin VZA Yöntemi ile Belirlenmesi ve Tobit Model Uygulaması", **Yönetim Bilimleri Dergisi**, Cilt.11, Sayı.21, 2013, s. 76.

²⁵ Barış Taş: "AB Uyum Sürecinde Türkiye İçin Yeni Bir Bölge Kavramı: İstatistiki Bölge Birimleri Sınıflandırması (İBBS)", **Sosyal Bilimler Dergisi**, s. 187; Ümran Şengül v.d., **a.g.e.**, s. 76.

²⁶ Ahmet Öztürk: "Homojen ve Fonksiyonel Bölgelerin Tespiti ve Türkiye İçin İstatistiki Bölge Birimleri Önerisi", **DPT Uzmanlık Tezleri**, Yayın No: DPT: 2802, Ankara, 2009.

Tablo 4.5: Türkiye’de İBBS’ye Göre Sınıflandırılmış Bölgeler

Kod	Düzy-1 (12 Bölge)	Kod	Düzy 2 (26 Alt Bölge)	Düzy 3 (81 il)
TRA	Kuzeydoğu Anadolu	TRA1	Erzurum Alt Bölgesi	Erzurum, Erzincan, Bayburt
		TRA2	Ağrı Alt Bölgesi	Ağrı, Kars, Iğdır, Ardahan
TRB	Ortadoğu Anadolu	TRB1	Malatya Alt Bölgesi	Malatya, Elazığ, Bingöl, Tunceli
		TRB2	Van Alt Bölgesi	Van, Muş, Bitlis, Hakkari
TRC	Güneydoğu Anadolu	TRC1	Gaziantep Alt Bölgesi	Gaziantep, Adıyaman, Kilis
		TRC2	Şanlıurfa Alt Bölgesi	Şanlıurfa, Diyarbakır
		TRC3	Mardin Alt Bölgesi	Mardin, Batman, Şırnak, Siirt
TR1	İstanbul	TR10	İstanbul Alt Bölgesi	İstanbul
TR2	Batı Marmara	TR21	Tekirdağ Alt Bölgesi	Tekirdağ, Edirne, Kırklareli
		TR22	Balıkesir Alt Bölgesi	Balıkesir, Çanakkale
TR3	Ege	TR31	İzmir Alt Bölgesi	İzmir
		TR32	Aydın Alt Bölgesi	Aydın, Denizli, Muğla
		TR33	Manisa Alt Bölgesi	Manisa, Afyonkarahisar, Kütahya, Uşak
TR4	Doğu Marmara	TR41	Bursa Alt Bölgesi	Bursa, Eskişehir, Bilecik
		TR42	Kocaeli Alt Bölgesi	Kocaeli, Sakarya, Düzce, Bolu, Yalova
TR5	Batı Anadolu	TR51	Ankara Alt Bölgesi	Ankara
		TR52	Konya Alt Bölgesi	Konya, Karaman
TR6	Akdeniz	TR61	Antalya Alt Bölgesi	Antalya, Isparta, Burdur
		TR62	Adana Alt Bölgesi	Adana, Mersin
		TR63	Hatay Alt Bölgesi	Hatay, Kahramanmaraş, Osmaniye
TR7	Orta Anadolu	TR71	Kırıkkale Alt Bölgesi	Kırıkkale, Aksaray, Niğde, Nevşehir, Kırşehir
		TR72	Kayseri Alt Bölgesi	Kayseri, Sivas, Yozgat
TR8	Batı Karadeniz	TR81	Zonguldak Alt Bölgesi	Zonguldak, Karabük, Bartın
		TR82	Kastamonu Alt Bölgesi	Kastamonu, Çankırı, Sinop
		TR83	Samsun Alt Bölgesi	Samsun, Tokat, Çorum, Amasya
TR9	Doğu Karadeniz	TR90	Trabzon Alt Bölgesi	Trabzon, Ordu, Giresun, Rize, Artvin, Gümüşhane

İstatistiki bölge birimleri sınıflandırması AB ülkelerini ekonomik bölgelere bölmek için hiyerarşik bir sistemdir. Bunun anlamı, üye ülkeler ilk olarak düzey-1 bölgelerine, sonrasında düzey-1 bölgeleri düzey-2 bölgelerine ve son olarak da düzey-2 bölgeleri düzey-3 bölgelerine ayrılacaktır²⁷. Burada, AB üye ülkelerinde bölgesel istatistiklerin toplanması, geliştirilmesi ve uyumlaştırılması, AB bölgesel politika çerçevesinin belirlenmesi ve bölgelerin sosyo-ekonomik analizlerinde temel olması gibi amaçlardan söz edilebilir. Temel istatistiki göstergeler, nüfus, illerin gelişmişlik düzeyleri gibi kriterler göz önüne alındığında Türkiye için Düzey-1 12 tane, Düzey-2 26 tane ve Düzey-3 81 tane NUTS bölgesi oluşturulmuştur.²⁸

Türkiye’de İBBS bölgelerinin oluşturulmasının temel nedeni bölgelerin mevcut kapasitesinin optimum şekilde kullanılması belirleyen kalkınma ajanslarının kurulması mecburiyetidir. AB üye ülkelerinde kalkınma ajanslarının oluşturulması Düzey-2 bölgeleri baz alınarak gerçekleştirilmektedir. Diğer bir ifadeyle, AB üye ülkelerine ya da birliğin bütününe yönelik kalkınma ve gelişme planlarının hazırlanması Düzey-2 bölgeleri temel alınarak yapılmaktadır. Bu da Düzey-2 26 bölgenin İBBS içerisindeki konumunu farklılaştırmaktadır.

4.4. İBBS Düzey-2 Bölgelerinin Sağlık Göstergeleri Açısından İncelenmesi

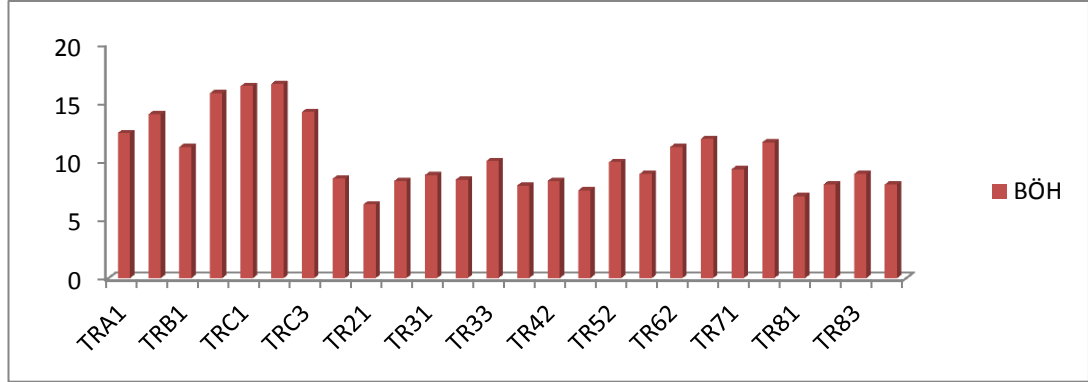
Ülkelerin kalkınma düzeyleri ile sağlık göstergeleri arasında pozitif bir ilişki vardır ve bunun yanında ülkelerin kalkınma düzeyleri de temelde bölgesel kalkınmaya bağlı olmaktadır²⁹. Bu nedenle İBBS düzey-2 bölgelerinin analizde kullanılacak sağlık göstergeleri açısından incelenmesi önemlidir.

²⁷ <http://ec.europa.eu/eurostat/web/nuts/overview>

²⁸ Taş, **a.g.e.**, s. 191.

²⁹ Yumuşak, Yıldırım, **a.g.e.**

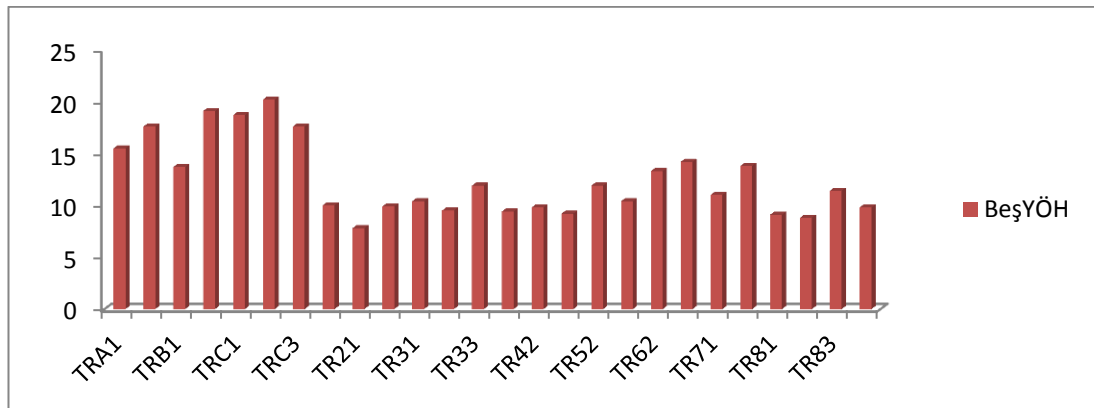
Şekil 4.7: İBBS Düzey-2 Bölgelerinde Bebek Ölüm Hızı Oranları (2015) *



*Kaynak: Bölgesel İstatistikler/ TÜİK

Yukarıdaki grafikte görüldüğü gibi; bebek ölüm hızı oranları bakımından TRB2 (15,8), TRC1 (16,4) ve TRC2 (16,6) düzey-2 bölgelerinin Türkiye ortalaması olan 7,6'dan oldukça yüksek olduğu görülmektedir. Binde 6,3 oranıyla TR21 Tekirdağ bölgesi en düşük bebek ölüm hızına sahiptir.

Şekil 4.8: İBBS Düzey-2 Bölgelerinde Beş Yaş Altı Bebek Ölüm Hızı Oranları (2015) *



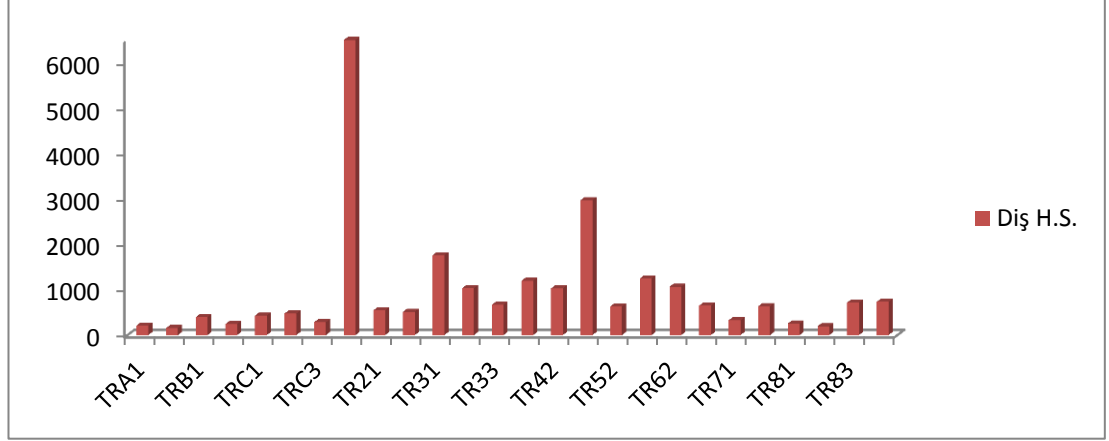
*Kaynak: Bölgesel İstatistikler/ TÜİK

Aynı şekilde bölgeler bazında beş yaş altı çocuk ölüm hızı³⁰ oranlarını incelediğimizde; TRC2 (20.2), TRB2 (19.1) düzey-2 bölgeleri başta olmak üzere

³⁰ Belli bir yıl içinde bin canlı doğum başına düşen beş yaş altı ölüm sayısı, $BYAÖH = \frac{\ddot{O}_{0-4}}{D} \times 1000$.

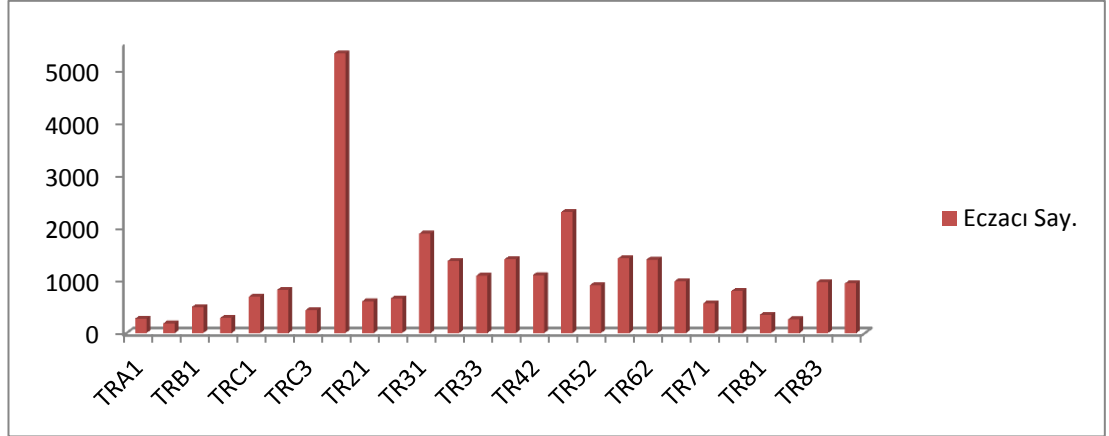
doğu bölgelerinde bu oranın 12.8 olan Türkiye ortalamasının oldukça üzerinde olduğu görülmektedir. Burada da binde 7.8 beş yaş altı ölüm oranıyla en düşük orana sahip bölge Tekirdağ bölgesidir.

Şekil 4.9: İBSS Düzey-2 Bölgelerinde Diş Hekimi Sayısı (2015) *



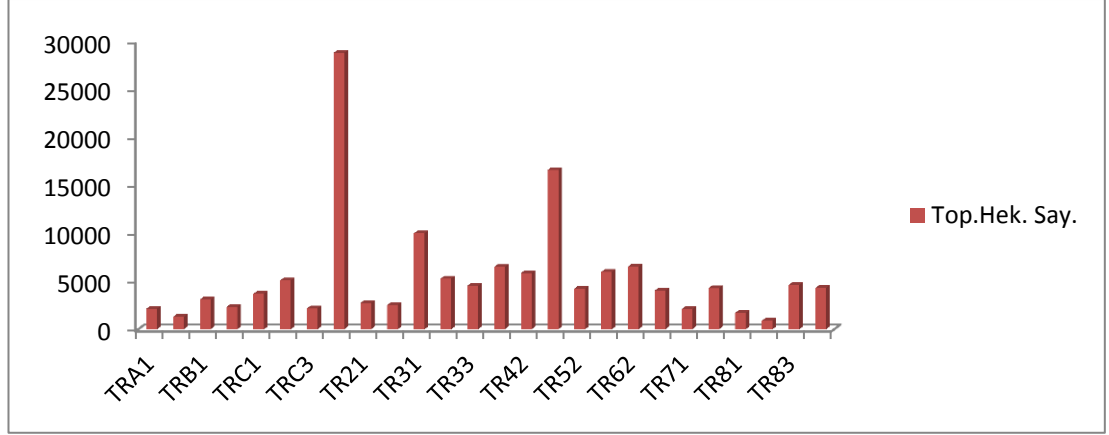
*Kaynak: Bölgesel İstatistikler/ TÜİK

Şekil 4.10: İBBS Düzey-2 Bölgelerinde Eczacı Sayısı (2015) *



*Kaynak: Bölgesel İstatistikler/TÜİK

Şekil 4.11: İBBS Düzey-2 Bölgelerinde Toplam Hekim Sayısı (2015)*

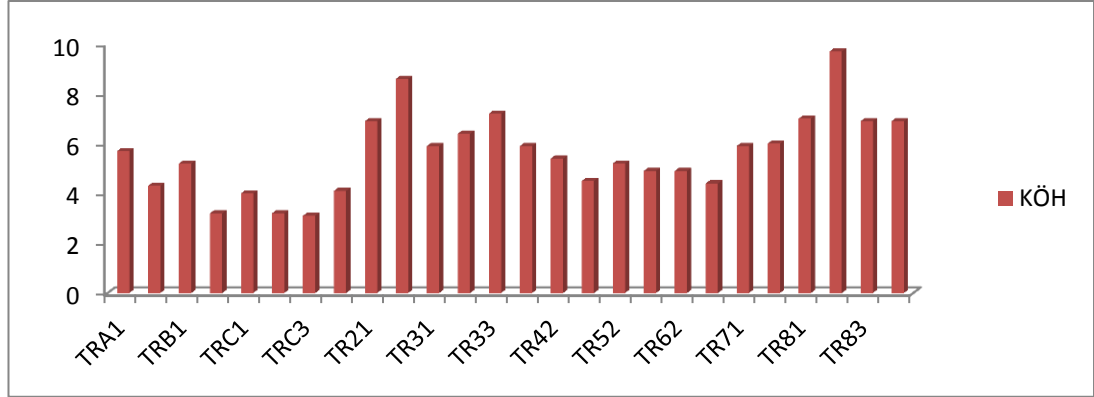


*Kaynak: Bölgesel İstatistikler/ TÜİK

Bölgelerin sağlık düzeyleri karşılaştırıldığında kullanılan bir diğer ölçüt sağlık personeli sayıları olmaktadır. Araştırmada yalnızca diş hekimi, eczacı ve hekim sayıları dikkate alınmıştır. Bu üç gösterge için bölgeler arası artış ve azalışların birbirlerine paralel bir şekilde gerçekleştiği görülmüştür. En uç noktanın İstanbul bölgesi olduğu bunu, nüfusa bağlı olarak, Ankara ve İzmir bölgelerinin izlediği söylenebilir.

Ölüm hızı toplumların gelişmişlik ve refah düzeyine, yaşa, cinsiyete, medeni duruma ve sosyoekonomik düzeylere bağlı olarak farklılaşmaktadır. Kaba ölüm hızı oranları incelendiğinde; TR82 (9.7) düzey-2 bölgesinin en yüksek orana sahip olduğu ve TRB2 (3.20), TRC2 (3.20), TRC3 (3.10) bölgelerinin ise en düşük KÖH oranlarına sahip olduğu görülmektedir.

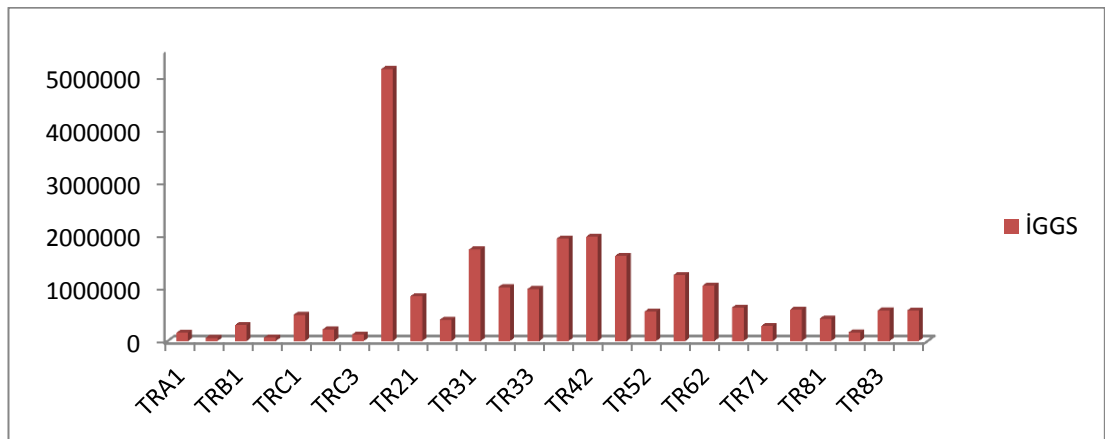
Şekil 4.12: İBBS Düzey-2 Bölgelerinde Kaba Ölüm Hızı Oranları (2015)*



*Kaynak: Bölgesel İstatistikler/ TÜİK

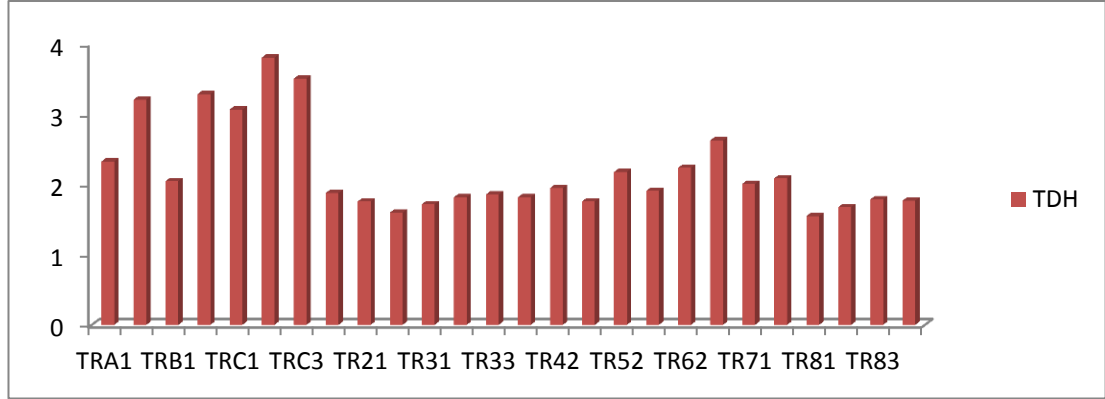
Bir işyerinde sigortalı olarak çalışan bir kişinin meslek hastalığından veya geçirdiği bir iş kazasından dolayı rapor alması, herhangi bir hastalıktan dolayı rapor alması ve sigortalı kadın çalışanların doğum iznine ayrılması gibi nedenlerle oluşan iş göremez gün sayısı bölgelere göre incelendiğinde en yüksek gün sayısının TR10 ile İstanbul bölgesinin olduğu görülmüştür. Bir sonraki en yüksek gün sayısı TR41 ve TR42 bölgelerine aittir. Burada, nüfus ve sanayileşme ile iş göremez gün sayısı arasında yakın bir ilişkinin olduğu söylemek yanlış olmayacaktır.

Şekil 4.13: İBBS Düzey-2 Bölgelerinde SGK İş Göremez Gün Sayısı (2015)*



*Kaynak: Sosyal Güvenlik Kurumu (SGK) Veri Tabanı

Şekil 4.14: İBBS Düzey-2 Bölgelerinde Toplam Doğurganlık Hızı (TDH) (2015)*



*Kaynak: Bölgesel İstatistikler/TÜİK

Şekil 4.14’de toplam doğurganlık hızı (TDH) İBBS Düzey-2 bölgelerine göre incelenmiştir. Bu oran 3.81 çocuk ile TRC2 bölgesi başta olmak üzere TRC3, TRB2, TRC1 bölgelerinde Türkiye geneli olan 2.14 çocuktan yüksek olduğu görülmüştür. TR81 bölgesi ise 1.55 çocuk ile en düşük TDH’na sahip bölge konumundadır.

4.5. Literatür İncelenmesi

Sağlık düzeylerinin karşılaştırılması ve tespit edilmesi gibi birçok araştırmada çok boyutlu ölçekleme ve kümeleme analizi gibi çok değişkenli istatistiksel yöntemlerden yararlanılmıştır. Bu çalışmalardan çoğu belli sayıda sağlık göstergesi açısından ülkeler arasındaki benzerlik ve farklılıkları ortaya çıkarma temeline dayanmaktadır. Son yıllarda yapılan birkaç çalışma aşağıda kısaca özetlenmiştir.

Sığırlı ve ark. (2006), AB’ye üye ülkeler ve üyeliğe aday ülkeler -Bulgaristan, Romanya, Türkiye- olmak üzere toplam 25 ülkeyi, sağlık göstergeleri açısından ÇBÖ yöntemi kullanarak karşılaştırmıştır. Çalışmalarında, ülkemizin AB üye ülkeler içerisindeki konumu, aralarındaki benzerlik ve farklılıkların ortaya konulması amaçlanmıştır. Doğuşta beklenen sağlıklı yaşam ümidi, kişi başına düşen yıllık TSH, TSH içinde genel yönetim sağlık harcaması payı gibi 9 değişken için ülkeler incelendiğinde; Türkiye’nin birinci boyutta özellikle 5 yaş altı çocuk ölüm hızı ve 15-59 yaş arası yetişkin ölüm hızı açısından AB ülkelerinden farklılaştığı ve bu

lkelerin olduka gerisinde olduėu sonucuna ulařılmıştır. İkinci boyutta lkelerin 3 farklı grup oluşturduėu, Trkiye'nin Slovakya, Macaristan ve ek Cumhuriyeti ile aynı grupta yer aldıėı gözlenmiştir.

Ersz (2009); Trkiye ile OECD lkelerini saėlık dzeyleri ve saėlık harcamaları aısından incelemiřtir. TSH ierisinde ila harcama oranı, 1 milyon kiřiye dřen MRI cihazı sayısı ve 15 yař zeri alkol tketim oranı gibi 14 saėlık gstergesi 30 lke iin elde edilmiř ve bu gstergeler aısından, B analizi kullanılarak, lkeler arasındaki benzerlikler ve farklılıkları tespit etmeye alışılmıştır. Orta uyuma sahip bir stress deėeri ile Trkiye'nin, Meksika gibi geliřmekte olan lkeler ile bir grup oluşturduėu ve kiři bařına dřen doktor sayısı ve kiři bařına saėlık harcaması gibi gstergeler aısından diėer OECD lkelerine oranla olduka dřk seviyede olduėu grlmřtir.

elik (2013), kmeleme analizini kullandıėı alışmasında, 81 ile ait saėlık gstergeleri aısından bu illerin birbirleriyle benzer yapıya sahip olanlarının sınıflandırılmasını ve en dřk seviyedeki illerin tespit edilmesi amalanmıştır. Deėiřkenler, farklı olarak, nfusa gre oranlama yapılarak kullanılmıştır. Tek Baėlantı, Ward ve k-ortalamlar kmeleme yntemleri uygulanan alışmada illerin 7 kmede sınıflandıėı ve Kilis'in diėer tm illerden ayrıřarak tek bařına bir kme oluşturduėu sonucuna ulařılmıştır. 2. basamakta birbirlerine en benzer iller Nevřehir ve Karaman olurken, 80. Basamakta en az benzeyen iller Adana ve Kilis olmuřtur.

Girginer (2013) alışmasında, 2010 Dnya Saėlık Raporundan elde ettiėi saėlık gstergeleri aısından Trkiye'nin AB lkelerine gre konumunu incelemiřtir. lkeler arasındaki benzerlik ve farklılıkları elde etmek amacıyla ok Boyutlu lekleme ve Hiyerarřik Olmayan Kmeleme Analizinden yararlanmıştır.

Boz ve ark. (2016), seilen saėlık gstergeleri bakımından OECD lkeleri arasındaki farklılıkları ve benzerlikleri grmek, buna neden olan deėiřkenleri tespit etmeyi amalamıştır. 34 OECD lkesinin dahil olduėu alışmada, ilk nce saėlık gstergelerinin tanımlayıcı istatistikleri incelenmiř ve sonrasında B ynteminden yararlanılmıştır. alışmanın sonucunda; ,Trkiye, Meksika, řili ve Kore Cumhuriyetinin birbirleriyle benzer olduėu grlmřtir. Ayrıca Trkiye, saėlık

harcamalarındaki büyüme oranı ve neonatal bebek ölüm hızı göstergeleri bakımından Japonya, İsviçre, Avusturalya gibi OECD ülkeleri ile en benzemezdir.

4.6. Sağlık Göstergelerinin İstatistiksel Analizi

Türkiye'nin her bölgesinde, gelişmişlik ve toplumsal yapısına uygun olarak, sağlık kuruluşu ve sağlık personel sayıları aynı değildir. Bu nedenle, bölgelerin sağlık göstergeleri bu değişkenlere göre farklılık göstermektedir. Hekim, eczacı ve sağlık personeli sayıları birbirine yakın olan bölgelerde ise doğal olarak birbirlerine benzemeleri beklenmelidir³¹.

Ülkelerin kalkınmışlık seviyeleri öncelikle bölgelerin kalkınmasına bağlı olduğundan, bölgelerin kalkınması ve bölgeler arasındaki farkların ortaya konulması önemlidir. Bu nedenle çalışmamızda, İBBS Düzey-2'ye göre ayrılan 26 bölgenin sağlık göstergeleri açısından birbirlerine göre konumlarının çok boyutlu ölçekleme ve kümeleme analizleri yardımıyla elde edilmesi, varsa benzerlik ve farklarının ortaya konulması amaçlanmıştır.

Tablo 4.6: Düzey-2 Bölgeleri İçin Analizde Kullanılacak Değişkenler

Değişkenler	
ES: Eczacı Sayısı	KÖH: Kaba Ölüm Hızı
THS: Toplam Hekim Sayısı	SGK: SGK İş Göremez Gün Sayısı
BÖH: Bebek Ölüm Hızı	TDH: Toplam Doğurganlık Hızı
DHS: Diş Hekimi Sayısı	TYS: Toplam Yatak Sayısı
BeşBÖH: Beş Yaş Altı Bebek Ölüm Hızı	TSH: Hanehalkı Bütçe Anketi Tüketim Harcaması içerisindeki sağlık harcaması payı

Bu çalışmada verilerin elde edilme düzeylerine bağlı olarak 2015 yılına ait eczacı sayısı, toplam hekim sayısı, bebek ölüm hızı, diş hekimisi sayısı, kaba ölüm hızı, beş yaş altı çocuk ölüm hızı, iş göremez gün sayısı, toplam doğurganlık sayısı, toplam yatak sayısı ve toplam sağlık harcaması gibi sağlık göstergelerinden yararlanılmıştır. Söz konusu değişkenler Tablo 4.10'da listelenmiştir. Düzey-2 Bölgelerine ait bu veriler TÜİK (Türkiye İstatistik Kurumu) Biruni veri tabanı

³¹ Çelik, a.g.e., s. 176.

bölgesel istatistikler alt başlığındaki çeşitli sayfalardan ve SGK (Sosyal Güvenlik Kurumu) istatistiksel veri tabanından alınmıştır.

4.6.1. Çok Boyutlu Ölçekleme Sonuçları

Çalışmanın birinci kısmında Tablo 4.10'de yer alan değişkenler ile birlikte 26 bölgeye çok boyutlu ölçekleme analizi uygulanmıştır. Burada birimler veya nesneler metrik olarak ölçüldüğünden ve analizde herhangi bir ağırlığa sahip olmadığından, ALSCAL algoritmasının kullanılması tercih edilmiştir. ALSCAL'da klasik ÇBÖ algoritması minimizasyon algoritmasının başlama noktası seçeneği olarak yer almaktadır³². Değişkenler oran ölçeği düzeyinde ölçüldüğünden, çözümleme ALSCAL algoritmasında oran ölçekli ÇBÖ modeli kullanılarak yapılmıştır. Analize başlamadan önce veriler arasındaki uzaklıklar için öklit uzaklığı kullanılmış ve z standart değişken dönüşümü uygulanmıştır.

Tablo 4.7: Üç Boyut İçin Elde Edilen İterasyonlar ve Kruskal Stress Değeri

Young's S-stress Formülü		
İterasyon	S-stress	Gelişim
1	0.02775	
2	0.02165	0.00609
3	0.02140	0.00025

For Matrix	
Stress= 0.03650	RSQ=0.99678

ÇBÖ analizinin etkinliği Kruskal stress istatistiğine uygun olarak ölçülür. Kruskal stress istatistiği, konfigürasyon ölçüleri ile tahmini konfigürasyon ölçüleri arasındaki oranın karekökü alınarak hesaplanır ve bu değer konfigürasyon uzaklıkları ve veri uzaklıkları arasındaki uygunluğu gösterir. Genellikle istenilen sıfır değerine yakın olan stress değerlerinin elde edildiği boyut çözümlerine ulaşmaktır. ALSCAL algoritmasının sonuçları Tablo 4.7'de gösterilmiştir. k=3 boyut için stress değerinin gelişimi 0.001'den küçük olduğu değere kadar iterasyon devam ettirilmiştir. Üçüncü iterasyonda 0.00025 değerine ulaşılmış ve iterasyon bu noktada durdurulmuştur. Düşük bir Kruskal stress istatistiği değeri (0.03650) ile 3-boyutlu uzayda oldukça iyi

³² Akkücüç, a.g.e., s. 313.

uyumlu bir çözüm elde edilmiştir. Analizde yer alan verilerin ÇBÖ modeli ile açıklanabilme yüzdesini veren RSQ değeri incelendiğinde, % 99.6 açıklayıcılık yüzdesiyle analizin güvenilir olduğu sonucuna ulaşılmıştır.

26 bölgenin sağlık göstergelerine göre belirlenen üç boyutlu çözüm koordinatları Tablo 4.8’de gösterilmiştir.

Tablo 4.8: ÇBÖ Sonuçlarına Göre Üç Boyutlu Uzay İçin Uyarıcı Koordinatlar

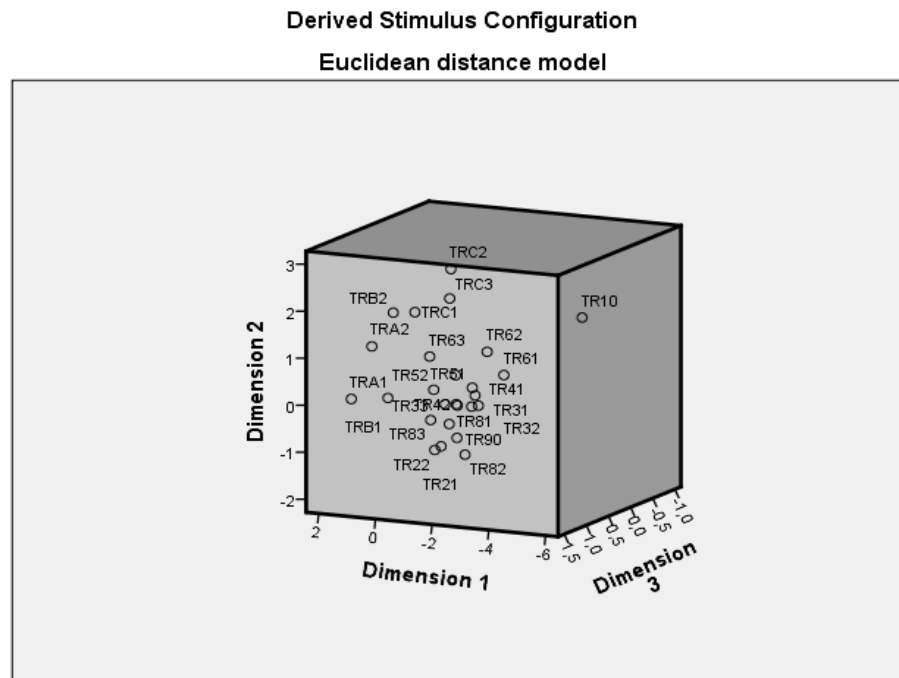
Düzyey-2/Bölge	Boyut-1	Boyut-2	Boyut-3
TRA1/ Erzurum Alt Bölgesi	1.2571	0.1026	1.3792
TRA2/ Ağrı Alt Bölgesi	1.6478	0.9254	0.6642
TRB1/ Malatya Alt Bölgesi	0.7440	-0.0370	0.8793
TRB2/ Van Alt Bölgesi	1.5536	1.4847	0.2394
TRC1/ Gaziantep Alt Bölgesi	1.1490	1.4310	0.0094
TRC2/ Şanlıurfa Alt Bölgesi	1.1302	2.0451	-0.7916
TRC3/ Mardin Alt Bölgesi	1.2523	1.3914	-0.8430
TR10/ İstanbul Alt Bölgesi	-5.0522	1.7520	0.2037
TR21/ Tekirdağ Alt Bölgesi	-0.0253	-1.2916	0.1699
TR22/ Balıkesir Alt Bölgesi	0.2276	-1.3908	0.1501
TR31/ İzmir Alt Bölgesi	-1.0860	-0.1824	0.0819
TR32/ Aydın Alt Bölgesi	-0.4645	-0.6178	-0.3916
TR33/ Manisa Alt Bölgesi	-0.1098	-0.5388	-0.1444
TR41/ Bursa Alt Bölgesi	-0.8361	-0.4645	0.0005
TR42/ Kocaeli Alt Bölgesi	-0.5912	-0.3546	0.1910
TR51/ Ankara Alt Bölgesi	-1.9232	0.2601	0.6819
TR52/ Konya Alt Bölgesi	0.2475	-0.1077	0.1622
TR61/ Antalya Alt Bölgesi	-0.6114	-0.1407	-0.8724
TR62/ Adana Alt Bölgesi	-0.3039	0.4049	-0.6924
TR63/ Hatay Alt Bölgesi	0.4903	0.5606	0.0987
TR71/ Kırıkkale Alt Bölgesi	0.4799	0.5839	-0.2425
TR72/ Kayseri Alt Bölgesi	0.2860	0.0009	-0.3585
TR81/ Zonguldak Alt Bölgesi	0.2300	-1.3266	-0.3525
TR82/ Kastamonu Alt Bölgesi	0.4149	-1.8090	-0.6514
TR83/ Samsun Alt Bölgesi	0.0165	-0.6556	0.3794
TR90/ Trabzon Alt Bölgesi	-0.1232	-0.8577	0.0495

Uyarıcı koordinatlar tablosuna göre, birinci boyutta TRA1, TRA2, TRB2, TRC1, TRC2 ve TRC3 bölgelerinin pozitif ve 1'in üzerinde değerler ile en yüksek değerlere sahip olması, bu bölgelerin incelenen sağlık göstergeleri açısından birbirleriyle benzer konumda olduklarını gösterir. Bu bölgeler diğer bölgelere oranla yüksek değerler aldıklarından birinci boyutta en önemli ayrıştırıcılardır. TR10, TR21, TR31, TR32, TR33, TR41, TR42, TR51, TR61 ve TR90 bölgelerinin ise negatif yüke sahip olduğu görülmüştür. Birinci boyutta en farklı bölgenin -5.0522 negatif değeriyle TR10 bölgesinin olduğu, dolayısıyla TR10'un tek başına diğerlerinden farklı bir konumda bulunduğu söylenebilir. İkinci boyutta; TRB2, TRC1, TRC3, TR10 bölgelerinin 1'in üzerinde pozitif değerlere sahip oldukları görülmüştür. 2'nin üzerinde bir değere sahip olan TRC2 bölgesi ikinci boyutta en önemli ayrıştırıcıdır. Aynı şekilde üçüncü boyut için de TRA1 bölgesinin 1'in üzerinde pozitif bir değerle en önemli ayrıştırıcı olduğu, diğer bölgelerin sıfır etrafında değerler aldığı görülmüştür.

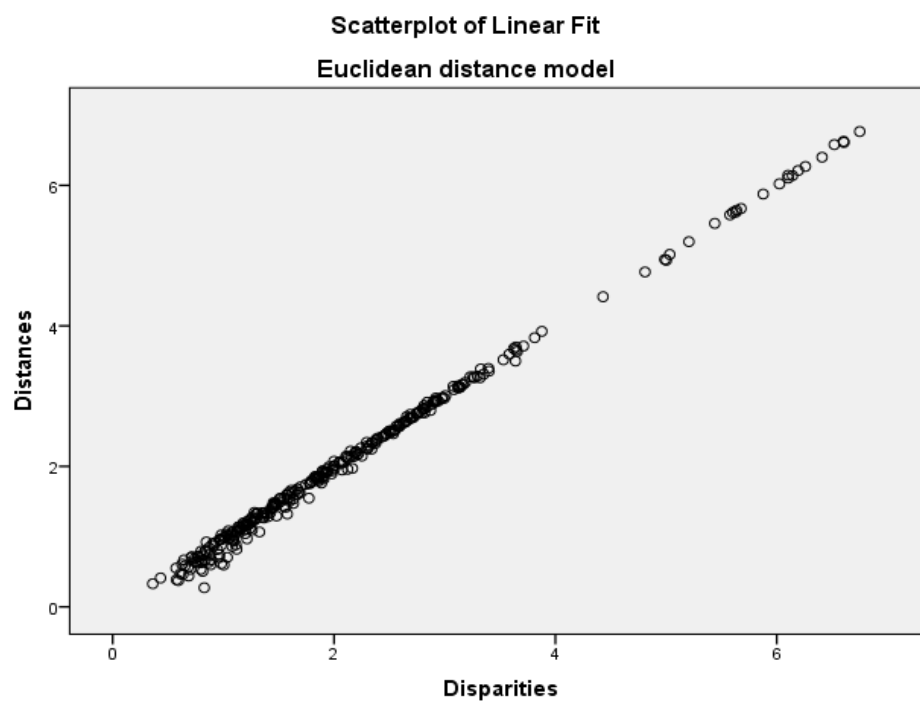
ÇBÖ analizinde benzer olan bölgelerin birbirlerine yakın konumlanması beklenmektedir. Öklid uzaklık modeli incelendiğinde, 26 bölgenin analizde yer alan sağlık göstergeleri açısından üç boyutlu uzayda üç farklı grup oluşturdukları görülmüştür. TR10 İstanbul bölgesi diğer tüm bölgelerden en uzakta ve en farklı olarak tek başına bir grup oluşturmuştur. TRA2, TRB2, TRC1, TRC2 ve TRC3 bölgeleri bir diğer grubu oluştururken; TR31, TR51, TR21 gibi büyük şehirleri içeren diğer bölgeler üçüncü grubu oluşturmaktadır.

Analizde yer alan verilere uygulanan monotonik dönüşümle elde edilen farklar ile türetilmiş uzaklıklar arasındaki ilişkiyi gösteren Shepard diyagramı Şekil.16'deki gibi elde edilmiştir. Üç boyutlu gösterim için elde edilen bu diyagramda farklılıklar ve uzaklıklar arasında doğrusal bir ilişkinin olduğu görülmektedir. Bu istenilen bir durumdur ve doğrusal bir fonksiyon elde edilmesi sonucun uygun olduğunun göstergesi olarak önemlidir.

Şekil 4.15: Üç Boyutlu Uzay İçin Türetilmiş Uyarıcılar Gösterimi

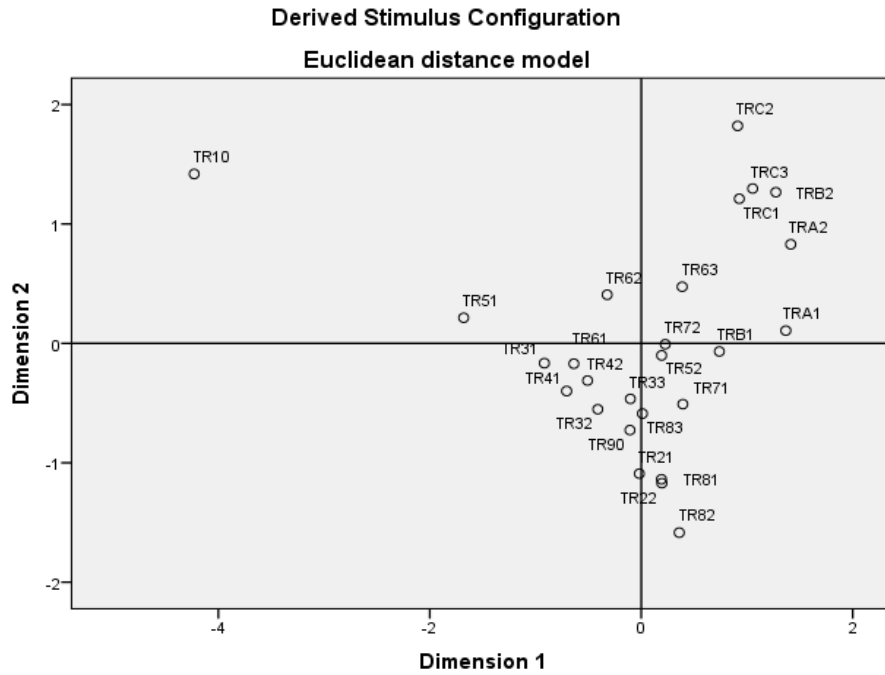


Şekil 4.16: Üç Boyutlu Uzay İçin Serpilme Diyagramı



k=3 boyut için ÇBÖ analiz sonuçlarının daha iyi çıkmasına rağmen iki boyutlu uzayda görsel açıdan bölgelerin birbirlerine göre konumları ve oluşan gruplar daha belirgindir. Bu nedenle iki boyutlu uzay için elde edilen grafiksel gösterimine de uygulamada yer verilmiştir. İki boyutlu gösterime ait stress değeri incelendiğinde; 0.10234 ile yaklaşık olarak iyi uyuma sahip olduğu sonucuna ulaşılmıştır. Ayrıca, R=0.97947 olarak elde edilmiştir.

Şekil 4.17: İki Boyutlu Uzay İçin Türetilmiş Uyarıcılar Gösterimi



İki boyutlu uzay için elde edilen ÇBÖ analiz sonuçları incelendiğinde, üç boyutta elde edilen ÇBÖ analiz sonuçları ile benzer sonuçlara ulaşıldığı görülmüştür. Birbirlerine benzer olan bölgelerin bir arada gruplandığı Şekil 4.17’de, TRC2, TRC3, TRC1, TRB2 ve TRA2 bölgelerinin diğer bölgelerden farklılaştığı, İstanbul bölgesinin ise diğer tüm bölgelerden farklılaşarak ayrı bir konumda bulunduğu gözlenmiştir. Ayrıca üçüncü grupta yer alan TR51 bölgesinin gruptan farklı eğilimler sergilediği söylenebilir

Ayrıca çalışmada bölgeler açısından sağlık göstergelerinin nasıl gruplandığını görmek amacıyla sağlık göstergelerine ÇBÖ uygulanmıştır. İki ve üç boyutlu uzay

için analiz sonuçları elde edilmiş, üç boyutlu uzay için elde edilen sonuçlar daha güvenilir bulunduğundan yorumlama üç boyut için yapılmıştır.

k=3 boyut için stress değeri 0.06478 olarak elde edilmiş ve tahmini uzaklık değerleri ile konfigürasyon değerleri arasında orta derecede bir uyum olduğu görülmüştür. Ayrıca RSQ değeri (0.99639) ile verilerin ÇBÖ modeli ile güvenilir biçimde açıklanabildiği sonucuna ulaşılmıştır.

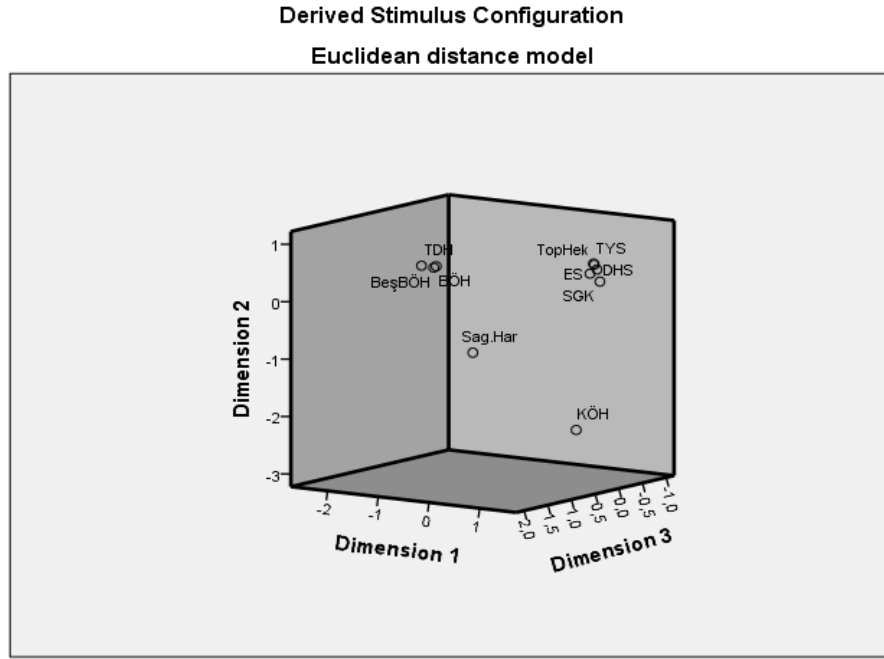
Tablo 4.9: Üç Boyutlu Uzayda Değişkenler İçin Elde Edilen İterasyonlar ve Kruskal Stress Değeri

Young's S-stress Formülü		
İterasyon	S-stress	Gelişim
1	0.02230	
2	0.01985	0.00245
3	0.01972	0.00013

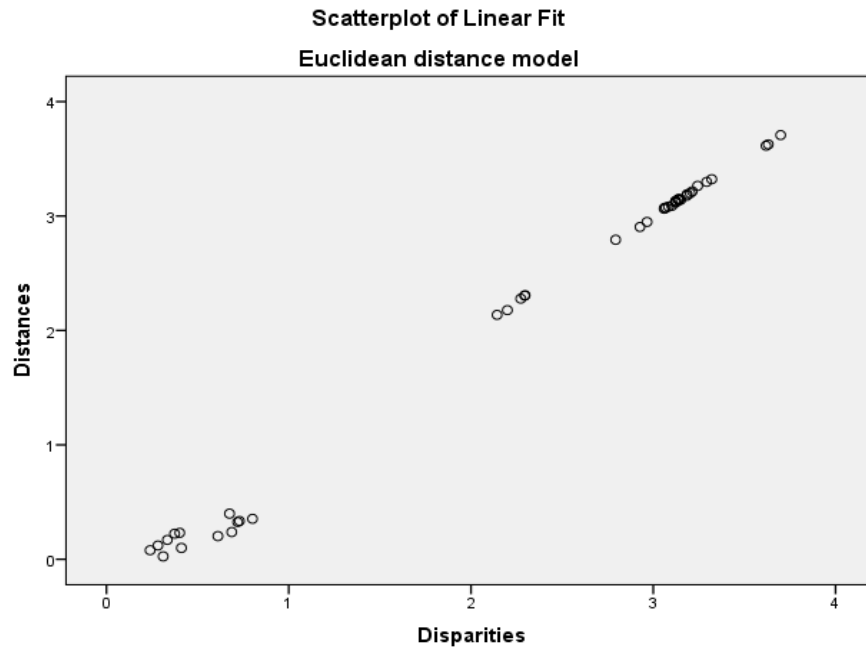
For Matrix	
Stress= 0.06478	RSQ=0.99638

Üç boyutlu uzayda Öklid uzaklık modeli incelendiğinde değişkenlerin 4 farklı grup oluşturduğu görülmüştür. Toplam hekim sayısı, toplam yatak sayısı, eczacı sayısı, diş hekimi sayısı, İş göremez gün sayısı değişkenleri bir grupta, toplam doğurganlık hızı, bebek ölüm hızı, beş yaş altı bebek ölüm hızı gibi demografik değişkenler ise diğer bir grupta konumlanmıştır. Ayrıca sağlık harcamaları payı ve kaba ölüm hızı değişkenleri diğerlerinden farklılaşarak ayrı birer grup oluşturmuştur.

Şekil 4.18: Üç Boyutlu Uzayda Değişkenler İçin Türetilmiş Uyarıcılar Gösterimi



Şekil 4.19: Üç Boyutlu Uzayda Değişkenler İçin Serpilme Diyagramı



Son olarak, deęişkenlerin monotonik dönüşümleri ile oluşturulan farklar ve türetilen uzaklıklar arasındaki ilişki incelenmiştir. Şekil 4.19'daki Shepard diyagramına göre uzaklıklar ve farklılıklar arasında yaklaşık olarak doğrusal bir ilişkinin var olduğu görülmüştür.

4.6.2. Kümeleme Analizi Sonuçları

Çalışmanın ilk bölümünde 26 bölgenin sağlık göstergeleri açısından birbirlerine göre konumları ÇBÖ analizi yardımıyla ortaya konulmuştur. Çalışmanın ikinci bölümünde ise verilere kümeleme analizi uygulanmış ve sonuçlar elde edilmiştir.

Küme sayısı önceden bilinmedięi için, öncelikle hiyerarşik kümeleme yöntemlerinin uygulanması bu aşamada önemlidir. Burada hiyerarşik kümeleme yöntemleri olarak tam bağlantı, ortalama bağlantı ve en küçük varyans/WARD kümeleme yöntemlerinden; hiyerarşik olmayan kümeleme yöntemleri olarak ise k-ortalamlar kümeleme yönteminden yararlanılmıştır.

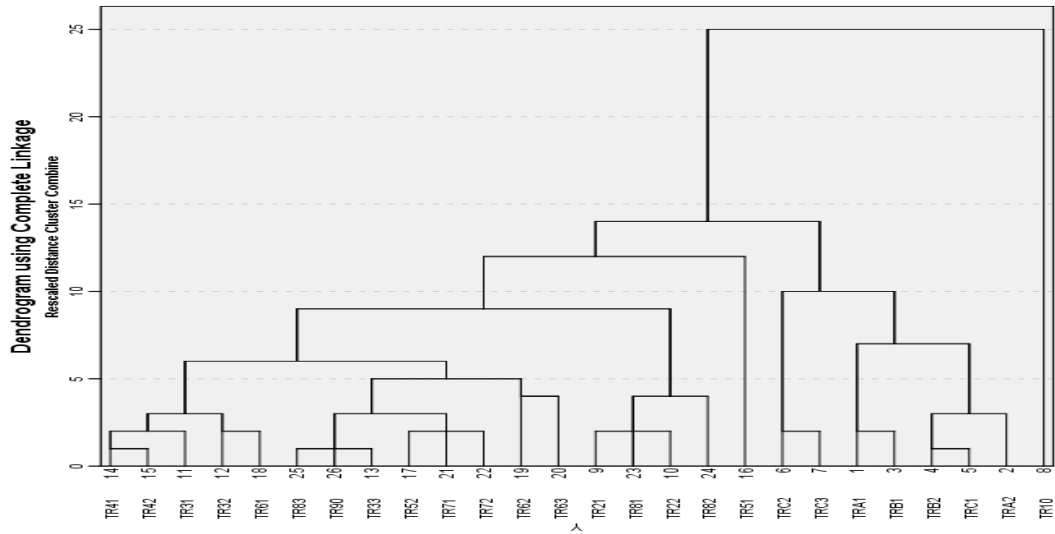
Hiyerarşik kümeleme yöntemleri için birbirlerine en benzer bölgelerin eşleştirildięi birleştirici tabloları elde edilmiştir. Üçüncü bölümde ayrıntılı bir şekilde anlatıldığı gibi, bu yöntemde kümeleme işlemi özelden genele doğru hiyerarşik bir düzende gitmekte ve gözlemlerin tümünü kapsayan tek ve büyük bir küme oluşana kadar bu işlem devam etmektedir. İlk olarak en yakın iki gözlemin bir küme oluşturması beklendiğinden gözlem sayısının 1 eksięi kadar küme meydana gelecektir. İlk aşamada birbirlerine en yakın bölgeler bir küme oluşturarak, sonrasında bu oluşan kümelere en yakın bölgelerden oluşan kümelerin eklenmesiyle üçüncü bir küme elde edilmiş olunacaktır.

En uzak komşuluk kümeleme yöntemine göre elde edilen uzaklıklar tablosu incelendiğinde ilk basamakta TR41 ve TR42 (0.644) bölgeleri birbirlerine en yakın bölgeler oldukları görülmüştür. Aynı şekilde, TR83 ve TR90 bölgelerinin de birbirlerine benzer olduęu söylenebilir. TRA1 ve TR10 (12.005) bölgeleri ise 25. basamakta birleşerek birbirlerine en uzak olan bölgeler konumundadır.

Tablo 4.10: En Uzak Komşuluk Yöntemi ile Elde Edilen Birleştirici Kümeleme Tablosu

Sıra	Küme-1	Küme-2	Uzaklık Katsayıları
1	Bursa Alt Bölgesi (TR41)	Kocaeli Alt Bölgesi (TR42)	.644
2	Samsun Alt Bölgesi (TR83)	Trabzon Alt Bölgesi (TR90)	.769
3	Manisa Alt Bölgesi (TR33)	Samsun Alt Bölgesi (TR83)	1.055
4	Van Alt Bölgesi (TRB2)	Gaziantep Alt Bölgesi (TRC1)	1.086
5	Konya Alt Bölgesi (TR52)	Kırıkkale Alt Bölgesi (TR71)	1.145
6	Tekirdağ Alt Bölgesi (TR21)	Zonguldak Alt Bölgesi (TR81)	1.159
7	Aydın Alt Bölgesi (TR32)	Antalya Alt Bölgesi (TR61)	1.320
8	İzmir Alt Bölgesi (TR31)	Bursa Alt Bölgesi (TR41)	1.416
9	Konya Alt Bölgesi (TR52)	Kayseri Alt Bölgesi (TR72)	1.420
10	Erzurum Alt Bölgesi (TRA1)	Malatya Alt Bölgesi (TRB1)	1.451
11	Tekirdağ Alt Bölgesi (TR21)	Balıkesir Alt Bölgesi (TR22)	1.472
12	Şanlıurfa Alt Bölgesi (TRC2)	Mardin Alt Bölgesi (TRC3)	1.487
13	Manisa Alt Bölgesi (TR33)	Konya Alt Bölgesi (TR52)	1.897
14	İzmir Alt Bölgesi (TR31)	Aydın Alt Bölgesi (TR32)	1.917
15	Ağrı Alt Bölgesi (TRA2)	Van Alt Bölgesi (TRB2)	1.928
16	Adana Alt Bölgesi (TR62)	Hatay Alt Bölgesi (TR63)	2.036
17	Tekirdağ Alt Bölgesi (TR21)	Kastamonu Alt Bölgesi (TR82)	2.359
18	Manisa Alt Bölgesi (TR33)	Adana Alt Bölgesi (TR62)	2.749
19	İzmir Alt Bölgesi (TR31)	Manisa Alt Bölgesi (TR33)	3.107
20	Erzurum Alt Bölgesi (TRA1)	Ağrı Alt Bölgesi (TRA2)	3.391
21	Tekirdağ Alt Bölgesi (TR21)	İzmir Alt Bölgesi (TR31)	4.470
22	Erzurum Alt Bölgesi (TRA1)	Şanlıurfa Alt Bölgesi (TRC2)	5.180
23	Tekirdağ Alt Bölgesi (TR21)	Ankara Alt Bölgesi (TR51)	6.039
24	Erzurum Alt Bölgesi (TRA1)	Tekirdağ Alt Bölgesi (TR21)	6.897
25	Erzurum Alt Bölgesi (TRA1)	İstanbul Alt Bölgesi (TR10)	12.005

Şekil 4.20: En Uzak Komşuluk Yöntemi ile Oluşan Dendogram



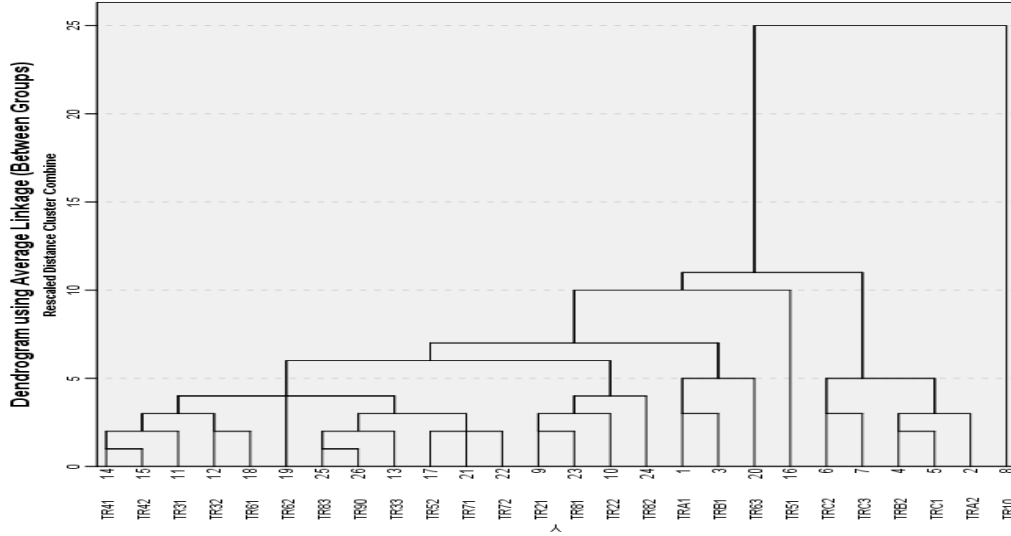
En uzak komşuluk kümeleme analizine göre elde edilen ağaç diyagramı incelendiğinde, gözlemlerin üç grupta kümelendiği görülmektedir. Şekil 4.20'ye göre TR10 tek başına bir küme oluşturduğu, TRC2, TRC3, TRA1, TRB1, TRB2, TRC1 ve TRA2'nin ayrı bir grupta kümelendiği, geri kalan bölgelerin tamamının ise tek bir grupta kümelendiği söylenebilir.

En uzak komşuluktan sonra ortalama bağlantı yöntemi ile elde edilen birleştirici kümeleme analizine ait sonuçlar Tablo 4.10'da verilmiştir. Buna göre birbirine en yakın bölgeler yine TR41 ve TR42 (0.644) bölgeleri olmuştur. En uzak bölgeler de yine TRA1 ve TR10 (10.121) bölgeleridir.

Tablo 4.11: Ortalama Bağlantı Yöntemi ile Elde Edilen Birleştirici Kümeleme Tablosu

Sıra	Küme-1	Küme-2	Uzaklık Katsayıları
1	Bursa Alt Bölgesi (TR41)	Kocaeli Alt Bölgesi (TR42)	.644
2	Samsun Alt Bölgesi (TR83)	Trabzon Alt Bölgesi (TR90)	.769
3	Manisa Alt Bölgesi (TR33)	Samsun Alt Bölgesi (TR83)	1.037
4	Van Alt Bölgesi (TRB2)	Gaziantep Alt Bölgesi (TRC1)	1.086
5	Konya Alt Bölgesi (TR52)	Kırıkkale Alt Bölgesi (TR71)	1.145
6	Tekirdağ Alt Bölgesi (TR21)	Zonguldak Alt Bölgesi (TR81)	1.159
7	İzmir Alt Bölgesi (TR31)	Bursa Alt Bölgesi (TR41)	1.222
8	Aydın Alt Bölgesi (TR32)	Antalya Alt Bölgesi (TR61)	1.320
9	Konya Alt Bölgesi (TR52)	Kayseri Alt Bölgesi (TR72)	1.334
10	Erzurum Alt Bölgesi (TRA1)	Malatya Alt Bölgesi (TRB1)	1.451
11	Tekirdağ Alt Bölgesi (TR21)	Balıkesir Alt Bölgesi (TR22)	1.459
12	Şanlıurfa Alt Bölgesi (TRC2)	Mardin Alt Bölgesi (TRC3)	1.487
13	Manisa Alt Bölgesi (TR33)	Konya Alt Bölgesi (TR52)	1.533
14	İzmir Alt Bölgesi (TR31)	Aydın Alt Bölgesi (TR32)	1.628
15	Ağrı Alt Bölgesi (TRA2)	Van Alt Bölgesi (TRB2)	1.651
16	Tekirdağ Alt Bölgesi (TR21)	Kastamonu Alt Bölgesi (TR82)	1.883
17	İzmir Alt Bölgesi (TR31)	Adana Alt bölgesi (TR62)	1.997
18	İzmir Alt Bölgesi (TR31)	Manisa Alt Bölgesi (TR33)	2.072
19	Erzurum Alt Bölgesi (TRA1)	Hatay Alt Bölgesi (TR63)	2.347
20	Ağrı Alt Bölgesi (TRA2)	Şanlıurfa Alt Bölgesi (TRC2)	2.424
21	Tekirdağ Alt Bölgesi (TR21)	İzmir Alt Bölgesi (TR31)	2.571
22	Erzurum Alt Bölgesi (TRA1)	Tekirdağ Alt bölgesi (TR21)	3.265
23	Erzurum Alt Bölgesi (TRA1)	Ankara Ant Bölgesi (TR51)	4.198
24	Erzurum Alt Bölgesi (TRA1)	Ağrı Alt Bölgesi (TRA2)	4.714
25	Erzurum Alt Bölgesi (TRA1)	İstanbul Alt Bölgesi (TR10)	10.121

Şekil 4.21: Ortalama Bağlantı Yöntemi ile Oluşan Dendogram



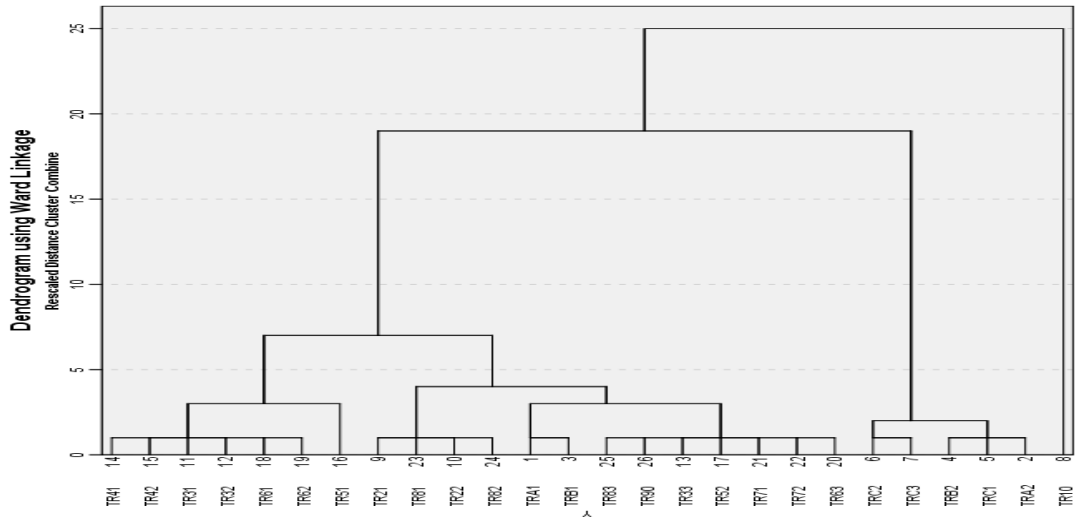
Şekil 4.21’da ortalama bağlantı kümeleme yöntemi ile elde edilen ağaç diyagramı gösterilmektedir. Bu diyagrama göre bölgelerin yaklaşık olarak üç grupta kümelendiği söylenebilir. TR10 bölgesi diğer tüm bölgelerden farklı olarak tek başına bir küme oluşturmuştur. TRA2, TRB2, TRC1, TRC2 ve TRC3 gibi doğu bölgelerinin birlikte ayrı bir küme oluşturduğu, diğer kalan tüm bölgelerin de üçüncü bir grupta kümelendiği sonucuna ulaşılmıştır. Ayrıca TR51 bölgesinin grupta yer alan diğer bölgelerden biraz da olsa farklılaştığı görülmüştür.

Karesel Öklid uzaklığı ile elde edilen uzaklıkların kullanıldığı Ward kümeleme yönteme göre sonuçlar Tablo 4.12’de gösterilmiştir. Birleştirici kümeleme tablosundaki sonuçlar incelendiğinde birbirlerine en benzer bölgelerin ilk basamakta yer alan TR41 ve TR42 (0.207) bölgelerinin olduğu görülmüştür. Birbirlerine en uzak ya da benzemez bölgeler ise 25. basamaktaki TRA1 ve TR10 (250.000) bölgeleridir. Aynı şekilde TR83 ve TR90 (0.503) bölgelerinin de birbirlerine benzer ve TRA1 ve TRA2 (155.842) ise birbirlerinden çok farklı olduğu söylenebilir.

Tablo 4.12: WARD Yöntemi ile Elde Edilen Birleştirici Kümeleme Tablosu

Sıra	Küme-1	Küme-2	Uzaklık Katsayıları
1	Bursa Alt Bölgesi (TR41)	Kocaeli Alt Bölgesi (TR42)	.207
2	Samsun Alt Bölgesi (TR83)	Trabzon Alt Bölgesi (TR90)	.503
3	Van Alt Bölgesi (TRB2)	Gaziantep Alt Bölgesi (TRC1)	1.093
4	Manisa Alt Bölgesi (TR33)	Samsun Alt Bölgesi (TR83)	1.711
5	Konya Alt Bölgesi (TR52)	Kırıkkale Alt Bölgesi (TR71)	2.367
6	Tekirdağ Alt Bölgesi (TR21)	Zonguldak Alt Bölgesi (TR81)	3.038
7	Aydın Alt Bölgesi (TR32)	Antalya Alt Bölgesi (TR61)	3.909
8	İzmir Alt Bölgesi (TR31)	Bursa Alt Bölgesi (TR41)	4.860
9	Konya Alt Bölgesi (TR52)	Kayseri Alt Bölgesi (TR72)	5.833
10	Erzurum Alt Bölgesi (TRA1)	Malatya Alt Bölgesi (TRB1)	6.886
11	Şanlıurfa Alt Bölgesi (TRC2)	Mardin Alt Bölgesi (TRC3)	7.992
12	Balıkesir Alt Bölgesi (TR22)	Kastamonu Alt Bölgesi (TR82)	9.124
13	Aydın Alt Bölgesi (TR32)	Adana Alt Bölgesi (TR62)	10.687
14	Ağrı Alt Bölgesi (TRA2)	Van Alt Bölgesi (TRB2)	12.358
15	Konya Alt Bölgesi (TR52)	Hatay Alt Bölgesi (TR63)	14.064
16	Tekirdağ Alt Bölgesi (TR21)	Balıkesir Alt Bölgesi (TR22)	16.412
17	İzmir Alt Bölgesi (TR31)	Aydın Alt Bölgesi (TR32)	19.830
18	Manisa Alt Bölgesi (TR33)	Konya Alt Bölgesi (TR52)	23.698
19	Ağrı Alt Bölgesi (TRA2)	Şanlıurfa Alt Bölgesi (TRC2)	29.547
20	İzmir Alt Bölgesi (TR31)	Ankara Alt Bölgesi (TR51)	37.501
21	Erzurum Alt Bölgesi (TRA1)	Manisa Alt Bölgesi (TR33)	47.499
22	Erzurum Alt Bölgesi (TRA1)	Tekirdağ Alt Bölgesi (TR21)	61.789
23	Erzurum Alt Bölgesi (TRA1)	İzmir Alt Bölgesi (TR31)	84.810
24	Erzurum Alt Bölgesi (TRA1)	Ağrı Alt Bölgesi (TRA2)	155.842
25	Erzurum Alt Bölgesi (TRA1)	İstanbul Alt Bölgesi (TR10)	250.000

Şekil 4.22: WARD Yöntemi ile Oluşan Dendrogram



Ward kümeleme yöntemi ile oluşturulan ağaç diyagramı incelendiğinde bölgelerin kabaca üç grup içinde kümelendiği görülür. Burada TR10 ile İstanbul

bölgesi tek başına bir küme oluşturmuştur. TRA2, TRC1, TRB2, TRC3 ve TRC2'nin bir küme, geri kalan tüm bölgeler ise diğer bir grupta kümelenmiştir.

Son olarak hiyerarşik olmayan kümeleme yöntemi uygulanmıştır. 26 bölgenin sağlık göstergeleri açısından üç kümede gruplandığı hiyerarşik kümeleme yöntemleri ile elde edilmişti. Burada, küme sayısının önceden belirlendiği hiyerarşik olmayan kümeleme analizi yöntemlerinden k-ortalamlar kümeleme yöntemi $k=3$ küme için elde edilmiştir ve sonuçlar Tablo 4.13'de gösterilmiştir. Ancak oluşan 3 küme içerisindeki bölgelerin diğer yöntemle elde edilen farklı olduğu görülmüştür.

k-ortalamlar yönteminde sınıflandırma 3 küme ile yapıldığında, İstanbul'un diğer tüm bölgelerden farklılaşarak ayrı bir küme oluşturduğu görülmüştür. Ayrıca; İzmir, Ankara ve Antalya alt bölgeleri ile bunlara komşu bölgelerin ayrı bir kümede birleştiği ve sağlık göstergeleri açısından birbirleri ile benzer oldukları söylenebilir.

Tablo 4.13: k-Ortalamlar ($k=3$) Yöntemi ile Elde Edilen Kümeler

Düzy-2 /Bölgeler	Küme	Düzy-2 /Bölgeler	Küme
Erzurum Alt Bölgesi (TRA1)	2	Bursa Alt Bölgesi (TR41)	3
Ağrı Alt Bölgesi (TRA2)	2	Kocaeli Alt Bölgesi (TR42)	3
Malatya Alt Bölgesi (TRB1)	2	Ankara Alt Bölgesi (TR51)	3
Van Alt Bölgesi (TRB2)	2	Konya Alt Bölgesi (TR52)	2
Gaziantep Alt Bölgesi (TRC1)	2	Antalya Alt Bölgesi (TR61)	3
Şanlıurfa Alt Bölgesi (TRC2)	2	Adana Alt Bölgesi (TR62)	3
Mardin Alt Bölgesi (TRC3)	2	Hatay Alt Bölgesi (TR63)	2
İstanbul Alt Bölgesi (TR10)	1	Kırıkkale Alt Bölgesi (TR71)	2
Tekirdağ Alt Bölgesi (TR21)	2	Kayseri Alt Bölgesi (TR72)	2
Balıkesir Alt Bölgesi (TR22)	2	Zonguldak Alt Bölgesi (TR81)	2
İzmir Alt Bölgesi (TR31)	3	Kastamonu Alt Bölgesi (TR82)	2
Aydın Alt Bölgesi (TR32)	3	Samsun Alt Bölgesi (TR83)	2
Manisa Alt Bölgesi (TR33)	3	Trabzon Alt Bölgesi (TR90)	2

k-Ortalamlar kümeleme analizinde son küme merkezleri, kümeler ve değişkenler için önemli bir çıktıdır. $k=3$ kümeli sınıflandırma için elde edilen küme merkezlerine bakıldığında, iş göremez gün sayısı dışında kalan tüm değişkenler

açısından en iyi konumda olan, bir numaralı kümede yer alan İstanbul bölgesidir. İş göremez gün sayısı dikkate alındığında ise en iyi kümenin üç numaralı küme olduğu görülmüştür. Ayrıca tüm değişkenler göz önüne alındığında en kötü konumda bulunan küme iki numaralı küme olmuştur.

Tablo 4.14: k- Ortalamalar Yönteminde Son Küme Merkezleri

Sağlık Göstergeleri	1	2	3
Eczacı Sayısı	5312.00	602.71	1496.50
Toplam Hekim Sayısı	28797.00	3013.06	7655.00
Bebek Ölüm Hızı	8.50	11.16	8.88
Diş Hekimi Sayısı	6497.00	434.00	1369.88
Kaba Ölüm Hızı	4.10	5.66	5.64
Beş Yaş Altı ÇÖH	10.00	13.54	10.49
İş Göremez Gün Sayısı	5151807.00	382636.94	1447125.75
Toplam Doğurganlık Hızı	1.88	2.37	1.89
Toplam Yatak Sayısı	34089.00	5354.29	10567.00
Toplam Sağlık Harcaması	2.40	1.81	2.10

k-Ortalamalar yönteminde; oluşan kümeler arasında değişkenler açısından fark olup olmadığının değerlendirilmesinde ve analize hangi değişkenlerin etki ettiğini belirlemek amacıyla varyans analizi sonuçlarından yararlanılmaktadır³³. Tablo 4.14'daki sonuçlara göre, eczacı sayısı, toplam hekim sayısı, diş hekimi sayısı, İş göremez gün sayısı ve toplam yatak sayısı değişkenleri açısından kümeler arasında anlamlı bir fark bulunmuştur (p anlamlılık düzeyi 0.05'ten küçük olanlar). Dolayısıyla bu değişkenlerin analizde etkili ve anlamlı olduğu söylenebilir. Ayrıca F değerlerine bakıldığında iş göremez gün sayısı (137.849), eczacı sayısı (113.349) ve diş hekimi sayısı (100.868) değişkenlerinin en etkili değişkenler olduğu görülmüştür.

Sonuç olarak çok boyutlu ölçekleme ve kümeleme analizi birlikte incelendiğinde; analizlerin birbirlerini destekleyici sonuçlar verdiği görülmüştür. Sağlık göstergeleri bakımından Düzey-2 bölgelerinin iki analiz türünde de yaklaşık olarak üç farklı grup oluşturdıkları söylenebilir. Farklı kümeleme analizi

³³ Şeref Kalaycı, a.g.e., s. 361

yaklaşımlarının uygulanmasıyla bölgelerin gruplar içindeki dağılımları da farklılaştığı görülmüştür.

Tablo 4.15: k- Ortalamalar Varyans Analizi Tablosu

ANOVA TABLOSU						
Sağlık Göstergeleri	Küme Kareler O.	df	Hata Kareler O.	df	F	Sig.
Eczacı Sayısı	11579365.928	2	102156.414	23	113.349	0.000
Toplam Hekim Sayısı	342464364.010	2	5868030.997	23	58.361	0.000
Bebek Ölüm Hızı	16.055	2	8.151	23	1.970	0.162
Dış Hekimi Sayısı	18352615.255	2	181947.603	23	100.868	0.000
Kaba Ölüm Hızı	1.159	2	2.714	23	0.427	0.657
Beş Yaş Altı ÇÖH	28.417	2	11.719	23	2.425	0.111
İş Göremez Gün Sayısı	12510934054643.9	2	90758164460.889	23	137.849	0.000
Toplam Doğurganlık Hızı	0.700	2	0.384	23	1.826	0.184
Toplam Yatak Sayısı	426121644.312	2	6300398.327	23	67.634	0.000
Toplam Sağlık Harcaması	0.355	2	0.116	23	3.063	0.066

Son olarak, İstanbul alt bölgesinin nüfus bakımından diğer bölgelerden oldukça fazla olduğu göz önüne alınmış ve analiz bir de İstanbul alt bölgesi hariç olarak gerçekleştirilmiştir. Uygulanan çok boyutlu ölçekleme analizi sonuçlarına göre, 25 bölgenin (0.04311) stress değeri ile üç boyutlu uzayda üç farklı grup oluşturdukları görülmüştür. Ankara alt bölgesi diğer tüm bölgelerden ayrı olarak tek başına bir grupta yer alırken, TRA1,TRA2,TRB2,TRC1,TRC2,TRC3 gibi doğu bölgelerinin ayrı bir grupta konumlanmıştır. Diğer tüm bölgeler ise üçüncü grubu oluşturmaktadır. Analizde $R^2 = 0.99369$ olarak bulunmuştur. Buna bağlı olarak üç boyutlu uzayda elde edilen modelin verileri çok iyi bir şekilde temsil ettiği söylenebilir.

5. SONUÇ

Ulusal kalkınmanın temelde bölgesel kalkınmaya dayandığı söylenebilir. Günümüzde sağlık göstergeleri de ülkelerin ve dolayısıyla temelde bölgelerin kalkınma seviyelerini belirleyen faktörler içerisinde yer almaktadır. Bu nedenle, bölgelerin mevcut kapasitelerinin optimum düzeyde kullanımını amaçlayan İstatistiki Bölge Birimleri Sınıflaması (İBBS) Düzey-2 bölgelerinin çalışmada kullanılması tercih edilmiştir. Sağlık göstergeleri açısından bölgeler arasındaki farklılıkların ortaya konulması ve sonucunda bu farklılıkları giderici politikaların uygulanmasıyla bölgelerin kalkınması sağlanabilir.

Bu çalışmada, İstatistiki bölge birimleri sınıflamasına göre ayrılan Düzey-2/26 bölgenin seçilen sağlık göstergeleri açısından birbirlerine göre konumlarının belirlenmesi; ÇBÖ ve kümeleme analizleri yardımıyla bölgeler arasında var olan benzerlik ve farklılıkların tespit edilmesi amaçlanmıştır. Burada; eczacı sayısı, toplam hekim sayısı, bebek ölüm hızı, diş hekimi sayısı, kaba ölüm hızı, beş yaş altı çocuk ölüm hızı, İş göremez gün sayısı, toplam doğurganlık sayısı, toplam yatak sayısı, toplam sağlık harcaması olmak üzere toplam 10 adet sağlık göstergesi kullanılmıştır.

ÇBÖ ve kümeleme analizleri nesneler arasında var olan benzerlikler yardımıyla veri kümesinin temel boyutlarını ortaya çıkarmayı amaçlamaktadır. Her iki yöntem sonuçları incelendiğinde birbirlerini destekleyici nitelikte olduğu görülmüştür. Analiz sonuçlarına göre, 26 istatistiki bölge içerisinde İstanbul alt bölgesi her iki yöntemde de diğer bölgelerden anlamlı bir şekilde ayrıştığı söylenebilir.

Çalışmanın ilk bölümünde, ÇBÖ analizi olarak ALSCAL algoritmasından yararlanılmıştır. Analiz 2 ve 3 boyutlu olarak gerçekleştirilmiş, 3 boyutlu sonuçların daha güvenilir olmasından dolayı bu boyutta sonuçlar yorumlanmıştır. Ancak 2 boyutlu sonuçlara da yer verilmiştir.

ÇBÖ analizi sonuçlarına göre Düzey-2 bölgelerinin 3 boyutlu uzayda üç farklı grupta konumlandıkları görülmüştür. İstanbul alt bölgesi diğer tüm bölgelerden ayrılarak tek başına bir grup oluşturmuştur. Güneydoğu Anadolu bölgesi, Van Alt Bölgesi ve Ağrı Alt bölgesinin birbirlerine yakın şekilde konumlandığı; İzmir, Ankara, Tekirdağ gibi büyük şehirleri de içeren diğer bölgelerin ise üçüncü bir grupta konumlandığı söylenebilir. İstanbul alt bölgesinin özellikle diş hekimi, eczacı ve toplam hekim sayısı; iş göremez gün sayısı ve toplam yatak sayısı bakımından diğer bölgelerden önemli derecede farklılaştığı görülmüştür. Temel etken olarak İstanbul alt bölgesinin diğer bölgelerden çok daha fazla nüfusa sahip olması gösterilebilir. Aynı şekilde ikinci grubu oluşturan doğu bölgeleri incelendiğinde; özellikle BÖH, TDH, Beş yaş BÖH oranlarının diğer bölgelerden fazla olduğu söylenebilir. Buna karşılık KÖH oranlarının diğer bölgelere oranla daha düşük olduğu görülmüştür.

Çalışmanın ikinci bölümünde, kümeleme analizinden yararlanılmıştır. Hiyerarşik kümeleme yöntemleri olarak; tam bağlantı, ortalama bağlantı ve WARD yönteminden, hiyerarşik olmayan kümeleme yöntemleri olarak ise k-ortalamlar kümeleme yöntemi İBBS Düzey-2 verilerine uygulanmıştır.

Hiyerarşik kümeleme analizinde her üç yöntem sonuçlarına göre bölgelerin üç farklı grupta kümelendiği sonucuna ulaşılmıştır. ÇBÖ'de olduğu gibi İstanbul bölgesi diğer tüm bölgelerden farklılaşmış ve tek başına bir küme oluşturmuştur. Ortalama bağlantı ve WARD yöntemleri sonuçlarına göre Güneydoğu Anadolu Bölgesi, Van alt bölgesi ve Ağrı alt bölgeleri ikinci grubu oluştururken, en uzak komşuluk yönteminde Erzurum ve Malatya alt bölgelerinin de ikinci gruba dahil olduğu görülmüştür. Son olarak k-ortalamlar kümeleme analizine göre; eczacı sayısı, toplam hekim sayısı, diş hekimi sayısı, iş göremez gün sayısı ve toplam yatak sayısı değişkenlerinin analiz için etkili ve anlamlı olduğu sonucuna ulaşılmıştır.

KAYNAKÇA

- Ağgün, Fikri: “Çok Boyutlu Ölçekleme Analizinin İncelenmesi ve Bir Uygulama”, Yüzüncü Yıl Üniversitesi, Zootekni Anabilim Dalı, Yüksek Lisans Tezi, Van, 2011.
- Akın, Besim H.; Eren, Özge: “OECD Ülkelerinin Eğitim Göstergelerinin Kümeleme Analizi ve Çok Boyutlu Ölçekleme Analizi İle Karşılaştırmalı Analizi”, Ocak, 2012, s. 175-181.
- Akdur, Recep: “Türkiye’de Sağlık Hizmetleri ve Avrupa Topluluğu Ülkeleri ile Kıyaslanması”, A.Ü.T.F. Halk Sağlığı Anabilim Dalı, Ankara, 1999.
- Akküçük, Ulaş: “Bir Çok Boyutlu Ölçekleme Tekniği Olarak Torgersen Ölçekleme Yöntemi ve Temel Bileşenler Analizi ile Karşılaştırması”, **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi**, Sayı 25, Aralık, 2009, s.311-322.
- Alphar, Reha: **Uygulamalı Çok Değişkenli İstatistiksel Yöntemler**, Ankara, Detay Yayıncılık, 2013.
- Anderberg, Michael R.: **Cluster Analysis for Applications: Probability and Mathematical Statistics, Air Force Systems Command United States Air Force**, Academic Press New York.
- Anderson, Theodore W.: **An Introduction to Multivariate Statistical Analysis**, Third Edition, John Wiley and Sons, 2003.
- Atalay, Kumru Didem: “Kümeleme Analizi”, <http://tip.baskent.edu.tr/kw/upload/600/dosyalar/cg/sempozyum/ogrsmpzsnm13/13.P9.pdf>, (çevrimiçi), 18.09.2016.
- Bartholomew, David J.; Steele, Fiona; Moustaki, Irini; Galbraith, Jane I.: **Analysis of Multivariate Social Science Data**, Second Edition, Taylor & Francis Group, LLC, 2008.
- Boz, Canser; Sur, Haydar; Söyük, Selma: “The Similarities and Differences Analysis of OECD Countries in Terms of Health System Indicators”, **ACU Sağlık Bil Derg.**, 3, 2016, s. 154-164.

- Chapter 435: “Multidimensional Scaling”, **NCSS Statistical Software**, https://ncss-wpengine.netdnssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Multidimensional_Scaling.pdf, (Çevrimiçi), 14.08.2015.
- Chatfield, C.; Collins, A.J.: **Introduction to Multivariate Analysis**, Springer Science Business Media, 1980.
- Cox, Trevor F.; Cox, Michael A.A.: **Multidimensional Scaling**, Chapman & Hall/CRC, Second Edition, 2000.
- Çelik, Şenol: “Kümeleme Analizi İle Sağlık Göstergelerine Göre Türkiye’deki İllerin Sınıflandırılması”, **Doğuş Üniversitesi Dergisi**, 14, (2), 2013, s. 175-194.
- Çelik, Yusuf: “Türkiye Sağlık Harcamalarının Analizi ve Sağlık Harcama Düzeyinin Uygunluğunun Değerlendirilmesi”, **Sosyal Güvenlik Dergisi**, Cilt.1, Sayı.1, Haziran, 2011.
- Ersöz, Filiz: “Türkiye İle OECD Ülkelerinin Sağlık Düzeyleri ve Sağlık Harcamalarının Analizi”, **İstatistikçiler Dergisi**, 2, s. 95-104, 2008.
- Ersöz, Filiz: “OECD’ye Üye Ülkelerin Seçilmiş Sağlık Göstergelerinin Kümeleme ve Ayırma Analizi ile Karşılaştırılması”, **Türkiye Klinikleri J Med Sci**, 29 (6), 2009, s.1650-9.
- Everitt, B.: **An R and S-Plus Companion to Multivariate Analysis**, Springer-Verlag London Limited, 2005.
- Everitt B.S. et. al.: **Cluster Analysis**, 5th Edition, John Wiley and Sons, 2011.
- Girginer, Nuray: “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, (323), s. 55-72, 2013.
- Gore, Paul A.: “Cluster Analysis”, Department of Psychology, Southern Illinois University,
- Günay Atbaş, Azize Celile: “Kümeleme Analizinde Küme Sayısının Belirlenmesi Üzerine Bir Çalışma”, Ankara Üniversitesi, İstatistik Anabilim Dalı, Yüksek Lisans Tezi, Ankara, 2008.

- Gündüz, Selim: “Uzaklık Fonksiyonlarının Çok Boyutlu Ölçekleme Algoritmalarındaki Etkinliğinin İncelenmesi ve Uygulamalar”, Çukurova Üniversitesi, İstatistik Anabilim Dalı, Yüksek Lisans Tezi, Adana, 2011.
- Gürçaylılar Yenidoğan, Tuğba: “Pazarlama Araştırmalarında Çok Boyutlu Ölçekleme Analizi: Üniversite Öğrencilerinin Marka Algısı Üzerine Bir Araştırma”, **Akdeniz İ.İ.B.F. Dergisi** (15), 2008, s.138-169.
- Hout, Michael C.; Papesh, Megan H.; Goldinger, Stephen D.: “Multidimensional Scaling”, **Wiley Interdiscip Rev. Cogn. Sci.**, January 1, 4(1), 2013.
- Johnson, Richard A.; Wichern, Dean W. **Applied Multivariate Statistical Analysis**, Pearson Prentice Hall, 2007.
- Kalaycı, Şeref: **SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri**, 6. Baskı, Asil Yayın, 2014.
- Kruskal, Joseph B.; Wish, Myron: **Multidimensional Scaling: Series: Quantitative Applications In The Social Science**, Sage Publications, Inc., 1978.
- Kruskal, J. B.: “Multidimensional Scalling By Optimizing Goodness of Fit To A Nonmetric Hypothesis”, **Psychometrika**, vol.29, No.1, March, 1964.
- Kruskal, J. B.: “Nonmetric Multidimensional Scaling: A Numerical Method”, **Psychometrika**, vol.29, No.2, June, 1964.
- Mukhopadhyay, Parimal: “Multivariate Statistical Analysis”, World Scientific, (Çevrimiçi).
- Nemlioğlu, Karun A.: **Benzerlik Analizleri: Varsayımlara Uymayan Kategorik Verilerde Çok Değişkenli Analiz**, İstanbul, Beşir Kitapevi, Temmuz 2005.
- Orhunbilge, Neyran: **Çok Değişkenli İstatistik Yöntemler**, İstanbul Üniversitesi, Yayın No: 4942, İşletme Fakültesi Yayın No: 286, İstanbul, 2010.
- Özdamar, Kazım: **Paket Programlar İle İstatistiksel Veri Analizi: Minitab 16-IBM SPSS 21**, Nisan Kitapevi, 2013.
- Özpınar, Esra: “Bebek Ölüm Hızı Türkiye Geneline Azalırken

Neden Suriyeli Nüfusun Fazla Olduğu Yerlerde Artıyor?”, **Türkiye Ekonomi Politikaları Araştırma Vakfı**, No: 201619, Mayıs, 2016.

- Öztürk, Ahmet: “Homojen ve Fonksiyonel Bölgelerin Tespiti ve Türkiye İçin İstatistiki Bölge Birimleri Önerisi”, **DPT Uzmanlık Tezleri**, Yayın No: DPT: 2802, Ankara, 2009.
- Quinn, Gerry P.; Keough, Michael J.: **Experimental Design and Data Analysis**, Cambridge University Press, New York, 2002.
- Sharma, Subhash: **Applied Multivariate Techniques**, John Wiley and Sons, 1996.
- Sığırlı, Deniz; Ediz, Bülent; Cangür, Şengül; Ercan, İlker; Kan, İsmet: “Türkiye ve Avrupa Birliği’ne Üye Ülkelerin Sağlık Düzeyi Ölçütlerinin Çok Boyutlu Ölçekleme Analizi İle İncelenmesi”, **İnönü Üniversitesi Tıp Fakültesi Dergisi**, 2006, s. 81-85.
- Stratejik Plan 2013-2017: T.C.Sağlık Bakanlığı, https://sgb.saglik.gov.tr/content/files/stratejikplan_20132017/index.html
- Şengül, Ümran; Eslemian, Seyedhadi; Eren, Miraç: “Türkiye’de İstatistiki Bölge Birimleri Sınıflamasına Göre Düzey İki Bölgelerinin Ekonomik Etkinliklerinin VZA Yöntemi ile Belirlenmesi ve Tobit Model Uygulaması”, **Yönetim Bilimleri Dergisi**, Cilt:11, Sayı:21, ss.75-99, 2013.
- Taş, Barış: “AB Uyum Sürecinde Türkiye İçin Yeni Bir Bölge Kavramı: İstatistiki Bölge Birimleri Sınıflandırması (İBBS)”, **Sosyal Bilimler Dergisi**, s.185-198.
- T.C.Sağlık Bakanlığı Tarihçe: <https://www.saglik.gov.tr/TR,11492/tarihce.html>
- Timm, Neil T.: **Applied Multivariate Analysis**, Springer, 2002.
- Ünlükaplan, Yüksel: “Çok Değişkenli İstatistiksel Yöntemlerin Peyzaj Ekolojisi Araştırmalarında Kullanımı”, Çukurova Üniversitesi, Peyzaj Mimarlığı Anabilim Dalı, Doktora Tezi, Adana, 2008.
- Yiğit, Esra: “Çok Boyutlu Ölçekleme Yöntemlerinin İncelenmesi ve Bir Uygulama”, Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı.

- Yumuşak, İbrahim Güran; Yıldırım, Durmuş Çağrı: “Sağlık Harcamaları İktisadi Büyüme İlişkisi Üzerine Ekonometrik Bir İnceleme”, **The Journal of Knowledge Economy & Knowledge Management**, 2009, Volume IV Spring, s. 57-70.
- Young, Forrest W.: **Multidimensional Scaling: History, Theory, and Applications**, ed. Robert M. Hamer, Lawrence Erlbaum Associates, New Jersey, 1987.
- Young, Forrest W. ; Takane, Yoshio; Lewyckyj, Rostyslaw: “ALSCAL: A Nonmetric Multidimensional Scaling Program With Several Individual-Differences Options”, **Behavior Research Methods & Instrumentation**, Vol.10 (3), 1978, s. 451-453.
- Wickelmaier, Florian: “An Introduction to MDS”, Sound Quality Research Unit, Aalborg University, Denmark, May 4, 2003.
- http://www-users.cs.umn.edu/~han/dmclass/cluster_survey_10_02_00.pdf

Ek-1: Analizde Kullanılan Sağlık Göstergeleri Değişkenleri

Bölgeler	ES	THS	BÖH	DHS	BeşBÖH
TRA1	274	2140	12,40	210	15,50
TRA2	187	1300	14,00	155	17,60
TRB1	496	3101	11,20	394	13,70
TRB2	290	2327	15,80	241	19,10
TRC1	691	3690	16,40	428	18,70
TRC2	823	5085	16,60	475	20,20
TRC3	437	2171	14,20	283	17,60
TR10	5312	28797	8,50	6497	10,00
TR21	608	2730	6,30	545	7,80
TR22	660	2520	8,30	512	9,90
TR31	1891	9996	8,80	1759	10,40
TR32	1373	5273	8,40	1032	9,50
TR33	1088	4534	10,00	666	11,90
TR41	1403	6491	7,90	1194	9,40
TR42	1102	5845	8,30	1027	9,80
TR51	2300	16595	7,50	2971	9,20
TR52	905	4248	9,90	630	11,90
TR61	1419	5977	8,90	1246	10,40
TR62	1396	6529	11,20	1064	13,30
TR63	987	4036	11,90	646	14,20
TR71	564	2090	9,30	327	11,00
TR72	801	4277	11,60	636	13,80
TR81	348	1704	7,00	255	9,10
TR82	264	888	8,00	194	8,80
TR83	966	4615	8,90	716	11,40
TR90	945	4300	8,00	731	9,80

*ES: Eczacı Sayısı, THS: Toplam Hekim Sayısı, BÖH: Bebek Ölüm Hızı, DHS: Diş Hekimi Sayısı, BeşBÖH: Beş Yaş Altı Bebek Ölüm Hızı

Ek-1: Devam

Bölgeler	KÖH	SGK	TDH	TYS	TSH
TRA1	5,70	156882	2,33	4412	0,90
TRA2	4,30	59314	3,21	2101	1,30
TRB1	5,20	302766	2,05	6447	1,30
TRB2	3,20	66016	3,29	4538	1,60
TRC1	4,00	495028	3,07	6071	1,80
TRC2	3,20	222061	3,81	7652	2,30
TRC3	3,10	122969	3,51	3769	2,30
TR10	4,10	5151807	1,88	34089	2,40
TR21	6,90	850814	1,76	4796	1,80
TR22	8,60	403705	1,60	4362	1,80
TR31	5,90	1734293	1,72	11466	2,00
TR32	6,40	1021689	1,82	7813	2,20
TR33	7,20	986463	1,86	8857	2,00
TR41	5,90	1942202	1,82	10611	2,00
TR42	5,40	1975031	1,95	9156	1,90
TR51	4,50	1614092	1,76	18109	1,80
TR52	5,20	556306	2,18	6947	1,80
TR61	4,90	1251479	1,91	8230	2,50
TR62	4,90	1051757	2,24	10294	2,40
TR63	4,40	634463	2,63	6969	1,80
TR71	5,90	284578	2,01	3655	2,00
TR72	6,00	600183	2,09	7491	2,10
TR81	7,00	422112	1,55	3240	2,10
TR82	9,70	165595	1,68	1947	2,20
TR83	6,90	582822	1,79	8489	1,70
TR90	6,90	579214	1,77	8137	1,90

*KÖH: Kaba Ölüm Hızı, SGK: Sosyal Güvenlik Kurumu İş Göremez Gün Sayısı, TDH: Toplam Doğurganlık Hızı, YYS: Toplam Yatak Sayısı, TSH: Hanehalkı Bütçe Anketi Tüketim Harcaması İçerisindeki Sağlık Harcaması Payı.

Ek-2: Kümeleme Analizi Sonuçları Birleştirici Kümeleme Tabloları

En Uzak Komşuluk Yöntemi ile Birleştirici Kümeleme Tablosu

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	14	15	,644	0	0	8
2	25	26	,769	0	0	3
3	13	25	1,055	0	2	13
4	4	5	1,086	0	0	15
5	17	21	1,145	0	0	9
6	9	23	1,159	0	0	11
7	12	18	1,320	0	0	14
8	11	14	1,416	0	1	14
9	17	22	1,420	5	0	13
10	1	3	1,451	0	0	20
11	9	10	1,472	6	0	17
12	6	7	1,487	0	0	22
13	13	17	1,897	3	9	18
14	11	12	1,917	8	7	19
15	2	4	1,928	0	4	20
16	19	20	2,036	0	0	18
17	9	24	2,359	11	0	21
18	13	19	2,749	13	16	19
19	11	13	3,107	14	18	21
20	1	2	3,391	10	15	22
21	9	11	4,470	17	19	23
22	1	6	5,180	20	12	24
23	9	16	6,039	21	0	24
24	1	9	6,897	22	23	25
25	1	8	12,005	24	0	0

Ek-2: Devam

Ortalama Bağlantı Yöntemi ile Birleştirici Kümeleme Tablosu

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	14	15	,644	0	0	7
2	25	26	,769	0	0	3
3	13	25	1,037	0	2	13
4	4	5	1,086	0	0	15
5	17	21	1,145	0	0	9
6	9	23	1,159	0	0	11
7	11	14	1,222	0	1	14
8	12	18	1,320	0	0	14
9	17	22	1,334	5	0	13
10	1	3	1,451	0	0	19
11	9	10	1,459	6	0	16
12	6	7	1,487	0	0	20
13	13	17	1,533	3	9	18
14	11	12	1,628	7	8	17
15	2	4	1,651	0	4	20
16	9	24	1,883	11	0	21
17	11	19	1,997	14	0	18
18	11	13	2,072	17	13	21
19	1	20	2,347	10	0	22
20	2	6	2,424	15	12	24
21	9	11	2,571	16	18	22
22	1	9	3,265	19	21	23
23	1	16	4,198	22	0	24
24	1	2	4,714	23	20	25
25	1	8	10,121	24	0	0

Ek-2: Devam

Ward Yöntemi ile Birleştirici Kümeleme Tablosu

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	14	15	,207	0	0	8
2	25	26	,503	0	0	4
3	4	5	1,093	0	0	14
4	13	25	1,711	0	2	18
5	17	21	2,367	0	0	9
6	9	23	3,038	0	0	16
7	12	18	3,909	0	0	13
8	11	14	4,860	0	1	17
9	17	22	5,833	5	0	15
10	1	3	6,886	0	0	21
11	6	7	7,992	0	0	19
12	10	24	9,124	0	0	16
13	12	19	10,687	7	0	17
14	2	4	12,358	0	3	19
15	17	20	14,064	9	0	18
16	9	10	16,412	6	12	22
17	11	12	19,830	8	13	20
18	13	17	23,698	4	15	21
19	2	6	29,547	14	11	24
20	11	16	37,501	17	0	23
21	1	13	47,499	10	18	22
22	1	9	61,789	21	16	23
23	1	11	84,810	22	20	24
24	1	2	155,842	23	19	25
25	1	8	250,000	24	0	0

EK-3: Üçüncü Bölümde Dendogram Örnekleri İçin Kullanılan Veri Kümesi

Avrupa'da Protein Tüketimi

Ülke Numarası	Rmeat	Wmeat	Eggs	Milk	Fish	Cerl	Starch	Nuts	Fr-veg
1	10,1	1,4	0,5	8,9	0,2	42,3	0,6	5,5	1,7
2	8,9	14,0	4,3	19,9	2,1	28,0	3,6	1,3	4,3
3	13,5	9,3	4,1	17,5	4,5	26,6	5,7	2,1	4,0
4	7,8	6,0	1,6	8,3	1,2	56,7	1,1	3,7	4,2
5	9,7	11,4	2,8	12,5	2,0	34,3	5,0	1,1	4,0
6	10,6	10,8	3,7	25,0	9,9	21,9	4,8	0,7	2,4
7	8,4	11,6	3,7	11,1	5,4	24,6	6,5	0,8	3,6
8	9,5	4,9	2,7	33,7	5,8	26,3	5,1	1,0	1,4
9	18,0	9,9	3,3	19,5	5,7	28,1	4,8	2,4	6,5
10	10,2	3,0	2,8	17,6	5,9	41,7	2,2	7,8	6,5
11	5,3	12,4	2,9	9,7	0,3	40,1	4,0	5,4	4,2
12	13,9	10,0	4,7	25,8	2,2	24,0	6,2	1,6	2,9
13	9,0	5,1	2,9	13,7	3,4	36,8	2,1	4,3	6,7
4	9,5	13,6	3,6	23,4	2,5	22,4	4,2	1,8	3,7
15	9,4	4,7	2,7	23,3	9,7	23,0	4,6	1,6	2,7
16	6,9	10,2	2,7	19,3	3,0	36,1	5,9	2,0	6,6
17	6,2	3,7	1,1	4,9	14,2	27,0	5,9	4,7	7,9
18	6,2	6,3	1,5	11,1	1,0	49,6	3,1	5,3	2,8
19	7,1	3,4	3,1	8,6	7,0	29,2	5,7	5,9	7,2
20	9,9	7,8	3,5	24,7	7,5	19,5	3,7	1,4	2,0
21	13,1	10,1	3,1	23,8	2,3	25,6	2,8	2,4	4,9
22	17,4	5,7	4,7	20,6	4,3	24,3	4,7	3,4	3,3
23	9,3	4,6	2,1	16,6	3,0	43,6	6,4	3,4	2,9
24	11,4	12,5	4,1	18,8	3,4	18,6	5,2	1,5	3,8
25	4,4	5,0	1,2	9,5	0,6	55,9	3,0	5,7	3,2

Kaynak: Neil H.Timm, Applied Multivariate Analysis, p. 535.