

T.C.
MUĞLA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK ANABİLİM DALI

HİBRİD VE İNDİRGENMİŞ KÜMELEME ANALİZİ İLE TÜRKİYE' DEKİ
İLÇELERİN SOSYO-EKONOMİK ÖZELLİKLERİNE GÖRE
SINIFLANDIRILMASI

YÜKSEK LİSANS TEZİ

GÖKHAN SAYIN

OCAK 2010

MUĞLA

T.C.
MUĞLA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK ANABİLİM DALI

HİBRİD VE İNDİRGENMİŞ KÜMELEME ANALİZİ İLE TÜRKİYE' DEKİ
İLÇELERİN SOSYO-EKONOMİK ÖZELLİKLERİNE GÖRE
SINIFLANDIRILMASI

YÜKSEK LİSANS TEZİ

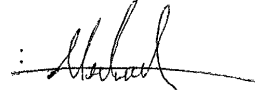
GÖKHAN SAYIN

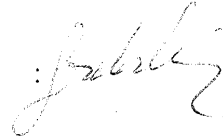
MUĞLA 2010

T.C.
MUĞLA ÜNİVERSİTESİ
Fen Bilimler Enstitüsü

Öğr. Gör. Dr. Mehmet KARAHASAN danışmanlığında **Gökhan SAYIN** tarafından hazırlanan “**Hibrid ve İndirgenmiş Kümeleme Analizi ile Türkiye’deki İlçelerin Sosyo-Ekonomik Özelliklerine göre Sınıflandırılması**” başlıklı tez, 15 Ocak 2010 tarihinde aşağıdaki juri tarafından İstatistik Anabilim Dalı’nda yüksek lisans tezi olarak oybirliği / oyeokluğu ile kabul edilmiştir.

Başkan : Yrd. Doç. Dr. Atilla GÖKTAŞ İmza : 

Üye : Öğr. Gör. Dr. Mehmet KARAHASAN İmza : 

Üye : Yrd. Doç. Dr. Serdar DENİZ İmza : 

ÖNSÖZ

Bu araştırma ile Türkiye’ de ilçeler bazında sosyo-ekonomik gelişmişlik düzeyleri araştırılmış ve çok değişkenli istatistiksel yöntemler aracılığı ile ilçeler kümelenmeye çalışılmıştır.

Yapılan bu araştırmanın temel amacı olarak, son dönemde kümeleme çalışmalarında hızla gelişen Hibrid Kümeleme Yöntemi’ nin bir uygulaması dikkate alınmış ve tanıtılmaya çalışılmıştır. Ayrıca hibrid kümeleme algoritması yaygın kullanılan diğer kümeleme araçları ile aynı veriye uygulanarak, sonuçların geçerlilikleri araştırılmıştır.

Bu tez çalışmasının her aşamasında, öneri ve tavsiyeleriyle yol gösteren tez danışmanım Öğr. Gör. Dr. Mehmet KARAHASAN’ a ve manevi destekleriyle beni hiçbir zaman yalnız bırakmayan aileme teşekkür ederim.

Gökhan SAYIN

MUĞLA - 2010

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ	II
İÇİNDEKİLER	III
ÖZET.....	VI
ABSTRACT.....	VIII
ŞEKİLLER DİZİNİ.....	X
ÇİZELGELER DİZİNİ	XI
SEMBOLLER DİZİNİ.....	XIV
1. GİRİŞ	1
2. KAVRAMSAL ÇERÇEVE VE MEVCUT ÇALIŞMALAR	5
2.1. BÜYÜME – KALKINMA – GELİŞME İLE İLGİLİ KAVRAMSAL ÇERÇEVE	5
2.2. YEREL EKONOMİK KALKINMA	7
2.3. YEREL BİRİMLERİN ÖNEMİ.....	8
2.4. SOSYO-EKONOMİK GELİŞİMİŞLİKTE KONUM FARKLILIKLARI	9
2.5. SOSYO-EKONOMİK GELİŞME FARKLILIKLARININ BÖLGEDE MEYDANA GETİRDİĞİ SORUNLAR.....	10
2.6. BÖLGESEL GELİŞME FARKLILIKLARINI GİDERME ÇALIŞMALARI.....	11
2.7. TÜRKİYE’ DE UYGULANAN BAŞLICA BÖLGESEL GELİŞME POLİTİKALARI.....	11
2.8. İLÇELERİN SOSYO-EKONOMİK GELİŞME DÜZEYLERİNİN BELİRLENMESİNİN ÖNEMİ.....	13
2.9. ULUSLAR ARASI KURUMLAR TARAFINDAN YAPILAN ARAŞTIRMALARDA TÜRKİYE’ NİN KONUMU	14
2.10. GELİŞİMİŞLİK FARKLILIKLARININ BELİRLENMESİNDE KULLANILAN DEĞİŞKENLER	16
2.11. SOSYO-EKONOMİK GELİŞME İLE İLGİLİ FARKLILIKLARIN TESPİTİNDE MEVCUT ÇALIŞMALAR.....	19
3. VERİ İNDİRGEME VE KÜMELEME ANALİZİ.....	25
3.1. TEMEL BİLEŞENLER ANALİZİ	25

3.1.1. Temel bileşenler analizi için yararlanılan matrisler	27
3.2. KÜMELEME ANALİZİ	27
3.2.1. Yaygın kümeleme analizi yöntemleri.....	29
4. HİBRİD (HYBRİD) KÜMELEME ANALİZİ	30
4.1. AŞAMALI KÜMELEME YÖNTEMİ.....	30
4.1.1. Aşamalı kümelemede birimleri birleştirme yöntemleri.....	31
4.2. AŞAMALI OLMAYAN KÜMELEME YÖNTEMİ.....	32
4.2.1. K-ortalamalar yöntemi (Mac Quennn's K-Means Method).....	33
4.3. HİBRİD KÜMELEME	34
4.3.1. Hibrid kümeleme algoritması.....	34
4.4. KÜME GEÇERLİLİK ÖLÇÜLERİ	36
4.4.1. Silhouette Genişliği.....	36
4.4.2. Davies–Bouldin İndeksi	36
5. MATERYAL VE YÖNTEM	38
5.1. VERİLER	38
5.2. YÖNTEM.....	40
6. ARAŞTIRMA BULGULARI	43
6.1. VERİLERİN ÖN İNCELEMESİ.....	43
6.1.1. Tanımlayıcı istatistikler ve verilerin dağılımı.....	43
6.1.2. Değişkenler arasındaki doğrusal ilişkiler.....	45
6.1.3. Verilere uygulanan dönüşümler.....	49
6.1.4. Aykırı gözlemlerin (ilçelerin) belirlenmesi	52
6.2. VERİ İNDİRGEME	52
6.2.1. Aykırı gözlemler üzerinde veri indirgeme	57
6.3. HİYERARŞİK KÜMELEME SONUÇLARI	60
6.3.1. Grup içi ortalama bağlantı yöntemi sonuçları	60
6.3.2. Ward' in yöntemine ilişkin sonuçları	63
6.3.3. Aykırı ilçeler için hiyerarşik kümeleme uygulaması	67
6.4. K-ORTALAMALAR KÜMELEME YÖNTEMİ.....	69
6.4.1. Kullanılan program yardımıyla küme merkezlerini kendi seçmesi.....	69
6.4.2. Başlangıç küme merkezlerinin rastgele seçilmesi.....	73

6.4.3. Verilerden rastgele örneklem seçerek küme merkezlerinin belirlenmesi.	76
6.5. HİBRİD KÜMELEME YÖNTEMİ	79
6.5.1. Hiyerarşik grup içi ortalama bağlantı + K-ortalamlar	79
6.5.2. Hiyerarşik Ward' in yöntemi + K-ortalamlar.....	82
6.6. KÜMELEME SONUÇLARININ KARŞILAŞTIRILMASI	85
6.6.1. Amaç fonksiyonu kullanılarak karşılaştırma	85
6.6.2. Ortalama silhouette genişliği kullanılarak karşılaştırma.....	87
6.6.3. Davies-Bouldin indeksi kullanılarak karşılaştırma	88
6.7. BÖLGELERİN İLÇELERİN GELİŞMİŞLİK DÜZEYİNE GÖRE İNCELENMESİ	90
7. SONUÇLAR VE TARTIŞMA	94
KAYNAKLAR	97
EKLER.....	100
EK-1 DEĞİŞKENLERE İLİŞKİN ELDE EDİLEN VERİLER	101
EK-2 AMAÇ FONKSİYONU DEĞERLERİNİ BULAN MİNİTAB 14 MAKRO PROGRAMI	102
EK-3 ORTALAMA SİLHOUETTE GENİŞLİĞİNİ HESAPLAYAN MS-DOS QBASIC KODLARI	103
EK-4 DAVİES-BOULDİN İNDEKSİNİ HESAPLAYAN MİNİTAB 14 MAKRO PROGRAMI	105
EK-5 SEÇİLEN TEMEL BİLEŞENLERE AİT VERİLER	107
EK-6 SOSYO-EKONOMİK GELİŞME DÜZEYLERİNE GÖRE İLÇELERİN KÜMELENMİŞ LİSTESİ	108
ÖZGEÇMİŞ	109

HİBRİD VE İNDİRGENMİŞ KÜMELEME ANALİZİ İLE TÜRKİYE' DEKİ İLÇELERİN SOSYO-EKONOMİK ÖZELLİKLERİNE GÖRE

SINIFLANDIRILMASI

(Yüksek Lisans Tezi)

Gökhan SAYIN

MUĞLA ÜNİVERSİTESİ

FEN BİLİMLER ENSTİTÜSÜ

2010

ÖZET

Yerel bazda sosyo-ekonomik gelişmişlik düzeyinin ölçülmesi veya belirlenmesi, ilgili yerel yöre için idari, mali ve sosyal açılardan daha doğru kararlar verilmesini, daha gerçekçi projeler geliştirilmesini ve kaynakların daha iyi dağıtılmasını sağlayabilir.

Kümeleme işlemi sosyo-ekonomik gelişmenin belirlenmesinde kullanılabilir bir araçtır. Aşamalı ve K-ortalamar kümeleme algoritmaları yaygın olarak kullanılan kümeleme araçlarından ikisidir. Aşamalı kümelemede algoritmanın hangi adımda durması gerektiğine dair kesin bir ölçüt olmaması ve bir kez bir kümeye atanan nesnenin, kümelemeyi iyileştirmek adına başka bir kümeye atanamaması dezavantajlar olarak görülmektedir. Öte yandan, K-ortalamar kümeleme algoritması önceden kümelerin sayısının belirlenmesini gerektirir, aykırı değerlere duyarlıdır ve sonuç başlangıç küme merkezlerine bağlıdır. Dolayısı ile yukarıda sözü edilen dezavantajların bazılarını kurtulmak için aşamalıdan elde edilen küme merkezlerinin K-ortalamar algoritmasında kullanıldığı bir hibrid yaklaşımı uygulanabilmektedir.

Bu çalışmada ilçelerin sosyo-ekonomik gelişmişlik seviyelerini tespit edilmesi ve bu tespit doğrultusunda gelişmişlik sıralamasının yapılması amaçlanmıştır. Ayrıca, yukarıda tanımlanan Hibrid Kümeleme Yönteminin tanıtımının, uygulamasının yapılması ve bulunan sonuçlar çerçevesinde Hibrid algoritmasının faydalarının belirtilmesi hedeflenmiştir.

İlçelerle ilgili TÜİK' in veri tabanında, 2008 yılına ilişkin ekonomik ve sosyal alanlardan seçilen demografik, eğitim, medeni durum ve tarım göstergelerinden oluşan veri seti kullanılmıştır. İlgili yıl itibariyle, mevcut idari yapı temelinde 81 ile bağlı 957 ilçe (hakkında ilgilenilen değişkenler bazında çoğunlukla veri bulunmayan 12 ilçe hariç) kapsanmış ve bu ilçelere ilişkin sosyal ve ekonomik göstergeler arasından seçilen 23 adet değişken dikkate alınmıştır. Değişken sayısı verilerin ön incelemesi sonucu 20' ye indirilmiştir.

Bu çalışmada verilere ilk önce değişken sayısını indirgemek amacıyla temel bileşenler analizi uygulanmıştır. Daha sonra ise yaklaşık %81' lik açıklama oranına sahip 5 temel bileşen üzerinden aşamalı, K-ortalamalar ve hibrid kümeleme algoritmaları Minitab 14 ve Spss 15 paket programları kullanılarak uygulanmıştır. Her bir algoritma için küme sayısı olarak K=4, 5, 6 ve 7 alınmıştır. Elde edilen kümeleme sonuçları amaç fonksiyonunun değeri, ortalama silhouette genişliği ve Davies-Bouldin İndeksi açısından karşılaştırılmıştır.

Karşılaştırma sonuçlarına göre, K=6 ile grup içi bağlantı yöntemiyle aşamalı kümeleme ve K-ortalamalar yöntemini kullanan Hibrid algoritması sonuçlarının en iyi olduğu kararına varılmıştır. Bu sonuçlara göre Türkiye' deki coğrafi bölgelerin gelişmişlik sıralaması yapılmıştır.

Anahtar Kelimeler: Sosyo-ekonomik gelişme, İlçe, Hiyerarşik kümeleme, K-ortalama, Hibrid kümeleme yöntemi

Sayfa âdeti: 109

Tez yöneticisi: Öğr. Gör. Dr. Mehmet KARAHASAN

**CLASSIFYING THE COUNTIES OF TURKEY ACCORDING TO THEIR
SOCIO-ECONOMIC CHARACTERISTICS BY HYBRID AND REDUCED
CLUSTERING ANALYSIS**

(M. Sc. Thesis)

Gökhan SAYIN

**MUĞLA UNIVERSITY
INSTITUTE of SCIENCE and TECHNOLOGY**

2010

ABSTRACT

Measuring or determining level of socio-economic development on a local basis can enable better decisions to be made, more realistic projects to be developed, and resources to be allocated more efficiently for that area in terms of administrative, financial, and social aspects.

The process of clustering is a tool which can be used in determining socio-economic development. Hierarchical and K-means clustering algorithms are two of widely used clustering algorithms. It is considered as disadvantages of hierarchical clustering that there is no certain criteria about in which step to stop algorithm and that once an object has been assigned to a cluster, it cannot be reassigned any other cluster in the name of improving clustering. On the other hand, K-means algorithm requires to determine the number of clusters in advance, and it is sensitive to outliers, and the results depend on initial cluster centers. Thus, in order to eliminate some of disadvantages mentioned above, a hybrid approach in which the cluster centers obtained from hierarchical algorithm are used in K-means algorithm can be applied.

In this study, it is aimed first to determine socio-economic levels of the counties in Turkey, and then to make ranking of them according to the socio-

economic status determined. In addition, it is objective of this study to introduce and apply the hybrid clustering method described above and indicate the benefits of the hybrid algorithm in the context of the results obtained.

A data set consisting of demographic, education, civil status and agriculture indicators chosen from economic and social domain for the year of 2008 was used in the database of Turkish Statistics Institution for the counties. On the basis of the administrative structure of Turkey as of the year, 957 counties of 81 provinces in Turkey (except for the 12 counties about which data were not available) were covered and 23 variables for the counties were taken into consideration. As a result of preliminary data analysis the number of variables was reduced to 20.

In the study, first principal components analysis was applied to the data in order to reduce the number of variables. By using statistical packages Minitab 14 and SPSS15, hierarchical, k-means and hybrid clustering algorithm were applied to five principal components selected. The number of clusters was taken as $K=4, 5, 6$ and 7 for each of the algorithms. The results of clustering were compared in terms of values of the objective function, average silhouette width and Davies-Bouldin index.

According to results of the comparison, it was concluded that the results of the hybrid algorithm with $K=6$ which uses hierarchial algorithm with within average linkage and k-means algorithm together are best. Then, based on the results, the ranking of status of development for the geographic regions of Turkey was made.

Key Words: Socio-economic development, Counties, Hierarchical clustering, K-means, Hybrid cluster analysis

Page number: 109

Adviser: Öğr. Gör. Dr. Mehmet KARAHASAN

ŞEKİLLER DİZİNİ

<u>Şekil No</u>	<u>Sayfa No</u>
ŞEKİL 4. 1 İKİ BOYUTLU KÜME ŞEMASI VE AĞAÇ DALI ŞEMASI ÖRNEĞİ	31
ŞEKİL 6. 1 BAZI DEĞİŞKENLERE AİT HİSTOGRAMLAR	46
ŞEKİL 6. 2 BAZI DEĞİŞKENLER ARASINDAKİ SAÇILIM GRAFİKLERİ	47
ŞEKİL 6. 3 DÖNÜŞÜMÜ YAPILMIŞ BAZI DEĞİŞKENLERE AİT HİSTOGRAMLAR.....	51
ŞEKİL 6. 4 TEMEL BİLEŞENLERE İLİŞKİN SCREE PLOT.....	54
ŞEKİL 6. 5 AYKIRI GÖZLEMLER İÇİN BİLEŞENLERE AİT SCREE PLOT.....	58
ŞEKİL 6. 6 KÜMELEME YÖNTEMLERİNE GÖRE AMAÇ FONKSİYONU DEĞERLERİNİN ÇİZGİ GRAFİKLERİ	86
ŞEKİL 6. 7 KÜMELEME YÖNTEMLERİNE GÖRE ORTALAMA SİLHOETTE GENİŞLİĞİ DEĞERLERİNİN ÇİZGİ GRAFİKLERİ.....	88
ŞEKİL 6. 8 KÜMELEME YÖNTEMLERİNE GÖRE DAVİES-BOULDİN İNDEKS DEĞERLERİNİN ÇİZGİ GRAFİKLERİ.....	89
ŞEKİL 6. 9 TÜRKİYE' DEKİ COĞRAFİ BÖLGELERE GÖRE İLÇELERİN GELİŞMİŞLİK DÜZEYLERİNİN DAĞILIMININ ÇUBUK GRAFİĞİ.....	92

ÇİZELGELER DİZİNİ

<u>Çizelge No</u>	<u>Sayfa No</u>
ÇİZELGE 2. 1 UNDP' NİN 2008 YILI İNSANİ GELİŞİM GÖSTERGELERİNDE TÜRKİYE' NİN 177 ÜLKE ARASINDAKİ SIRALAMASI	15
ÇİZELGE 2. 2 SOSYO-EKONOMİK GELİŞİMİŞLİĞİNİ YANSITAN DEĞİŞKENLER (DPT).....	17
ÇİZELGE 5. 1 ÇALIŞMADA KULLANILAN DEĞİŞKENLER	39
ÇİZELGE 6. 1 VERİLERİN DAĞILIMINA İLİŞKİN TANIMLAYICI ÖZELLİKLER	44
ÇİZELGE 6. 2 KORELASYON MATRİSİ.....	48
ÇİZELGE 6. 3 KORELASYON MATRİSİNDEKİ KISALTMALARIN AÇIKLAMALARI.....	49
ÇİZELGE 6. 4 DÖNÜŞÜM YAPILAN DEĞİŞKENLERİN LİSTESİ	50
ÇİZELGE 6. 5 AYKIRI İLÇELER TABLOSU	53
ÇİZELGE 6. 6 TEMEL BİLEŞENLER ANALİZİ SONUCU	54
ÇİZELGE 6. 7 DEĞİŞKENLERE AİT VARYANS AÇIKLANMA ORANLARI.....	55
ÇİZELGE 6. 8 DEĞİŞKENLERE AİT DÖNDÜRÜLMÜŞ BİLEŞEN MATRİSİ	56
ÇİZELGE 6. 9 TEMEL BİLEŞENLERİN ANLAMI	56
ÇİZELGE 6. 10 BİLEŞENLER KATSAYI MATRİSİ.....	57
ÇİZELGE 6. 11 AYKIRI GÖZLEMLER İÇİN BİLEŞENLERE AİT AÇIKLAMA YÜZDELERİ	58
ÇİZELGE 6. 12 AYKIRI GÖZLEMLER İÇİN DEĞİŞKENLERE AİT VARYANS AÇIKLANMA ORANI.....	59
ÇİZELGE 6. 13 GRUP İÇİ ORTALAMA BAĞLANTI YÖNTEMİNE GÖRE KÜME BÜYÜKLÜKLERİ VE KÜME MERKEZLERİ	61
ÇİZELGE 6. 14 GRUP İÇİ ORTALAMA BAĞLANTI YÖNTEMİNE GÖRE KÜMELER ARASI ÖKLİD UZAKLIKLARI	62
ÇİZELGE 6. 15 WARD YÖNTEMİNE GÖRE KÜME BÜYÜKLÜKLERİ VE KÜME MERKEZLERİ	64
ÇİZELGE 6. 16 WARD YÖNTEMİNE GÖRE KÜMELER ARASI ÖKLİD UZAKLIKLARI	66
ÇİZELGE 6. 17 AYKIRI İLÇELER İÇİN GRUP İÇİ ORTALAMA BAĞLANTI YÖNTEMİNE GÖRE KÜME BÜYÜKLÜKLERİ VE KÜME MERKEZLERİ.....	67
ÇİZELGE 6. 18 AYKIRI İLÇELER İÇİN KÜMELEME SONUÇLARI	68

ÇİZELGE 6. 19 SPSS 15 PROGRAMININ BAŞLANGIÇ KÜME MERKEZLERİNİ KENDİ SEÇTİĞİ K-ORTALAMALAR KÜMELEME YÖNTEMİNE GÖRE KÜME BÜYÜKLÜKLERİ VE KÜME MERKEZLERİ.....	70
ÇİZELGE 6. 20 SPSS 15 PROGRAMININ BAŞLANGIÇ KÜME MERKEZLERİNİ KENDİ SEÇTİĞİ K-ORTALAMALAR KÜMELEME YÖNTEMİNE GÖRE KÜMELER ARASI ÖKLİD UZAKLIKLARI	72
ÇİZELGE 6. 21 BAŞLANGIÇ KÜME MERKEZLERİNİN RASTGELE SEÇİLDİĞİ K- ORTALAMALAR KÜMELEME YÖNTEMİNE GÖRE KÜME BÜYÜKLÜKLERİ VE KÜME MERKEZLERİ.....	74
ÇİZELGE 6. 22 BAŞLANGIÇ KÜME MERKEZLERİNİN RASTGELE SEÇİLDİĞİ K- ORTALAMALAR KÜMELEME YÖNTEMİNE GÖRE KÜMELER ARASI ÖKLİD UZAKLIKLARI	75
ÇİZELGE 6. 23 VERİLERDEN RASTGELE ÖRNEKLEM SEÇEREK BAŞLANGIÇ KÜME MERKEZLERİNİN BELİRLENDİĞİ K-ORTALAMALAR KÜMELEME YÖNTEMİNE GÖRE KÜME BÜYÜKLÜKLERİ VE KÜME MERKEZLERİ.....	77
ÇİZELGE 6. 24 VERİLERDEN RASTGELE ÖRNEKLEM SEÇEREK BAŞLANGIÇ KÜME MERKEZLERİNİN BELİRLENDİĞİ K-ORTALAMALAR KÜMELEME YÖNTEMİNE GÖRE KÜMELER ARASI ÖKLİD UZAKLIKLARI.....	78
ÇİZELGE 6. 25 HİYERARŞİK GRUP İÇİ ORTALAMA BAĞLANTI + K-ORTALAMA İLE HİBRİD KÜMELEME YÖNTEMİNE GÖRE KÜME BÜYÜKLÜKLERİ VE KÜME MERKEZLERİ	80
ÇİZELGE 6. 26 HİYERARŞİK GRUP İÇİ ORTALAMA BAĞLANTI + K-ORTALAMA İLE HİBRİD KÜMELEME YÖNTEMİNE GÖRE KÜMELER ARASI ÖKLİD UZAKLIKLARI.....	81
ÇİZELGE 6. 27 HİYERARŞİK WARD BAĞLANTI + K-ORTALAMA İLE HİBRİD KÜMELEME YÖNTEMİNE GÖRE KÜME BÜYÜKLÜKLERİ VE KÜME MERKEZLERİ.....	83
ÇİZELGE 6. 28 HİYERARŞİK WARD BAĞLANTI + K-ORTALAMA İLE HİBRİD KÜMELEME YÖNTEMİNE GÖRE KÜMELER ARASI ÖKLİD UZAKLILARI.....	84
ÇİZELGE 6. 29 KÜMELEME YÖNTEMLERİNE GÖRE AMAÇ FONKSİYONU DEĞERLERİ.....	85
ÇİZELGE 6. 30 KÜMELEME YÖNTEMLERİNE GÖRE ORTALAMA SİLHOUETTE GENİŞLİĞİ DEĞERLERİ	87
ÇİZELGE 6. 31 KÜMELEME YÖNTEMLERİNE GÖRE DAVİES-BOULDİN İNDEKSİ DEĞERLERİ	89

ÇİZELGE 6. 32 TÜRKİYE’DEKİ COĞRAFİ BÖLGELERE GÖRE İLÇELERİN GELİŞMİŞLİK DÜZEYLERİNİN DAĞILIMI.....	91
ÇİZELGE 6. 33 COĞRAFİ BÖLGELERE GÖRE İLÇELERİN GELİŞMİŞLİK DÜZEYİ FARKLILIKLARININ ANLAMLILIK TESTLERİ.....	93

SEMBOLLER DİZİNİ

AB	Avrupa Birliđi
ABD	Amerika Birleşik Devletleri
ADNKS	Adrese Dayalı Nüfus Kayıt Sistemi
BYKP	Beş Yıllık Kalkınma Planı
ÇBÖ	Çok Boyutlu Ölçekleme
DAP	Dođu Anadolu Projesi
DB	Davies - Bouldin
DNA	Deoksiribonükleik Asit
DOKAP	Dođu Karadeniz Bölgesel Gelişme Planı
DPT	Devlet Planlama Teşkilatı
FCA	Fuzzy Cluster Analysis
GAP	Güneydođu Anadolu Projesi
GBM	Geleneksel Bölge Merkezleri
GRASP	Greedy Randomized Adaptive Search Prosedure
GSMH	Gayri Safi Milli Hâsıla
GSYİH	Gayri Safi Yurtiçi Hâsıla
HDI	Human Development Index
İGE	İnsani Gelişim İndeksi
KA	Kümeleme Analizi
KANN	Kohonen Artificial Neural Networks
KMO	Kaiser-Meyer-Olkin Measure of Sampling Adequacy
KOB	Katılım Ortaklığı Belgesi
KOBİ	Küçük ve Orta Büyüklükteki İşletmeler
KOSGEB	Küçük ve Orta İşletmeleri Geliştirme ve Destekleme İdaresi Başkanlığı
KÖY	Kalkınmada Öncelikli Yöre
KSS	Küçük Sanayi Siteleri
OSB	Organize Sanayi Bölgeleri
PSA	Particle Swarm Optimization
SOM	Self-Organizing Map
SPSS	Statistical Package for the Social Sciences

TBA	Temel Bileşenler Analizi
TÜİK	Türkiye İstatistik Kurumu
UNDP	United Nation Development Programme
YSO	Yeni Sanayi Odaları

1. GİRİŞ

Sosyo-ekonomik gelişme bir toplumun demografik yapısının, kırsal nüfusun, doğum ve ölüm oranlarının, yatay ve dikey hareketliliğin, refah düzeyinin, şehirleşme hızının, aile büyüklüğünün, gelir dağılımının ekonomik gelişmeye paralel değişikliklere uğraması olarak tanımlanmaktadır. Söz konusu değişimin bölgesel bazda geliştirilmesi ülkelerin planlama ve yatırım projelerinde önem kazanmaktadır.

Belediyeçilik anlayışının gelişmesiyle sosyo-ekonomik gelişme bölgesel yapıdan yerel yapıya kaymaktadır. Yapısal olarak il yönetim sisteminin temel parçası olan ilçeler, kamu hizmetlerinin halka götürülmesinde ve ülke genelinde sosyal ve ekonomik yaşamın geliştirilmesinde son derece önemli bir işleve sahip olmuştur. Bu nedenle Türk yönetim ve planlama sisteminde ilçelere özel bir önem verilmiştir. Örneğin Birinci BYKP' nda ilçeler, hizmetlerin halka götürüldüğü temel hizmet birimleri olarak kabul edilmiştir. İlçeler, kamu kurumlarının teşkilatlarının bulunduğu, mevcut nüfus, alan, mali kaynaklar, personel, araç açısından sahip olduğu imkânlar nedeniyle mahalli sorunların çözümlendiği toplum kalkınmasının yerel yapı taşları olarak değerlendirilmiştir.

Türk planlama sisteminde, ulusal kaynakların en yüksek ekonomik ve sosyal faydayı sağlayacak şekilde geliştirilmesi ve bölgeler arası dengesizliklerin en aza indirilmesi temel amaçlardan başlıcaları olmuştur. Kalkınma planlarında; bölgesel gelişme politikalarının uygulanmasında; sürdürülebilirlik, bölgeler arası bütünleşme, sosyal ve ekonomik dengelerin sağlanması, yaşam kalitesinin iyileştirilmesi, fırsat eşitliği, kültürel gelişme ve katılımcılık ilkelerinin esas alınacağı önemle belirtilmiştir (Dinçer ve Özaslan, 2004).

Bölgesel dengesizlik sorunu 1950'li yıllardan bu yana sosyal, siyasal ve ekonomik yönlerden çeşitli platformlarda tartışılmıştır. Konunun kapsamlı oluşu kadar, ortaya konulmasındaki güçlükler ve bölgesel düzeyde veri sağlama sorunu bu konuda istatistiksel ve ekonometrik analizler yapılmasını yakın zamana kadar kısıtlamış gibi görünmektedir. Ancak, son yıllarda bu alana yönelik veri oluşturulmaya başlaması ve uluslararası literatürde yeniden canlanmaya başlayan tartışmalar, bölgesel veya yerel kalkınma konusundaki çalışmalarını yeniden özendirilmiş görünmektedir (Öztürk, 2005).

Türkiye coğrafyasında ve mekânlar arasındaki gelişmişlik denkliğini sağlamak maksadı, hem ülke için hem de AB uyum sürecinde olduğumuz şu son zamanlarda son derece önemli görülmektedir. Bu kapsamda ülkenin yıl içi planlama çalışmalarında beşeri ve fiziki kaynakların dağıtımında tutarlı olması, kentsel gelişmenin dikkate alınması kısaca; mekân ve sosyal yapı ile ilişkinin sağlanması önemle belirtilen hedefler arasında yer almaktadır.

Benzer şekilde adaylık aşamasında olduğumuz Avrupa Birliği'nde de bölgesel dengelerin sağlanması konusu önem taşımaktadır. Zira AB, üye ülkelerdeki bölgesel gelişme politikalarına büyük önem vermekte, Birlik üyesi ülkelerdeki bölgesel gelişmişlik farklılıklarını gidermek, mekânsal ve sosyal uyumu sağlamak için bu amaç doğrultusunda oluşturduğu fonlar yoluyla geniş finansal olanaklar sağlamaktadır. Kaynakların yönlendirileceği bölgelerin belirlenmesinde de bölgesel istatistikler önemli bir yer tutmaktadır.

Bu çalışmada, ilçelerin değerlendirilmesinde ekonomik ve sosyal alanlardan seçilen demografik, eğitim, medeni durum ve tarım göstergelerinden oluşan veri seti kullanılmıştır. Mevcut idari yapı temelinde 81 ile bağlı 957 ilçenin kapsandığı araştırmada sosyal ve ekonomik göstergeler arasından seçilen 23 adet değişken kullanılmaktadır. Bu çalışmanın kullandığı göstergeler dışında sosyo-ekonomik gelişmeyi ölçmede kullanılan başka değişkenlere ilişkin verilere de ilçeler bazında ulaşılmaya çalışılmıştır. Ancak TÜİK ve öteki kuruluşların veri tabanlarında güncel olarak ilçeler temelinde detaylı ve geniş bir veri dağarcığına ulaşılamamıştır.

Çalışmanın amaçlarını maddeler halinde sunmak gerekirse:

- İlçelerin sosyo-ekonomik gelişmişlik seviyelerini tespit etmek ve bu tespit doğrultusunda gelişmişlik sıralamalarını yapmak,
- Son zamanlarda bilimsel araştırmalarda uygulanan bir *Hibrid Kümeleme Yöntemi* ile ilgili bilgi sunmak ve bu Hibrid algoritmasının faydalarını ortaya çıkarmak,
- İlgili hibrid algoritması ile oluşturulan grupların, yaygın olarak kullanılan kümeleme yöntemlerine göre oluşturulan gruplara göre daha iyi sonuçlar verip vermediğini saptamak.

Hibridleme yöntemi en az iki yöntemin uygun biçimde sentezlenerek ortaya çıkması olarak tanımlanmaktadır. Başlıca özelliği mevcut yöntemlerin özünde olan

dezavantajlarını azaltmak ya da mümkünse yok etmek için diğer bir yöntem(ler)den faydalanmaya çalışmasıdır.

Bu araştırmada aşamalı ve K-ortalama kümeleme yöntemleri hibridlemeye çalışılmaktadır. Aşamalı ve K-ortalama kümeleme iki ana analiz aracıdır. Bununla birlikte, her ikisi de tabii olan dezavantajlara sahiptir. Aşamalı kümelemede algoritmanın hangi adımda durması gerektiğine dair kesin bir ölçüt olmamakta ve bir kez bir kümeye atanan nesne, kümelemeyi iyileştirmek adına başka bir kümeye atanmamaktadır. K-ortalama kümeleme algoritması ise, küme sayısının önceden belirlenmiş olmasını gerektirir. Ayrıca K-ortalama algoritmasının sonuçları başlangıç küme merkezlerinden etkilenmektedir. Aynı zamanda, bu yöntem aykırı değerlerde duyarlıdır. Dolayısıyla yukarıda sözü edilen bazı dezavantajlardan kurtulmak ve ikisinin avantajlarını kullanabilmek adına bir hibrid yaklaşımı ele alınacaktır.

Çalışma şu bölümlerden oluşmaktadır:

2. Bölümde: büyüme, gelişme ve kalkınma ile ilgili temel kavramlar tartışılmış, geçmişten günümüze kadar olan gelişmelere değinilmiş ve bu kapsamda bölgesel gelişme politikaları ve bölgesel gelişmişlik farklılıkları özetlenmiştir. Ayrıca uluslararası kuruluşlar tarafından Türkiye'nin dünyadaki yeri üzerinde durulmuştur. Ek olarak literatürde sosyo-ekonomik gelişme farklılıklarını belirlerken yapılan çalışmalara yer verilmiştir.

Sosyo-ekonomik gelişme ile ilgili temel kavramlar verildikten sonra, 3. Bölümde analiz aşamasında kullanılan yöntemlere giriş yapılmıştır. Bu amaçla ilk olarak veri indirgemesini yardımcı olan Temel Bileşenler Analizi hakkında genel bilgi verilmiş ve ardından da kullanılan kümeleme yöntemlerine değinilmiştir.

4. Bölümde ise çalışmamızın asıl amaçlarından olan Karışım (Hibrid) Kümeleme hakkında tanımlayıcı, teknik bilgilere yer verilmiş ve hibridlemede kullanılacak yöntemlerin neler olduğu hakkında bilgiler sunulmuştur. Ek olarak bir Hibrid kümeleme algoritmasını kullanmış bilimsel araştırmalara değinilmiş ve oluşturulan kümelerin geçerliliğini test etmekte kullanılan yöntemler hakkında bilgi verilmiştir.

Analize geçmeden önceki 5. Bölümde ise veriler hakkında şu bilgiler ortaya konmuştur: kullanılan ilçelerin boyutu, analizde kullanılacak değişkenlerin neler

olduđu ve neden arařtırmaya dâhil edildiđi hakkında genel bilgi. Ayrıca, kullanılan yöntemlerin nasıl Hibrid kümeleme için sentezleneceđine yer verilmiřtir.

“Arařtırma Bulguları” adını taşıyan 6. Bölümde, verileri tanımak maksadı ile tanımlayıcı istatistiklere yer verilmiř ve bu veriler üzerindeki yapılan indirgemenin sonuçları gösterilmiřtir. Son ařama olarak, elde edilen verilere birer birer řu kümeleme yöntemler uygulanmıřtır:

- Hiyerarřik kümeleme
- K-ortalamlar kümeleme
- Karıřım (Hibrid) kümeleme

Son bölümde ise arařtırma sonucu elde edilen bulgular deđerlendirilmiř ve öneriler sıralanmıřtır.

2. KAVRAMSAL ÇERÇEVE ve MEVCUT ÇALIŞMALAR

2.1. Büyüme – Kalkınma – Gelişme ile İlgili Kavramsal Çerçeve

Büyüme, ülke ekonomisinde işgücünün çoğalması, üretim araçları ve GSMH'nin artması vb. genel verilerin yükselmesidir. Diğer bir ifade ile ekonominin fiziksel olarak gövdesel genişliğe uğraması anlamına gelir; *Kalkınma*, bir ülkenin ekonomik, sosyal, kültürel ve teknolojik açıdan ilerlemesi ya da ekonomideki niteliksel gelişmelerdir. Ekonomik kalkınma toplumun yaşam standartlarında, üretilen malların kalitesinde veya üretim organizasyonunda iyileşmeler yaşanan bir ortamı ifade etmektedir; *Gelişme* ise gelişmekte olan ülkenin toplam geliri ile kişi başına gelirin büyüme süreci ve bu arada ülke ekonomisinin yapısal gelişmesi, ekonomik gelişme olarak nitelendirilir. Ekonomik gelişmede temel amaç, genel yaşam düzeyinin yükselmesidir. Ancak bu konuda kullanılan nicel ölçülerin yanında ülke düzeyinde bazı niteliksel dönüşümlerin gerçekleştirilmiş olması da göz önüne alınmaktadır.

Yakın tarihe kadar, “büyüme”, “kalkınma” ve “gelişme” kavramları, denk oluşumları içerdiği düşüncesiyle, eşanlamlarda kullanılmıştır. Ayrıca İkinci Dünya Savaşı'nda yaşanan gelişmelerden sonra, ulusal kalkınma sürecinde iktisadi büyüme olgusunun tek başına yeterli olmadığı, bunun yaşamın diğer alanlarındaki sosyal, kültürel, çevresel ve mekânsal boyutlarla dengelenmesini öne çıkarmıştır. Bir yönüyle “sürdürülebilir gelişme” anlayışının yolunu açan bu gelişmeler kalkınma anlayışını “iktisadi büyüme”den “toplumsal gelişme”ye doğru yönlendirmiştir. Bu tür yaklaşımlar, özellikle 1970'li yıllardan sonra, kalkınma ve modernleşme literatürüne de yansyarak, söz konusu kavramların saydamlaşmasının yolunu açmıştır. Bugüne gelindiğinde, bu kavramların aynı kavramlar olmadığı gerçeği kabul görmekte ve daha farklı anlamlarda kullanılmaktadır.

Kalkınmadaki asıl amaç, üretim ve istihdam yapısını, tarımdan ziyade, sanayi ve hizmetler sektörleri doğrultusunda dönüştürmek olunca, ülke refahındaki değişimlerin temel göstergesi olarak “fert başına ulusal gelir” kullanıldı. Fakat 1960'lardan sonra meydana gelen gelişmeler bu yaklaşımın yetersizliğini vurgulaması ile 1970'lerde kalkınma kavramının yeniden tanımlanmasına ihtiyaç duyulmuştur. Gelişmeyi, insani, sosyal, kültürel, çevresel ve mekânsal boyutlarıyla da tanımlama amacı taşıyan yeni yaklaşım, ekonomik büyüme kavramı yanında; yoksulluk,

işsizlik, gelir dağılımı ve bölgesel dengesizliklerin de gelişme tanımlarının içinde değerlendirilmesini öne çıkarmıştır (Dinçer vd., 2003).

Gelir artışlarıyla ölçülen ekonomik *büyüme*, toplumsal gelişmişliği yeterince açıklayamayan, yalnızca talep genişlemesini açıklayabilen dar kapsamlı bir parametredir. Büyüme kavramı için en belirgin özelliğin “bir ekonominin üretim kapasitesinde, sayısal olarak ölçülebilen genişleme veya miktar artışı” olduğu söylenebilir. Ekonomik kalkınma kavramı, niteliksel değişmeye işaret etmektedir. Bu durumda ekonomik kalkınma hem daha fazla çıktı hem de teknik ve kurumsal yapıdaki değişimleri kapsamaktadır. Büyüme ile bir karşılaştırma yapılarak denebilir ki, ekonomik büyüme daha çok aynı şeydeki basit artış sürecini, ekonomik kalkınma ise daha fazla ve farklı olanın yer aldığı yapısal değişme sürecini ifade eder (Flammang, 1979). Ekonomik kalkınma sadece ekonomik boyutlarla sınırlanmayan, toplumu sosyolojik, psikolojik ve politik tüm boyutlarıyla kuşatan karmaşık bir süreçtir (Yavilioğlu, 2002).

Gelişme kavramı ise zaman içinde farklı anlamlar ifade ede gelmiştir. Gelişme kavramı, 19. yüzyılda “ekonomik büyüme” anlamına gelirken, temel ölçütler milli ya da kişisel gelir, yaratılan katma değer, sanayi sektöründe üretim/çalışan hacmi, vb. olarak tanımlanmaktaydı. 20. yüzyılın ilk çeyreğinde “sosyal refah” içeriğine kavuşan gelişme kavramı, gelir yanında insanların/toplumların sahip oldukları kolaylıkları (fiziksel ve sosyal altyapı, vb.) da kapsar hale gelmiştir. 20. yüzyılın son çeyreğinde ise gelişme “yaşam kalitesi” ile ölçülmeye başlanmıştır. Yaşam kalitesi, özellikle nitelikli doğal, fiziksel, sosyal ve kültürel çevrenin varlığı ve tüketilmesi/tüketilme olanağına kavuşulması anlamını taşımaktadır (Üstümişik, 2007).

Yukarıda bahsedilen tanım ve gelişmeler ile şu anlaşılmaktadır: Gelişme olgusu, ülke genelinde toplumsal, siyasal, kültürel ve benzeri kurumlardaki yapısal değişimleri kapsayan çok boyutlu bir öze sahiptir. Bu anlayışla gelişme; ülkenin ekonomik, sosyal, siyasal ve kültürel yapılarındaki ilerlemeyi kapsamakta ve bir bütün oluşturmaktadır. Belirtilen kapsamda bu çalışmaya konu olan sosyo-ekonomik gelişme olgusu da; kişi başına düşen milli gelirin artırılması şeklinde özetlenebilecek iktisadi büyüme kavramıyla beraber, yapısal ve insani gelişmeyi içine alan ve ölçebildiğimiz bütün sosyal değişkenleri içermektedir (Dinçer vd., 2003).

Özetlemek gerekirse, gelişme için milli gelirin artması yeterli değildir. Bunun yanında örneğin okuryazarlık oranının, sanayinin istihdam ve milli gelirdeki paylarının vb. artması gerekmektedir (Manisalı, 1975). Genel olarak ülkelerin veya coğrafi alanların milli gelirlerindeki artış oranları, büyüme hızı olarak kabul edilmektedir. Fakat bu oranın düşüklüğü ve yüksekliği ülkelerin gelişme veya kalkınma sürecinde hangi aşamada bulunduğu hakkında bilgi vermez. Ülkelerin veya coğrafi alanların kalkınmalarını tamamlayıp tamamlayamadıklarına karar verebilmek için birtakım sosyo-ekonomik faktörlere ve bu faktörlerdeki değişmelere bakmak ve bunları tümünü bir arada düşünerek bir değerlendirme yapmak gerekmektedir (Dinçer vd., 2003).

2.2. Yerel Ekonomik Kalkınma

Yerel ekonomik kalkınma kavramının zaman içinde anlamının değişikliğe uğradığını görüyoruz. Örneğin, Moriarty (1980) yerel ekonomik kalkınmayı, dış sanayi yatırımlarının bir bölgeye çekilmesi olarak tanımlamıştır (Teitz, 1994). Geleneksel olarak da yerel ekonomik kalkınma, önemli ölçüde istihdam olanağı sağlayacak büyük sanayi yatırımlarını bir yöreye çekmeyi amaçlayan teşviklerden oluşan bir strateji olarak görülmüştür. Günümüzde ise yerel ekonomik kalkınma, çok daha profesyonelce yürütülen, yerel ölçekte kamu, özel ve sivil toplum kuruluşları arasında sıkı işbirliğini öngören, yerel ekonominin kendi kapasitesinin daha çok geliştirilmesini hedefleyen, yeni girişimciler yaratarak ve var olan firmaları daha da geliştirerek ekonomik büyümeyi amaçlayan bir politika olarak görülmektedir (Teitz, 1994). Bu süreçte, özellikle yerel kamu otoriteleri, bölgedeki hizmetlerin dağıtımının planlamasında ve eşgüdümünde liderlik görevi üstlenir, yerel stratejik kalkınma planları hazırlar ve genel kamu yararı için zorunlu altyapı yatırımlarını gerçekleştirirler (Gül, 2004; DPT, 2004).

Yerel ekonomik kalkınma planlamasında güdülen genel amaçlar şöyle sıralanabilir (Blakely, 1994):

- Ekonomik refahı ve yaşam kalitesini yükseltmek: Refah ve yaşam kalitesi, ekonomik olduğu kadar, çevresel ve toplumsal bir olgudur. Çevreye ve zamana duyarlılık açısından daha hızlı akan kent-içi ulaşım planlaması, engelliler için daha kullanılabilir kaldırım ve toplu taşıma, çalışan kadınlar için çocuk bakımevleri, daha temiz ve düzenli çevre gibi.

- Kaliteli istihdam olanakları ve özel girişimler yaratabilen, dışarıdan yeni yatırımlar çekebilen, dışa bağımlılığı az, ekonomik ve istihdam yapısında çeşitliliği ve kendi kendine yeterliliği (*'self-sufficiency'*) olan bir ekonomik yapı oluşturmak.

- Küresel ve ulusal ekonomide yaşanan dönüşümlerin yerel ekonomi, işletmeler ve topluluk üzerinde oluşturabileceği olumsuzlukları hafifletmek.

- Kaynakların, mal ve hizmetlerin etkin pazarlanmasına olanak sağlamak.

Bunun yanında, yerel ekonomik kalkınma çabalarının daha ayrıntılı bazı hedeflerden de bahsetmek gerekir (Blakely, 1994).

- Yerel ya da bölgesel ekonominin, insan ve doğal kaynak gizilgücünü daha iyi kullanarak, göreceli ekonomik üstünlüğünü ve rekabet gücünü artırmak;

- Yerel olarak üretilen kendine özgü ürünlerin ve kaynakların daha iyi pazarlanmasını sağlayarak yeni fırsatlar yakalamak ve yerel-bölgesel ekonomiye dışarıdan daha fazla kaynak girmesini sağlamak;

- Yerel halka daha fazla istihdam olanakları sunmak ve uzun dönemli kariyer fırsatları yaratmak;

- Dışlanmış grupların yerel ekonomiye katılımını desteklemek ve yerel ekonomik kalkınma getirilerinden yararlanmada adaleti gözetmek;

- İşletmelerin daha verimli çalışabilmelerini sağlamak ve yeni işletmeleri çekebilmek için kentin altyapısını iyileştirmek ve kentte yaşam kalitesini artırmak.

2.3. Yerel Birimlerin Önemi

Küresel rekabetin yarattığı baskı ortamı; hem ekonomik bir analiz birimi olarak “mekân” ın, hem de yeni birer ekonomik aktör olarak “yerel birimler” in önemini artırmaktadır. Diğer taraftan, yerel girişimcilik, yerel kaynaklar, bilgi ve beceriler stoku ve diğer yerel potansiyellere de bağımlı olarak belirli sektörlerde uzmanlaşma becerisi gösterebilen yöreler, dünya ekonomisi içerisinde karşılaştırmalı bir üstünlük ve beraberinde rekabet avantajı elde edebilmektedirler. Tüm bu gelişmeler; girişimcilik ve özellikle yerel düzeyde ortak girişimcilik (*partnership*) olgularının, ortaya çıkan küresel rekabet ortamında belirleyici bir değişken olarak ekonomik analizde yerini almasını sağlamıştır. Yerel girişimcilik kabiliyeti yüksek yöreler, daha da önemlisi kamu, özel ve gönüllü sektör kuruluşlarından oluşan yerel kurumların belirli stratejiler ve politikalar doğrultusunda organize olabildiği ve ortak

hareket ettiği kentler, diğerleri ile olan rekabette öne çıkmaktadır. Bu nedenle; kamu, özel ve gönüllü sektörlerin etkin bir işbirliği ve ortaklık modelleri çerçevesinde organize olabildiği yörelerimizin yerel ekonomik gelişme sürecini hızlandıracaklarını öngörebiliriz.

Ekonomik istikrarı sağlamak, bölgeler arası gelişmişlik farklarını gidermek ve diğer sosyal adaletçi politikaları uygulamak halen merkezi hükümete düşen önemli görevlerdir. Ancak, değişen ulusal ve küresel ekonomik koşullar karşısında, istihdam yaratarak ve vergilendirilebilir yatırımları teşvik ederek, ekonomik ve sosyal refahı artırmayı hedefleyen ekonomik kalkınma girişimlerini sadece merkezi hükümetlerden beklemek gerçekçi ve geçerli bir yaklaşım niteliğini kaybetmeye başlamıştır (Dulupcu, 2002; Gül, 1998). Artık, ekonomik kalkınma yerel olmak; yereli de içine almak, yani yerel yönetim birimlerini, kenttaşları, yerel oda ve birlikleri, işletmeleri vb. de kapsamak durumundadır. Bu amaç doğrultusunda gelişmiş batı ülkelerinde yerel ekonomik kalkınma konusunda önemli yerel pratik ve teorik bilgi birikimi oluşmuştur. AB, bölgesel ve yerel ekonomik girişimlerin desteklenmesini ilke olarak benimsemekte ve girişimlere kaynak sağlamada, yerel düzeyde ekonomik kalkınma örgütlenmesinin sağlanmış olmasını, önemli koşullardan biri olarak aramaktadır (DPT, 2004).

2.4. Sosyo-Ekonomik Gelişmişlikte Konum Farklılıkları

Elde edilen kaynakların (doğal ya da toplumsal) dağılımı, mutlak olarak eşitlik arz etmediğinden, sosyo-ekonomik gelişmede dengeli bir süreçten geçmemektedir. Böylece ekonomik ve sosyal gelişme konum bakımından farklılık göstermektedir.

Bölge yaşamında meydana gelen bu farklılıklar ile gelişmemiş mekânlardaki halkın gelişmiş alanlara büyük oranda göç etmesine neden olmuştur. Bununla birlikte, bölgesel dengesizlikler birçok ülkede ulusal bir sorun haline gelmiş ve toplumların ekonomik ve siyasi gündemlerinde önemli bir detay olmuştur. 2. Dünya Savaşı sonrasında giderek büyüyen bir sorun haline gelmesiyle beraber, soruna yönelik kamuoyu duyarlılığının da artması, çözüme yönelik kurumsal çalışmaları ve çözüme yönelik kamu müdahalesini arttırmıştır.

Kalkınma sürecinin ilk kısmında, gelişme merkezleri konumundaki büyük kentler, sundukları mal, hizmet ve istihdam olanakları nedeniyle, üretim faktörleri için çekim odakları olmaktadır. Ancak, nüfusun düzensiz ve hızlı birikimi nedeniyle

dışsal yararlar, zamanla negatif dışsallıklara dönüşerek, büyük kentlerde yaşamının maliyetini artırmaktadır. Ülkenin diğer yerleşim birimlerinden bu tür kentsel kutuplara doğru, başta insan gücü göçü olmak üzere, yoğun bir faktör hareketliliği meydana gelmektedir. Bu akım belirli bir aşamada dengelenmediği takdirde, gelişme kutuplarındaki optimal ölçek aşılmakta ve yeni tür metropol sorunları gündeme gelmektedir.

2.5. Sosyo-Ekonomik Gelişme Farklılıklarının Bölgede Meydana Getirdiği Sorunlar

Sosyo-ekonomik gelişmişlik farklılıklarından kaynaklanan yoğun göçler büyük bir nüfus baskısı oluşturarak ülke genelinde tüm mekânsal birimleri etkileyen yerleşme/kentleşme sorunlarına neden olmaktadır. Ülke ölçeğinde yaygınlaşan yerleşme sorunları ise zamanla nitelik değiştirerek iktisadi, toplumsal ve siyasal yapılara da nüfuz etmekte ve derinleşmektedir. Göç dalgalarıyla büyüyen kentlerde beliren; eğitim ve sağlık hizmetleri yetersizlikleri ile arsa ve konut ihtiyacı, su, enerji, altyapı ve genel olarak belediye hizmetlerinin yetersizliği, trafik sıkışıklığı, kalabalık, gürültü ve çevre kirliliği gibi büyük kent sorunları, kamu yatırım ihtiyacını da artırarak, kamu maliyesine ek bir yük getirmektedir. Böylece; bölgelerarası sosyo-ekonomik gelişmişlik farklılıkları, sadece geri kalmış yöreler açısından değil, gelişmiş bölgeler açısından da düzeltilmesi gereken bir sorun olmaktadır.

Bölgelerarası dengesizliğin diğer boyutunda ise, geri kalmış bölgelerin sorunları bulunmaktadır. Meydana gelen göçler sayesinde göç edilen bölgelerde yerel efektif talep hacmini daraltırken, diğer yandan da yerel gelişmenin temel unsurları olan genç iş gücü ve sermayenin bölge dışına akması anlamına gelmektedir. Böylece zaten geri kalmış bölgeyi giderek durgunluğa itmekte ve kaçınılmaz olarak kamu müdahalesini gerekli kılmaktadır. Böylece, her iki durumda da kamu müdahalesi zorunlu olmaktadır. Bir yandan büyük kentlerin altyapı tesislerinin yenilenmesi ve genişletilmesi, diğer yandan geri kalmış bölgelerde gelişmeyi başlatmak ve hızlandırmak için kamu yatırımlarını artırmak, toplumsal sorumluluk haline gelmektedir.

Kısaca, ülkemizde bölgelerarası gelir dağılımı eşitsizlikleri giderek büyümektedir. Kent-kır arasında farklılaşma büyümekte, kırsal alanlarda yaşam

standartı hızla düşmektedir. Artan nüfus kent merkezlerine itilmekte, bu merkezlerin aldığı yüksek göçler nedeniyle hızlı ve kontrolsüz biçimde büyüyen bu yerleşme yerlerinde en temel kentsel hizmetler (ulaşım, su, kanalizasyon vb.) dahi karşılanamamaktadır. Gelir dağılımı eşitsizliklerinin (yoksulluğun) artması gerilimi artırmakta, işsizliğin yüksek olduğu bu mekânlarda suç işleme oranları artış göstermektedir (DPT, 2004).

2.6. Bölgesel Gelişme Farklılıklarını Giderme Çalışmaları

Bu kapsamda devlet, alternatif gelişme stratejileri sayesinde dengeli bir yerleşme ve doğal olarak dengeli bir kalkınma politikasını benimsemiştir. Gelişme farklılıklarını giderme amacıyla, gelişmiş ülkeler sosyal adalete dönük politikalara yön verirken, gelişmekte olan ülkeler kaynak yaratıcı alanlara birikimlerini yönlendirerek, yatırımları ve milli geliri artırmaya çalışmaktadır.

Geri kalmış yörelere, bu yörelerden elde edilen bütçe gelirin çok üzerinde harcama yapmak, planlı dönemin yatırım ve harcama politikalarının temel özelliklerinden biri olmuştur. Kamu yatırım politikaları yanında, özel sektörü bu yörelere çekebilmek için yapılan devlet yardımları ve uygulanan personel politikaları da, bölgeler arası dengesizliği gidermede kullanılan başlıca kamu araçları olmuştur.

Kamu sektörü açısından, dengeli bir bölgesel gelişmeden amaç, üretim ve hizmet yatırımlarının dağılımını en iyi yansıtan bir yerleşme sisteminin kurulması olmalıdır. Ayrıca, bu tür bir mekânsal dağılım deseni; nüfus ve gelirin dengeli dağılımı kadar, kaynakların sürdürülebilir kalkınmayı gerçekleştirmek amacıyla dengeli kullanımına da en iyi zemini hazırlamaktadır. Böylece, ihtiyaçlar ile kaynaklar arasında uzun süreli denge kurulması olarak da nitelendirebileceğimiz istikrarlı kalkınma, ulusal boyutun yanı sıra bölgesel düzeyde de önem kazanmaktadır.

Sosyal ve ekonomik dengesizliği gidermede devlet iktisadi rasyonellerinin yanı sıra toplumsal rasyonellerle de kaynak dağılımına yön verme ihtiyacı duymaktadır. Bu kapsamda özel sektör yatırımlarını teşvik etmektedir.

2.7. Türkiye’ de Uygulanan Başlıca Bölgesel Gelişme Politikaları

1960 yılından günümüze uzanan planlı dönemde ise planlama, imar planlamasının dar kapsamından çıkarılarak; fiziksel, sosyal ve ekonomik boyutların

da içerildiği bütüncül bir yaklaşım çerçevesinde değerlendirilmeye başlanmıştır. Hazırlanan tüm kalkınma planlarında bölgeler arası dengesizliklerin giderilmesi öncelikli alanlardan biri olmuştur.

AB ile ilişkiler sürecinde, Avrupa Komisyonu tarafından yayınlanan 2002 yılı İlerleme Raporu ve 2003 yılı Katılım Ortaklığı Belgesi'nde (KOB) yer alan Ulusal Kalkınma Planı'nın hazırlık çalışmalarında, Türkiye'nin ulusal ve bölgesel gelişme stratejilerini yeniden belirlemesi gerektiği belirtilmektedir. AB' ye katılım öncesi işbirliği çerçevesinde ve 2004 - 2006 döneminde bölgesel farklılıkların giderilmesi ve uzun vadeli stratejilerin oluşturulması hedefleri doğrultusunda, bölgesel gelişme fonlarından ve uyum fonlarından sağlanacak desteklerden yararlanılmasına yönelik olarak, Türkiye' nin bölgesel gelişme faaliyetlerini yönlendirmesi beklenen İl Gelişme Planları Başbakanlık Devlet Planlama Teşkilatı Müsteşarlığı'nın koordinasyonunda sürdürülmektedir.

Ülkemizde bölgesel gelişmenin hızlandırılması ve bölgeler arası dengesizliklerin giderilmesinde temel araç olan bölgesel gelişme planlarının yanı sıra; yatırım teşvikleri, kalkınmada öncelikli yöre (KÖY) politikaları, organize sanayi bölgeleri (OSB) ve küçük sanayi siteleri (KSS) ile kırsal kalkınma projeleri gibi çeşitli kalkınma araçları kullanılmıştır.

Bunların en başında gelen *Bölgesel Gelişme Planları* hem kalkınma planlarının sektörel öncelikleri ile mekânsal boyutların bütünleştirilmesi, hem de bölgeler arası gelişmişlik farklılıklarını azaltmak ve sürdürülebilir kalkınmayı gerçekleştirmek amacıyla hazırlanmıştır. Çeşitli dönemlerde hazırlanan bölgesel gelişme planlarının başlıcaları: Doğu Marmara Planlama Projesi, Antalya Projesi, Çukurova Bölgesi Projesi, Güneydoğu Anadolu Projesi (GAP), Zonguldak-Bartın-Karabük Bölgesel Gelişme Projesi, Doğu Anadolu Projesi (DAP), Doğu Karadeniz Bölgesel Gelişme Planı (DOKAP) ve Yeşilirmak Havza Gelişim Projesi'dir. Ancak son yıllara kadar anılan bölgesel kalkınma projelerinden GAP dışında kalanlar kapsamlı bir uygulama şansı bulamamıştır. Ayrıca, AB ile uyum sürecinin hızlandığı son yıllarda AB finansman desteği ile yürürlüğe konulmakta olan bölgesel kalkınma programlarının hazırlıklarına da başlanmıştır.

Diğer bir bölgesel gelişme aracı olan *Yatırım Teşvikleri* uygulamaları ise, Cumhuriyet'in başından beri bir araç olarak kullanılmıştır. Teşvik politikalarının

amaçlarından biri, ekonomik ve sosyal açılardan geri kalmış yörelerin kalkınmasını hızlandırmak amacıyla bu yörelere ilişkin özel teşvik politikalarının uygulanması olmuştur.

Bu çerçevede, KÖY' ler belirlenerek bu yörelere yönelik özendirici politikalar devreye sokulmuştur. KÖY politikası kapsamında getirilen kamu destekleri sadece teşvik tedbirleriyle sınırlı kalmamış, ek olarak KÖY' lerde çalışanların ücretlerini nispi olarak iyileştiren uygulamalar, tarımsal ve mesleki amaçlı kredi destekleri, Kamu Ortaklığı Fonundan yatırımlara sağlanan finansman desteği, DPT Bütçesinden KÖY' lerdeki yerel idarelerce sürdürülen projelere sağlanan finansman imkânları gibi olanaklarda sunulmuştur.

Türkiye'de hem kentsel ve bölgesel gelişme, hem de sanayi kuruluşları için altyapı geliştirme aracı olarak kullanılan bir diğer mekânsal düzenleme aracı ise; orta ölçekli sanayiler için OSB, küçük ölçekli sanayiler için ise KSS kurmak olmuştur. Planlı dönemin başlangıcından günümüze değin uygulanmakta olan OSB ve KSS politikaları, yerel sermaye birikiminin il dışına çıkmasını engellemede ve yerel ekonomiye transferinde önemli bir kalkınma aracı olarak işlev görmüştür.

Türkiye'de uygulanmış ve uygulanmakta olan *Kırsal Kalkınma Projelerinin* başlıca amaçları ise geri kalmış yörelerde tarımsal faaliyetlerin ve gelirin artırılması yoluyla refahın yükseltilmesi olmuştur. Kırsal Kalkınma Projeleri, Dünya Bankası kredi desteği ile 1970'li yılların sonlarında DPT tarafından başlatılmış ve şu faaliyet alanlarını içermiştir: Tarım ve hayvancılığın geliştirilmesi, sulama, köy yolları yapımı, orman yolları yapımı, içme suyu göletleri, içme suyu sağlanması, tarımsal ve hayvansal üretimin artırılması, ağaçlandırma. Tamamlanan ve devam eden Kırsal Kalkınma Projeleri ise şunlardır: Çorum-Çankırı Kırsal Kalkınma Projesi, Erzurum Kırsal Kalkınma Projesi, Bingöl-Muş Kırsal Kalkınma Projesi (1990-1999), Yozgat Kırsal Kalkınma Projesi (1991-2001) ve Ordu-Giresun Kırsal Kalkınma Projesi (1995-2003).

2.8. İlçelerin Sosyo-Ekonomik Gelişme Düzeylerinin Belirlenmesinin Önemi

Yapılan analiz sonuçlarından görülebileceği gibi ülkemizin en önemli yapısal sorunlarından biride mekânlar arası sosyo-ekonomik gelişmelerin dengesiz oluşudur. 21. y.y. da gelişen bir ülke olmadaki en önemli adımı bu problemi optimum duruma

getirerek atmalıdır. Mevcut sorunların uygun politikalarla giderilmesinde ilk adım, şüphesiz, sağlıklı verilerin elde edilmesidir.

Her ülkenin gelişimi, dinamik ve karmaşık bir süreçtir. Bu süreç, planlı gelişme süreçlerini geç başlatan az gelişmiş ülkelerde çok daha karmaşıktır. Bu ülkelerde sosyal ve coğrafi alanların sosyo-ekonomik gelişimini sağlayacak gelişme yönteminin seçiminin yanında, dengeli ekonomik kalkınmayı sağlayacak kaynak dağılımının araştırılması da zorunludur (Basu, 2000).

Daha önce bu problemin giderilmesi amacıyla yapılmış planların ne kadar faydalı olup olmadığını ve gelişen dünyada ülkemizdeki mekânların ne şekilde etkilendiklerini bilimsel yollarla temin edilmiş olunacaktır.

Zamanla meydana gelen dünyadaki değişmelere paralel Türkiye’deki sosyo-ekonomik değişkenler belirlenmiş ve bu değişkenler aracılığı ile Avrupa Birliği sınıflandırmaları çerçevesinde yapılan “İstatistikî Bölge Birimleri” ne göre sosyo-ekonomik gelişmişlik endeksi sıralaması yapılmıştır.

Ülkemizde mekânlar arasındaki sosyo-ekonomik gelişme düzeyleri en net biçimde görülerek, Metropol olarak nitelendirdiğimiz bazı şehirlerde dahi gelişmemiş yaşam yerlerinin varlığının söz konusu olup olmadığı bu çalışma ile tespit edilmeye çalışılacaktır.

Kentleşmenin giderek hızlandığı ve kentleşme sorunlarının arttığı günümüzde bu mekânlar artık yönetilemez hale gelmişlerdir. Plan fikri olmayan kamu yöneticilerinin, yerel yönetimlerin ve girişimcilerin işi çok zor olacaktır. İlçelerin sosyo-ekonomik açıdan gelişmişlik seviyelerinin belirlenmesi planlama kararlarının alanla bütünleştirilmesine yardımcı olacağı gibi, ilçelerde zaman içinde meydana gelen sosyo-ekonomik gelişimlerin izlenmesi, kamu kaynaklarının rasyonel dağılımı, teşvik sistemi politikalarının belirlenmesi ve kalkınma öncelikli mekânların saptanması gibi birçok teknik konuda karar vericilere yardımcı olacaktır.

2.9. Uluslar Arası Kurumlar Tarafından Yapılan Araştırmalarda Türkiye’ nin Konumu

UNDP insani gelişme kavramı aracılığıyla bazı gelişmekte olan ülkeler için Gayri Safi Milli Hâsıladaki büyüme ile toplumsal refah arasındaki ters orantıya dikkat çekmiştir. Gelişmekte olan ülkelerde hızlı büyümenin gelir dağılımındaki

eşitsizliklere neden olabileceğini vurgulamıştır (UNDP, 2001). UNDP’ nin gelir, eğitim ve sağlık değişkenleri ile yapmış olduğu HDI çalışmasında, Türkiye ile aynı ekonomik büyüklük kümesinde bulunan ülkelerden Arjantin, HDI sıralamasında 35., Güney Kore ise 31. sırada yer almıştır. HDI sıralamasında Türkiye, insani gelişme indeksi sıralamasında 73. sıradaki Kazakistan’ın altında 85. sırada yer almıştır.

Kişi başına geliri Türkiye’den düşük olan ülkelerden Kazakistan, Filipinler, Jamaika ve Peru gibi ülkelerin bile HDI sıralamasında daha üst sıralarda yer alması, bu ülkelerde ekonomik göstergeler ile sosyal göstergeler arasındaki bağıntının daha güçlü olduğu, Türkiye’deki ekonomik büyümenin ise sosyal göstergelere yeteri kadar yansımadağı konusunda ipuçları vermiştir. Aynı şekilde Türkiye HDI sıralamasında bütün Avrupa Birliği (AB) üyeleri ve adayları içinde de alt sırada yer almıştır (Göçer ve Çıracı, 2003).

UNDP 2008 “İnsani Geliştirme Göstergelerine” göre çeşitli alanlarda Türkiye’ nin başka ülkeler arasındaki sıralaması Çizelge 2.1’deki gibidir:

Çizelge 2.1 UNDP’ nin 2008 yılı insani gelişim göstergelerinde Türkiye’ nin 177 ülke arasındaki sıralaması

Ölçütler	Sıra
Genç işsizliği oranı, toplam (15-24 yaş arası iş gücünün yüzdesi olarak) (2006)	10
Genç işsizliği oranı (erkek oranının yüzdesi olarak kadın oranı) (2006)	11
Sağlık alanında kamu harcamaları (GSYİH’ nin yüzdesi olarak) (2004)	41
İnternet kullanıcıları (1000 kişi başına düşen sayı olarak) (2005)	52
Bilim, mühendislik üretim ve inşaat dallarında okuyan yükseköğretim öğrencileri (toplam yükseköğretim öğrencilerinin yüzdesi olarak) (1999-2005)	54
Gençliğin okuryazarlık oranı (15-24 yaş yüzdesi olarak) (2005)	66
Yükseköğretimde brüt okullaşma oranı, kadın (%)	74
Ortaöğretimde net okullaşma oranı (%) (2005)	74
İlk, orta ve yükseköğretimde, erkeklerde birleşik brüt okullaşma oranı (%) (2005)	77
Gençliğin okuryazarlık oranı, kadınlarda (15-24 yaş yüzdesi olarak) (2005)	78
İnsani gelişme endeksi (2005)	84

Çizelge 2.1 UNDP' nin 2008 yılı insani geliştirme göstergelerinde Türkiye' nin sıralaması (devamı)

Ölçütler	Sıra
Gençliğin okuryazarlık oranı (kadın oranının erkek oranına rasyonu) (2005)	85
Eğitim endeksi	104
Eğitim alanında kamu harcamaları (GSYİH' nin yüzdesi olarak) (2002-2005)	105
İlk, orta ve yükseköğretimde, birleşik brüt okullaşma oranı (%) (2005)	108
Yükseköğretimde brüt okullaşma oranı (kadın oranının erkek oranına rasyonu) (2005)	109
Ortaöğretimde brüt okullaşma oranı, kadınlarda (%) (2005)	110
İlk, orta ve yükseköğretimde, kadınlarda birleşik brüt okullaşma oranı (%) (2005)	115
Ortaöğretimde brüt okullaşma oranı (kadın oranının erkek oranına rasyonu) (2005)	138

2.10. Gelişmişlik Farklılıklarının Belirlenmesinde Kullanılan Değişkenler

Dünyada da bölgeler ve yerleşimler arası gelişme farklılıklarını ölçmek üzere çeşitli çalışmalar yürütülmektedir. Ülkeleri gelişmişlik düzeyi açısından değerlendiren Birleşmiş Milletler Kalkınma Programı İnsani Gelişme Raporu'nda kullanılan *insani gelişim endeksi (İGE)*, *kişi başına düşen gelir*, *eğitim düzeyi*, *sağlık (ortalama yaşam süresi)* gibi üç temel alandaki verilerin birlikte değerlendirilmesinden oluşmaktadır.

Avrupa Birliği'nde bölgeler arası gelişmişlik farklılıklarının değerlendirilmesi ve bölgesel politikaların uygulanması açısından önem taşıyan başlıca göstergeler; *satın alma gücü paritesi cinsinden kişi başı gayri safi yurt içi hâsıla*, *demografik yapı*, *işsizlik oranı*, *gayri safi yurt içi hâsılanın ve istihdamın sektörel dağılımı ile diğer ekonomik ve sosyal gelişmeyi yansıtan istatistikî verilerdir*.

Ülkemizde, Avrupa Birliğine uyum sürecinde öngörülen kriterler baz alınarak en son (2003/il – 2004/ilçe) yapılan araştırmada sosyo-ekonomik gelişme olarak kullanılan değişkenler; *demografik*, *istihdam*, *eğitim*, *sağlık*, *üretim*, *mali ve diğer refah* ana başlıkları altında toplanmıştır.

DPT tarafından yapılan son araştırmalarda ülkenin sosyo-ekonomik gelişmişliğini yansıtan değişkenler Çizelge 2.2' de verilmiştir.

Çizelge 2.2 Sosyo-ekonomik gelişmişliğini yansıtan değişkenler (DPT)

DEMOGRAFİK GÖSTERGELERİ	<ul style="list-style-type: none"> • Toplam Nüfus Kişi • Şehirleşme Oranı • Yıllık Ortalama Nüfus Artış Hızı • Nüfus Yoğunluğu • Yıllık Net Göç Oranı • Doğurganlık Oranı • Ortalama Hane halkı Büyüklüğü 	<p>Kişi Yüzde Binde Kişi/Km2 Binde Çocuk sayısı Kişi</p>
İSTİHDAM GÖSTERGELERİ	<ul style="list-style-type: none"> • Tarım İşkolunda Çalışanların Toplam İstihdama Oranı • Sanayi İşkolunda Çalışanların Toplam İstihdama Oranı • Ticaret İşkolunda Çalışanların Toplam İstihdama Oranı • Mali Kurumlar İşkolunda Çalışanların Toplam İstihdama Oranı • Ücretli Çalışanların Toplam İstihdama Oranı • Ücretli Çalışan Kadınların Toplam İstihdama Oranı • İşverenlerin Toplam İstihdama Oranı 	<p>Yüzde Yüzde Yüzde Yüzde Yüzde Yüzde Yüzde</p>
EĞİTİM GÖSTERGELERİ	<ul style="list-style-type: none"> • Okur-Yazar Nüfus Oranı • Okur-Yazar Kadın Nüfusun Toplam Kadın Nüfusa Oranı • Üniversite Bitirenlerin Okul Bitirenlere Oranı • İlkokullar Okullaşma Oranı • Ortaokullar Okullaşma Oranı • Liseler Okullaşma Oranı 	<p>Yüzde Yüzde Yüzde Yüzde Yüzde Yüzde</p>
SAGLIK GÖSTERGELERİ	<ul style="list-style-type: none"> • Bebek Ölüm Oranı • Onbin Kişiye Düşen Hekim Sayısı • Onbin Kişiye Düşen Diş Hekimi Sayısı • Onbin Kişiye Düşen Eczane Sayısı • Onbin Kişiye Düşen Hastane Yatağı Sayısı 	<p>Binde Adet Adet Adet HastaneYatağı</p>
SANAYİ GÖSTERGELERİ	<ul style="list-style-type: none"> • Organize Sanayi Bölgesi Parsel Sayısı • İmalat Sanayi İşyeri Sayısı • İmalat Sanayi Yıllık Çalışanlar Ortalama Sayısı • İmalat Sanayi Kurulu Güç Kapasite Miktarı • Fert Başına İmalat Sanayi Elektrik 	<p>Parsel Adet Adet Beygir Gücü</p>

	<i>Tüketimi</i>	<i>Kws</i>
	• <i>Fert Başına İmalat Sanayi Katma Değer</i>	<i>TL</i>
<i>TARIM GÖSTERGELERİ</i>	• <i>Kırsal Nüfus Başına Tarımsal Üretim Değeri</i>	<i>TL</i>
	• <i>Tarımsal Üretim Değerinin Türkiye İçindeki Payı</i>	<i>Yüzde</i>
<i>İNŞAAT GÖSTERGELERİ</i>	• <i>Daire Sayısı</i>	<i>Adet</i>
	• <i>Şehirsal Nüfus Başına Konut Alanı</i>	<i>M2</i>
<i>MALİ GÖSTERGELER</i>	• <i>Gayri Safi Yurt İçi Hâsıla İçindeki Payı</i>	<i>Yüzde</i>
	• <i>Fert Başına GSYİH</i>	<i>TL</i>
	• <i>Banka Şube Sayısı</i>	<i>Adet</i>
	• <i>Fert Başına Banka Mevduatı</i>	<i>TL</i>
	• <i>Toplam Banka Mevduatı İçindeki Payı</i>	<i>Yüzde</i>
	• <i>Toplam Banka Kredileri İçindeki Payı</i>	<i>Yüzde</i>
	• <i>Kırsal Nüfus Başına Tarımsal Kredi Miktarı</i>	<i>TL</i>
	• <i>Fert Başına Sınai, Ticari ve Turizm Kredileri Miktarı</i>	<i>TL</i>
	• <i>Fert Başına Belediye Giderleri</i>	<i>TL</i>
	• <i>Fert Başına Genel Bütçe Gelirleri</i>	<i>TL</i>
	• <i>Fert Başına Gelir ve Kurumlar Vergisi Miktarı</i>	<i>TL</i>
	• <i>Fert Başına Kamu Yatırımları Miktarı</i>	<i>TL</i>
	• <i>Fert Başına Teşvik Belgeli Yatırım Tutarı</i>	<i>TL</i>
	• <i>Fert Başına İhracat Miktarı</i>	<i>ABD Doları</i>
	• <i>Fert Başına İthalat Miktarı</i>	<i>ABD Doları</i>
<i>ALTYAPI GÖSTERGELERİ</i>	• <i>Kırsal Yerleşmelerde Asfalt Yol Oranı</i>	<i>Yüzde</i>
	• <i>Yeterli İçme suyu Götürülen Nüfus Oranı</i>	<i>Yüzde</i>
	• <i>TCK Asfalt Yol Oranı</i>	<i>Yüzde</i>
<i>DİĞER REFAH GÖSTERGELERİ</i>	• <i>Onbin Kişiye Düşen Özel Otomobil Sayısı</i>	<i>Adet</i>
	• <i>Onbin Kişiye Düşen Motorlu Kara Taşıtı Sayısı</i>	<i>Adet</i>
	• <i>Fert Başına Elektrik Tüketim Miktarı</i>	<i>Kws</i>
	• <i>Fert Başına Telefon Kontör Değeri</i>	<i>Adet</i>
	• <i>Onbin Kişiye Düşen Faks Sayısı</i>	<i>Adet</i>

2.11. Sosyo-Ekonomik Gelişme ile İlgili Farklılıkların Tespitinde Mevcut Çalışmalar

DPT Araştırması (2004): “İlçelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması” çalışmasında temel bileşenler analizi tekniği kullanılmaktadır. 2000 yılındaki idari yapı esas alınarak 81 ilin toplam 872 ilçenin kapsandığı araştırmada; sosyal (demografik, istihdam, eğitim, sağlık, alt yapı, diğer refah) ve ekonomik (imalat, inşaat, tarım, mali) alanlardan seçilen 32 adet değişken kullanılmaktadır. Göstergelerin duyarlılığını artırdığı durumlarda değişkenler; orantılı ya da fert başına düşen değerler olarak, bunun dışında ise ilçeler itibarıyla toplam mutlak büyüklükler olarak kullanılmış ve denge sağlanmaya çalışılmıştır.

Çalışmanın öncelikli amacı, ilçelerin sosyo-ekonomik gelişmişlik seviyelerini tespit etmek ve bu tespit doğrultusunda gelişmişlik sıralamalarını yapmaktır. İkinci amaç ise, aynı veri setini kullanarak, benzer özellikleri taşıyan ilçe gruplarını tespit etmektir.

Çıkan sonuçlara göre, birinci derecede gelişmiş ilçeler kademesindeki ilçeler, çoğunlukla ülkenin batı kesimlerinde yer almakla birlikte, ülke coğrafyasına benek benek dağılmış durumdadır. Bu ilçeler gelişme çizgilerine göre kendi içlerinde üç farklı gruba ayrılmaktadır. İlk grup, İstanbul, İzmir ve Adana gibi geleneksel bölge merkezlerinin (GBM) ardbölgelerinde yer alan ilçelerden oluşmaktadır. Turizm merkezi olarak ortaya çıkan Antalya-Merkez İlçe ise birinci kademe de yer alan ilçeler içerisinde turizme dayalı olarak farklı bir gelişme çizgisi izlemektedir. Birinci kademedeki üçüncü ilçe grubu ise son dönemlerde Yeni Sanayi Odakları (YSO) olarak tanımlanan ilçelerden oluşmaktadır.

İkinci derecede gelişmiş ilçeler kademesi ise ülkenin Batı ve Orta Anadolu kesimlerinde yoğunlaşmıştır. Ülkenin doğu kesimlerinde yer alan ikinci derecede gelişmiş ilçelerin ise, özellikle Merkez ilçelerden oluşmakla birlikte az gelişmiş doğu bölgelerinde gelişme adacıkları olarak ortaya çıktıkları gözlenmiştir. Üçüncü kademedeki ilçeler ise; ülkenin Orta ve Kuzeydoğu Anadolu Bölgelerinde kümelenmiştir. Dördüncü kademe de yer alan ilçeler ise Orta Anadolu'dan Kuzeydoğu Anadolu'ya doğru uzanan bir aksta yoğunlaşmıştır. Beşinci kademedeki ilçeler çoğunluğu Orta ve İç Doğu Anadolu'da olmak üzere, ülke coğrafyasına kısım kısım dağılmıştır. En az gelişmiş ilçe grubu olan altıncı kademe ise, ağırlıklı olarak

Doğu ve Güneydoğu Anadolu Bölgelerindeki ilçelerden oluşmaktadır (Dinçer ve Özaslan, 2004).

Özmen (1998): İl olmaya aday ilçelerin il olabilme şansları araştırılarak halen il olan ilçeler ile karşılaştırılması, sosyo-ekonomik gelişmişliklerine göre sıralanması ve ilçeler arasındaki ilişki yapısının mümkün olduğunca az boyutlu görüntüsü ortaya konularak gelişmişlik düzeylerinin belirlenmesi amaçlanmaktadır. Bu amaçla; 1989 yılından sonra il olan ilçeler, ilçe kapsamında tutularak il olmaya aday ilçelerle birlikte 45 ilçe araştırmada dikkate alınmıştır.

Çalışmada, ilçelerin sosyo-ekonomik yapısal özelliklerini gösterdiği varsayılan 17 değişken belirlenmiştir. Bu değişkenlerin yanı sıra diğer araştırmalarda kullanılmayan ve diğer değişkenler kadar önemli olduğu düşünülen “en yakın ile uzaklık” değişkeni de değerlendirmeye katılarak toplam 18 değişken üzerinden çözümlenmeler yapılmıştır.

Çalışmanın ilk aşamasında, değişkenler arasındaki ilişkinin derecesini ve yönünü belirlemek amacıyla korelasyon matrisi incelenmiştir. İkinci aşamada, değişkenler arasındaki bağımlılık yapısını ortadan kaldırarak, değişkenlerin bileşkesi olarak ifade edilebilecek değişkenler (temel değişkenler) elde etmek amacıyla TBA uygulanmıştır. İlçelerin konumlarını ve ilçeler arası ilişki yapısını mümkün olduğunca az boyutla görüntüsünün ortaya konması açısından Çok Boyutlu Ölçekleme yöntemi kullanılmıştır. Ayrıca ÇBÖ yönteminden elde edilen görüntüye göre ilçelerin gelişmişlik düzeyleri üç grupta toplanarak, üç kümeden oluşan Kümeleme Analizi (KA) sonuçları ile karşılaştırılmıştır.

A. İ. Özdemir ve A. Altıparmak (2005): “Sosyo-ekonomik göstergeler açısından illerin gelişmişlik düzeyinin karşılaştırmalı analizi” adlı çalışmayı, 2000 yılı genel nüfus sayımı verileri baz alınarak yapılmış ve Türkiye’deki 81 ili sosyal ve ekonomik göstergeler açısından analiz edilmiştir. Analizde çok değişkenli istatistik tekniklerinden faktör analizi kullanılmıştır. Sosyal değişkenlere göre yapılan faktör analizi sonucunda, *sağlık göstergeleri*, *eğitim göstergeleri*, *ilk ve orta öğretim okullaşma oranı* şeklinde isimlendirilen üç faktör elde edilmiştir. Bu üç faktör toplam değişimin %75,9’ unu açıklamaktadır. Ekonomik değişkenlerle yapılan faktör analizi sonrasında da, *mali göstergeler*, *imalat sanayi göstergeleri* şeklinde isimlendirilen iki faktör elde edilmiştir. Bu faktörler toplam değişimin %75,18’ini

açıklamaktadır. İllerin, elde edilen bu faktörlerdeki faktör yüklerine göre sosyal ve ekonomik gelişmişlik açısından yerleri belirlenmiştir.

K. Göcer ve H. Cıracı (2003): Türkiye’ de kentlerin sosyal ve ekonomik göstergeleri arasındaki ilişki’ yi araştırmışlardır. Bu çalışmada sosyal ve ekonomik göstergeler arası ilişkilerin illere göre nasıl değiştiğini açıklamak amaçlanmıştır. Sosyo-ekonomik yapıları ortaya çıkarmak için faktör analizi kullanılmıştır. 80 il bazında 30 değişken, 3 faktör altında toplanmıştır. Araştırmada, birinci faktör sonucuna göre; bazı kentler ekonomik olarak daha gelişmiş iken, bazı kentler sosyal olarak daha ileri düzeyde çıkmıştır. İkinci faktör sonucuna göre; sanayi üretiminde verimlilik önem kazanmaktadır. Üçüncü faktör sonucunda ise, GAP bölgesine yapılan yatırımların etkisi görülmüştür.

A. S. Albayrak ve arkadaşları (2004): Temel bileşenler analizi yardımı ile Türkiye’ de coğrafi bölgelere göre illerin sosyo-ekonomik gelişmişlik düzeyleri incelenmiştir. Araştırma coğrafi, nüfus, eğitim ve kültür, sağlık, istihdam, sosyal güvenlik, mali ve finansal, imalat sanayi, tarım, dış ticaret, enerji, konut ve altyapı gibi farklı alanlardan seçilen, fakat temel bileşenler analizinin iç varsayımlarına uygun olan aynı göstergeler iki ayrı zaman kesitinde kullanılarak uygulanmaktadır.

A. S. Albayrak (2005): “Türkiye’ de illerin sosyoekonomik gelişmişlik düzeylerinin çok değişkenli istatistiksel yöntemlerle incelenmesi” adlı çalışmanın birinci amacı, Türkiye’ de illerin sosyoekonomik gelişmişlik düzeylerini belirleyen varsayımsal yapıları çok değişkenli bir yaklaşımla incelemektir. Söz konusu varsayımsal yapılar öncelikle çok değişkenli istatistik yöntemlerden uygun olan açıklayıcı faktör analiziyle araştırılmaktadır. Çalışmanın ikinci amacı ise, faktör analizi sonuçlarına diskriminant analizi uygulanarak iller önceden belirlenen gelişmişlik gruplarına göre sınıflandırılmaktadır. Araştırma coğrafi, nüfus, eğitim ve kültür, sağlık, istihdam, sosyal güvenlik, mali ve finansal, imalat sanayi, tarım, dış ticaret, enerji, konut ve altyapı gibi farklı alanlardan seçilen, fakat faktör analizinin iç varsayımlarına uygun (faktörleştirilebilen) olan aynı göstergeler iki ayrı zaman kesitinde kullanılarak uygulanmaktadır.

Kaygısız ve arkadaşları (2005): İller bazında gelişmişlik düzeyini etkileyen faktörleri inceleyerek kümeleme yapmışlardır. Bu çalışmanın amacı, birbirleri ile sebep sonuç ilişkisi içinde olduğu düşünülen değişkenler arasındaki ilişkileri

gösteren path diyagramlarının oluşturulması, değişkenler arasındaki ilişkilerin derecesini gösteren korelasyon katsayılarının kendisini oluşturan direkt etkiler, dolaylı etkiler ve bileşik path katsayılarına ayrılarak analiz edilmesi ve analiz sonuçlarının doğru bir şekilde yorumlanması işlemlerini kapsayan Path Analizi metodunu tanıtarak gelişmişlik düzeyi üzerinde etkili olduğu düşünülen değişkenlerin ne kadarının gelişmişlik düzeyi üzerinde anlamlı olduğunu belirleyip, anlamlı bulunan değişkenlerin etkilerini direkt ve dolaylı etkiler şeklinde sayısal değerlerle ayrı ayrı ifade etmek ve son olarak da kümeleme analizi metodu ile aynı yapıyı gösteren homojen il gruplarını belirlemektir.

Çalışmada 2000 yılındaki idari yapı esas alınarak 81 il incelemeye alınmıştır. Gelişmişlik düzeyi belirlenmeye çalışılırken, bağımlı değişken olarak kişi başına GSYİH, bağımsız değişkenler olarak ise gelişmişlik düzeyini etkileyeceği düşünülen 17 adet sosyal ve ekonomik değişken kullanılmıştır. Analizde kullanılan değişkenler Türkiye İstatistik Kurumu, Devlet Planlama Teşkilatı ve çeşitli resmi kuruluşların yayınlarından alınmıştır.

Dura ve arkadaşları (2004): Beşeri sermaye göstergeleri açısından 2001 yılı itibarıyla Türkiye'nin Avrupa Birliği karşısındaki gelişme seviyesi incelenmiştir. Yapılan ampirik çalışmalar, beşeri sermayenin ekonomik büyümeyi ve ekonomik kalkınmayı hızlandırdığını ortaya koymuştur. Beşeri sermaye göstergesi olarak kullanılan 16 değişken Cinsiyete Dayalı Kalkınma İndeksi göstergeleri, Cinsiyete Dayalı Yetki İndeksi göstergeleri ve diğer beşeri sermaye göstergeleri Nüfus artış hızı, Her bin canlı doğumda ölüm oranı, Kamu sektörü eğitim harcamalarının GSMH' daki payı, kişi başına sağlık harcamaları, (dolar) bir anlamda da sosyo-ekonomik gelişmeyi gösterdiğinden bu çalışmada incelenmiştir.

Dura, bu değişkenleri çok değişkenli istatistiksel analiz yöntemlerinden kümeleme ve çok boyutlu ölçekleme analizlerine tabi tutarak, beşeri sermaye açısından Türkiye ile benzer gelişme seviyesine sahip olan AB ülkeleri belirlenmeye çalışmıştır. Analize 25 AB üyesi ülke ile Türkiye dâhil edilmiştir. Analize dâhil edilen AB ülkeleri; Belçika, Çek Cumhuriyeti, Danimarka, Almanya, Estonya, Yunanistan, İspanya, Fransa, İrlanda, İtalya, Güney Kıbrıs Letonya, Litvanya, Lüksemburg, Macaristan, Malta, Hollanda, Avusturya, Polonya, Portekiz, Slovenya, Slovak Cumhuriyeti, Finlandiya, İsveç ve İngiltere'dir. Türkiye de dâhil olmak üzere

kümeleme analizi uygulanan 26 ülkenin ele alınan göstergeler bakımından beş farklı kümede toplandıkları gözlemlenmiştir. Tek başına bir kümede yer alan Türkiye'nin beşeri sermaye göstergeleri açısından hiçbir AB ülkesi ile aynı gelişme seviyesini paylaşmadığı ortaya çıkmıştır. ÇBÖ analizi sonuçları da aynı sonucu desteklemiştir. Türkiye ile ilgili analizde ortaya çıkan başka önemli bir nokta, ele alınan göstergeler bakımından kadınlar ve erkekler arasında kadınlar aleyhine gözlenen farklılıktır.

Turanlı ve arkadaşları (2006): Avrupa Birliği aday ve üye ülkelerinin ekonomik benzerliklerini ortaya koymak üzere kümeleme analizi kullanmıştır. Çalışmada ülkelerin GSMH, enflasyon oranı, işsizlik oranı, internet kullanım oranı, ömür boyu eğitim indeksi ve ithalat, ihracat oranı bakımından ekonomik benzerlikleri incelenmiştir. Analiz sonucunda Belçika, Danimarka, Almanya, İspanya, Fransa, İrlanda, İtalya, Hollanda, Lüksemburg, Avusturya, Finlandiya, İsveç ve İngiltere'nin bir küme Çek Cumhuriyeti, Estonya, Yunanistan, Kıbrıs Rum Kesimi, Malta, Litvanya, Letonya, Macaristan, Polonya, Portekiz, Slovenya, Slovakya, Bulgaristan, Romanya, Hırvatistan ve Türkiye'nin diğer bir küme olmak üzere yirmi dokuz ülkenin iki kümede toplandığı görülmüştür.

DPT Araştırması (2003): B. Dinçer ve diğerleri tarafından hazırlanan yayında "İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması" yapılmıştır. Çalışmanın başlıca amaçları şunlardır:

- Sosyal ve ekonomik alanlardan seçilen değişkenler temel alınarak, illerin sosyoekonomik gelişmişlik düzeylerinin belirlenmesi ve bu belirleme doğrultusunda sıralamalarının yapılması,
- Aynı özellikleri taşıyan il gruplarının belirlenmesi,
- Coğrafi bölgelere göre gelişmişlik sıralamalarının belirlenmesi,
- İstatistikî bölge birimlerinin sosyo-ekonomik gelişmişlik düzeylerinin belirlenmesi,
- Eğitim, sağlık ve sanayi gibi sosyal ve ekonomik sektörlere göre farklı mekân birimlerinin (iller, coğrafi bölgeler ve istatistikî bölge birimleri) sosyo-ekonomik gelişmişlik düzeylerinin ölçülmesidir.

Araştırma, mevcut idari yapı çerçevesinde 81 ili kapsamaktadır. Ayrıca aynı veri tabanı ve araştırma tekniği kullanılarak 7 coğrafi bölge ile istatistikî bölge birimleri araştırma kapsamında yer almaktadır.

Araştırmaya, ülke genelinde 81 ilin sosyo-ekonomik gelişmişlik seviyelerini yansıttığı varsayılan ve gelişmişliğin neden ve/veya sonucu olarak ortaya çıkan 100'e yakın göstergenin çeşitli kurum ve kuruluşlardan temin edilmesi ile başlanmıştır. Bu aşamada, göstergelerin tutarlılığı ve güvenilirliği konusunda yapılan ön inceleme ve değerlendirmeler sonunda, il ölçeğinde ancak 58 göstergenin kullanılabilir nitelikte olduğu görülmüştür.

Analiz kısmında daha tutarlı sonuçlar verebilen, istatistikî anlamlılık testleri yapılabilen ve birçok ülkede değişik araştırmacılar tarafından benzer amaçlarla yaygın olarak kullanılan, temel bileşenler analizi tekniği kullanılmış ve kabul görmüştür.

81 il arasında sosyo-ekonomik yönden en gelişmiş ilin İstanbul olduğu tespit edilmiştir. İstanbul'u takiben ikinci sırada Ankara, üçüncü sırada ise İzmir yer almaktadır. 81 il içerisinde en geri kalmış ilimiz Muş'tur. Bu ili Ağrı, Bitlis ve Şırnak takip etmektedir. Araştırmaya göre sosyo-ekonomik yönden en gelişmiş illerin Marmara ve Ege Bölgeleri'nde, en geri kalmış illerin ise Doğu ve Güneydoğu Anadolu bölgelerinde toplandığı tespit edilmiştir. Marmara Bölgesi en yüksek indeks değerine sahip iken Doğu Anadolu en düşük indeks değerine sahip bulunmaktadır. Ege ve İç Anadolu Bölgelerinin sosyo-ekonomik gelişmişlik indeks değerleri birbirine yakın bulunmaktadır. Karadeniz Bölgesi de, Doğu ve Güneydoğu Anadolu'nun ardından sosyo-ekonomik gelişmenin en düşük olduğu üçüncü bölge durumundadır.

Kademeli il grupları içerisinde ise birinci derece gelişmiş iller arasında beş il bulunmaktadır: İstanbul, Ankara, İzmir, Kocaeli ve Bursa. İkinci derece gelişmiş iller arasında en başta Eskişehir olmak üzere 20 il yer almaktadır. Üçüncü, dördüncü ve beşinci derece gelişmiş il gruplarında ise sırasıyla 21, 19 ve 16 il bulunmaktadır.

3. VERİ İNDİRGEME ve KÜMELEME ANALİZİ

3.1. Temel Bileşenler Analizi

Temel Bileşenler Analizi orijinal p değişkenin varyans yapısını daha az sayıda ve bu değişkenlerin doğrusal bileşenleri olan yeni değişkenlerle ifade etme yöntemidir. Aralarında korelasyon bulunan p sayıda değişkenin açıkladığı yapıyı, aralarında korelasyon bulunmayan ve sayıca orijinal değişken sayısından daha az sayıda ($p > k$) orijinal değişkenlerin doğrusal bileşenleri olan değişkenlerle ifade etme yöntemine *temel bileşenler analizi* denir.

Bir veri kümesinde yer alan p değişken arasında önemli kovaryans ya da korelasyon saptanıyorsa bu verileri analizlerde doğrudan kullanmak yerine, bu verilerin doğrusal birleşimlerinden oluşan veriler kullanılabilir.

Temel bileşenler, bizzat kendileri bir sonuç olmaktan ziyade sonuç almayı sağlayıcı özelliğe sahiptirler. Çünkü temel bileşenler daha geniş incelemeler için bir ara adım özelliği taşırlar. Örneğin, çoklu regresyon analizi uygulanması istenilen fakat regresyon varsayımlarından birisi olan çoklu bağlantı olmaması koşulunun yerine gelmemesi nedeniyle regresyon analizinin uygulanamadığı durumlarda çoklu regresyona bir giriş değeri olabilirler ya da kümeleme analizine veri olabilirler. Ayrıca, ölçeklenmiş temel bileşenler, faktör analizi modelleri için kovaryans matrisinin bir ayrışımı olurlar.

Orijinal değişkenlerin ölçüm değerlerinin, değişim aralıklarının ve ölçü birimlerinin çok farklı olduğu durumlarda, değişken sayısının birim sayısından çok fazla olduğu ($n < p$) durumlarda korelasyon ya da kovaryans matrislerini tekil olmayan hale getirmek için veri indirgemesi yapmak ve temel bileşen skorları hesaplayarak kümeleme analizi uygulamak için bu yöntemden yararlanılır (Özdamar, 2002).

Değişkenler arasında tam bağımsızlık söz konusu olamayacağı için oluşan şeklin eksenleri birbirine dik olamayacaktır. Diğer taraftan bu eksenlerin birbirine dik olması daha fazla bilgi verecektir. Bu amaçla, bir dönüştürme yaparak noktaların sahip oldukları toplam varyansın ilk eksen boyunca değişmemesi sağlanarak yeni eksenlerin birbirine dik olması sağlanmaktadır. Harold Hotelling tarafından önerilen bu yöntemde Z standartlaştırılmış veri matrisi de kullanılmaktadır.

$T_{p \times p}$ dönüşüm matrisi olmak üzere,

$$Y_{p \times n} = T_{p \times p}' * Z_{p \times n} \quad (3.1)$$

biçiminde ifade edilir. Diğer bir ifadeyle, birbiriyle ilişkili Z değerlerinden birbiri ile ilişkili olmayan Y değerleri elde edilmektedir. Daha sonra bu bilgiler kullanılarak öz değerler bağıntısından yararlanarak önemli bileşenler elde edilir.

Daha detaylı matematiksel olarak anlatmak gerekirse, cebirsel olarak temel bileşenler X_1, X_2, \dots, X_p p tane rastgele değişkenin özel doğrusal kombinasyonlarıdır. Geometrik olarak, orijinal sistemdeki X_1, X_2, \dots, X_p koordinatlarını döndürmeyle elde edilen yeni koordinat seçimini bu koordinatlar temsil eder. Yeni eksenler daha basit bir yapı sağlar ve maksimum değişkenliği temsil eder.

Temel bileşenler X_1, X_2, \dots, X_p ' nin kovaryans matrisi Σ ' ya veya korelasyon matrisi ρ ' ya bağlıdır. Bu bileşenler çok değişkenli normal dağılım varsayımı gerektirmez. Fakat çok değişkenli normal dağılım kitlesinden çıkarılan temel bileşenler sabit yoğunluklu elipsoidler bakımından yararlı yorumlar yapmamızı sağlar.

$X' = [X_1, X_2, \dots, X_p]$ rastgele vektörü Σ varyans kovaryans matrisine sahip ve öz değerleri $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ eşitsizliğini sağlasın. Temel bileşenler aşağıdaki gibi tanımlanan ilişkisiz doğrusal birleşimlerdir.

$$\begin{aligned} Y_1 &= \ell_1' X = \ell_{11} X_1 + \ell_{21} X_2 + \dots + \ell_{p1} X_p \\ Y_2 &= \ell_2' X = \ell_{12} X_1 + \ell_{22} X_2 + \dots + \ell_{p2} X_p \\ &\quad \cdot \quad \cdot \\ &\quad \cdot \quad \cdot \\ Y_p &= \ell_p' X = \ell_{1p} X_1 + \ell_{2p} X_2 + \dots + \ell_{pp} X_p \end{aligned} \quad (3.2)$$

$$\text{Var}(Y_i) = \ell_i' \Sigma \ell_i = \lambda_i \quad i = 1, 2, \dots, p \quad (3.3)$$

$$\text{Cov}(Y_i, Y_k) = \ell_i' \Sigma \ell_k = 0 \quad i, k = 1, 2, \dots, p, \quad i \neq k \quad (3.4)$$

Birinci temel bileşen maksimum varyanslı doğrusal kombinasyondur: $\ell_1' \ell_1 = 1$ kısıtı altında $\text{Var}(Y_1) = \ell_1' \Sigma \ell_1 = \lambda_1$ maksimumdur. İkinci temel bileşende

$l_2' l_2 = 1$ kısıtı ve $\text{Cov}(l_1' X, l_2' X) = 0$ kısıtı altında 2. en büyük varyansa sahiptir: $\text{Var}(Y_2) = l_2' \Sigma l_2 = \lambda_2$. Benzer biçimde $l_i' l_i = 1$ ve $k < i$ için,

$$\text{Cov}(l_i' X, l_k' X) = 0 \quad (3.5)$$

olmak üzere, i. temel bileşen i. en büyük varyansa sahiptir: $\text{Var}(Y_i) = l_i' \Sigma l_i = \lambda_i$. Ayrıca l_i' katsayı vektörü λ_i özdeğerine karşılık gelen i. özdeğer vektörü e_i ' dir.

3.1.1. Temel bileşenler analizi için yararlanılan matrisler

Temel bileşenler analizi hem orijinal verilerden hem de standartlaştırılmış verilerden hesaplanabilir. Bu verilerin kovaryans matrisinden ya da korelasyon matrisinden yararlanılarak uygulanabilir. Her iki matrise göre elde edilen temel bileşenler farklılık göstermektedir. Bu nedenle bu yöntemi uygulama esnasında verilere uygun bir matris seçmek gerekmektedir.

Temel bileşenler analizinde, araştırma kapsamına alınan değişkenlerin ölçü birimleri genellikle farklıdır. Tüm değişkenlerin ölçü birimleri aynı olsa da büyük varyansa sahip değişkenler temel bileşende daha büyük ağırlıklara, küçük varyansa sahip değişkenler ise daha küçük ağırlıklara sahip olur. Veri kümesinde yer alan p değişkenin ölçü birimleri çok farklı ise ve değişkenlerin değişim aralıkları oldukça değişik ve birbirinden büyük farklılıklar gösteriyorsa, korelasyon matrisinden yararlanmak uygundur. Diğer halde kovaryans matrisinden yararlanmak yerinde olur.

Yöntem uygulamasında en önemli adımdır ve bu sorunun tam bir cevabı yoktur. Literatürde yöntemlerden bazıları aşağıda belirtilmektedir:

- Birden büyük özdeğer sayısı kadar temel bileşen seçmek,
- Genel varyansın en az %67 sini açıklayan sayıda temel bileşen seçmek,
- Yamaç Eğimi Testi yaparak temel bileşen sayısına karar vermektir

(Özdamar, 2002).

3.2. Kümeleme Analizi

Kümeleme analizi, n sayıda birimin p sayıda değişkene ilişkin gözlemlerinin elde edilmesi ile oluşan n*p boyutlu X veri matrisinde yer alan ve doğal gruplamaları kesin olarak bilinmeyen birimleri, değişkenleri ya da hem birim hem de değişkenleri

bir arada ele alarak birbiri ile benzer olan alt kümelere (grup, sınıf) ayırmaya yardımcı olan yöntemler topluluğudur.

Kümeleme analizi, temel olarak dört değişik amaçla yapılır:

- n sayıda birimi, nesneyi, oluşumu, p değişkenden saptanan özelliklerine göre olabildiğince kendi içinde türdeş (homojen) ve kendi aralarında farklı (heterojen) alt gruplara ayırmak,
- p sayıda değişkeni, n sayıda birimde saptanan değerlere göre ortak özellikleri açıkladığı varsayılan alt kümelere ayırmak ve ortak faktör yapıları ortaya koymak,
- Hem birimleri hem de değişkenleri birlikte ele alarak ortak n birimi p değişkene göre ortak özellikleri alt kümelere ayırmak,
- Birimlerin, p değişkenden saptanan değerlere göre, izledikleri biyolojik ve tipolojik sınıflamayı ortaya koymak (taksonomik sınıflandırma yapmak).

Bu amaç doğrultusunda kümeleme analizi hemen hemen tüm bilim alanlarında yararlanılan bir yöntemdir. Gruplama, uzaklık ya da benzerlik ölçümlerinden yararlanılarak yapılır (Johnson and Wichern, 2002).

Uzaklık ölçüleri ya da benzerlik ölçüleri, veri matrisinde yer alan değişkenlerin ölçü birimlerine göre de farklılık göstermektedir. Eğer değişkenler oransal ya da aralıklı ölçekle elde edilmiş değerler ise uzaklık (distance) ya da ilişki (correlation) türü ölçülerden yararlanır. Ölçümler sayısal değerler olarak yapılmış ise tercih edilen ölçüler kare uzaklık ölçüsü ya da Phi kare uzaklık ölçüsüdür. Eğer ikili (binary) gözlemlere göre ölçümler yapılmış ise birimler arasındaki benzerlikleri belirlemede Öklid, kare Öklid, size difference, pattern difference, Lance and Williams difference, shape difference gibi benzerlik ya da farklılık ölçülerinden yararlanılmaktadır (Özdamar, 2002).

Özet olarak, kümeleme yöntemleri uzaklık matrisi ya da benzerlik matrisinden yararlanarak birimleri ya da değişkenleri kendi içinde homojen ve kendi aralarında heterojen uygun gruplara ayırır. Kümeleme analizinde temel amaç birimlerin veya değişkenlerin doğal gruplamasını keşfetmektir.

Yöntemin pek çok alanda uygulanabilir olması, çıkan kümelemede izlenen yeni yaklaşımların ortaya çıkmasına neden olmuştur.

3.2.1. Yaygın kümeleme analizi yöntemleri

Gözlem vektörlerini kümelemek için yaygın olarak kullanılan iki yaklaşım aşamalı (hierarchical) ve parçalayıcı (partitioning) kümeleme yöntemleridir. Aşamalı kümeleme, her gözlemi tek bir küme olduğunu varsayarak başlar ve n gözlemi içeren bir tek kümeyle sona erer. Her adımda, bir gözlem ya da gözlemlerin bir kümesi, diğer kümeyle eklenerek yeni bir küme oluşturur. Bu uygulamanın terside yapılabilir. Parçalayıcı kümeleme, basit olarak gözlemleri g adet kümeyle ayırır. Bu yöntem ilk parçalama ya da küme merkezi ile başlar ve daha sonra uygun kritere göre gözlemleri yeniden tahsis eder.

Birleştirici olan aşamalı kümeleme için bazı dezavantajlar vardır. Mesela, erken bir adımda yanlış olarak gruplanmış olan birimler ya da değişkenler yeniden tahsis edilemez ve kümeler arasındaki benzerlik ölçümü için farklı bir benzerlik ölçütü kullanıldığında farklı sonuçlara götürebilir (Gan vd, 2007). Parçalayıcı kümeleme yöntemlerinden yaygın olarak kullanılanlardan biri, K-ortalama yöntemidir. K-ortalama kümeleme algoritması ise, küme sayısını önceden belirlenmiş olmasını gerektirir. Ayrıca K-ortalama algoritmasının sonuçları başlangıç küme merkezlerinden etkilenmektedir (Özdamar, 2002).

Araştırmalardaki birçok kümeleme yöntemi, ya nesnelere belli sayıda kümeyle en iyi şekilde bölmeyi arayan parçalayıcı yöntemler olarak ya da kümelerin iç içe girmiş bir dizisini üreten aşamalı yöntemler olarak tanımlanabilir (Chipman ve Tibshirani, 2006).

4. HİBRİD (HYBRID) KÜMELEME ANALİZİ

Hibrid kümeleme algoritmaları, iki ya da daha fazla kümeleme yönteminin uygun bir şekilde birleştirilerek bir arada kullanılmasıyla oluşturulan bir yöntemdir. Hibrid kümeleme analizindeki en temel amaç, klasik kümeleme algoritmalarının dezavantajlarını gidermektir. Bu bölümde bu algoritmaların hibrid kümelemeyi gerçekleştirmek için nasıl birleştirilebildiği hakkında detaylı bilgiler verilecektir.

İlk olarak hibrid kümeleme için gruplamada en çok kullanılan aşamalı kümeleme ve parçalı kümeleme yöntemleri üzerinde tartışılacaktır.

4.1. Aşamalı Kümeleme Yöntemi

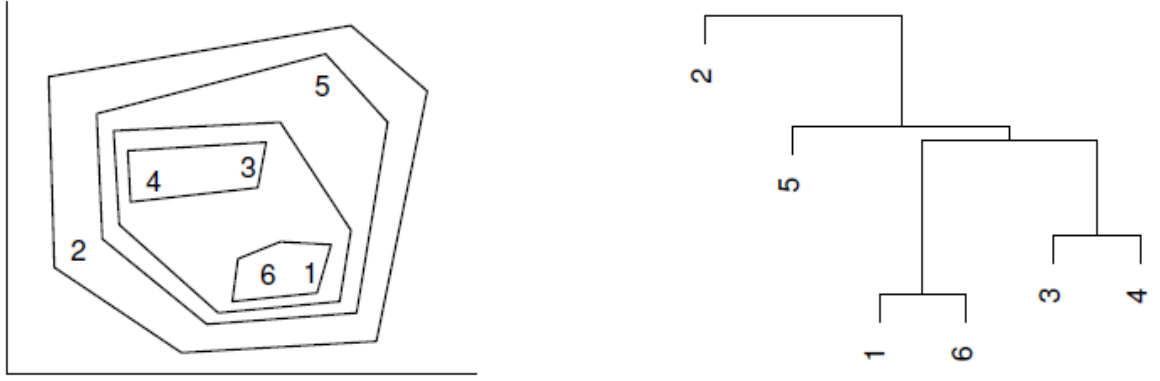
Aşamalı kümeleme yöntemleri, birimleri birbirleri ile değişik aşamalarda bir araya getirerek ardışık biçimde kümeler belirlemeyi ve bu kümelere girecek elemanların hangi uzaklık (ya da benzerlik) düzeyinde küme elemanı olduğunu belirlemeye yönelik yöntemlerdir.

Veri matrisinde birimlerin (ya da değişkenlerin) başlangıç aşamasında kaç küme oluşturduğuna ve küme elemanlarını belirlemede başlangıçta hangi kriterin seçildiğine göre aşamalı yöntemler iki ana gruba ayrılır:

- Birleştirici Aşamalı Kümeleme Yöntemleri (Agglomerative Hierarchical Clustering Prosedures): Başlangıçta tüm birimlerin ayrı birer küme oluşturduğunu kabul ederek, n birimi aşamalı olarak sırasıyla $n, n-1, n-2, \dots, n-r, \dots, 3, 2, 1$ kümeye yerleştirmeyi amaçlayan bir yaklaşımdır.

- Ayırıcı Aşamalı Kümeleme Yöntemleri (Divisive Hierarchical Clustering Prosedures): Başlangıçta tüm birimlerin bir küme oluşturduğunu kabul ederek birimleri aşamalı olarak n birimi sırasıyla $1, 2, 3, \dots, n-r, \dots, n-2, n-1, n$ kümeye ayırmayı amaçlayan bir yaklaşımdır.

Hem ayırıcı hem de birleştirici metodun sonuçları *ağaç dalı şeması* olarak bilinen iki boyutlu bir diyagram formunda gösterilebilir. *Ağaç dalı şeması*, peş peşe seviyelerde yapılan birleştirmeleri ya da ayrıştırmaları gösterir (Johnson and Wichern, 2002). Görsel olarak göstermek gerekirse, Şekil 4.1' de sol tarafta bulunan iki boyutlu ve sağ tarafta bulunan ise ilgili ağaç dalı şeması (dendrogram) içinde 6 noktası verilen bir aşamalı kümeleme örneğini temsil eder.



Şekil 4.1 İki boyutlu küme şeması ve ağaç dalı şeması örneği

Literatürde, aşamalı kümeleme yöntemi denildiğinde birleştirici aşamalı kümeleme yöntemleri anlaşılmaktadır. Çünkü birleştirici aşamalı kümeleme yöntemleri birimlerin birbirleriyle hangi aşamada ve hangi benzerlik düzeyinde ortak özelliklere sahip kümeler oluşturduklarını göstermeleri açısından yaygın olarak kullanılmaktadır. Bu yöntemi sözel olarak biraz daha açmak gerekirse, her birim başlangıçta tek başına birer küme kabul edilir. Daha sonra birbirleri ile yüksek derecede benzerlik gösteren iki birim bir küme oluşturur. Bu kümeye değişik benzerlik düzeyinde diğer birimler eklenerek birimlerin tümü bir kümede toplanacak biçimde birbirleri ile birleştirilirler. Ayırıcı aşamalı kümeleme, birleştirici aşamalı kümelemenin tersidir.

Aşamalı kümeleme yöntemi iki önemli kararı içerir. İlk karar, uygun bir uzaklık ölçümünün seçimidir. Bir kez bu seçildikten sonra, tüm ikişerli uzaklıkları içeren bir uzaklık matrisi, her bir gözlemin kendisinin bir küme olduğu ikinci aşama için başlangıç noktası olarak hesaplanabilir. İkinci karar ise kümeleme ya da bağlantı yöntemlerinin seçimidir.

4.1.1. Aşamalı kümelemede birimleri birleştirme yöntemleri

Aşamalı kümeleme yöntemlerinde, birimlerin birbirleri ile birleştirilmesinde değişik yaklaşımlar uygulanmaktadır. Bu nedenle değişik isimlerle anılan yöntemler vardır. Bu yöntemlerden yaygın olarak kullanılan ve kabul görmüş olanlardan bazıları aşağıda belirtilmektedir.

Tek bağlantı: En yakın komşu yaklaşımı olarak da bilinen tek bağlantı kümeleme yöntemi bir A kümesindeki birim (değişken) ile bir B kümesindeki birim (değişken)

arasındaki minimum farka bağlıdır (Larose, 2005). Bu yöntemin dezavantajı, her kümedeki birimlerin birçoğu birbirine çok uzak olmasına rağmen, birbirine yakın olan yalnız birimlerden dolayı kümeler birlikte olmaya zorlanır. Tek bağlantı yöntemi üstünkörü, dağınık kümeler oluşturan uzun sınırlar meydana getirir. Daha çok sayısal sınıflandırma biliminde kullanılır (Fielding, 2007).

Ortalama bağlantı: İki gruptaki birimleri arasındaki tüm uzaklıkların ortalamasını kullanan yöntemdir (Chipman ve Tibshirani, 2006). Ortalama bağlantı hesaplamak için uygun pek çok metotlar vardır. Bunlar ağırlıklandırılmamış ikili-grup ortalaması, ağırlıklandırılmış ikili-grup ortalaması, ağırlıklandırılmamış ikili-grup merkezi ve ağırlıklandırılmış ikili-grup merkezleridir (Rao ve Srinivas, 2008).

Tam bağlantı: En uzak-komşu yaklaşımı olarak da bilinen tam bağlantı kümeleme yöntemi bir A kümesindeki birim (değişken) ile B kümesindeki birim (değişken) arasındaki maximum farka bağlıdır (Larose, 2005). Bu yöntem ise küçük ve birbirine yakın kümeleri oluşturmak için kullanılır. Genellikle büyük veri kümelerine uygulamak için uygun değildir. (Rao ve Srinivas, 2008).

Küresel ortalama bağlantı: Kilit merkezleri kullanan ağırlaştırılmış ikili grup metodu olarak da bilinen bu yöntem, grup merkezleri (grup ortalaması gibi) arasındaki uzaklığı kullanır.

Ward bağlantı: İki küme birleştiğinde, bilgi kaybı ya da amaç fonksiyonunun değerindeki değişim diğer kümelerle olan ilişkiye değil, sadece birleşen iki küme arasındaki ilişkiye bağlı olan yöntemdir. Bu yöntem küme yapısını iyileştirmede başarılıdır ve oldukça eşit büyüklükteki küresel yapıya sahip kümeleri oluşturmaya yönelir.

Yukarıda bahsedilen yöntemlerin yanında *Mcquitty bağlantı*, *Ortanca bağlantı* gibi uygulamada pek rastlanmayan kümeleme yöntemleri de mevcuttur.

4.2. Aşamalı Olmayan Kümeleme Yöntemi

Aşamalı kümeleme yönteminde hem birimler hem de değişkenler birbirleriyle değişik benzerlik düzeylerinde gruplar meydana getirirken, bu yöntemde sadece birimlerin uygun oldukları kümelerde toplanmaları hedeflenmektedir. Ayrıca oluşan bu kümelerin, kendi içinde homojen ve kendi aralarında heterojen olmaları amaçlanmaktadır. Parçalı kümeleme yöntemi ağaç dalı yapısının hesaplama

açısından engelleyici olduğu büyük veri kümelerini içeren uygulamalarda avantajlara sahiptir (Abonyi ve Feil, 2007).

Bu kümeleme çalışması benzerlik matrisi ya da uzaklık matrisi yerine X veri matrisi ya da dönüştürülmüş z veri matrisini ele almaktadır. Aşamalı olmayan yöntemler içinde yaygın olarak kullanılan yöntemlerden bir kaç, K-ortalama ve yığma kümelemedir. Bu araştırmada K-ortalama kümeleme yöntemi ele alınmaktadır.

4.2.1. K-ortalamalar yöntemi (Mac Quennn's K-Means Method)

K-ortalamalar yöntemi aşağıdaki adımları izleyerek birimleri kümelere ayırır.

1. Birimleri K adet başlangıç kümesine dağıtır veya K adet küme merkezi belirler.

2. Birimlerin listesinden birimler sırayla alınır ve uzaklığı en yakın kümeye atar. Buradaki uzaklık genellikle ya standartlaştırılmış ya da standartlaştırılmamış değerlerden hesaplanan Öklid uzaklığı kullanılarak elde edilir.

3. Bir kez atama gerçekleştiğinde, yeni birimi kabul eden kümeler ve birimi kaybeden kümeler için küme merkezlerinin yarı ortalamaları tekrar hesaplanır.

4. Oluşan kümelere birimlerin yeniden atanmalarına gerek kalmayınca kadar 2. adım tekrar edilir (Johnson ve Wichern, 2002; Fielding, 2007). Bu algoritma her gözlem ile onun küme merkezi arasındaki Öklid uzaklığının karesinin toplamını en küçük kılan K kümeyi bulmayı hedefler.

K-ortalama yöntemini içeren ya da parçalama yaklaşımını kullanan iterasyon yöntemlerinin pek çoğu, kümeler içindeki değişimi en aza indirmek için hata kareler ölçütünü kullanır.

Bir X (K kümeyi içeren) örnek setinin $V = \{v_i / i=1, \dots, K\}$ kümelemesi için hata kareler kriteri

$$J(X; V) = \sum_{i=1}^K \sum_{j \in i} \|x_j^{(i)} - v_i\|^2 \quad (4.1)$$

dir. Burada $x_j^{(i)}$ i. kümeye ait j. birimdir ve v_i i. kümenin merkezidir (Abonyi ve Feil, 2007).

Estivill – Castro ve Yang (2004) K-ortalama yönteminin ana problemlerini listelemişlerdir: zayıf bir optimuma yakınsama eğilimi; ölçeklemeye ve öteki

dönüşümlere duyarlılık; ölçüm hatalarına ve aykırı değerlere duyarlılık; yanlışlık (yani yanlış parametre değerlerine yakınsama). Yine de, en önemli problem K ' nın uygun bir değerini hesaplamaktır.

Algoritma en iyi parçalanmayı bulmasını garanti etmemesi nedeniyle, aynı veriyi kullanan iki sürecin farklı kümeleme sonuçları üreteceği oldukça olasıdır. Değişik K değerleri için K -ortalamalar yöntemi gerçekleştirildiğinde en iyi K değeri ortaya çıkan kümeler içindeki varyansın toplamını en küçük kılan K değeridir. K -ortalamanın performansını geliştirmek için ortaya çıkan alternatif bir yaklaşım, başlangıç kümelerine birimleri dağıtmak için aşamalı kümeleme analizinden gelen sonuçları kullanmaktır.

K -ortalama kümelemesinin büyük bir avantajı mevcut olan kümelere yeni birimleri aktarabiliyor olunmasıdır.

4.3. Hibrid Kümeleme

Parçalayıcı kümeleme algoritması küme sayısı, küme merkezi, ... gibi başlangıç parametreleri tahmininden yüksek oranda etkilenirken, hiyerarşik kümeleme algoritması başlangıç değeri ve yerel minimumdan etkilenmez. Parçalayıcı kümeleme algoritmaları özellik vektörlerinin (feature vectors) bir kümeden diğerine amaç fonksiyonunu minimize etmek için hareket edebilen sağduyuya sahiptir. Buna karşın, aşamalı kümeleme algoritmasında herhangi bir adımda bir kümeye atanan özellik vektörü bir gruptan diğerine hareket edemez. Bu nedenle bu iki yöntemi bir arada kullanan bir hibrid kümeleme algoritması düşüncesi doğar: aşamalı kümeleme algoritmasından elde edilen küme merkezleri parçalayıcı kümeleme algoritmasını başlatmak için kullanılır.

Yukarıda bahsedilen yöntemleri kullanan hibridlemenin yanı sıra uygun yöntemleri sentezleyerek hibridleme çalışmasını uygulamak mümkündür. Bu kapsamda son yıllarda pek çok araştırma mevcuttur.

4.3.1. Hibrid kümeleme algoritması

1. adım: $Y = \{ y_i / i=1, \dots, n \}$, p -boyutlu uzayda n özellik vektörünün bir kümesini gösterebilir. Ayrıca x_i , n -boyutlu uzayda yeniden ölçeklendirilmiş i . özellik vektörünü

gösterebilirsin. Burada x_i , y_i ölçümleri tarafından aşağıdaki matematiksel eşitlik aracılığı ile elde edilir.

$$x_{ij} = \frac{w_j}{\sigma_j} [f(y_{ij})], \quad i=1, \dots, n; j=1, \dots, p \text{ için} \quad (4.2)$$

Burada,

$f(\cdot)$: dönüşüm fonksiyonu,

y_{ij} : p-boyutlu vektör y_i ' deki j. değişkenin değeri,

x_{ij} : y_{ij} ' nin yeniden ölçeklendirilen değeri,

w_j : j değişkenine tahsis edilmiş ağırlığı,

σ_j : j değişkenin standart sapmasıdır.

2. adım: Birleştirici aşamalı kümeleme algoritmasının n-K adımında oluşan K kümenin merkezi, K-ortalamalar yöntemini başlatmak için kullanılır. K-ortalama yöntemi Eşitlik 4.3' teki amaç fonksiyonunun (objective function) değerini en küçük kılmak için bir kümeden diğerine özellik vektörlerinin hareket etmesini sağlayan tekrarlayıcı bir süreçtir.

$$F = \sum_{k=1}^K \sum_{j=1}^p \sum_{i=1}^{n_k} d^2(x_{ij}^k - x_{.j}^k) \quad (4.3)$$

Eşitlik 4.3' teki tanımları ifade etmek gerekirse;

K: küme sayısını,

N_k : k. kümedeki özellik vektörlerinin sayısı,

x_{ij}^k : k. küme tahsis edilmiş i özellik vektöründe j özelliğinin yeniden ölçeklendirilmiş değerini,

$x_{.j}^k$: k. kümedeki j özelliğinin örneklem ortalama değerini,

$$x_{.j}^k = \frac{\sum_{i=1}^{n_k} x_{ij}^k}{n_k} \quad (4.4)$$

ifade etmektedir. Eşitlik 4.3' deki F' yi en küçükleyerek, her özellik vektörünün ait olduğu küme merkezinden uzaklığı en küçüklenir. Biz Öklid ya da Mahalanobis gibi

uygun bir uzaklık ölçümü $d(\cdot)$ kullanarak kümelerin global şekli ya da büyüklüğü hakkında bilgileri dahil etme seçeneğine sahibiz (Rao ve Srinivas, 2008).

4.4. Küme Geçerlilik Ölçüleri

4.4.1. Silhouette Genişliği

Bir özellik vektörünün silhouette genişliği, o özellik vektörünün kendi bulunduğu kümedeki özellik vektörlerine, diğer kümelerdeki özellik vektörlerine kıyasla ne kadar benzediğinin bir ölçüsüdür. k kümesindeki i . özellik vektörü için $s(i)$ ile gösterilen silhouette genişliği aşağıdaki gibi tanımlanmıştır:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (4.5)$$

Burada $a(i)$, k . küme içerisinde i . özellik vektöründen tüm diğer özellik vektörlerine olan ortalama uzaklıktır. $b(i)$ ise i . özellik vektörünün bir diğer j . kümedeki ($j=1, \dots, K$; $j \neq k$) tüm özellik vektörlerine olan en küçük ortalama uzaklıktır. Bu genişlik -1 ile 1 arasında değer alır.

Eğer $s(i)$ 1 yakınsa, i . özellik vektörünün uygun bir kümede bulunduğunu söyleyebiliriz. Diğer bir deyişle, $s(i)$ -1 ' e yakın olduğu zaman i . özellik vektörünün yanlış sınıflandığı sonucuna varırız. Yaklaşık olarak bu ölçü 0 ise, i . özellik vektörünün iki kümeye eşit yakınlıkta olduğunu gösterir. Verilen bir K değeri için silhouette genişliğinin genel ortalaması, veri kümesindeki tüm özellik vektörlerinin silhouette genişliğinin ortalamasıdır. En büyük silhouette genişliği genel ortalamasına sahip olan bölünüm, verilerin en iyi bölünümü olarak alınır (Rao ve Srinivas, 2008).

4.4.2. Davies–Bouldin İndeksi

Davies–Bouldin indeksi kümeler içindeki yayılım toplamının kümeler arasındaki ayrışmaya oranının bir fonksiyonudur.

N özellik vektörünün verilmiş bir kümesinin $\{C_1, C_2, \dots, C_K\}$ gibi K kümeye öyle bir biçimde parçalandığı düşünölsün ki her C_k kümesi N_k özellik vektörüne sahip ve her özellik vektörü de sadece bir kümede olsun. Böylelikle de,

$\sum_{k=1}^K N_k = N$ dir. k . küme içindeki yayılım $S_{k,q}$ aşağıdaki denklem kullanılarak

hesaplanır ve C_j ve C_k kümelerini tanımlayan küme merkezleri arasındaki t. dereceden Minkowski uzaklığı Eşitlik 4.6' dadır.

$$S_{k,q} = \left(\frac{1}{N_k} \sum_{x_i \in C_k} \|x_i - z_k\|_2^q \right)^{\frac{1}{q}} \quad (4.6)$$

$$d_{jk,t} = \|z_j - z_k\|_t \quad (4.7)$$

Burada z_k k. kümenin merkezini temsil eder ve $S_{k,q}$ 'de onların ortalamasına göre k. kümedeki noktaların Öklid uzaklıklarının q. momentinin q. köküdür. Araştırmacılar arasında q=1 ve t=2 değerleri sıklıkla benimsenen değerlerdir. Davies–Bouldin indeksi ise Eşitlik 4.8' den hesaplanır. DB için küçük bir değer, merkezlerin birbirinden uzak olduğu ve özellik vektörlerinin de küme merkezlerine yakın olduğu iyi bir parçalanmayı gösterir (Rao ve Srinivas, 2008).

$$DB = \frac{1}{K} \sum_{k=1}^K \max_{j, j \neq k} \left\{ \frac{S_{k,q} + S_{j,q}}{d_{jk,t}} \right\} \quad (4.8)$$

5. MATERYAL VE YÖNTEM

Çalışma, 2008 yılındaki idari yapı esas alınarak 81 ilin toplam 957 ilçesini kapsamaktadır. Hakkında veri bulunamayan 12 ilçe çalışmaya dâhil edilmemiştir.

5.1. Veriler

DPT tarafından sosyo-ekonomik gelişmişliği ölçtüğü kabul edilen Çizelge 2.2' deki değişkenlerin birçoğuna 2008 yılı itibariyle ilçeler bazında ulaşılamamıştır. İlçelere ait ulaşılan değişkenler Çizelge 5.1' de verilmektedir.

81 ile ait 957 ilçe için Çizelge 5.1' deki değişkenlere ilişkin elde edilen veriler Ek-1' de verilmektedir. Bazı merkez ilçelere (12 ilçeye) ilişkin birçok değişkenin verilerine ulaşılamamıştır. Bu nedenle bu ilçeler analize dâhil edilmemiştir.

Demografik değişkenler ilçe nüfus toplamları ve alınan göç oranından oluşmaktadır. İlçelere göre nüfus toplamları, ilçe merkezinde ve bu ilçelere bağlı belde ve köylerindeki toplam nüfusu ve alınan göç oranı da bir ilçede ikamet edipte orada doğmamış olanların toplam ilçe nüfusuna oranını göstermektedir.

Araştırma kapsamında eğitim göstergeleri, ilçedeki çeşitli eğitim düzeyindeki kişilerin sayısının ilçe nüfusuna oranlamasıyla elde edilmiştir.

Tarım değişkenlerinden üretilen süt miktarı (ton olarak), yeterli miktarda süt temin edilen hayvanlardan sığır (yerli), sığır (kültür), sığır (melez) ve koyun (yerli) sütleri ayrı ayrı ilçeler bazında düzenlenmiş ve toplanmıştır. Üretilen et miktarında ise her bir ilçedeki büyükbaş hayvanlardan; inek, boğa, dana, düve, tosun küçükbaşlardan; oğlak, koyun, kuzu-toklu, keçi etleri toplanmıştır. Traktör sayısı, ekilen hektar başına traktör sayısını temsil etmektedir

Tahıllar toplamı, ülkede en çok üretimi olan patates-kuru baklagiller-yenilebilir kök ve yumrular, şeker imalatında kullanılan bitkiler ve tahıllar ayrı ayrı ilçeler bazında elde edilmiş ve toplanmıştır. Sebzeler toplamı ise kök ve yumru sebzeler ve meyvesi için yetiştirilen sebzeler ilgili veri tabanından ayrı ayrı temin edilmiş ve ilçeler bazında toplanmıştır. Meyveler toplamı (üretimi ton olarak); çay, turunç, üzüm, zeytin, muz, incir ve diğer meyveler ilçeler bazında TÜİK' in veri tabanından temin edilmiştir.

Çizelge 5.1 Çalışmada kullanılan değişkenler

DEĞİŞKEN	YILI	KAYNAK	BİRİM
<u>DEMOGRAFİK GÖSTERGELER</u>			
1. İlçelere Göre (Merkez + belde(köy)) Nüfus toplamları	2008	TÜİK-ADNKS	Kişi
2. Alınan Göç Oranı	2008	TÜİK-ADNKS	Yüzde
<u>EĞİTİM DURUMU</u>			
3. Okuma Yazma Bilmeyenlerin Oranı	2008	TÜİK-ADNKS	Yüzde
4. Okuma Yazma Bilenlerin Oranı	2008	TÜİK-ADNKS	Yüzde
5. Bilinmeyenlerin Oranı	2008	TÜİK-ADNKS	Yüzde
6. Okuma Yazma Bilen Fakat Bir Okul Bitirmeyenlerin Oranı	2008	TÜİK-ADNKS	Yüzde
7. İlkokul Mezunu Oranı	2008	TÜİK-ADNKS	Yüzde
8. İlköğretim Mezunu Oranı	2008	TÜİK-ADNKS	Yüzde
9. Ortaokul veya Dengi Okul Mezunu Oranı	2008	TÜİK-ADNKS	Yüzde
10. Lise veya Dengi Okul Mezunu Oranı	2008	TÜİK-ADNKS	Yüzde
11. Yüksekokul veya Fakülte Mezunu Oranı	2008	TÜİK-ADNKS	Yüzde
12. Yüksek Lisans Mezunu Oranı	2008	TÜİK-ADNKS	Yüzde
13. Doktora Mezunu Oranı	2008	TÜİK-ADNKS	Yüzde
<u>MEDENİ DURUM</u>			
14. Hiç Evlenmemişler	2008	TÜİK-ADNKS	Yüzde
15. Evli Olanlar	2008	TÜİK-ADNKS	Yüzde
16. Boşanmış Olanlar	2008	TÜİK-ADNKS	Yüzde
17. Eşi Ölü Olanlar	2008	TÜİK-ADNKS	Yüzde
<u>TARIM</u>			
18. Üretilen Süt (Ton) Miktarı	2008	TÜİK-VERİ TABANI	Ton
19. Tahıllar Üretim Toplamı	2008	TÜİK-VERİ TABANI	Ton
20. Sebzeler Üretim Toplamı	2008	TÜİK-VERİ TABANI	Ton
21. Meyveler Üretim Toplamı	2008	TÜİK-VERİ TABANI	Ton
22. Üretilen Et (Büyükbaş + Küçükbaş) Miktarı	2008	TÜİK-VERİ TABANI	Ton
23. Ekilen Alan Başına Traktör Sayısı	2008	TÜİK-VERİ TABANI	Adet

Maalesef ülkemizde genel yönetim açısından en küçük istatistiksel bilgi toplama ünitesi olarak ilçe düzeyinin temel alınması ilkesi henüz benimsenmemiştir. Ülkenin istatistiksel analiz yapan devlet kurumu olarak Türkiye İstatistik Kurumu, 81 ilden sadece 25 tanesinde mevcuttur. Birçok kurum il ve bölge düzeyinde örgütlendiğinden istatistiksel çalışmalar da genelde il düzeyinde yapılmaktadır. Bu nedenle ilçe bazında veri temininde güçlükler yaşanmaktadır. TÜİK tarafından en son yapılan detaylı araştırma 2000 genel nüfus sayımında yapılmış ve bugüne kadar ki olan bütün bilimsel araştırmalar bu yılın veri tabanından yararlanmıştır. Dolayısı ile sosyo-ekonomik gelişmişliği ölçen pek çok veriye güncel olarak ulaşamamaktadır.

İlçelerdeki gelişme dinaminizmini yansıtan değişkenlerin temin edilememesi, doğal olarak çalışmanın sonuçlarını etkileyen ve sınırlayan önemli bir unsur olmuştur. Bu çalışma için resmi veya özel birçok kurumun veri tabanı araştırılmış, ancak ilçeler bazında bilgiye ulaşamamıştır.

5.2. Yöntem

Bu çalışmada verilerin elde edilışinden analizine kadar izlenen yöntem aşama aşama aşağıda belirtilmektedir.

Verilerin toplanması ve kalitesinin kontrolü: Çizelge 5.1' de belirtilen değişkenlerden bir çoğuna ilişkin veriler TÜİK' in veri tabanlarında hazır olmadığından, bu değişkenlere ilişkin verileri oluşturmak için, veri tabanları tek tek her ilçe için incelenmiş ve her bir değişken için farklı Excel 2003 dosyalarına dikkatlice kaydedilmiştir. Verilerin Çizelge 5.1' deki yapısıyla elde edilebilmesi için gerekli her türlü işlem Excel programı aracılığıyla gerçekleştirilmiştir.

Veriler daha sonra satırlar ilçeler, sütunlar değişkenler olmak üzere bir Excel dosyasına aktarılmış, ilçelerin ait oldukları illeri ve coğrafi bölgeleri gösteren iki yeni değişken oluşturulmuştur. Sonrasında ise veriler SPSS 15 paket programına ilişkin bir veri dosyasına aktarılmıştır. Verilerde tutarsızlık olup olmadığını ve kayıp gözlem durumunu belirlemek için frekans dağılım tabloları oluşturulmuş ve incelenmiştir. Tespit edilen aksaklıklar düzeltilmiştir. Ayrıca birçok değişkene ilişkin veri bulunmayan 12 ilçe dışlanmıştır.

Verilerin ön incelemesi ve aykırı gözlemlerin tespiti: Verileri tanımlamak amacıyla ortalama ortanca, standart sapma, en küçük ve en büyük gözlem değerleri

elde edilmiştir. Değişkenlerin dağılımları hakkında bilgi edinmek için histogramlar çizilmiş; değişkenler arasındaki ilişkileri görmek için saçılım çizimleri elde edilmiştir. Ayrıca ilişki (korelasyon) matrisi elde edilmiştir. Bütün bu işlemler için SPSS 15 paket programı kullanılmıştır.

Verilerin incelemesi sonucu benzer bilgiyi veren bazı değişkenler birleştirilmiş, bazı değişkenler dışlanmış ve değişken sayısı 20' ye inmiştir. Ayrıca dağılımları aşırı derecede sağa çarpık olan değişkenlere doğal logaritma dönüşümü uygulanmıştır.

Dikkate alınan ve bazıları dönüştürülmüş 20 değişkene göre aykırı ilçelerin tespiti için ilk önce saçılım grafikleri incelenmiştir. Daha sonra her değişkenin 10. yüzdeliği ve 90. yüzdeliği bulunup, ilçelerin 20 değişken üzerinden kaç değişken için 10. yüzdeliğin altında veya 90. yüzdeliğin ötesinde değer aldığı tespit edilmiştir.

Ayrıca aykırı ilçeleri çok boyutlu uzayda tespit etmenin bir yolu olarak Mahalanobis uzaklığının karesi kullanılmıştır. Mahalanobis uzaklıkları Minitab 14 ve SPSS 15 programı yardımıyla bulunmuştur.

Verilerin indirgenmesi: Temel bileşenler analizi kullanılarak kümeleme aşamasında değişken sayısının azaltılması amaçlanmıştır. Temel bileşenlerin aykırı değerlerden etkilenmemesi amacıyla ve elde edilecek temel bileşenlerin ileride aykırı değerlere duyarlı kümeleme algoritmalarında kullanılacağından dolayı, aykırı ilçelerin ayrı olarak indirgenip analiz edilmesine karar verilmiştir. Verilerin temel bileşenler analizine uygun olup olmadığına karar verebilmek için KMO istatistiği ve Bartlett' in küresellik testi sonuçları incelenmiştir.

Temel bileşenlerin, değişkenlerin farklı ölçüm birimlerinden etkilenmesini önlemek amacıyla korelasyon matrisi dikkate alınmıştır. Temel bileşenlerin yorumlanabilmesi amacıyla dik döndürme yöntemi olarak Varimax seçilmiş; temel bileşenlere ilişkin veriler ise regresyon yöntemiyle elde edilmiştir. Yine bu işlemler için SPSS 15 paket programı kullanılmıştır.

İlçelerin kümeleneşmesi: Seçilen bu temel bileşenlere ilişkin verilerle, ilk önce $K=4, 5, 6, 7$ küme sayısı için, çeşitli bağ yöntemlerini kullanan Hiyerarşik kümeleme yöntemi uygulanmıştır. Daha sonra aynı K değerleri için, 3 değişik başlangıç küme merkezi belirleme yöntemini kullanan K -ortalamalar kümeleme algoritması uygulanmıştır. Ayrıca yine $K=4, 5, 6, 7$ Hiyerarşik kümelemeden elde edilen küme

merkezlerini, K-ortalama algoritmasında başlangıç küme merkezi olarak kullanan bir Hibrid kümeleme algoritması kullanılmıştır. Aykırı ilçelerin kümelenmesi için sadece Hiyerarşik kümeleme yöntemi kullanılmıştır.

Tüm kümeleme sonuçları için ilçelerin kümelere dağılımı, küme merkezleri ve bu merkezler arasındaki Öklid uzaklıkları bulunmuştur. Ayrıca kümelerin oluşmasında hangi temel bileşenlerin etkili olduğu konusunda fikir edinmek amacıyla çok değişkenli varyans analizi sonuçları da incelenmiştir. Bütün bu işlemler SPSS 15 programı aracılığıyla gerçekleştirilmiştir.

Kümeleme sonuçlarının karşılaştırılması: Kümeleme sonuçlarının karşılaştırılması ve uygun bir K değerinin seçilmesi amacıyla, K-ortalamar algoritmalarında kullanılan amaç fonksiyonunun değeri, ortalama silhouette genişlikleri ve Davies-Bouldin İndeks değerleri her K değeri için tüm yöntemler için hesaplanmıştır. Bu amaçla Minitab 14' te oluşturulan makro programlar ve MS-DOS QBasic programlama diliyle oluşturulan programlar kullanılmıştır. Bu programlara ilişkin kodlar Ek-2, Ek-3 ve Ek-4' te verilmektedir.

Elde edilen amaç fonksiyonunun değerleri, ortalama silhouette genişlikleri ve Davies-Bouldin İndeks değerleri çizelgeler ve çizgi (line) grafikleri halinde sunulmuştur.

İlçelerin gelişmişlik düzeylerine göre tasnif edilmesi: Bir önceki adımda “en iyi” olarak seçilen kümeleme yöntemi ve K değeri için bulunmuş küme üyelikleri, aykırı ilçelerin kümelenmesi sonucu elde edilen küme üyelikleriyle birlikte dikkate alınarak, gelişmişlik düzeyleri sıralanmıştır. Böylece ilçelerin gelişmişlik düzeylerine göre listesi elde edilmiştir.

Ayrıca çeşitli gelişmişlik düzeylerine sahip ilçelerin Türkiye'nin coğrafi bölgelerine frekans ve oransal dağılımı çapraz tablo ile incelenmiş ve bu tablodaki bilgiler çubuk grafiği ile ifade edilmiştir. Bölgeler arası gelişmişlik sıralaması da yaklaşık olarak yapılmış ve gelişmişlik açısından bölgeler arası farklılıkların anlamlı olup olmadığı Ki-kare testi ve olabilirlik oranı ile sınımlanmıştır.

6. ARAŞTIRMA BULGULARI

Bu bölümde, Türkiye’deki ilçelerin 2008 yılı itibariyle sosyo-ekonomik gelişmişlik düzeylerini ölçmek ve kümelemek amacıyla elde edilebilen, Materyal ve Yöntem bölümünde tanımlanmış 23 değişkene ilişkin 957’ şer gözlemden oluşan verilerin (Ek 1) analizi yapılacaktır.

Bu amaçla Alt bölüm 6.1’de verilerin ön incelemesi, Alt bölüm 6.2’de verilerin boyut indirilmesi, Alt bölüm 6.3’te verilerin aşamalı (hiyerarşik) kümeleme yöntemi ile kümelenebilmesi, Alt bölüm 6.4’te verilerin K-ortalamlar kümeleme yöntemiyle kümelenebilmesi yapılacaktır.

Ayrıca Alt bölüm 6.5’te verilerin Alt bölüm 4.5’te tanımlanan karışım (hybrid) kümeleme algoritması ile kümelenebilmesi yapılacak, Alt bölüm 6.6’da ise daha önceki alt bölümlerde yapılan kümeleme sonuçları karşılaştırılarak, en iyi kümeleme sonucu seçilecektir. Son olarak Alt bölüm 6.7’de en iyi kümeleme için Türkiye’nin coğrafi bölgeleri, ilçelerin gelişmişlik düzeylerine göre incelenecektir.

6.1. Verilerin Ön İncelemesi

6.1.1. Tanımlayıcı istatistikler ve verilerin dağılımı

Bu alt bölümde ilk olarak verilerin dağılımlarına ilişkin tanımlayıcı istatistikler elde edilmiştir (Çizelge 6.1). Çizelge 6.1 incelendiğinde; ilçelerin yarısının nüfuslarının 30203 kişiden az olduğu, ilçelerde okuma bilmeyenlerin oranının ortalama % 10 civarında, ilk öğretim mezunlarının %13, orta öğretim mezunlarının oranının % 13 olduğu görülmektedir. Ayrıca ilçelerin yarısında yüksek öğrenim görenlerin oranlarının % 2.6’ dan düşük olduğu görülmektedir.

Medeni duruma göre ilçeler incelendiğinde ise ilçelerdeki hiç evlenmeyenlerin oranının ortalama % 26 civarında, evlenenlerin % 65, boşananların % 2 ve eşi ölenlerin % 7 olduğu görülmektedir.

Çizelge 6.1’deki tarımsal değişkenler incelendiğinde, ilçelerin yarısında üretilen tahıllar toplamı 17097 tondan, sebzeler toplamı 2648 tondan ve meyveler toplamı 3233 tondan azdır. Tarımda makineleşme adına ilçelerin yarısında 10 hektar başına traktör sayısının 3.5’den az olduğu anlaşılmaktadır. Ayrıca ilçelerin yarısında dıştan göç edenlerin oranı % 8.5’den azdır.

Çizelge 6.1 Verilerin dağılımına ilişkin tanımlayıcı özellikler

Değişkenler	En Küçük Değer	En Büyük Değer	Ortalama	Standart Sapma	Ortanca
Nüfus Toplamı	1853	785330	74730.5	118125.4	30203
Okuma Bilmeyen	0.013	0.271	0.097	0.046	0.092
Okuma Bilen	0.443	0.879	0.753	0.076	0.772
Bilinmeyen	0.014	0.231	0.057	0.026	0.052
Okuma B. Okul B.	0.081	0.401	0.198	0.053	0.184
İlkokul Mezunu	0.028	0.523	0.282	0.090	0.288
İlköğretim Mezunu	0.049	0.157	0.094	0.015	0.094
Ortaokul Mezunu	0.004	0.077	0.034	0.014	0.034
Lise Mezunu	0.018	0.299	0.112	0.047	0.105
Yo/Fakülte Mezunu	0.004	0.239	0.031	0.025	0.025
Yüksek Lisans Mezunu	0.000	0.045	0.002	0.003	0.001
Doktora Mezunu	0.000	0.013	0.001	0.001	0.0001
Hiç Evlenmemiş	0.140	0.617	0.266	0.070	0.250
Evli	0.354	0.872	0.648	0.054	0.660
Boşandı	0.001	0.076	0.019	0.012	0.018
Eşi Öldü	0.024	0.144	0.067	0.022	0.066
Süt Toplam	0.000	193423	12244,3	16373	6816.4
Tahıllar Toplam	0.000	1231263	51227,7	99375.2	17097
Sebzeler Toplam	0.000	813874	25315,7	69484.3	2648
Meyveler Toplam	0.000	838374	16246,8	48516.4	3233
Et Toplam	0.000	22465.3	443.3	1557.7	48.3
Göç Oranı	0.012	0.970	0.173	0.205	0.086
Hektar Başına Traktör	0.000	12.308	0.543	0.806	0.347

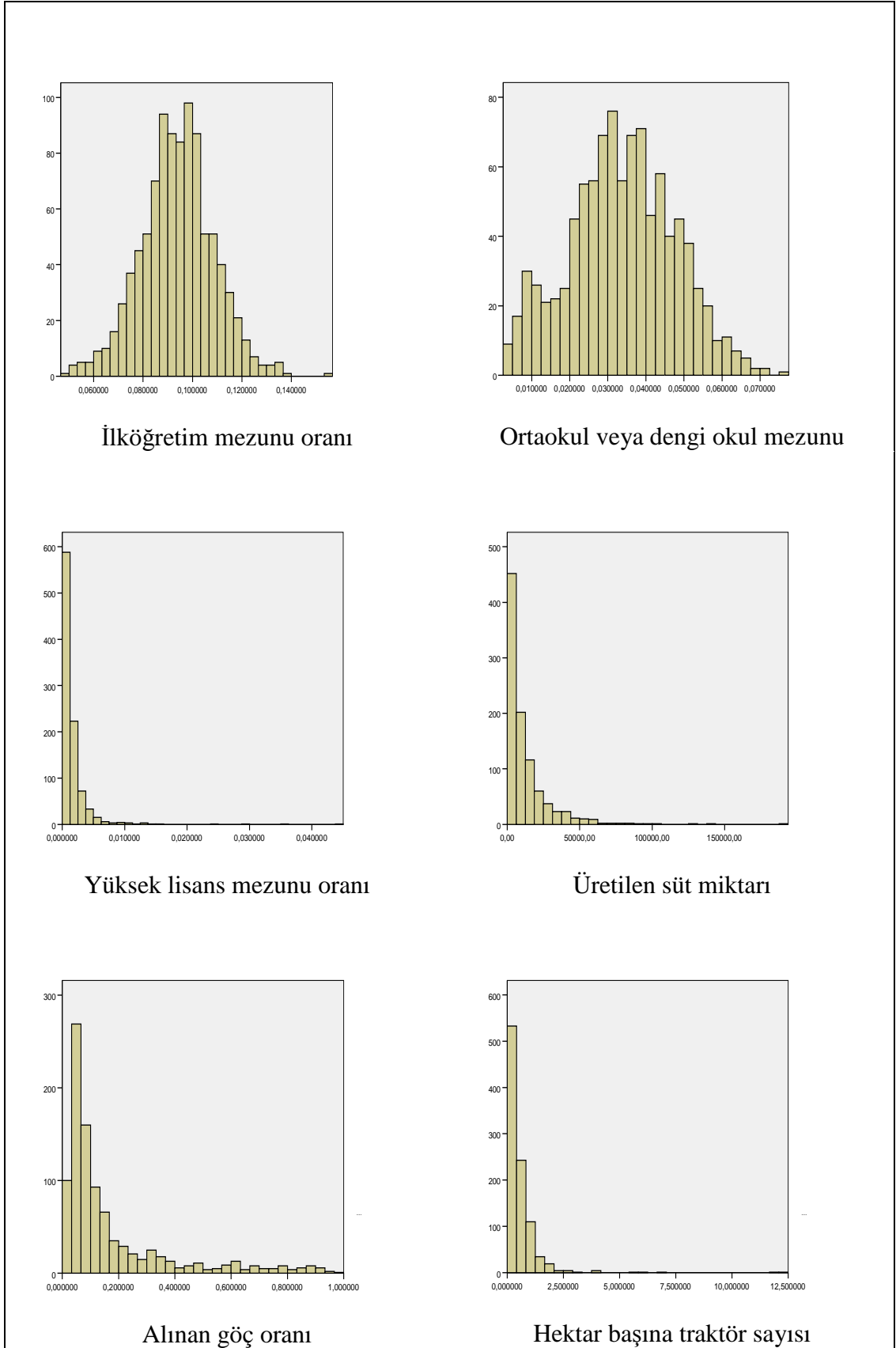
Bu değişkenlerin dağılımları bireysel olarak incelendiğinde ise, bazı değişkenlerin (ilköğretim mezunu oranı, ortaokul veya dengi okul mezunu, ... gibi) normale yakın dağıldığı saptanırken, bazı değişkenlerin ise (örneğin; nüfus toplamı, yüksek lisans mezunu, üretilen süt miktarı, ... gibi) ciddi anlamda özellikle sağa

çarpık olduğu tespit edilmiştir. Şekil 6.1' de bu değişkenlere ilişkin histogramlardan bazıları verilmektedir.

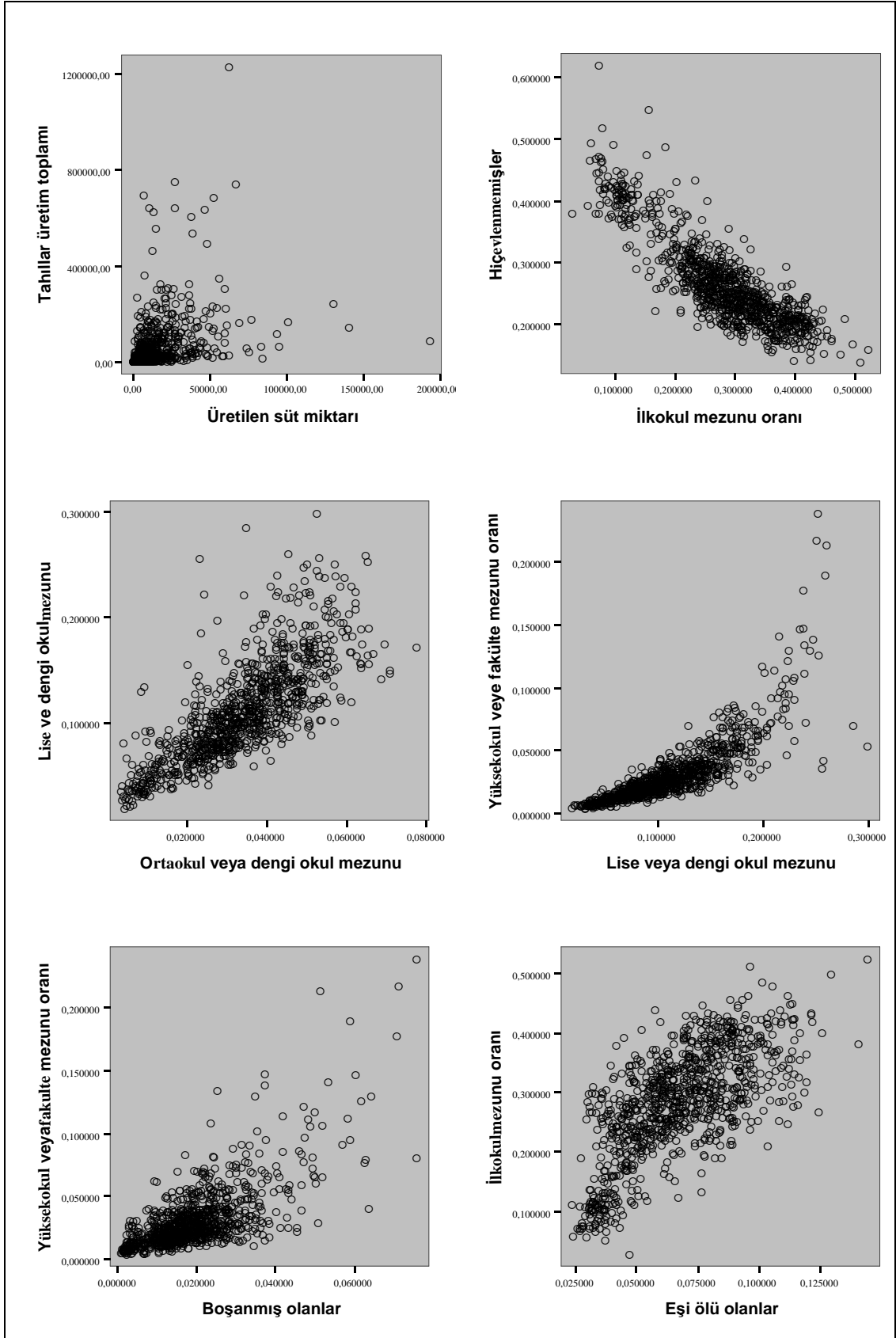
6.1.2. Değişkenler arasındaki doğrusal ilişkiler

Tüm değişkenlerin birbiriyle olan saçılım çizimleri incelenmiş ve bu çizimlerden bazıları Şekil 6.2' de verilmiştir. Değişkenler arasında dikkate değer bir doğrusal ilişki olduğu izlenimi Şekil 6.2' den edinilmektedir. Değişkenler arasında var olan doğrusal ilişki derecesi Çizelge 6.2' de verilmektedir. Bazı değişkenler arasında örneğin evli olanların oranı ile boşananların oranı arasında beklenildiği gibi yüksek bir ilişki mevcuttur.

Şekil 6.2' de dikkati çeken bazı hususlar vardır. Bunlardan ilki hiç evlenmeyenlerle ilkokul mezunu olanların ters yönlü bir doğrusal ilişkiye sahip olmalarıdır. Böyle bir ilişki ilkokul mezunlarının genellikle evlenme eğiliminde olduklarını belirtmektedir. Bir başka husus ise boşanmış olanlarla yüksek okul-fakülte mezunları arasındaki aynı yönlü doğrusal ilişkidir. Yüksekokul-fakülte mezunu oranı arttıkça boşanma oranının arttığı görülmektedir. Tuhaf diğer bir ilişkide aynı yönlü olan eşi ölenlerle ilkokul mezunu olanların oranı arasındadır.



Şekil 6.1 Bazı değişkenlere ait histogramlar



Şekil 6.2 Bazı değişkenler arasındaki saçılım grafikleri

Çizelge 6.2 Korelasyon matrisi

	NT	OKBYN	OKB	BYN	OBOB	IM	IOM	OOM	LM	YFM	YLM	DM	HEVS	EV	BOS	OL	ST	TT	SEBT	MT	ETT	GO	HBT
NT	1,00	-,38	,14	,07	-,07	-,20	-,06	,25	,37	,45	,42	,44	,13	-,07	,31	-,42	,09	,09	,14	,13	,26	,62	,12
OKBYN	-,38	1,00	-,82	,23	,62	-,41	,04	-,65	-,69	-,62	-,41	-,37	,34	-,35	-,59	,10	-,03	-,12	-,14	-,12	-,19	-,60	-,23
OKB	,14	-,82	1,00	-,61	-,84	,75	-,02	,70	,62	,53	,29	,25	-,67	,58	,60	,37	,03	,08	,13	,12	,13	,35	,25
BYN	,07	,23	-,61	1,00	,34	-,60	-,17	-,21	-,08	-,04	,07	,06	,54	-,54	-,10	-,33	-,06	-,12	-,12	-,12	-,07	,11	-,10
OBOB	-,07	,62	-,84	,34	1,00	-,70	,10	-,74	-,65	-,55	-,33	-,28	,67	-,51	-,66	-,50	,03	-,01	-,06	-,07	-,08	-,32	-,23
IM	-,20	-,41	,75	-,60	-,70	1,00	-,08	,32	,03	-,03	-,16	-,17	-,86	,76	,32	,69	,09	,11	,14	,09	,04	-,07	,23
IOM	-,06	,04	-,02	-,17	,10	-,08	1,00	-,19	-,12	-,22	-,19	-,19	,16	-,04	-,25	-,24	-,01	,08	-,02	,02	,03	-,11	-,08
OOM	,25	-,65	,70	-,21	-,74	,32	-,19	1,00	,77	,57	,35	,29	-,43	,33	,63	,20	-,13	-,04	-,03	,04	,09	,49	,14
LM	,37	-,69	,62	-,08	-,65	,03	-,12	,77	1,00	,82	,57	,52	-,07	-,01	,59	-,07	-,09	-,04	,00	,05	,12	,60	,12
YFM	,45	-,62	,53	-,04	-,55	-,03	-,22	,57	,82	1,00	,84	,80	-,08	-,01	,70	-,10	-,04	-,05	,04	,05	,12	,60	,14
YLM	,42	-,41	,29	,07	-,33	-,16	-,19	,35	,57	,84	1,00	,92	,06	-,14	,54	-,14	-,08	-,06	-,04	-,03	,05	,52	,12
DM	,44	-,37	,25	,06	-,28	-,17	-,19	,29	,52	,80	,92	1,00	,07	-,13	,48	-,16	-,06	-,04	-,04	-,02	,07	,44	,07
HEVS	,13	,34	-,67	,54	,67	-,86	,16	-,43	-,07	-,08	,06	,07	1,00	-,93	-,39	-,69	-,06	-,10	-,08	-,05	-,06	,05	-,20
EV	-,07	-,35	,58	-,54	-,51	,76	-,04	,33	-,01	-,01	-,14	-,13	-,93	1,00	,23	,45	,09	,15	,11	,08	,10	-,03	,18
BOS	,31	-,59	,60	-,10	-,66	,32	-,25	,63	,59	,70	,54	,48	-,39	,23	1,00	,13	-,09	,00	,07	,07	,07	,56	,22
OL	-,42	,10	,37	-,33	-,50	,69	-,24	,20	-,07	-,10	-,14	-,16	-,69	,45	,13	1,00	,00	-,06	-,06	-,06	-,11	-,37	,07
ST	,09	-,03	,03	-,06	,03	,09	-,01	-,13	-,09	-,04	-,08	-,06	-,06	,09	-,09	,00	1,00	,41	,27	,09	,31	-,18	,05
TT	,09	-,12	,08	-,12	-,01	,11	,08	-,04	-,04	-,05	-,06	-,04	-,10	,15	,00	-,06	,41	1,00	,30	,15	,24	-,11	-,02
SEBT	,14	-,14	,13	-,12	-,06	,14	-,02	-,03	,00	,04	-,04	-,04	-,08	,11	,07	-,06	,27	,30	1,00	,33	,13	,02	,10
MT	,13	-,12	,12	-,12	-,07	,09	,02	,04	,05	,05	-,03	-,02	-,05	,08	,07	-,06	,09	,15	,33	1,00	,11	,02	,06
ETT	,26	-,19	,13	-,07	-,08	,04	,03	,09	,12	,12	,05	,07	-,06	,10	,07	-,11	,31	,24	,13	,11	1,00	,11	,07
GO	,62	-,60	,35	,11	-,32	-,07	-,11	,49	,60	,60	,52	,44	,05	-,03	,56	-,37	-,18	-,11	,02	,02	,11	1,00	,24
HBT	,12	-,23	,25	-,10	-,23	,23	-,08	,14	,12	,14	,12	,07	-,20	,18	,22	,07	,05	-,02	,10	,06	,07	,24	1,00

Çizelge 6.2' de gösterilen kısaltmaların anlamı rahat bir şekilde incelenmek amacı ile aşağıda tablo halinde verilmiştir.

Çizelge 6.3 Korelasyon matrisindeki kısaltmaların açıklamaları

Değişkenler	Açıklamaları	Değişkenler	Açıklamaları
NT	Nüfus Toplamı	HEVS	Hiç Evlenmemiş Oranı
OKBYN	Okuma Bilmeyen Oranı	EV	Evli Oranı
OKB	Okuma Bilen Oranı	BOS	Boşandı Oranı
BYN	Bilinmeyen Oranı	OL	Eşi Öldü Oranı
OBOB	Okuma Bilen Okul Bitirmeyen	ST	Süt Toplam
IM	İlkokul Mezunu Oranı	TT	Tahıllar Toplam
IOM	İlköğretim Mezunu Oranı	SEBT	Sebzeler Toplam
OOM	Ortaokul Mezunu Oranı	MT	Meyveler Toplam
LM	Lise Mezunu Oranı	ETT	Et Toplam
YFM	Yüksekokul Fakülte Mezunu	GO	Göç Oranı
YLM	Yüksek Lisans Mezunu Oranı	HBT	Hektar Başına Traktör Sayısı
DM	Doktora Mezunu Oranı		

Çizelge 6.1' de verilen değişkenlerin bazılarının aynı tür bilgiyi veya benzer bir bilgiyi verdiği dikkate çekmektedir. Bu nedenle ortaokul mezunlarının oranı ile ilköğretim mezunlarının oranı birleştirilmiştir. Benzer biçimde yüksek lisans mezunu olanların oranı ile doktora mezunlarının oranı birleştirilmiştir. Ayrıca okuma bilen okul bitirmeyenlerin oranları analizden dışlanmıştır, çünkü bu bilgi okuma bilenlerin oranı içinde verilmektedir. Dolayısıyla analize 20 değişken ve 957 gözlem üzerinden devam edilmektedir.

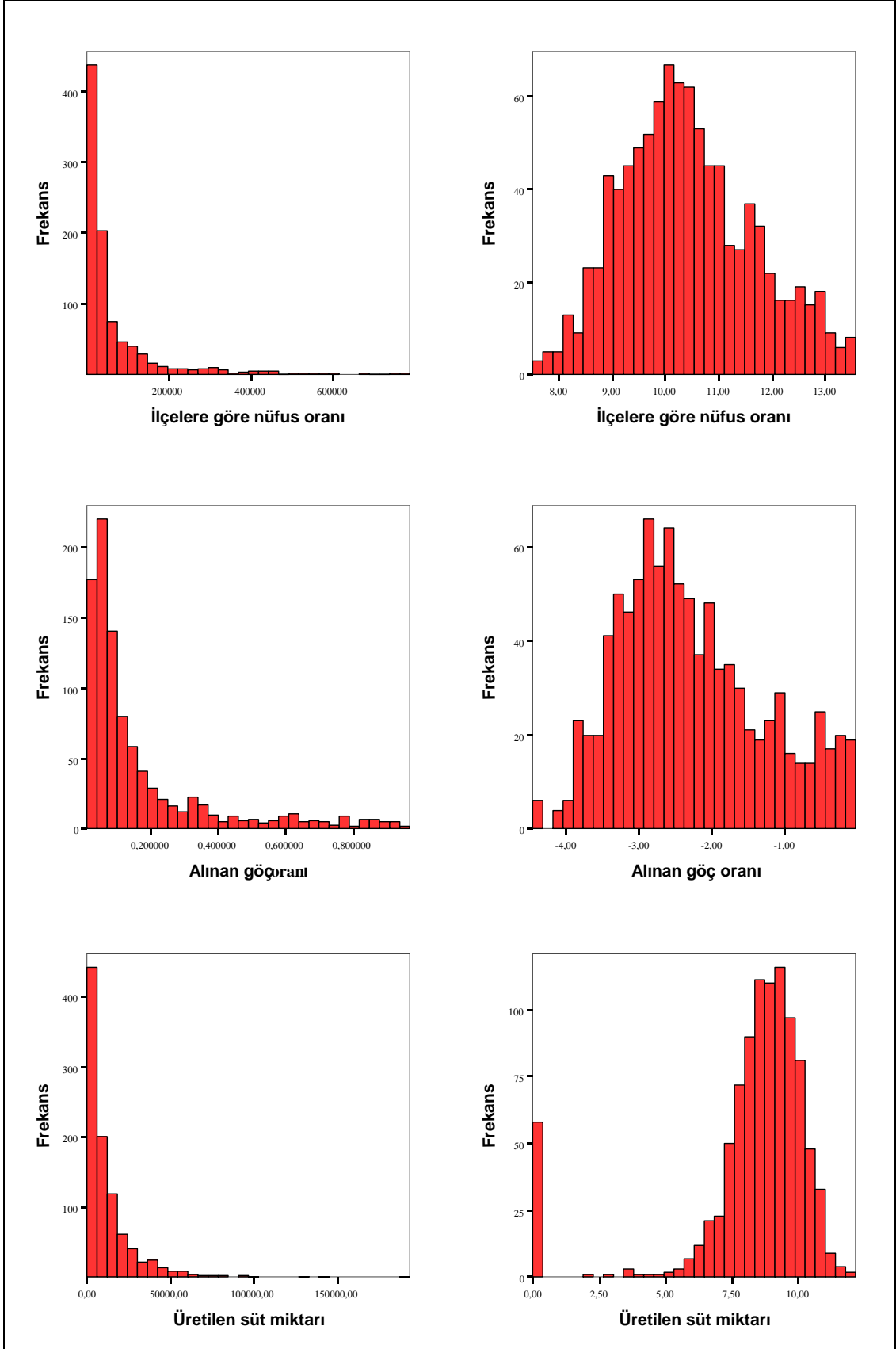
6.1.3. Verilere uygulanan dönüşümler

Bazı değişkenlerde görülen aşırı sağa çarpıklığı gidermek ve aykırı gözlemlerin etkisini hafifletmek amacıyla, bu değişkenlere doğal logaritma dönüşümü yapılmıştır. Gözlemlerin 0 değeri aldığı durumlarda, çok küçük bir değer 0' a eklenerek dönüşüm yapılabilmiştir. Dönüşüm yapılan değişkenlerin listesi Çizelge 6.4' de verilmektedir.

Çizelge 6.4 Dönüşüm yapılan değişkenlerin listesi

Orijinal Değişkenler	Dönüştürülmüş Değişkenler
Nüfus Toplamı	Ln Nüfus Toplamı
Yüksekokul Fakülte Mezunu	Ln Yüksekokul Fakülte Mezunu
Yüksek Lisans+Doktora Mezunu	Ln Yüksek Lisans+Doktora Mezunu
Boşandı	Ln Boşandı
Süt Toplamı	Ln Süt Toplamı
Tahıllar Toplamı	Ln Tahıllar Toplamı
Sebzeler Toplamı	Ln Sebzeler Toplamı
Meyveler Toplamı	Ln Meyveler Toplamı
Et Toplamı	Ln Et Toplamı
Göç Oranı	Ln Göç Oranı
Hektar Başına Traktör Oranı	Ln Hektar Başına Traktör Oranı

Dönüştürülmüş bazı değişkenlere ilişkin histogramlar Şekil 6.3' te verilmektedir. Bu histogramlardan dönüşümlerin dağılımın çarpıklığını giderme adına etkinliği görülmektedir.



Şekil 6.3 Dönüşümü yapılmış bazı değişkenlere ait histogramlar

6.1.4. Aykırı gözlemlerin (ilçelerin) belirlenmesi

Dikkate alınan bazıları dönüştürülmüş 20 değişkene göre aykırı ilçelerin tespiti için ilk önce saçılım grafikleri incelenmiştir. Daha sonra her değişkenin 10. yüzdeliği ve 90. yüzdeliği bulunup, ilçelerin 20 değişken üzerinden kaç değişken için 10. yüzdeliğin altında veya 90. yüzdeliğin ötesinde değer aldığı tespit edilmiştir.

Ayrıca aykırı ilçeleri çok boyutlu uzayda tespit etmenin bir yolu olarak Mahalanobis uzaklığının karesi kullanılmıştır (Mahalanobis, 1936).

$$D = (X - \bar{X})'S^{-1}(X - \bar{X}) \quad (6.1)$$

Büyük örneklem için bu uzaklığın karesinin p serbestlik derecesi ile yaklaşık olarak ki-kare (χ_p^2) dağılımlı olduğu söylenebilmektedir. Bu nedenle her ilçe için bu uzaklığın karesi Minitab 14 ve SPSS 15 paket programları kullanılarak hesaplanmış ve $\chi_{20,0.99}^2 = 37.5662$ tablo değeri ile karşılaştırılmıştır. İlgili tablo değerinden büyük olan uzaklıklara sahip ilçeler belirlenmiş ve 20 değişken açısından teker teker incelenmiştir. Bu işlemler sonucu elde edilen ilçeler Çizelge 6.5' de verilmektedir.

6.2. Veri İndirgeme

Bu alt kesimde temel bileşenler analizi kullanılarak kümeleme aşamasında değişken sayısının azaltılması amaçlanmıştır. Temel bileşenlerin aykırı değerlerden etkilenmemesi amacıyla ve elde edilecek temel bileşenlerin ileride aykırı değerlere duyarlı kümeleme algoritmalarında kullanılacağından dolayı, Çizelge 6.5' deki aykırı ilçelerin ayrı olarak analiz edilmesine karar verilmiştir. İçinde Çizelge 6.4' deki dönüştürülmüş değişkenlerinde bulunduğu 20 değişkene ilişkin 882 gözlem (ilçe) kullanılarak temel bileşenler analizi SPSS 15 programı aracılığıyla gerçekleştirilmiştir.

Temel bileşenler analizinin değişkenlerin ölçüm biriminden etkilenmemesi için korelasyon matrisi kullanılmış ve döndürme yöntemi olarak Varimax tercih edilmiştir.

KMO istatistiği 0.684 olarak bulunmuştur. Bu değer 1' e yakın olması faktör analizinin veriler için uygun olduğu kanısını oluşturmaktadır. Korelasyon matrisinin birim matris olup olmadığını denetleyen Bartlett' in küresellik testi için test istatistiği 23376.575, serbestlik derecesi 190 ve anlamlılık düzeyi 0.000 olarak elde edilmiştir.

Buradan verilerin temel bileşenler analizi yapılması için elverişli olduğu kanaatine ulaşılır.

Çizelge 6.5 Aykırı ilçeler tablosu

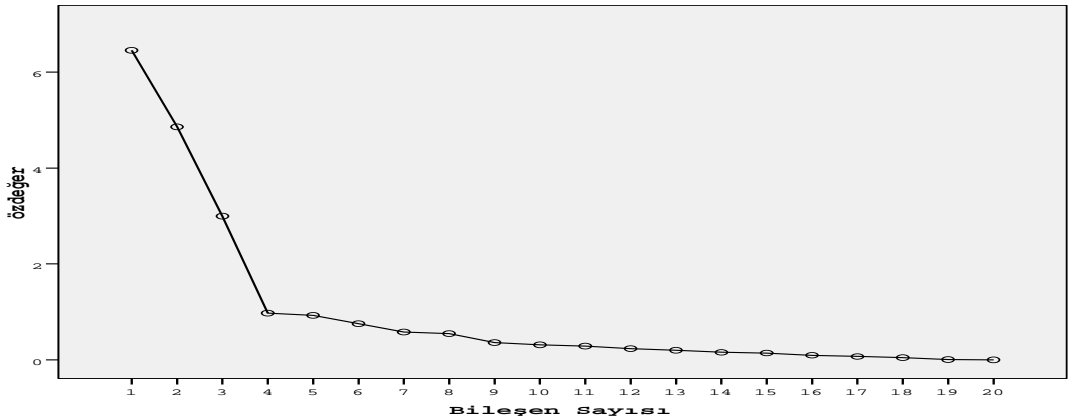
	İL	İLÇE	İL	İLÇE	İL	İLÇE
1	İstanbul	Adalar	İstanbul	Esenler	Karabük	Ovacık
2	Şanlıurfa	Akçakale	Ankara	Etimesgut	Mardin	Ömerli
3	Ankara	Altındağ	İstanbul	Eyüp	Ağrı	Patnos
4	Konya	Altıneki	İstanbul	Fatih	Siirt	Pervari
5	Yalova	Armutlu	Şırnak	Güçlükön	Kastamonu	Pınarbaş
6	İstanbul	Arnavutk	İstanbul	Güngören	Tunceli	Pülümür
7	Mersin	Aydıncık	Ağrı	Hamur	İstanbul	Sarıyer
8	İstanbul	Bağcılar	Şanlıurfa	Harran	Şırnak	Silopi
9	Van	Bahçesaray	Rize	Hemşin	İstanbul	Sultanbeyli
10	İstanbul	Bakırköy	Bitlis	Hizan	Diyarbakır	Sur
11	İstanbul	Başakşehir	Tunceli	Hozat	Hakkâri	Şemdinli
12	Van	Başkale	İstanbul	Kadıköy	İstanbul	Şişli
13	İstanbul	Bayrampaşa	İzmir	Karşıyaka	Erzurum	Tekman
14	İstanbul	Beşiktaş	Ankara	Keçiören	Ağrı	Tutak
15	Çanakkale	Bozcaada	Burdur	Kemer	Isparta	Uluborlu
16	Ankara	Çankaya	Kırklareli	Kofçaz	İstanbul	Ümraniye
17	Rize	Çayeli	İzmir	Konak	İstanbul	Üsküdar
18	Hakkâri	Çukurca	Antalya	Konyaaltı	Şanlıurfa	Viranşehir
19	Ardahan	Damal	Muş	Malazgirt	Kırıkkale	Yahşihan
20	Mardin	Dargeçit	Ağrı	Merkez	Konya	Yalılıyü
21	İzmir	Dikili	Hakkâri	Merkez	Bingöl	Yayladere
22	Sinop	Dikmen	Tunceli	Merkez	Bingöl	Yedisu
23	Kocaeli	Dilovası	Mardin	Midyat	Ankara	Yenimahalle
24	Van	Erciş	Eskişehir	Mihalgazi	Mardin	Yeşilli
25	Siirt	Eruh	Bursa	Nilüfer	Bursa	Yıldırım

Temel bileşen analizi sonucu 5 tane bileşen seçilmiştir (Çizelge 6.6). Çizelge 6.6' den görüldüğü gibi temel bileşenler arasından ilk 3 bileşene ilişkin öz değer 1 değerini aşmaktadır. Bu 3 bileşen ile toplam varyansın sadece % 71.544' ü açıklanmaktadır.

Çizelge 6.6 Temel bileşenler analizi sonucu

Bileşen	Başlangıç Özdeğerler		
	Öz değer	Varyans Yüzdesi	Toplam Yüzde
1	6.456	32.280	32.280
2	4.856	24.280	56.560
3	2.997	14.985	71.544
4	0.976	4.880	76.424
5	0.926	4.629	81.053

Ancak 4. ve 5. temel bileşenlere ilişkin öz değerlerin 1 değerine oldukça yakın olması; temel bileşenlerin kümeleme analizinde kullanılacak olması ve dolayısıyla mümkün olan en fazla bilgiyle kümeleme analizine gitmek adına, bu bileşenlerin de analize katılması uygun görülmüştür. Sonuçta 5 bileşen ile kümelemeye gidilmesine karar verilmiştir. Elde edilen bu 5 bileşen 20 değişkene ilişkin toplam değişkenliğin % 81.053' ünü açıklamaktadır. Şekil 6.4' de 5 temel bileşenin seçilmesinin uygun olduğunu desteklemektedir.



Şekil 6.4 Temel bileşenlere ilişkin scree plot

Ayrıca her bir değişkendeki değişimin bu 5 temel bileşence açıklanma miktarı Çizelge 6.7’ de verilmektedir. Her bir değişkenin varyansının açıklanma oranı incelendiğinde; en küçüğü 0.647, en büyüğü ise 0.939 olan bu değerlerin kabul edilebilir olduğu kanısına varılmıştır.

Çizelge 6.7 Değişkenlere ait varyans açıklanma oranları

	Açıklanma Oranı
ln_nuf_top	0.879
okuma_bilmeyen	0.795
okuma_bilen	0.937
bilinmeyen	0.753
ilkokul_mez	0.907
orta_ilköğretim	0.778
lise_mez	0.888
ln_yüksekokul_fakülte_mez	0.902
ln_yuklis_dok	0.788
hic_evlenmemiş	0.939
evli	0.850
ln_bosandı	0.750
eşi_öldü	0.871
ln_süt_top	0.840
ln_tahıllar_top	0.792
ln_sebzeler_top	0.749
ln_meyveler_top	0.647
ln_et_top	0.739
ln_göç_oranı	0.752
ln_hektar_bas_traktör	0.654

Değişkenler ile temel bileşenler arasındaki ilişki katsayısını gösteren faktör yükleri Çizelge 6.8’ de verilmektedir. Çizelge 6.8’ deki yüklere göre, temel bileşenlerin anlamı yaklaşık olarak Çizelge 6.9’ daki gibi ifade edilebilmektedir.

Çizelge 6.8 Değişkenlere ait döndürülmüş bileşen matrisi

	Bileşen				
	1	2	3	4	5
ln_yüksekokul_fakülte_mez	0,923	0,184	-0,011	0,046	0,119
lise_mez	0,903	-0,022	-0,101	-0,001	0,248
ln_yuklis_dok	0,867	-0,043	-0,099	0,157	-0,018
ln_göç_oranı	0,804	-0,007	-0,148	0,288	-0,022
okuma_bilmeyen	-0,732	-0,363	0,041	-0,306	-0,180
ln_bosandı	0,673	0,527	0,031	-0,122	0,061
hic_evlenmemiş	-0,135	-0,931	-0,113	0,200	0,035
evli	0,058	0,913	0,101	0,050	0,020
ilkokul_mez	0,119	0,911	0,159	-0,190	0,042
okuma_bilen	0,609	0,683	0,057	-0,004	0,311
bilinmeyen	0,058	-0,638	-0,173	-0,112	-0,548
ln_süt_top	-0,219	-0,005	0,881	-0,070	-0,104
ln_tahıllar_top	-0,228	0,081	0,856	-0,008	-0,015
ln_sebzeler_top	0,045	0,142	0,837	0,149	0,061
ln_meyveler_top	0,046	0,098	0,776	-0,008	0,181
ln_hektar_bas_traktör	-0,082	0,091	0,774	-0,151	-0,128
ln_et_top	0,390	0,223	0,598	0,417	-0,074
ln_nuf_top	0,407	-0,165	0,058	0,825	-0,044
eşi_öldü	-0,054	0,598	0,125	-0,692	-0,124
orta_ilköğretim	0,394	0,040	-0,079	-0,024	0,784

Çizelge 6.9 Temel bileşenlerin anlamı

Bileşenler	Anlamı
1	Lise ve Üzeri Mezunu Eğitim Düzeyi
2	Medeni Durum
3	Tarım
4	Nüfus
5	İlköğretim Mezunu Eğitim Düzeyi

Seçilen 5 temel bileşenin değerlerini hesaplamak için Çizelge 6.10' daki bileşenler katsayı matrisi kullanılmıştır. Elde edilen veriler Ek-5' de verilmektedir.

Çizelge 6.10 Bileşenler katsayı matrisi

	Bileşen				
	1	2	3	4	5
ln_nuf_top	-0,035	0,076	-0,016	0,548	-0,113
okuma_bilmeyen	-0,083	-0,085	0,026	-0,161	-0,018
okuma_bilen	0,063	0,116	-0,007	-0,009	0,144
bilinmeyen	0,197	-0,173	0,018	-0,216	-0,486
ilkokul_mez	-0,031	0,238	-0,026	0,005	-0,073
orta_ilköğretim	0,005	-0,128	0,024	-0,126	0,727
lise_mez	0,246	-0,144	0,041	-0,220	0,118
ln_yüksekokul_fakülte_mez	0,239	-0,058	0,039	-0,142	-0,039
ln_yuklis_dok	0,238	-0,079	0,019	-0,075	-0,149
hic_evlenmemiş	0,018	-0,255	0,042	-0,005	0,156
evli	-0,109	0,311	-0,070	0,226	-0,096
ln_bosandı	0,166	0,059	0,010	-0,146	-0,095
eşi_öldü	0,096	0,060	0,029	-0,433	-0,160
ln_süt_top	0,023	-0,073	0,247	-0,086	-0,045
ln_tahıllar_top	-0,020	-0,036	0,226	-0,014	0,034
ln_sebzeler_top	0,008	-0,020	0,220	0,065	0,056
ln_meyveler_top	0,030	-0,080	0,225	-0,075	0,186
ln_et_top	0,047	0,066	0,136	0,252	-0,157
ln_göç_oranı	0,180	-0,023	-0,015	0,062	-0,154
ln_hektar_bas_traktör	0,070	-0,061	0,221	-0,154	-0,101

6.2.1. Aykırı gözlemler üzerinde veri indirgeme

Bu alt bölümde de Çizelge 6.5' de verilen aykırı ilçeler için temel bileşenler analizi yapılmaktadır. Temel bileşenler analizinin değişkenlerin ölçüm biriminden

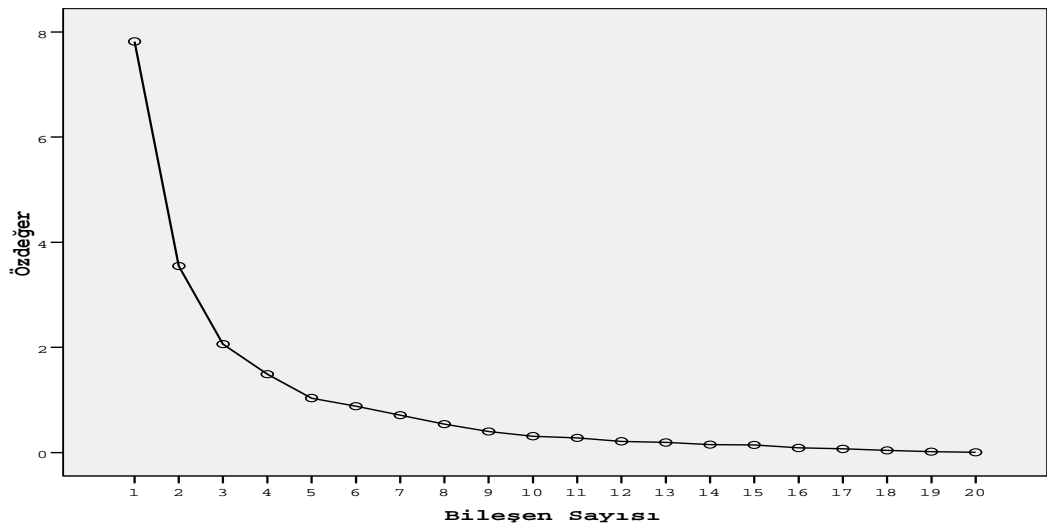
etkilenmemesi için yine korelasyon matrisi kullanılmaktadır. Ayrıca bu durum için yine Varimax döndürme yöntemi uygulanmıştır.

KMO istatistiği 0.737 olarak elde edilmiştir. Bu değerin 1' e yakın olması faktör analizinin veriler için uygun olduğunu belirtmektedir. Bartlett'in küresellik testi için test istatistiği 1629.72, serbestlik derecesi 190 ve anlamlılık düzeyi p 0.000 olarak bulunmuştur. Buradan temel bileşenler analizinin yapılmasının uygun olduğu kararına ulaşılır.

Aykırı 75 gözlem için 5 tane temel bileşenin uygun olduğu ve bu temel bileşenlerin değişkenliğin % 79.768' ini açıklayabildiği tespit edilmiştir. Çizelge 6.11' den görüldüğü gibi 5 bileşene ilişkin öz değerler 1 değerini aşmaktadır.

Çizelge 6.11 Aykırı gözlemler için bileşenlere ait açıklama yüzdeleri

Bileşen	Başlangıç Özdeğerler		
	Toplam	Varyans Yüzdesi	Toplam Yüzde
1	7.816	39.082	39.082
2	3.548	17.740	56.822
3	2.064	10.322	67.144
4	1.489	7.443	74.587
5	1.036	5.181	79.768



Şekil 6.5 Aykırı gözlemler için bileşenlere ait scree plot

Bu sonuçlar aykırı gözlemleri dışlayarak yapılan analiz sonuçları ile benzer bir yapı göstermektedir. Ayrıca Şekil 6.5' de 5 temel bileşen seçimini desteklemektedir.

Ayrıca bu 5 temel bileşenin her bir değişkeni açıklama miktarları Çizelge 6.12' de verilmektedir. Çizelge 6.12' e göre en küçük açıklanma oranı 0.537, en büyüğü ise 0.950' dir. Burada verilmemesine rağmen bileşen puanı katsayıları yardımıyla 5 temel bileşene ilişkin veriler oluşturulmuştur.

Çizelge 6.12 Aykırı gözlemler için değişkenlere ait varyans açıklanma oranı

	Açıklanma Oranı
ln_nuf_top	0.870
okuma_bilmeyen	0.772
okuma_bilen	0.950
bilinmeyen	0.537
ilkokul_mez	0.879
orta_ilköğretim	0.668
lise_mez	0.902
ln_yüksekokul_fakülte_mez	0.897
ln_yuklis_dok	0.736
hic_evlenmemiş	0.902
evli	0.846
ln_bosandı	0.872
eşi_öldü	0.752
ln_süt_top	0.777
ln_tahıllar_top	0.744
ln_sebzeler_top	0.713
ln_meyveler_top	0.806
ln_et_top	0.832
ln_göç_oranı	0.833
ln_hektar_bas_traktör	0.665

6.3. Hiyerarşik Kümeleme Sonuçları

Uygulama aşamasının bu bölümünde, SPSS 15 paket programı aracılığıyla EK-5' de verilen temel bileşenlere ait veriler kullanılarak 882 ilçe birleştirici aşamalı kümeleme yöntemi ile kümelere ayrılacaktır. Aykırı gözlem ilan edilen 75 ilçe ise ayrıca kümelenecektir.

Hiyerarşik yöntemin sonuçlarına geçmeden önce verilerin kaç kümeyle ayrılacağı araştırılmıştır. Bu kapsamda geçmişte yapılmış ilgili çalışmalarda dikkate alınarak ilçelerin ilk başta 3 ila 10 küme sayısı arasında gruplara ayrılacağına karar verilmiştir.

Daha sonra tam bağlantı, tek bağlantı, grup içi ortalama bağlantı, gruplar arası ortalama bağlantı, Ward' ın yöntemi, medyan, küme merkezi (centroid) gibi, gözlemleri veya kümeleri birleştiren birçok yöntem; Öklid uzaklığının karesi, blok, Chebychev gibi çeşitli uzaklıklarla teker teker 3 ila 10 arasındaki küme sayıları için uygulanmıştır.

Sonuçlar incelendiğinde büyük yığılmalara sahip kümelere ve/veya küçük kümelere neden olan yöntemler ve uzaklıklar dışlanmıştır: Grup içi ortalama yöntemi ile Ward' ın yönteminin Öklid uzaklığının karesi için $K=4, 5, 6$ ve 7 küme sayısı için iyi sonuçlar verdiği görülmüştür. Dolayısı ile tüm kümeleme çalışmaları bu iki yöntem, Öklid uzaklığının karesi ve $K=4, 5, 6$ ve 7 küme sayılarına tabi tutularak gerçekleştirilmiştir. Gözlem sayısının fazla olması nedeniyle sonuçları sunarken ağaç dalı şeması (dendrogram) verilememiştir.

Bu alt bölümde bulunan küme merkezleri, hibrid kümeleme algoritmasında K -ortalamlar yöntemine başlangıç küme merkezleri olarak kullanılacaktır.

6.3.1. Grup içi ortalama bağlantı yöntemi sonuçları

Grup içi bağlantı yöntemi kullanılarak $K=4, 5, 6$ ve 7 küme sayıları için gerçekleştirilen hiyerarşik kümelemeye ilişkin küme merkezleri ve küme büyüklükleri Çizelge 6.13' de verilmektedir. Ayrıca küme merkezlerinin birbirine olan Öklid uzaklıkları da bulunmuştur (Çizelge 6.14).

Çizelge 6.13 Grup içi ortalama bağlantı yöntemine göre küme büyüklükleri ve küme merkezleri

		Frekans	1. temel bileşen	2. temel bileşen	3. temel bileşen	4. temel bileşen	5. temel bileşen
K=4	1	491	0,376	0,327	0,391	0,078	-0,286
	2	269	-0,469	-0,106	-0,151	-0,594	0,788
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	79	-1,229	-1,809	0,043	0,781	-0,965
Toplam		882					
K=5	1	450	0,407	0,374	0,462	0,237	-0,143
	2	269	-0,469	-0,106	-0,151	-0,594	0,788
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	79	-1,229	-1,809	0,043	0,781	-0,965
	5	41	0,028	-0,188	-0,383	-1,667	-1,849
Toplam		882					
K=6	1	450	0,407	0,374	0,462	0,237	-0,143
	2	269	-0,469	-0,106	-0,151	-0,594	0,788
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	22	-1,383	-1,707	-0,309	0,789	-2,484
	5	57	-1,170	-1,849	0,179	0,778	-0,379
	6	41	0,028	-0,188	-0,383	-1,667	-1,849
Toplam		882					
K=7	1	450	0,407	0,374	0,462	0,237	-0,143
	2	140	-0,188	-0,690	0,020	-0,538	1,116
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	22	-1,383	-1,707	-0,309	0,789	-2,484
	5	57	-1,170	-1,849	0,179	0,778	-0,379
	6	129	-0,774	0,528	-0,338	-0,655	0,432
	7	41	0,028	-0,188	-0,383	-1,667	-1,849
Toplam		882					

Küme sayısı $K=7$ iken 3 büyük kümenin varlığı söz konusu iken, küme sayısı azaldıkça 2 büyük kümenin varlığını devam ettirdiği görülmektedir. $K=7$ ' den $K=6$ ' ya geçerken 2 ve 6 numaralı kümeler birleşmektedirler. Çizelge 6.14' den bu kümelerin merkezlerinin birbirlerine en yakın olduğu gözükmektedir. Ancak küme merkezleri birbirine en yakın olmamalarına rağmen, $K=6$ ' dan $K=5$ ' e geçerken 4 ve 5 numaralı kümelerin birleştikleri görülmektedir. Bir sonraki adımda $K=4$ ' e geçerken 1 ve 5 numaralı kümeler birbirlerine en yakın küme merkezlerine sahip olmamalarına rağmen birleşmektedir. Genel olarak birleşen kümelerin, merkezlerinin birbirlerine en yakın olmasa bile yakın olduğu dikkat çekmektedir.

Temel bileşenlerin Alt bölüm 6.2' deki isimlendirilmeleri dikkate alındığında, her kümenin merkezi, o kümeyi oluşturan ilçelerin hangi özellikleri taşıdıkları hakkında bilgi vermektedir. Örneğin $K=4$ alındığında öteki kümelere kıyasla, 3. kümedeki ilçelerin lise ve üstü eğitim oranının ve nüfuslarının yüksek olduğu, 1. kümedeki ilçelerin medeni durum oranlarında ve tarımsal üretimde iyi olduğu ve 2. kümedeki ilçelerde ilköğretim veya ortaokul mezunu oranının daha fazla olduğu söylenebilir.

Çizelge 6.14 Grup içi ortalama bağlantı yöntemine göre kümeler arası Öklid uzaklıkları

Küme	1	2	3	4	
1	0	1,675	4,260	2,868	
2		0	4,283	2,914	
3			0	4,860	
4				0	
Küme	1	2	3	4	5
1	0	1,714	4,268	2,933	2,777
2		0	4,283	2,914	2,901
3			0	4,860	4,955
4				0	3,342
5					0

Tablo 6.14 Grup içi ortalama bağlantı yöntemine göre kümeler arası Öklid uzaklıkları (devamı)

Küme	1	2	3	4	5	6	
1	0	1,714	4,268	3,731	2,805	2,777	
2		0	4,283	4,006	2,625	2,901	
3			0	5,197	4,861	4,955	
4				0	2,175	3,278	
5					0	3,557	
6						0	
Küme	1	2	3	4	5	6	7
1	0	1,969	4,268	3,731	2,805	1,786	2,777
2		0	4,468	4,159	2,511	1,562	3,245
3			0	5,197	4,861	4,227	4,955
4				0	2,175	3,996	3,278
5					0	2,965	3,557
6						0	2,719
7							0

6.3.2. Ward' ın yöntemine ilişkin sonuçları

Ward' ın yöntemi kullanılarak gerçekleştirilen hiyerarşik kümeleme elde edilen kümeleme sonuçları Çizelge 6.15' de gösterilmektedir. Küme merkezlerinin Öklid uzaklıkları ise Çizelge 6.16' da verilmektedir.

Grup içi ortalama bağlantı yönteminden farklı olarak $K=7$ için 4 büyük kümenin varlığı dikkati çekmektedir. Daha sonraki aşamalarda 3 büyük kümeye ve nihayetinde grup içi ortalama yöntemine benzer olarak $K=4$ için 2 büyük kümeye ulaşılmıştır.

Çizelge 6.15 Ward yöntemine göre küme büyüklükleri ve küme merkezleri

		Frekans	1. temel bileşen	2. temel bileşen	3. temel bileşen	4. temel bileşen	5. temel bileşen
K=4	1	416	0,381	0,332	0,420	0,143	-0,371
	2	326	-0,305	0,072	-0,076	-0,544	0,716
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	97	-1,009	-1,779	0,049	0,598	-0,863
Toplam		882					
K=5	1	124	1,462	-0,201	0,554	0,589	0,031
	2	326	-0,305	0,072	-0,076	-0,544	0,716
	3	292	-0,078	0,560	0,364	-0,046	-0,542
	4	43	0,902	0,244	-3,605	1,393	0,111
	5	97	-1,009	-1,779	0,049	0,598	-0,863
Toplam		882					
K=6	1	124	1,462	-0,201	0,554	0,589	0,031
	2	326	-0,305	0,072	-0,076	-0,544	0,716
	3	243	-0,120	0,703	0,509	0,289	-0,333
	4	43	0,902	0,244	-3,605	1,393	0,111
	5	97	-1,009	-1,779	0,049	0,598	-0,863
	6	49	0,130	-0,149	-0,358	-1,711	-1,580
Toplam		882					
K=7	1	124	1,462	-0,201	0,554	0,589	0,031
	2	205	-0,683	0,441	-0,244	-0,512	0,543
	3	121	0,335	-0,553	0,208	-0,598	1,008
	4	243	-0,120	0,703	0,509	0,289	-0,333
	5	43	0,902	0,244	-3,605	1,393	0,111
	6	97	-1,009	-1,779	0,049	0,598	-0,863
	7	49	0,130	-0,149	-0,358	-1,711	-1,580
Toplam		882					

$K=7$ ' den $K=6$ ' ya geçerken 2 ve 3 numaralı kümelerin birleştiği Çizelge 6.15' den ve birleşen bu kümelerin merkezlerinin en yakın olmasa bile yine yakın oldukları Çizelge 6.16' den görülmektedir. Bir sonraki aşamada $K=6$ ' dan $K=5$ ' e geçerken 3 ve 6 numaralı kümelerin birleşmektedir. Bu iki kümenin merkezlerinin uzaklığı 4. en küçük uzaklıktır. Son aşamada $K=4$ ' e geçerken 1 ve 3 numaralı kümeler birleşmektedir. Bu iki kümenin merkezinin uzaklığı 2. en küçük uzaklıktır.

Temel bileşenlerin isimlendirilmeleri ve küme merkezleri dikkate alındığında, Ward' ın yöntemi grup içi ortalama bağlantı yöntemine benzer kümeleme sonuçları vermiştir. Örneğin $K=4$ alındığında yine $K=4$ alındığında öteki kümelere kıyasla, 3. kümedeki ilçelerin lise ve üstü eğitim oranının ve nüfuslarının en yüksek olduğu, 1. kümedeki ilçelerin medeni durum oranlarında ve tarımsal üretimde iyi olduğu ve 2. kümedeki ilçelerde ilköğretim veya ortaokul mezunu oranının daha fazla olduğu söylenebilir.

Çizelge 6.16 Ward yöntemine göre kümeler arası Öklid uzaklıkları

Küme	1	2	3	4			
1	0	1,563	4,276	2,643			
2		0	4,251	2,783			
3			0	4,765			
4				0			
Küme	1	2	3	4	5		
1	0	2,314	1,930	4,298	3,108		
2		0	1,522	4,251	2,783		
3			0	4,395	2,638		
4				0	4,765		
5					0		
Küme	1	2	3	4	5	6	
1	0	2,314	1,884	4,298	3,108	3,241	
2		0	1,604	4,251	2,783	2,637	
3			0	4,428	2,746	2,665	
4				0	4,765	4,878	
5					0	3,158	
6						0	
Küme	1	2	3	4	5	6	7
1	0	2,672	1,970	1,884	4,298	3,108	3,241
2		0	1,568	1,539	4,204	2,888	2,640
3			0	2,113	4,503	2,877	2,910
4				0	4,428	2,746	2,665
5					0	4,765	4,878
6						0	3,158
7							0

6.3.3. Aykırı ilçeler için hiyerarşik kümeleme uygulaması

Analizden en iyi verimi almak için dışlanan 75 gözlem (ilçe) bu bölümde ayrı olarak analize tabi tutularak ve elde edilen sonuçlar daha sonra diğer 882 gözlem için elde edilen sonuçlarla birleştirilecektir.

Bu 75 ilçe, az gelişmiş ve çok gelişmiş olarak dikkate alınabileceğinden, kümelemede küme sayısının 2 olmasına karar verilmiştir.

Aykırı ilçeler hem grup içi ortalama bağlantı hem de Ward' ın yöntemine tabi tutulmuştur. Ancak Ward' ın yönteminin oluşturduğu kümelerin içinde bir arada olmaması gereken ilçelerin yan yana olması nedeni ile bu yöntemin sonuçları dikkate alınmamıştır. Grup içi ortalama bağlantı yöntemi ile elde edilen iki küme arasındaki uzaklık 1.88 bulunmuş ve oluşan küme merkezleri de Çizelge 6.17' de verilmiştir. Bu kümeleri oluşturan ilçeler Çizelge 6.18' de verilmektedir.

Çizelge 6.17 Aykırı ilçeler için grup içi ortalama bağlantı yöntemine göre küme büyüklükleri ve küme merkezleri

		Frekans	1. temel bileşen	2. temel bileşen	3. temel bileşen	4. temel bileşen	5. temel bileşen
K=2	1	45	-0,405	-0,160	0,481	-0,341	0,164
	2	30	0,608	0,240	-0,722	0,511	-0,246

Çizelge 6.17' ye bakıldığında 2. kümedeki ilçelerin lise ve üstü eğitim düzeyinde, nüfusta, medeni durumda üstün oldukları görülmektedir. Buna karşın 2. kümedeki ilçelerin tarımda ve ortaokul mezunu oranında 1. kümeye nazaran daha iyi olduğu görülmektedir.

Çizelge 6.18 Aykırı ilçeler için kümeleme sonuçları

K=1				K=2				
	İL	İLÇE		İL	İLÇE		İL	İLÇE
1	Şanlıurfa	Akçakale	31	Ağrı	Patnos	1	İstanbul	Adalar
2	Konya	Altıneki	32	Siirt	Pervari	2	Ankara	Altındağ
3	Yalova	Armutlu	33	Kastamonu	Pınarbaşı	3	İstanbul	Arnavutköy
4	Mersin	Aydıncık	34	Tunceli	Pülümür	4	İstanbul	Bağcılar
5	Van	Bahçesaray	35	Şırnak	Silopi	5	İstanbul	Bakırköy
6	Van	Başkale	36	Hakkâri	Şemdinli	6	İstanbul	Başakşehir
7	Çanakkale	Bozcaada	37	Erzurum	Tekman	7	İstanbul	Bayrampaşa
8	Rize	Çayeli	38	Ağrı	Tutak	8	İstanbul	Beşiktaş
9	Hakkâri	Çukurca	39	Isparta	Uluborlu	9	Ankara	Çankaya
10	Ardahan	Damal	40	Şanlıurfa	Viranşehir	10	İzmir	Dikili
11	Mardin	Dargeçit	41	Kırıkkale	Yahşihan	11	Kocaeli	Dilovası
12	Sinop	Dikmen	42	Konya	Yalıhüyük	12	İstanbul	Esenler
13	Van	Erciş	43	Bingöl	Yayladere	13	Ankara	Etimesgut
14	Siirt	Eruh	44	Bingöl	Yedisu	14	İstanbul	Eyüp
15	Şırnak	Güçlükonak	45	Mardin	Yeşilli	15	İstanbul	Fatih
16	Ağrı	Hamur				16	İstanbul	Güngören
17	Şanlıurfa	Harran				17	İstanbul	Kadıköy
18	Rize	Hemşin				18	İzmir	Karşıyaka
19	Bitlis	Hizan				19	Ankara	Keçiören
20	Tunceli	Hozat				20	İzmir	Konak
21	Burdur	Kemer				21	Antalya	Konyaaltı
22	Kırklareli	Koçaz				22	Bursa	Nilüfer
23	Muş	Malazgirt				23	İstanbul	Sarıyer
24	Ağrı	Merkez				24	İstanbul	Sultanbeyli
25	Hakkâri	Merkez				25	Diyarbakır	Sur
26	Tunceli	Merkez				26	İstanbul	Şişli
27	Mardin	Midyat				27	İstanbul	Ümraniye
28	Eskişehir	Mihalgazi				28	İstanbul	Üsküdar
29	Karabük	Ovacık				29	Ankara	Yenimahalle
30	Mardin	Ömerli				30	Bursa	Yıldırım

6.4. K-Ortalamlar Kümeleme Yöntemi

Burada 882 ilçe K-ortalamlar yöntemi kullanılarak SPSS 15 programı aracılığıyla K=4, 5, 6 ve 7 küme sayıları için kümelenecektir. K-ortalamlar kümeleme analizinde başlangıç küme merkezleri 3 farklı yöntem ile belirlenmiştir:

- SPSS 15 programının küme merkezlerini kendi seçmesi
- Verilerin rastgele olarak bölünmesi ile küme merkezlerinin belirlenmesi
- Verilerin rastgele seçilen bir bölümünün K-ortalamlar algoritmasına tabi tutularak, verilerin bütünü için küme merkezlerinin bulunması

6.4.1. Kullanılan program yardımıyla küme merkezlerini kendi seçmesi

SPSS 15 programı verilen bir K değeri için birbirinden mümkün olduğu kadar uzak K adet gözlemi verilerin içinden seçerek, K-ortalamlar algoritmasında başlangıç küme merkezleri olarak kullanmaktadır. Bu yöntem kullanılarak elde edilen kümeleme sonuçları Çizelge 6.19' da, bulunan küme merkezlerinin Öklid uzaklığı Çizelge 6.20' de verilmektedir.

Alt bölüm 6.3' teki hiyerarşik kümeleme sonuçlarının tersine, ilçe sayılarının kümelere dağılımı daha dengeli olma eğilimindedir: K=4 için 3 büyük kümenin varlığı söz konusu iken K=5, 6, 7 için 4 büyük kümenin varlığı dikkati çekmektedir.

Kümelemede etkili olan temel bileşenler hakkında fikir edinmek için çok değişkenli Anova testleri incelenmiş, genel olarak tarımı temsil eden bileşenin en etkin, medeni durumu temsil eden bileşenin 2. en etkin etken oldukları görülmüştür. Ayrıca kümelemede bu 2 etkenden sonra, lise ve üstü eğitim düzeyi ve orta öğretim eğitim düzeyini temsil eden bileşenlerin etkili olduğu tespit edilmiştir.

Çizelge 6.19 Spss 15 programının başlangıç küme merkezlerini kendi seçtiği k-ortalamar kümeleme yöntemine göre küme büyüklükleri ve küme merkezleri

		Frekans	1. temel bileşen	2. temel bileşen	3. temel bileşen	4. temel bileşen	5. temel bileşen
K=4	1	395	0,510	0,413	0,514	0,300	-0,204
	2	117	-0,919	-1,492	-0,024	0,298	-1,143
	3	326	-0,412	0,003	-0,134	-0,658	0,643
	4	44	0,922	0,232	-3,555	1,390	0,104
Toplam		882					
K=5	1	43	0,903	0,245	-3,606	1,393	0,111
	2	196	1,107	-0,156	0,618	0,661	0,085
	3	219	-0,304	-0,352	-0,119	-0,667	0,939
	4	333	-0,253	0,768	0,178	-0,300	-0,409
	5	91	-1,155	-1,745	0,007	0,622	-0,998
Toplam		882					
K=6	1	190	1,141	-0,154	0,623	0,617	-0,003
	2	166	-0,249	-0,663	-0,061	-0,579	1,101
	3	227	-0,532	0,859	0,312	0,273	0,137
	4	43	0,903	0,245	-3,605	1,393	0,111
	5	175	0,050	0,402	-0,117	-1,125	-0,719
	6	81	-1,264	-1,685	-0,042	0,662	-1,141
Toplam		882					
K=7	1	79	-1,081	-1,714	0,170	0,729	-0,305
	2	43	0,902	0,244	-3,605	1,393	0,111
	3	191	-0,628	0,796	0,264	0,298	0,355
	4	34	-0,775	-1,315	-0,333	-0,135	-2,445
	5	183	1,145	-0,066	0,646	0,664	0,023
	6	165	-0,082	-0,494	-0,138	-0,870	1,024
	7	187	-0,016	0,595	0,037	-0,791	-0,742
Toplam		882					

K=4 için 1 ve 3 numaralı kümelerin büyük, 2 numaralı kümenin orta büyüklükte ve 4 numaralı kümenin küçük bir küme olduğu dikkat çekmektedir (Çizelge 6.19). Çizelge 6.20' den 4 numaralı küçük kümenin merkezinin diğer kümelerin merkezine en uzak olduğu, 2 büyük kümenin merkezlerinin birbirlerine en yakın olduğu ve orta büyüklükteki 2 numaralı kümenin büyük kümelerin merkezlerine yaklaşık olarak eşit uzaklıkta oldukları görülmektedir.

K=4 için küme merkezleri temel bileşenlerin ismi dikkate alınarak incelendiğinde ise, 4 numaralı en küçük kümenin lise ve üstü eğitim seviyesinde en iyi, nüfusta en kalabalık, medeni durum açısından ikinci ve tarım açısından en kötü olduğu söylenebilir. Orta büyüklükteki 2 numaralı kümenin merkezi dikkate alındığında ise, bu kümenin ortaokul, lise ve üstü eğitim seviyesinde ve medeni durum açısından en kötü, tarımda düşük ancak nüfusta kalabalık olduğu görülmektedir.

K=4 için en büyük küme olan 1 numaralı kümenin, lise ve üstü seviyesinde ikinci en iyi, medeni durum ve tarımda en iyi, nüfus açısından ikinci en kalabalık olduğu dikkat çekmektedir. Bu durumun aksine, 2. en büyük küme olan 3 numaralı kümenin lise ve üstü seviyesinde ikinci en kötü, medeni durumda iyi olmadığı ve tarımda 2. en kötü, nüfus açısından en az kalabalık ve ortaokul mezunu oranının en fazla olduğu görülür.

Tarımın büyük ilçelerde fazla yapılmadığı dikkate alınırsa, bu kümelerin sosyo ekonomik gelişmişlik sıralaması en yüksekte en düşüğe doğru 4, 1, 3, 2 numaralı kümeler olarak oluşabilir.

Benzer yorumlar K=5, 6 ve 7 içinde yapılabilir ancak burada bunlara değinilmeyecektir. Dikkat edilirse bazı kümeler K=4' ten başlayarak K=7' ye kadar varlıklarını korumaktadırlar. Ayrıca bazı kümelerin merkezleri arasında ayırt edici net bir uzaklık varken, bazı küme merkezleri arasında da benzer uzaklıklar vardır.

Küme sayısı K arttıkça, ilçelerin kümelere daha dengeli dağıldığı görülmekte ve kümelerdeki ilçeler incelendiğinde ise bu ilçelerin birbirine daha denk olduğu izlenimi yüzeysel olarak edinilmektedir.

Çizelge 6.20 Spss 15 programının başlangıç küme merkezlerini kendi seçtiği k-ortalamalar kümeleme yöntemine göre kümeler arası Öklid uzaklıkları

Küme	1	2	3	4			
1	0	2,616	1,753	4,247			
2		0	2,571	4,645			
3			0	4,245			
4				0			
Küme	1	2	3	4	5		
1	0	4,310	4,347	4,366	4,803		
2		0	2,250	2,016	3,031		
3			0	1,816	2,845		
4				0	2,890		
5					0		
Küme	1	2	3	4	5	6	
1	0	2,305	2,015	4,326	2,366	3,142	
2		0	2,048	4,425	2,200	2,941	
3			0	4,363	1,850	2,987	
4				0	4,467	4,818	
5					0	3,076	
6						0	
Küme	1	2	3	4	5	6	7
1	0	4,759	2,672	2,416	2,831	2,629	2,999
2		0	4,346	4,984	4,333	4,415	4,443
3			0	3,587	2,069	1,985	1,691
4				0	3,599	3,712	2,774
5					0	2,380	2,207
6						0	2,086
7							0

6.4.2. Başlangıç küme merkezlerinin rastgele seçilmesi

Bu uygulamada önce veriler, Minitab 14 programından elde edilen kesikli tekdüze dağılımlı rastgele sayılar aracılığıyla yaklaşık olarak eşit $K=4, 5, 6, 7$ rastgele gruba ayrılmıştır. Daha sonra bu grupların merkezleri K -ortalamalar algoritmasında başlangıç küme merkezleri olarak kullanılmışlardır. Bu yöntemle elde edilen sonuçlar Çizelge 6.21 ve Çizelge 6.22' de verilmektedir.

Burada Alt bölüm 6.4.1' deki kümeleme sonuçlarına paralel sonuçlar elde edilmiştir. Özellikle $K=5$ ve $K=7$ için aynı kümeleme sonuçları elde edilmiştir. Dolayısıyla başlangıç merkezlerini farklı yolla almanın K -ortalamalar algoritmasında büyük değişikliklere neden olmadığı görülmüştür. Ayrıca incelenen çok değişkenli Anova testlerinde genellikle kümelemede yine sırasıyla 3, 2, 1, 5 ve 4. temel değişkenlerin en önemli rolü oynadığı görülmüştür.

Küme merkezlerini ve uzaklıklarını dikkate alan yorumlar Alt bölüm 6.4.1' deki yorumlara benzer olacağından, burada bu yorumlara değinilmeyecektir.

Çizelge 6.21 Başlangıç küme merkezlerinin rastgele seçildiği K-ortalamlar kümeleme yöntemine göre küme büyüklükleri ve küme merkezleri

		Frekans	1. temel bileşen	2. temel bileşen	3. temel bileşen	4. temel bileşen	5. temel bileşen
K=4	1	327	-0,409	0,014	-0,132	-0,660	0,637
	2	397	0,514	0,406	0,502	0,280	-0,231
	3	44	0,922	0,232	-3,555	1,389	0,104
	4	114	-0,970	-1,547	0,002	0,382	-1,064
Toplam		882					
K=5	1	196	1,106	-0,155	0,618	0,660	0,085
	2	91	-1,154	-1,744	0,006	0,621	-0,998
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	333	-0,252	0,768	0,178	-0,300	-0,408
	5	219	-0,303	-0,351	-0,119	-0,667	0,938
Toplam		882					
K=6	1	190	1,123	-0,157	0,645	0,651	0,011
	2	207	-0,628	0,787	0,274	0,316	0,261
	3	199	0,023	0,478	-0,018	-0,912	-0,767
	4	43	0,902	0,244	-3,605	1,393	0,111
	5	166	-0,179	-0,608	-0,112	-0,726	1,076
	6	77	-1,261	-1,787	-0,027	0,687	-1,130
Toplam		882					
K=7	1	191	-0,628	0,796	0,264	0,298	0,355
	2	34	-0,775	-1,315	-0,333	-0,135	-2,445
	3	183	1,145	-0,066	0,646	0,664	0,023
	4	79	-1,081	-1,714	0,170	0,729	-0,305
	5	165	-0,082	-0,494	-0,138	-0,870	1,024
	6	187	-0,016	0,595	0,037	-0,791	-0,742
	7	43	0,902	0,244	-3,605	1,393	0,1112
Toplam		882					

Çizelge 6.22 Başlangıç küme merkezlerinin rastgele seçildiği K-ortalamalar kümeleme yöntemine göre kümeler arası Öklid uzaklıkları

Küme	1	2	3	4			
1	0	1,746	4,246	2,599			
2		0	4,243	2,640			
3			0	4,667			
4				0			
Küme	1	2	3	4	5		
1	0	3,031	4,310	2,016	2,250		
2		0	4,803	2,890	2,845		
3			0	4,366	4,347		
4				0	1,816		
5					0		
Küme	1	2	3	4	5	6	
1	0	2,068	2,260	4,341	2,347	3,179	
2		0	1,782	4,345	2,013	3,034	
3			0	4,448	2,160	3,079	
4				0	4,419	4,864	
5					0	3,072	
6						0	
Küme	1	2	3	4	5	6	7
1	0	3,587	2,069	2,672	1,985	1,691	4,346
2		0	3,599	2,416	3,712	2,774	4,984
3			0	2,831	2,380	2,207	4,333
4				0	2,629	2,999	4,759
5					0	2,086	4,415
6						0	4,443
7							0

6.4.3. Verilerden rastgele örneklem seçerek küme merkezlerinin belirlenmesi

K-ortalamlar kümeleme yönteminde son olarak uygulanan bu aşamada ise, Minitab 14 programında üretilen kesikli tekdüze dağılımlı rastgele sayılar kullanılarak verilerin yarısı yani 441 ilçe rastgele seçilmiştir. Daha sonra bu 441 ilçe K-ortalamlar algoritmasına tabi tutularak K=4, 5, 6, 7 için küme merkezleri elde edilmiştir. Elde edilen bu küme merkezleri verilerin bütünü yani 882 ilçe için K-ortalamlar algoritmasında K=4, 5, 6, 7 için başlangıç küme merkezleri olarak kullanılmıştır. Bu işlemler neticesinde elde edilen sonuçlar Çizelge 6.23 ve Çizelge 6.24' de verilmektedir.

Burada K=4, 5, 6 için elde edilen sonuçlar Alt bölüm 6.4.2' deki kümeleme sonuçları ile tamamen aynıdır. K=7 için ise benzerdir. Bu nedenle bir kez daha başlangıç küme merkezlerinin farklı bir yolla seçilmesi, kümeleme sonuçlarında dikkate değer farklılıklara yol açmamıştır. Bu nedenle başlangıç küme merkezlerini seçmenin her üç yöntemi, benzer sonuçlara neden olmuştur.

Çizelge 6.23 Verilerden rastgele örneklem seçerek başlangıç küme merkezlerinin belirlendiği K-ortalamlar kümeleme yöntemine göre küme büyüklükleri ve küme merkezleri

		Frekans	1. temel bileşen	2. temel bileşen	3. temel bileşen	4. temel bileşen	5. temel bileşen
K=4	1	44	0,921	0,232	-3,555	1,389	0,104
	2	327	-0,409	0,014	-0,132	-0,660	0,637
	3	397	0,514	0,406	0,502	0,280	-0,231
	4	114	-0,970	-1,546	0,002	0,382	-1,064
Toplam		882					
K=5	1	333	-0,252	0,768	0,178	-0,300	-0,408
	2	91	-1,154	-1,744	0,006	0,621	-0,998
	3	196	1,106	-0,155	0,618	0,660	0,085
	4	219	-0,303	-0,351	-0,119	-0,667	0,938
	5	43	0,902	0,244	-3,605	1,393	0,111
Toplam		882					
K=6	1	77	-1,26171	-1,78767	-0,02781	0,687998	-1,13046
	2	190	1,123617	-0,15792	0,64548	0,651613	0,011518
	3	207	-0,62836	0,787898	0,274513	0,316009	0,261559
	4	43	0,902763	0,244618	-3,60558	1,393388	0,111295
	5	166	-0,17974	-0,60893	-0,11262	-0,726	1,076223
	6	199	0,023886	0,478016	-0,01803	-0,91254	-0,76746
Toplam		882					
K=7	1	199	0,035	0,470	-0,022	-0,912	-0,770
	2	206	-0,636	0,790	0,269	0,302	0,258
	3	77	-1,261	-1,787	-0,027	0,687	-1,130
	4	23	1,747	-0,198	-2,999	0,986	-0,223
	5	187	1,089	-0,141	0,691	0,654	0,029
	6	24	0,291	0,560	-3,774	1,727	0,339
	7	166	-0,179	-0,608	-0,112	-0,726	1,076
Toplam		882					

Çizelge 6.24 Verilerden rastgele örneklem seçerek başlangıç küme merkezlerinin belirlendiği K-ortalamalar kümeleme yöntemine göre kümeler arası Öklid uzaklıkları

Küme	1	2	3	4			
1	0	4,246	4,243	4,667			
2		0	1,746	2,599			
3			0	2,640			
4				0			
Küme	1	2	3	4	5		
1	0	2,890	2,016	1,816	4,366		
2		0	3,031	2,845	4,803		
3			0	2,250	4,310		
4				0	4,347		
5					0		
Küme	1	2	3	4	5	6	
1	0	3,179	3,034	4,864	3,072	3,079	
2		0	2,068	4,341	2,347	2,260	
3			0	4,345	2,013	1,782	
4				0	4,419	4,448	
5					0	2,160	
6						0	
Küme	1	2	3	4	5	6	7
1	0	1,781	3,077	4,018	2,257	4,727	2,160
2		0	3,034	4,249	2,050	4,394	2,010
3			0	4,618	3,178	5,020	3,072
4				0	3,773	2,040	4,104
5					0	4,724	2,341
6						0	4,643
7							0

6.5. Hibrid Kümeleme Yöntemi

Burada Alt bölüm 6.3' te 882 ilçe için elde edilen hiyerarşik kümeleme merkezleri, K-ortalamlar algoritmasında K=4, 5, 6, 7 için başlangıç küme merkezleri olarak kullanılarak hibrid kümeleme sonuçları elde edilmiştir. Bu işlemler SPSS 15 programı yardımıyla gerçekleştirilmiştir.

6.5.1. Hiyerarşik grup içi ortalama bağlantı + K-ortalamlar

Bu çalışmada ilk olarak grup içi ortalama bağlantı yönteminde Alt bölüm 6.3.1' de bulunan küme merkezleri, K-ortalamlar algoritmasında başlangıç küme merkezleri olarak kullanılmıştır. Bulunan sonuçlar K=4, 5, 6, 7 için Çizelge 6.25 ve Çizelge 6.26' da verilmektedir

Çizelge 6.25 incelendiğinde, K=4, 5 için elde edilen kümeleme sonuçlarının, Alt bölüm 6.4.1, 6.4.2 ve 6.4.3' teki K-ortalamlar algoritmasının sonuçlarıyla aynı veya çok benzer olduğu tespit edilmektedir. Ancak K=6, 7 için özellikle de K=6 için sonuçların farklılaştığı dikkati çekmektedir.

Çok değişkenli Anova tablolarına bakılarak, kümelemede daha önce olduğu gibi sırasıyla 3, 2 ve 1. temel bileşenlerinin daha etkili olduğu söylenebilir. Ayrıca küme merkezlerine ve bu merkezler arasındaki uzaklıklara bakılarak Alt bölüm 6.4.1' deki yorumlara benzer yorumlar yapılabilir ancak burada ayrıntılara girilmeyecektir.

Çizelge 6.25 Hiyerarşik grup içi ortalama bağlantı + K-ortalama ile hibrid kümeleme yöntemine göre küme büyüklükleri ve küme merkezleri

		Frekans	1. temel bileşen	2. temel bileşen	3. temel bileşen	4. temel bileşen	5. temel bileşen
K=4	1	397	0,514	0,406	0,502	0,280	-0,231
	2	327	-0,409	0,014	-0,132	-0,660	0,637
	3	44	0,921	0,232	-3,555	1,389	0,104
	4	114	-0,970	-1,546	0,002	0,382	-1,064
Toplam		882					
K=5	1	196	1,106	-0,155	0,618	0,660	0,085
	2	219	-0,303	-0,351	-0,119	-0,667	0,938
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	91	-1,154	-1,744	0,006	0,621	-0,998
	5	333	-0,252	0,768	0,178	-0,300	-0,408
Toplam		882					
K=6	1	188	1,138	-0,099	0,623	0,644	0,082
	2	210	-0,146	-0,288	-0,154	-0,878	0,861
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	63	-0,164	-0,625	-0,348	-1,077	-1,944
	5	93	-1,115	-1,613	0,185	0,803	-0,371
	6	285	-0,379	0,906	0,263	-0,011	-0,154
Toplam		882					
K=7	1	185	1,159	-0,090	0,630	0,610	0,025
	2	158	-0,156	-0,498	-0,105	-0,752	1,119
	3	43	0,902	0,244	-3,605	1,393	0,111
	4	30	-0,895	-1,409	-0,320	0,096	-2,477
	5	79	-1,081	-1,714	0,170	0,729	-0,305
	6	213	-0,553	0,868	0,307	0,302	0,183
	7	174	0,008	0,446	-0,081	-1,027	-0,730
Toplam		882					

Çizelge 6.26 Hiyerarşik grup içi ortalama bağlantı + K-ortalama ile hibrid kümeleme yöntemine göre kümeler arası Öklid uzaklıkları

Küme	1	2	3	4			
1	0	1,746	4,243	2,640			
2		0	4,246	2,599			
3			0	4,667			
4				0			
Küme	1	2	3	4	5		
1	0	2,250	4,310	3,031	2,016		
2		0	4,347	2,845	1,816		
3			0	4,803	4,366		
4				0	2,890		
5					0		
Küme	1	2	3	4	5	6	
1	0	2,284	4,315	3,161	2,792	1,982	
2		0	4,361	2,840	2,676	1,854	
3			0	4,779	4,741	4,370	
4				0	2,860	2,666	
5					0	2,758	
6						0	
Küme	1	2	3	4	5	6	7
1	0	2,344	4,330	3,660	2,827	2,021	2,319
2		0	4,421	3,884	2,577	2,046	2,102
3			0	5,014	4,759	4,361	4,453
4				0	2,342	3,582	2,939
5					0	2,719	3,032
6						0	1,803
7							0

6.5.2. Hiyerarşik Ward'ın yöntemi + K-ortalamlar

Hibrid kümelemenin son uygulaması olarak, Alt bölüm 6.3.2' de verilen Ward'ın küme merkezleri, K-ortalamlar algoritmasında başlangıç küme merkezleri olarak kullanılmıştır. Bulunan sonuçlar Çizelge 6.27 ve Çizelge 6.28' de verilmektedir.

Çizelge 6.27 incelendiğinde, K=4, 5 için elde edilen kümeleme sonuçlarının, Alt bölüm 6.4.1, 6.4.2 ve 6.4.3' teki K-ortalamlar algoritmasının sonuçlarıyla aynı veya çok benzer olduğu tespit edilmektedir. Ancak K=6 için elde edilen sonuçların sadece Alt bölüm 6.4.1' deki sonuçlarla; K=7 için elde edilen sonuçların, hem Alt bölüm 6.4.1 hem de Alt bölüm 6.4.2' deki sonuçlarla birebir aynı olduğu dikkati çekmektedir.

Netice olarak kullanılan hibrid kümeleme algoritmalarının genellikle K-ortalama algoritmalarının sonuçlarına benzer sonuçlar ürettikleri izlenimi edinilmektedir.

İki hibrid kümeleme algoritması sonuçları karşılaştırıldığında ise, K=4, 5 için sonuçların tamamıyla aynı olduğu; ancak özellikle K=6 için olmak üzere K=6, 7 için sonuçların farklılaştığı görülmektedir.

Ayrıca küme merkezlerine ve bu merkezleri arasındaki uzaklıklara kümeler hakkında çeşitli yorumlar yapılabilir ancak burada yorumlara girilmeyecektir.

Çizelge 6.27 Hiyerarşik ward bağlantı + K-ortalama ile hibrid kümeleme yöntemine göre küme büyüklükleri ve küme merkezleri

		Frekans	1. temel bileşen	2. temel bileşen	3. temel bileşen	4. temel bileşen	5. temel bileşen
K=4	1	397	0,514	0,406	0,502	0,280	-0,231
	2	327	-0,409	0,014	-0,132	-0,660	0,637
	3	44	0,921	0,232	-3,555	1,389	0,104
	4	114	-0,970	-1,546	0,002	0,382	-1,064
Toplam		882					
K=5	1	196	1,106	-0,155	0,618	0,660	0,085
	2	219	-0,303	-0,351	-0,119	-0,667	0,938
	3	333	-0,252	0,768	0,178	-0,300	-0,408
	4	43	0,902	0,244	-3,605	1,393	0,111
	5	91	-1,154	-1,744	0,006	0,621	-0,998
Toplam		882					
K=6	1	190	1,141	-0,154	0,623	0,617	-0,003
	2	166	-0,248	-0,662	-0,061	-0,579	1,101
	3	227	-0,531	0,858	0,312	0,273	0,137
	4	43	0,902	0,244	-3,605	1,393	0,111
	5	81	-1,264	-1,684	-0,042	0,662	-1,141
	6	175	0,049	0,401	-0,117	-1,125	-0,719
Toplam		882					
K=7	1	183	1,145	-0,066	0,646	0,664	0,023
	2	165	-0,082	-0,494	-0,138	-0,870	1,024
	3	79	-1,081	-1,714	0,170	0,729	-0,305
	4	191	-0,628	0,796	0,264	0,298	0,355
	5	43	0,902	0,244	-3,605	1,393	0,111
	6	34	-0,775	-1,315	-0,333	-0,135	-2,445
	7	187	-0,016	0,595	0,037	-0,791	-0,742
Toplam		882					

Çizelge 6.28 Hiyerarşik ward bağlantı + K-ortalama ile hibrid kümeleme yöntemine göre kümeler arası Öklid uzaklıkları

Küme	1	2	3	4			
1	0	1,746	4,243	2,640			
2		0	4,246	2,599			
3			0	4,667			
4				0			
Küme	1	2	3	4	5		
1	0	2,250	2,016	4,310	3,031		
2		0	1,816	4,347	2,845		
3			0	4,366	2,890		
4				0	4,803		
5					0		
Küme	1	2	3	4	5	6	
1	0	2,305	2,015	4,326	3,142	2,366	
2		0	2,048	4,425	2,941	2,200	
3			0	4,363	2,987	1,850	
4				0	4,818	4,467	
5					0	3,076	
6						0	
Küme	1	2	3	4	5	6	7
1	0	2,380	2,831	2,069	4,333	3,599	2,207
2		0	2,629	1,985	4,415	3,712	2,086
3			0	2,672	4,759	2,416	2,999
4				0	4,346	3,587	1,691
5					0	4,984	4,443
6						0	2,774
7							0

6.6. Kümeleme Sonuçlarının Karşılaştırılması

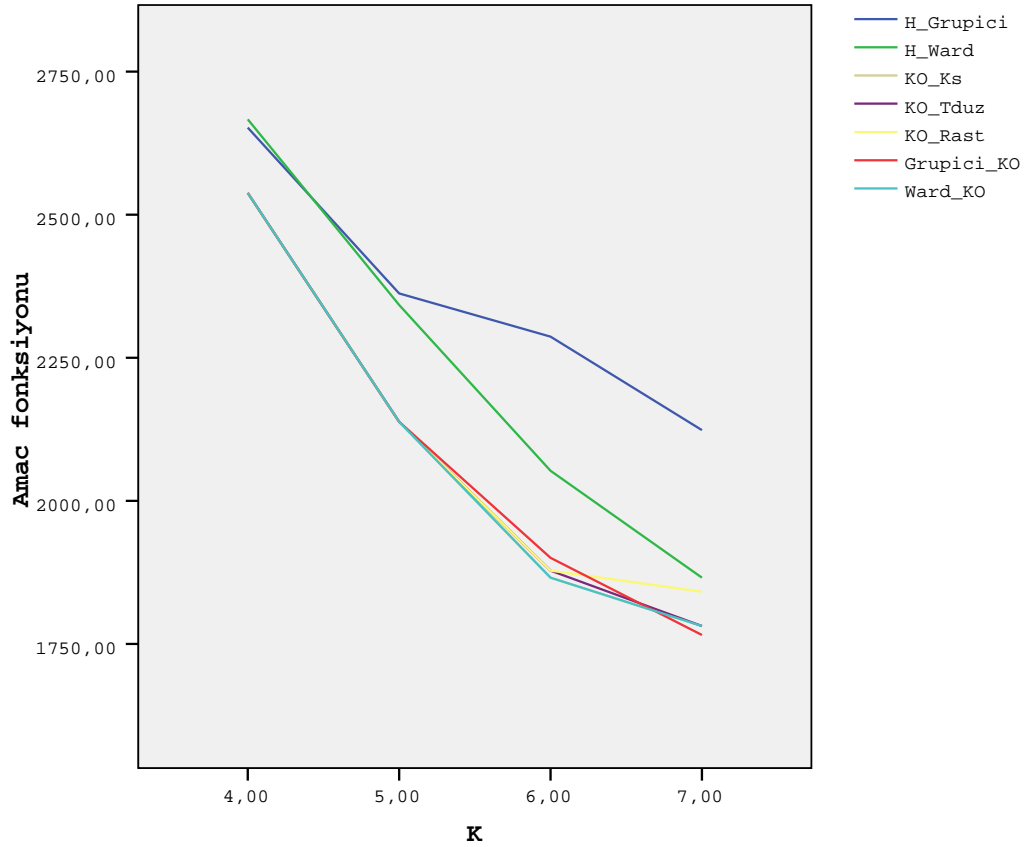
Burada daha önceki bölümlerde 882 ilçe üzerinde yapılan kümelemeler amaç fonksiyonu, ortalama silhouette genişliği ve Davies-Bouldin indeksi kullanılarak karşılaştırılacaktır.

6.6.1. Amaç fonksiyonu kullanılarak karşılaştırma

Alt bölüm 4.5, Eşitlik 4.3' te tanımlanan amaç fonksiyonu F' nin değerleri Minitab 14 programının komutları kullanılarak oluşturulmuş Bir Makro programı aracılığıyla (EK-2) hesaplanmıştır. Elde edilmiş sonuçlar Çizelge 6.29' da verilmektedir. Sonuçlara ilişkin çizgi grafikleri de Şekil 6.6' da gösterilmektedir.

Çizelge 6.29 Kümeleme yöntemlerine göre amaç fonksiyonu değerleri

AMAÇ FONKSİYONU							
Hiyerarşik Kümeleme			K-ortalamlar Kümeleme (KO)			Hibrid Kümeleme	
K	Grup İçi	Ward	K.S.	TDÜZ	RAST	Grup İçi+KO	Ward+KO
4	2652.28	2666.84	2537.54	2538.39	2538.39	2538.39	2538.39
5	2362.52	2342.67	2138.38	2138.38	2138.38	2138.38	2138.38
6	2287.40	2053.04	1866.09	1878.29	1878.29	1900.64	1866.09
7	2123.61	1866.07	1781.72	1781.72	1841.67	1766.07	1781.72



Şekil 6.6 Kümeleme yöntemlerine göre amaç fonksiyonu değerlerinin çizgi grafikleri

Çizelge 6.29' de, K.S. SPSS 15 programının başlangıç küme merkezlerini kendisi seçtiğini; TDÜZ, verilerin rastgele olarak K parçaya bölünerek, her parçanın merkezinin başlangıç küme merkezleri olarak kullanıldığını ve RAST ise rastgele 2' ye bölünen verilerin bir kısmına K-ortalamar yönteminin uygulanarak, verilerin bütünü için başlangıç küme merkezlerinin bulunduğunu göstermektedir.

F fonksiyonu küme sayısı K' nın azalan bir fonksiyonu olduğundan, Çizelge 6.29 incelendiğinde beklenildiği gibi K artarken her üç kümeleme yöntemi için F' nin değerleri azalmaktadır. K-ortalamar kümeleme ve Hibrid kümeleme algoritmalarında hiyerarşik kümeleme algoritmalarına üstünlüğü, K=4, 5, 6, 7 için F' nin değerlerinin daha düşük olmasıyla görülebilmektedir.

K-ortalamar kümeleme algoritmaları Hibrid kümeleme algoritmaları ile karşılaştırıldığında ise, K=4, 5, 6 için K-ortalamar algoritmalarının Hibrid kümeleme algoritmalarına kıyasla daha iyi veya eşit düzeyde bir performans gösterdiği, ancak K=7 için ise tam tersinin doğru olduğu görülmektedir.

Ayrıca K-ortalamlar ve Hibrid kümeleme algoritmalarının kendi içlerinde incelendiğinde, K=4, 5 için performanslarının aynı veya çok yakın olduğu fakat K=6, 7 için farklılıklar gösterdiği anlaşılmaktadır. Hiyerarşik kümeleme algoritmaları kendi içinde incelendiğinde ise, K=5' ten itibaren Ward' ın algoritmasının performansının grup içi ortalama uzaklığı kullanan algoritmaya daha üstün olduğu dikkat çekmektedir.

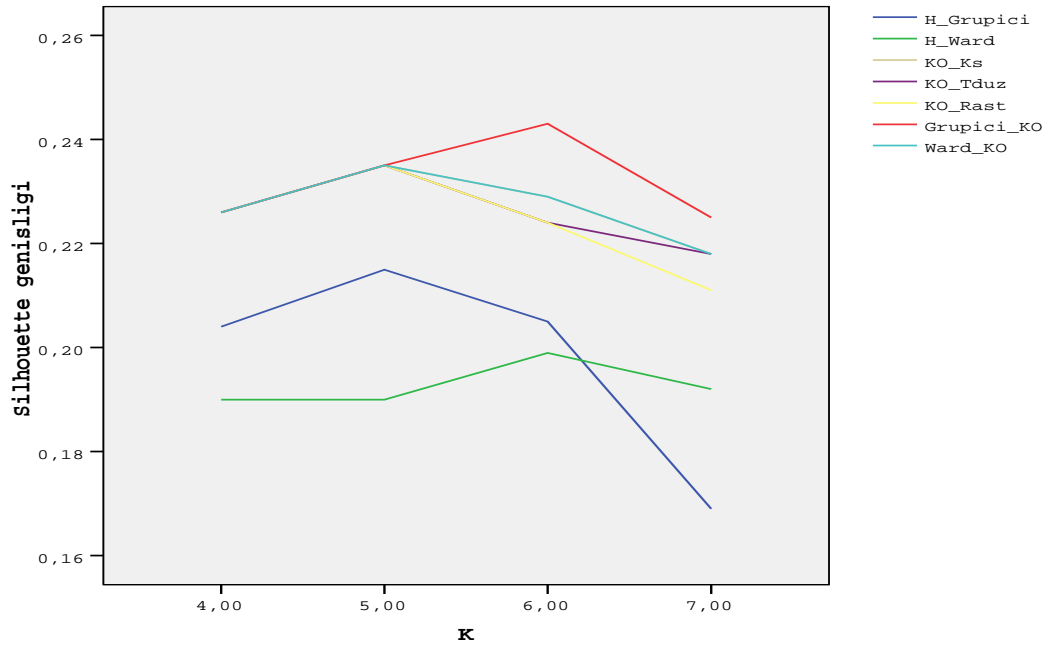
6.6.2. Ortalama silhouette genişliği kullanılarak karşılaştırma

Alt bölüm 4.7.1' de açıklanan silhouette genişlikleri Qbasic dilinde yazılan bir program (EK-3) aracılığıyla tüm kümelemeler için hesaplanmış ve her bir kümeleme için bulunan genişliklerin ortalaması Çizelge 6.30' da ve bunların görsel ifadeleri de Şekil 6.7' de verilmiştir.

Çizelge 6.30 Kümeleme yöntemlerine göre ortalama silhouette genişliği değerleri

ORTALAMA SILHOUETTE GENİŞLİĞİ							
Hiyerarşik Kümeleme			K-ortalamlar Kümeleme			Hibrid Kümeleme	
K	Grup İçi	Ward	K.S.	TDÜZ	RAST	Grup İçi+KO	Ward+KO
4	0.204	0.190	0.226	0.226	0.226	0.226	0.226
5	0.215	0.190	0.235	0.235	0.235	0.235	0.235
6	0.205	0.199	0.229	0.224	0.224	0.243	0.229
7	0.169	0.192	0.218	0.218	0.211	0.225	0.218

Çizelge 6.30 incelendiğinde ilk olarak bu ortalamaların birbirlerine yakın değerler oldukları göze çarpmaktadır. Büyük olan silhouette genişliği ortalamalarının verilerin daha iyi bir kümelemesini işaret ettiği dikkate alındığında, K-ortalamlar ve hibrid kümeleme algoritmalarının, hiyerarşik kümeleme algoritmalarına üstünlüğü görülmektedir.



Şekil 6.7 Kümeleme yöntemlerine göre ortalama silhouette genişliği değerlerinin çizgi grafikleri

K-ortalamlar ve Hibrid algoritmaları karşılaştırıldığında, K=4, 5 için K-ortalamlar algoritmalarının, K=6, 7 için ise Hibrid algoritmalarının daha iyi veya eşit düzeyde iyi kümeleme yaptıkları söylenebilmektedir. Hiyerarşik algoritmaları kendi içinde incelendiğinde, grup içi yöntemi için K=5' in, Ward' ın yöntemi için K=6' nın, verileri en iyi biçimde kümelendiği küme sayısı K olarak seçildiği görülmektedir.

Öteki algoritmalar da kendi içinde incelendiğinde, K-ortalamlar algoritmaları için K=5' in ve Hibrid algoritmalarında, grup içi+KO yöntemi için K=6' nın, Ward+KO yöntemi için K=5' in seçildiği görülür. Tüm algoritmalar üzerinden ortalama silhouette genişliğini en büyük yapan algoritmanın K=6 ile grup içi+KO Hibrid algoritması olduğu açıktır. Dikkati çeken bir diğer nokta ise, tüm kümeleme algoritmaları için seçilen K' nın civarında ortalama silhouette genişliğinin düşmesidir.

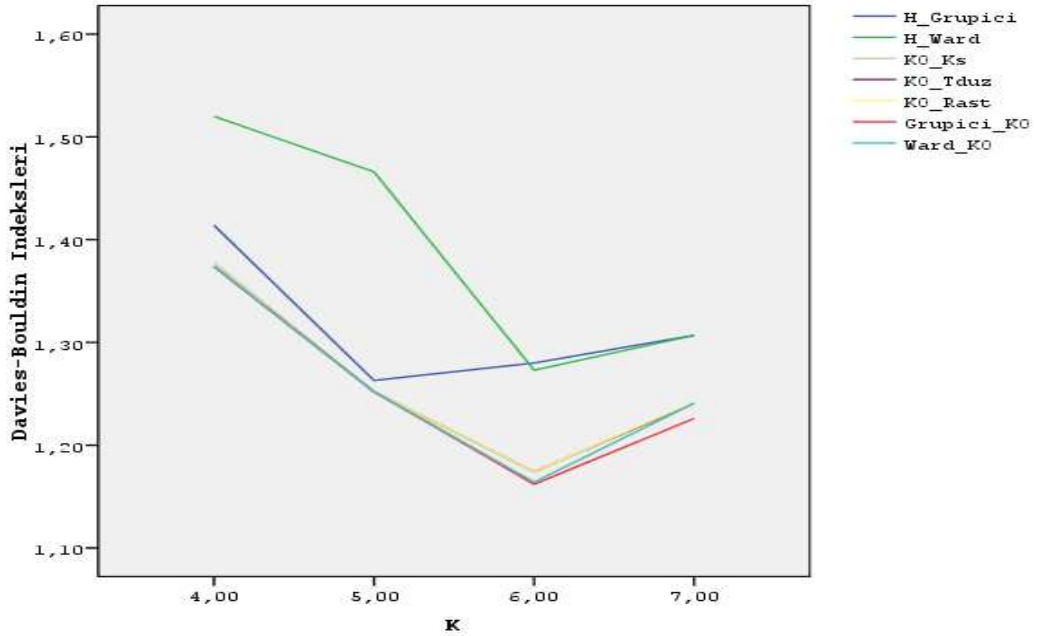
6.6.3. Davies-Bouldin indeksi kullanılarak karşılaştırma

Mintab 14 programında yazılan bir makro program (EK-4) yardımıyla Alt bölüm 4.7.2' de tanımlanan Davies-Bouldin indeksi tüm kümelemeler için hesaplanmıştır (Çizelge 6.31). Ayrıca bulunan indeks değerleri Şekil 6.8' de çizgi

grafikleri olarak ifade edilmektedir. Çizelge 6.31’ de ilk dikkati çeken husus Davies-Bouldin indeksi değerlerinin değişik K değerleri için ortalama silhouette genişliğinin değerlerine kıyasla daha ayırt edici olmasıdır.

Çizelge 6.31 Kümeleme yöntemlerine göre Davies-Bouldin indeksi değerleri

DAVIES – BOULDIN İNDEKSİ							
Hiyerarşik Kümeleme			K-ortalamalar Kümeleme			Hibrid Kümeleme	
K	Grup İçi	Ward	K.S.	TDÜZ	RAST	Grup İçi+KO	Ward+KO
4	1.414	1.520	1.378	1.374	1.374	1.374	1.374
5	1.263	1.466	1.252	1.252	1.252	1.252	1.252
6	1.280	1.273	1.164	1.174	1.174	1.162	1.164
7	1.307	1.307	1.241	1.241	1.242	1.226	1.241



Şekil 6.8 Kümeleme yöntemlerine göre Davies-Bouldin indeks değerlerinin çizgi grafikleri

Davies-Bouldin indeks değerlerinin küçük olanı, verilerin daha iyi bir kümelemesini gösterdiği dikkate alınır, K-ortalamalar ve Hibrid algoritmalarının tüm K değerleri için, hiyerarşik algoritmalara nazaran daha iyi bir kümeleme

yaptıkları Çizelge 6.31' dan görülmektedir. Ayrıca her K değeri için Hibrid algoritmalarının, K-ortalamlar algoritmalarından daha üstün veya eşit olduğu dikkati çekmektedir.

Çizelge 6.31' daki tüm Hibrid ve K-ortalamlar algoritmaları için K=6' nın, grup içi hiyerarşik algoritması için K=5' in ve Ward hiyerarşik uygulaması için K=6' nın, verilerin en iyi kümelendiği küme sayısı olarak seçildiği görülmektedir. Tüm algoritmalar üzerinden en küçük gösterge değerine sahip algoritma K=6 ile grup içi+KO Hibrid algoritmasıdır. Böylelikle hem ortalama silhouette genişliği hem de Davies-Bouldin indeksi aynı algoritmayı tercih etmişlerdir.

6.7. Bölgelerin İlçelerin Gelişmişlik Düzeyine Göre İncelenmesi

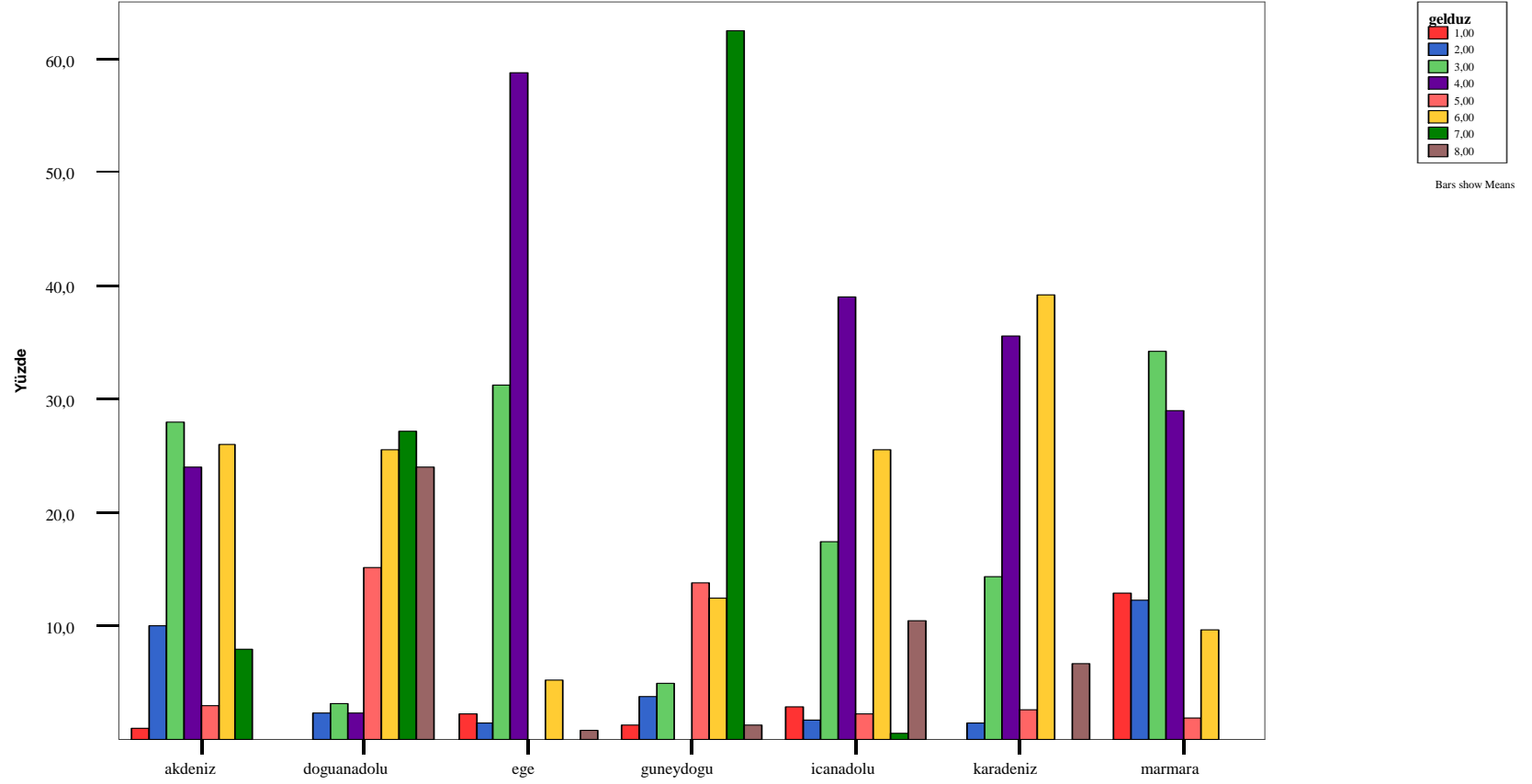
İlk olarak 882 ilçeye ilişkin K=6 için grup içi ortalama uzaklık+K-ortalamlar biçimindeki Hibrid kümelemesi ve Alt bölüm 6.3.3' teki 75 aykırı ilçenin hiyerarşik kümelemesi dikkate alınmış ve kümelerin 1' den 8' e kadar gelişmişlik düzeyleri belirlenmiştir.

Gelişmişlik sıralamaları oluşturulurken kümeler arasındaki uzaklıklar ve küme merkezleri dikkate alınmıştır. Daha sonra Türkiye'deki coğrafi bölgeler, ilçelerin gelişmişlik düzeylerine göre karşılaştırılmıştır. Elde edilen sonuçlar Çizelge 6.32 ve Şekil 6.9' da verilmektedir. Ayrıca 957 ilçe sosyo-ekonomik gelişme düzeyleri dikkate alınarak oluşturulan 8 kümenin listesi EK-6' da verilmektedir.

Çizelge 6.32' den Marmara Bölgesinin 1. düzeye sahip ilçelerin oranları açısından net üstünlüğü; Marmara ve Akdeniz Bölgelerinin 2. gelişmişlik düzeyindeki üstünlükleri; Marmara, Ege ve Akdeniz Bölgelerinin 3. gelişmişlik düzeyindeki üstünlükleri; Güneydoğu ve Doğu Anadolu Bölgelerindeki İlk 4 gelişmişlik düzeyindeki ilçelerinin oranlarının çok düşük olması ancak son 4 gelişmişlik düzeyinde ilgili oranların yüksek olması; Akdeniz, Karadeniz ve İç Anadolu bölgelerinde özellikle 6. gelişmişlik düzeyi başta olmak üzere son 3 gelişmişlik düzeyine sahip ilçelerin oranının yüksek olması dikkat çekmektedir.

Çizelge 6.32 Türkiye'deki coğrafi bölgelere göre ilçelerin gelişmişlik düzeylerinin dağılımı

		Gelişmişlik Düzeyleri															
		1. Düzey		2. Düzey		3. Düzey		4. Düzey		5. Düzey		6. Düzey		7. Düzey		8. Düzey	
Bölgeler	İlçe Sayısı	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde
Akdeniz	100	1	1	10	10	28	28	24	24	3	3	26	26	8	8	0	0
Doğu Anadolu	125	0	0	3	2.4	4	3.2	3	2.4	19	15.2	32	25.6	34	27.2	30	24
Ege	131	3	2.3	2	1.5	41	31.3	77	58.8	0	0	7	5.3	0	0	1	0.8
Güneydoğu Anadolu	80	1	1.3	3	3.8	4	5	0	0	11	13.8	10	12.5	50	62.5	1	1.3
İç Anadolu	172	5	2.9	3	1.7	30	17.4	67	39	4	2.3	44	25.6	1	0.6	18	10.5
Karadeniz	194	0	0	3	1.5	28	14.4	69	35.6	5	2.6	76	39.2	0	0	13	6.7
Marmara	155	20	12.9	19	12.3	53	34.2	45	29	3	1.9	15	9.7	0	0	0	0



Şekil 6.9 Türkiye’deki coğrafi bölgelere göre ilçelerin gelişmişlik düzeylerinin dağılımının çubuk grafiği

Bölgeler genel olarak değerlendirildiğinde en iyiden en kötüye doğru sıralamanın yaklaşık olarak şu şekilde olabileceği söylenebilir: Marmara Bölgesi, Ege, Akdeniz, İç Anadolu, Karadeniz, Doğu Anadolu, Güneydoğu Anadolu. Sonuç olarak gelişmişlik düzeylerinin bölgelere göre farklılaştığı söylenebilir. Bölgeler için 2008’ de oluşan bu durumun tesadüf eseri olup olmadığının testi (Çizelge 6.33), bölgelerle gelişmişlik düzeyleri arasındaki ilişkinin anlamlı olduğunu teyit etmektedir.

Çizelge 6.33 Coğrafi bölgelere göre ilçelerin gelişmişlik düzeyi farklılıklarının anlamlılık testleri

Test	Test İstatistiği Değeri	Serbestlik Derecesi	Anlamlılık Düzeyi
Pearson Ki-Kare	791,263	42	0.000
Olabilirlik Oranı	734.785	42	0.000

7. SONUÇLAR VE TARTIŞMA

Yerel bazda sosyo-ekonomik gelişmişlik düzeyinin ölçülmesi veya belirlenmesi, ilgili yerel yöre için idari, mali ve sosyal açılardan daha doğru kararlar verilmesini, daha gerçekçi projeler geliştirilmesini ve kaynakların daha iyi dağıtılmasını sağlayabilir. Bu amaçla bu çalışmada Türkiye' nin ilçeler bazında sosyo-ekonomik gelişmişlik düzeylerinin belirlenmesi ve bu işlem için yöntem olarak bir hibrid kümeleme algoritmasının tanıtılması hedeflenmiştir.

Klasik kümeleme algoritmaları arasında yer alan hiyerarşik kümeleme ve K-ortalamlar algoritmaları bazı dezavantaj ve avantajlara sahiptir. Örneğin hiyerarşik kümeleme için bazı avantajları dillendirmek gerekirse:

- İstenen ayrıntıda kümeleme yapılabilmesi esnekliği,
- Herhangi bir uzaklık ya da benzerlik ölçüsünün kullanılabilmesi,
- İster niceliksel ister niteliksel olsun herhangi bir özelliğin kullanılabilmesi.

Bazı dezavantajlar ise:

- Algoritmayı durdurma ölçütlerindeki muğlâklık,
- Bir kez oluşturulan kümenin daha sonra iyileştirilmek adına yeniden işleme

alınmaması.

Benzer biçimde K-ortalamlar algoritması için bazı avantajlar ise:

- Basit ve kolay anlaşılabilir olması,
- Varyans analizi temeline dayanması.

Bazı dezavantajlar ise:

- Hangi küme sayısı K'nın kullanımının iyi olduğunun bilinmemesi,
- Sonucun başlangıç küme merkezlerine bağlı olması,
- Aykırı değerlerden algoritmanın olumsuz etkilenmesi,
- Kullandığı amaç fonksiyonu için ulaştığı yerel minimumun genel

minimumdan çok uzak olabilmesi,

- Sadece niceliksel özelliklerin kullanılabilmesi.

Bu çalışma, yukarıdaki iki kümeleme yönteminin bazı dezavantajlarını gidermek amacıyla, bu yöntemleri birlikte kullanan bir hibrid kümeleme yaklaşımını incelemektedir: Hiyerarşik kümeleme sonucu elde edilen küme merkezleri, K-ortalamlar algoritmasında başlangıç küme merkezleri olarak kullanılmaktadır.

2008 yılı için TÜİK veri tabanına dayanarak, ülke genelinde verilerin tam olarak elde edilemediği 12 ilçe hariç, 81 ilin 957 ilçesine ilişkin sosyo-ekonomik durumu doğrudan ya da dolaylı olarak etkileyen 23 adet değişkene ilişkin veri toplanmıştır. Bu değişkenler TÜİK ADNSKS tarafından kayıt edilen demografik, eğitim, medeni durum ve tarım göstergelerinden oluşan değişkenlerdir. Maalesef ilçeler bazında başka değişkenlere ulaşılamamıştır: Sosyo-ekonomik düzeyi daha iyi ölçmeyi sağlayacak istihdam, sağlık, sanayi, inşaat, mali, altyapı ve diğer refah göstergeleri gibi değişkenler elde edilememiştir.

Verilerin ön incelemesi sonucu, araştırmada kullanılan değişken sayısı 20' ye indirilmiştir ve bazı değişkenlere de doğal logaritma dönüşümü uygulanmıştır. Ayrıca aykırı gözlemler tespit edilerek, bu gözlemlerin analizi ayrı olarak gerçekleştirilmiştir.

Çözümleme işleminde ise değişkenleri indirgemek için Temel Bileşenler Analizi kullanılmış ve verilerdeki toplam değişimin yaklaşık %81' ini açıklayan 5 temel bileşen seçilmiştir. Seçilen bu temel bileşenlere ilişkin verilerle $K=4, 5, 6, 7$ küme sayısı için, değişik bağlantı yöntemlerini kullanan Hiyerarşik ve değişik başlangıç küme merkezi seçme yöntemlerini uygulayan K-ortalama kümeleme yöntemleri uygulanmıştır. Ayrıca Hiyerarşik kümelemeden elde edilen küme merkezlerini, K-ortalama algoritmasında başlangıç küme merkezi olarak kullanan bir Hibrid kümeleme algoritması kullanılmıştır.

Ortalama silhouette genişlik değerleri ve Davies-Bouldin İndeks değerleri ile yapılan karşılaştırmalar neticesinde, K-ortalamlar ve Hibrid kümeleme algoritmalarının Hiyerarşik kümeleme algoritmasına göre daha başarılı olduğu görülmüştür. Ayrıca K-ortalamlar ve Hibrid kümeleme algoritmalarının genellikle benzer bir performans gösterdikleri dikkati çekmiştir. Ancak küme sayısı K artarken Hibrid kümeleme algoritmasının diğer iki kümeleme algoritmasına nazaran performansının daha iyi olduğu izlenimi edinilmiştir.

Kümeleme sonuçları içinden hem ortalama silhouette genişliğine hem de Davies-Bouldin indeksine göre, ilçeleri en iyi biçimde kümeleyen algoritma olarak $K=6$ için grup içi ortalama bağlantı yöntemini kullanan hiyerarşik kümeleme algoritması ve K-ortalamlar algoritmasından oluşan hibrid algoritması seçilmiştir. Ayrıca aykırı ilçeler, grup içi ortalama bağlantı yöntemini kullanan hiyerarşik

kümeleme algoritması ile 2 kümeye ayrılmıştır. Bu iki ayrı kümeleme sonucunun bir araya getirilmesiyle, ilçeler için 8 ayrı sosyo-ekonomik gelişmişlik düzeyi belirlenmiştir.

Bu düzeyler, birinci seviyeden, diğer bir deyişle en gelişmiş ilçe grubundan en geri kalmış ilçe grubuna doğru sekiz kademeli olarak oluşturulmuştur. Sosyo-ekonomik gelişme açısından Akdeniz Bölgesi'ndeki mevcut ilçelerin en üst seviyede olmasa da çoğunluğu ortanın üstünde olduğu, İç Anadolu Bölgesi'ndeki ilçelerin çoğunun ise orta düzeyde olduğu görülmüştür. Ancak İç Anadolu Bölgesi'nde her ne kadar ilçelerin büyük bir dilimi gelişmekte olsa da kayda değer derecede gelişmemiş ilçelerde bulunmaktadır. Marmara Bölgesi'ndeki ilçelerin % 88.4' ünün gelişmiş olduğu görülmüş ve hiç gelişmemiş ilçe bulunmamıştır.

Ancak Doğudaki Bölgelerde ise tam tersine mevcut ilçelerin düzeyleri oldukça düşüktür. Özellikle Güney Doğu Bölgesi'nde % 90.1 gibi bir oranla ilçelerin büyük çoğunluğu 5 ve üzerindeki seviyededir. Karadeniz Bölgesi ise çoğunluğu orta düzeyde gelişmişliğe sahip ilçeleri barındırmaktadır. Son olarak Ege Bölgesi'nde yok denecek kadar az gelişmiş ilçe mevcuttur. Ayrıca bu bölgede ilçelerin % 90.1' i kadar 3. ve 4. düzeyde olmasına rağmen, üst seviyelerde de ilçelerin bulunduğu tespit edilmiştir.

Sosyo-ekonomik gelişme alanında ilçe gelişme planlarına ilişkin diğer önemli bir konuda, nitelikli bilgi edinilmesine yönelik güçlüklerdir. Söz konusu güçlükler; tam ve eksiksiz olarak verilerin derlenmemiş olması, verilere ulaşılamaması ve uzun yıllar geniş çaplı bilginin toplanmaması olarak sıralanabilir. Bu konuda en son 2000 genel nüfus sayımında çalışma olduğu görülmüştür. Bu tür eksiklikler daha ayrıntılı değerlendirmenin yapılmasını engellemektedir. Bu nedenle ülkemizdeki bilgi derleme ve değerlendirme sistemlerinin geliştirilmesi büyük önem taşımaktadır.

KAYNAKLAR

Abonyi, J., Feil, B., 2007. Cluster Analysis For Data Mining and System Identification, Birkhauser Verlag, Germany, 303p.

Albayrak, A.S., Kalaycı, Ş., Karataş, A., 2004. Türkiye’ de Coğrafi Bölgelere İllerin Sosyo-Ekonomik Gelişmişlik Düzeylerinin Temel Bileşenler Analiziyle İncelenmesi, Süleyman Demirel Üniversitesi İ.İ.B.F., 9, (2): 101–130.

Albayrak, A.S., 2005. Türkiye’ de İllerin Sosyo-Ekonomik Gelişmişlik Düzeylerinin Çok Değişkenli İstatistik Yöntemlerle İncelenmesi, ZKÜ Sosyal Bilimler Dergisi, 1, (1): 153–177.

Basu, P.K., 2000. Conflicts and Paradoxes in Economic Development Tourism in Papua New Guinea, International Journal of Social Economics, 27, (7/8/9/10): 907–916.

Blakely, E.J., 1994. Planning Local Economic Development: Theory and Practice, Thousand Oaks, Sage Publications.

Chipman, H., Tibshirani, R., 2006. Hybrid Hierarchical Clustering with Applications to Microarray Data, Biostatistics, 7, (2): 286–301.

Dinçer, B., Özaslan, M., Kavasoğlu, T., 2003. İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması, DPT, Bölgesel Gelişme ve Yapısal Uyum Genel Müdürlüğü, Ankara, 250p.

Dinçer, B., Özaslan, M., 2004. İlçelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması, DPT, Bölgesel Gelişme ve Yapısal Uyum Genel Müdürlüğü, Ankara, 245p.

DPT, 2000. Uzun Vadeli Strateji ve Sekizinci Beş Yıllık Kalkınma Planı 2001-2005, Ankara, 266p.

Dulupcu, M.A., 2002. Accomodating Turkey in New Regionalism, Centre for Urban and Regional Studies, UK.

Dura, C., Atik, H., Türker, O., 2004. Beşeri Sermaye Açısından Türkiye’nin Avrupa Birliği Karşısındaki Kalkınma Seviyesi, 3. Ulusal Bilgi, Ekonomi Ve Yönetim Kongresi, Eskişehir, 13-20.

Fielding, A.H., 2007. Cluster and Classification Techniques for The Biosciences, Cambridge University Press, New York, 246p.

Flammang R.A., 1979. Economic Growth and Economic Development; Counterparts or Competitors?, Economic Development and Cultural Change, 28, (1): 50p.

Gan, G., Ma, C., Wu, J., 2007. Data Clustering Theory, Algorithms, and Applications, Asa-Siam Series, Philadelphia, 465p.

Göçer, K., Çıracı, H., 2003. Türkiye’ de Kentlerin Sosyal ve Ekonomik Göstergeleri Arasındaki İlişki, İTÜ Dergisi, 2, (1): 3–14.

Gül, H., 1998, Decentralization, Local Government Reforms and Local Autonomy: A Comparative Analysis of Recent Trends, Studies in Comparative Analysis, Rod Erakovich and Dick Grisby, 9-43.

Gül, H., 2004. Ekonomik Kalkınmada Yerel Alternatifler, Kentsel Ekonomik Araştırmalar Sempozyumu, 1, (16):201–219.

Johnson, R.A., Wichern, D.W., 2002. Applied Multivariate Statistical Analysis, Pearson Education Ltd., USA, 767p.

Kaygısız, Z., Saraçlı, S., Dokuzlar, K.U., 2005. İllerin Gelişmişlik Düzeyini Etkileyen Faktörlerin Path Analizi ve Kümeleme Analizi ile İncelenmesi, 7. Ulusal Ekonometri Ve İstatistik Sempozyumu, İstanbul Üniversitesi.

Larose, D.T., 2005. Discovering Knowledge in Data, John Wiley & Sons, Canada, 221p.

Mahalanobis, P.C., 1936. On the Generalised Distance in Statistics, Proceedings of the National Institute of Sciences of India, 2, (1): 49–55.

Manisalı, E., 1975. Gelişme Ekonomisi, İ.Ü. İktisat Fakültesi Yayını, İstanbul, 2.

Özdamar, K., 2002. Paket Programlar ile İstatistiksel Veri Analizi 1 – 2 , Kaan Kitabevi, Eskişehir, 686– 513.

Özdemir, A.İ., Altınparmak A., 2005. Sosyo-Ekonomik Göstergeler Açısından İllerin Gelişmişlik Düzeyinin Karşılaştırmalı Analizi, Erciyes Üniversitesi İ.İ.B.F. Dergisi, (24): 97–110.

Özmen, İ., 1998. İlçelerin Sosyo-Ekonomik Gelişmişlik Sıralaması ve Gruplandırılmasına İlişkin Bir Çalışma”, Hazine Dergisi, (11): 41-61.

Öztürk, L., 2005. Bölgelerarası Gelir Eşitsizliği: İstatistikî Bölge Birimleri Sınıflandırması’na Göre Eşitsizlik İndeksleri ile Bir Analiz, 1965–2001, Akdeniz İ.İ.B.F. Dergisi (10): 95–110.

Rao, A.R., Srinivas, V.V., 2008. Regionalization of Watersheds, Springer, USA, 241p.

Teitz, M.B., 1994. Changes in Economic Development Theory and Practice, International Regional Science Review, 16, (1): 101–106.

Turanlı, M., Özden, Ü.H., Türedi S., 2006. Avrupa Birliği’ ne Aday ve Üye Ülkelerin Ekonomik Benzerliklerinin Kümeleme Analiziyle İncelenmesi, İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi, 9, (1): 95–108.

Üstünişik, N.Z., Türkiye' deki İller ve Bölgeler Bazında Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması: Gri İlişkisel Analiz Yöntemi ve Uygulaması, Yüksek Lisans Tezi, Gazi Üniversitesi, 2007.

Yavilioğlu, C., 2002. Kalkınmanın Anlambilimsel Tarihi ve Kavramsal Kökenleri, C.Ü. İktisadi Ve İdari Bilimler Dergi, 3, (1): 59.

EKLER

EK-1 Değişkenlere ilişkin elde edilen veriler¹

İL	İLÇE	Bolge_kod	İL_kod	Nuf_top	Okuma_bilmeyen	Okuma_bilen	Bilinmeyen
Adana	Seyhan	akdeniz	1	752308	,070168	,749983	,073875
	Ceyhan	akdeniz	1	158833	,093816	,755051	,045740
	Feke	akdeniz	1	19952	,122394	,718023	,058290
	Karaisal	akdeniz	1	23481	,140624	,734849	,037264
	Karataş	akdeniz	1	22472	,083393	,762148	,049617
	Kozan	akdeniz	1	124669	,094185	,760004	,046026
	Pozanti	akdeniz	1	21963	,072030	,791923	,047762
	Saimbeyl	akdeniz	1	17671	,106502	,741554	,047762
	Tufanbey	akdeniz	1	19045	,118299	,718876	,074665
	Yumurtal	akdeniz	1	19625	,082191	,783796	,036943
	Yüreğir	akdeniz	1	416160	,100284	,729606	,046139
	Aladag	akdeniz	1	18249	,131788	,733739	,043345
	imamoğlu	akdeniz	1	31257	,101865	,751256	,038551
	Sarıçam	akdeniz	1	109290	,053912	,793577	,033965
	Çukurova	akdeniz	1	271344	,021194	,846232	,048466
Adıyaman	Merkez	guneydogu	2	256247	,102366	,732875	,043197
	Besni	guneydogu	2	80651	,125318	,709501	,048865
	Çelikhan	guneydogu	2	15540	,111969	,727542	,040541
	Gerger	guneydogu	2	25769	,185611	,554775	,113935
	Gölbasi	guneydogu	2	47599	,109771	,737452	,048677
	Kahta	guneydogu	2	116049	,137209	,652612	,066394
	Samsat	guneydogu	2	10599	,106519	,716200	,040192
.....

¹ 957 satır ve 27 sütundan oluşan veriler ekte verilen CD' de "Veriler.sav" isimli dosyadadır.

EK-2 Amaç fonksiyonu değerlerini bulan MİNİTAB 14 Makro programı

```

WOpen "C:\Users\Asus\Desktop\objective function\hibrid\hibrid.MTW".
DO K2=6:13
LET K3= K2-5
LET C19(K3) = 0
STATISTICS c1-c5;
  by CK2;
  MEANS C14-C18.
IF K2=6
  LET K6=4
ENDIF
IF K2=7
  LET K6=5
ENDIF
IF K2=8
  LET K6=6
ENDIF
IF K2=9
  LET K6=4
ENDIF
IF K2=10
  LET K6=5
ENDIF
IF K2=11
  LET K6=6
ENDIF
IF K2=12
  LET K6=7
ENDIF
IF K2=13
  LET K6=7
ENDIF
DO K4=1:882
  DO K5=1:K6
  IF CK2(K4)=K5
  LET C19(K3)=C19(K3)+ (C1(K4)-c14(K5))**2+ (C2(K4)-c15(K5))**2+ (C3(K4)-
c16(K5))**2+ (C4(K4)-c17(K5))**2+(C5(K4)-c18(K5))**2
  ENDIF
  ENDDO
  ENDDO
  ENDDO
  ENDMACRO

```

EK-3 Ortalama silhouette genişliğini hesaplayan MS-DOS QBASIC kodları

```

10 OPEN "hibrid.DAT" FOR INPUT AS #1
15 OPEN "simple.out" FOR OUTPUT AS #2
  OPTION BASE 1
30 DIM M(882, 14)
  REM COMMANDS 40 TO 80 ASSIGN EACH ENTRY OF MATRIX M
  REM TO ZERO.
40 FOR I = 1 TO 882
50  FOR J = 1 TO 14
60  LET M(I, J) = 0
70  NEXT J
80 NEXT I
  REM COMMANDS 90 TO 130 ASSIGN DATA IN asamali.dat TO
  REM MATRIX M
90 FOR I = 1 TO 882
100  FOR J = 1 TO 14
110  INPUT #1, M(I, J)
120  NEXT J
130 NEXT I
LET K = 14
IF K = 7 THEN LET KS = 4
IF K = 8 THEN LET KS = 5
IF K = 9 THEN LET KS = 6
IF K = 10 THEN LET KS = 4
IF K = 11 THEN LET KS = 5
IF K = 12 THEN LET KS = 6
IF K = 13 THEN LET KS = 7
IF K = 14 THEN LET KS = 7
DIM N(KS), OUM(882, KS), EK(KS - 1)
LET SLG = 0
FOR I = 1 TO 882
  FOR SAY = 1 TO KS
    LET OUM(I, SAY) = 0
    LET N(SAY) = 0
  NEXT SAY
  FOR J = 1 TO 882
    IF I > J OR I < J THEN
      LET ARA = M(J, K)
      LET N(ARA) = N(ARA) + 1
      LET UZ = (M(I, 2) - M(J, 2)) ^ 2 + (M(I, 3) - M(J, 3)) ^ 2 + (M(I, 4) - M(J, 4)) ^ 2
      + (M(I, 5) - M(J, 5)) ^ 2 + (M(I, 6) - M(J, 6)) ^ 2
      LET OUM(I, ARA) = OUM(I, ARA) + SQR(UZ)
    END IF
  NEXT J
  LET AA = 0
  FOR SAY = 1 TO KS

```


EK-3 Ortalama silhouette genişliğini hesaplayan MS-DOS QBASIC kodları (devamı)

```
LET OUM(I, SAY) = OUM(I, SAY) / N(SAY)
IF M(I, K) = SAY THEN
  LET A = OUM(I, SAY)
ELSE
  LET AA = AA + 1
  LET EK(AA) = OUM(I, SAY)
END IF
NEXT SAY
LET B = 100000000
FOR SAY = 1 TO KS - 1
  IF EK(SAY) < B THEN LET B = EK(SAY)
NEXT SAY
IF A < B THEN SG = (B - A) / B
IF A >= B THEN SG = (B - A) / A
PRINT I, SG
LET SLG = SLG + SG
NEXT I
LET SLG = SLG / 882
PRINT "SILHOUTTE GENISLIGI="; SLG
```

EK-4 Davies-Bouldin indeksini hesaplayan MİNİTAB 14 Makro programı

```

GMACRO
DAVIES_BOULDIN
WOpen "C:\Users\Asus\Desktop\objective function\hibrid\HIBRID.MTW".
LET K3=0
DO K2=6:13
LET K3= K3+1
STATISTICS c1-c5;
  by CK2;
  MEANS C14-C18.
IF K2=6
  LET K6=4
ENDIF
IF K2=7
LET K6=5
ENDIF
IF K2=8
  LET K6=6
ENDIF
IF K2=9
  LET K6=4
ENDIF
IF K2=10
  LET K6=5
ENDIF
IF K2=11
  LET K6=6
ENDIF
IF K2=12
LET K6=7
ENDIF
IF K2=13
LET K6=7
ENDIF
LET K15=18+K3
DO K10=1:K6
LET CK15(K10)=0
ENDDO
DO K4=1:882
  DO K5=1:K6
IF CK2(K4)=K5
LET K7=(C1(K4)-C14(K5))**2+ (C2(K4)-C15(K5))**2+ (C3(K4)-C16(K5))**2+
(C4(K4)-C17(K5))**2+(C5(K4)-C18(K5))**2
LET K8=CK2(K4)
LET CK15(K8)=CK15(K8)+ SQRT(K7)
ENDIF
ENDDO

```

EK-4 Davies-Bouldin indeksini hesaplayan MİNİTAB 14 Makro programı (devamı)

```
ENDDO
LET K20=K15+1
LET K21=K15+2
TALLY CK2;
COUNT;
STORE CK20-CK21.
DO K25=1:K6
LET CK15(K25)=CK15(K25)/CK21(K25)
ENDDO
ERASE CK20-CK21
DO K30=1:K6
DO K31=1:K6
IF K31 NE K30
LET K32=(C14(K30)-C14(K31))**2+ (C15(K30)-C15(K31))**2+ (C16(K30)-
C16(K31))**2+ (C17(K30)-C17(K31))**2+(C18(K30)-C18(K31))**2
LET K33=SQRT(K32)
LET CK20(K31)=( CK15(K30)+CK15(K31) ) / K33
ENDIF
ENDDO
LET CK21(K30)=MAX(CK20)
ENDDO
LET K35=MEAN(CK21)
NOTE DAVIES-BOULDIN GOSTERGE
PRINT K35
ERASE CK20-CK21
ENDDO
ENDMACRO
```

EK-5 Seçilen temel bileşenlere ait veriler ²

İl	İlçe	Bölge_kod	İl_kod	T.B. 1	T.B. 2	T.B. 3	T.B. 4	T.B. 5	
Adana	Seyhan	akdeniz	1	1,50729	-1,10619	1,31208	1,12234	-,49612	
	Ceyhan	akdeniz	1	,67275	-,70326	1,26433	,73661	,34217	
	Feke	akdeniz	1	-,89339	-,33533	-,11076	-,37841	1,09220	
	Karaisal	akdeniz	1	-,86893	-,18305	,57388	-,25640	,94999	
	Karataş	akdeniz	1	,23484	-,39375	,84906	-,42794	,47989	
	Kozan	akdeniz	1	,23150	-,41214	1,21599	,54631	,87806	
	Pozanti	akdeniz	1	,21593	,32857	-,10589	,11337	,62807	
	Saimbeyl	akdeniz	1	-1,09299	-,26751	-,05981	-,26259	1,58488	
	Tufanbey	akdeniz	1	-,42302	-,80826	,22615	-1,01890	1,01710	
	Yumurtal	akdeniz	1	,16690	-,08372	,76950	-,18655	,95746	
	Yüreğir	akdeniz	1	,49204	-,87518	1,66496	1,83495	,49492	
	Aladag	akdeniz	1	-,82553	-,86337	-,04979	-,53575	1,83609	
	imamoğlu	akdeniz	1	-,39641	-,16885	,48495	,05029	1,30007	
	Sarıçam	akdeniz	1	-,53120	,90537	-3,84415	2,07858	1,82702	
	Çukurova	akdeniz	1	2,18380	-,26546	-3,50080	1,02578	,66174	
	Adıyaman	Merkez	guneydogu	2	,16063	-1,45162	1,08013	1,15869	2,12452
		Besni	guneydogu	2	-,38475	-,83136	,70157	,81643	,72821
Çelikhan		guneydogu	2	-,40377	-1,37873	-,19078	-,50495	2,13575	
Gerger		guneydogu	2	-1,65618	-1,52700	-,06009	,13732	-2,36972	
Gölbasi		guneydogu	2	,13283	-,66399	,59178	,34429	1,09306	
Kahta		guneydogu	2	-1,07815	-1,45038	,76415	1,52380	,21322	
.....		

² Temel bileşene ait verilerin tamamı ekte verilen CD' deki "TBileşen.sav" dosyasındadır.

EK-6 Sosyo-ekonomik gelişme düzeylerine göre ilçelerin kümelenmiş listesi³

1. DÜZEY	2. DÜZEY	3. DÜZEY	4. DÜZEY	5. DÜZEY	6. DÜZEY	7. DÜZEY	8. DÜZEY
İSTANBUL	SAKARYA	ADIYAMAN	KASTAMONU	Şanlıurfa	AKSARAY	BİTLİS	BİNGÖL
Adalar	Adapazarı	Merkez	Abana	Akçakale	Ağaçören	Adilcevaz	Adaklı
ANKARA	MERSİN	K.MARAŞ	NEVŞEHİR	Konya	ELAZIĞ	BİTLİS	SİVAS
Altındağ	Akdeniz	Afşin	Acıgöl	Altineki	Ağın	Ahlat	Akıncılar
İSTANBUL	ANTALYA	AFYON	DENİZLİ	Yalova	KASTAMONU	HATAY	KARS
Arnavutköy	Aksu	Merkez	Acıpayam	Armutlu	Ağlı	Altınözü	Akyaka
İstanbul	SAKARYA	DÜZCE	BURDUR	Mersin	KONYA	IĞDIR	GİRESUN
Bağcılar	Arifiye	Akçakoca	Ağlasun	Aydıncık	Ahırlı	Aralık	Alucra
İstanbul	SAMSUN	MANİSA	MANİSA	Van	TRABZON	ELAZIĞ	MALATYA
Bakırköy	Atakum	Akhisar	Ahmetli	Bahçesaray	Akçaabat	Arıcak	Arapgir
İstanbul	İSTANBUL	AKSARAY	YOZGAT	Van	MALATYA	BATMAN	ARDAHAN
Başakşehir	Ataşehir	Merkez	Akdağmadeni	Başkale	Akçadağ	Merkez	Merkez
İstanbul	ERZURUM	KONYA	ORDU	Çanakkale	KIRŞEHİR	SİİRT	MALATYA
Bayrampaşa	Aziziye	Akşehir	Akkuş	Bozcaada	Akçakent	Baykan	Arguvan
İstanbul	DİYARBAKIR	SAKARYA	ANTALYA	Rize	KAYSERİ	ADIYAMAN	KARS
Beşiktaş	Bağlar	Akyazı	Akseki	Çayeli	Akkışla	Besni	Arpaçay
Ankara	İSTANBUL	ANTALYA	ANKARA	HAKKARİ	DENİZLİ	BATMAN	YOZGAT
Çankaya	Bahçelievler	Alanya	Akyurt	Çukurca	Akköy	Beşiri	Aydıncık
İzmir	KOCAELİ	İZMİR	ÇORUM	ARDAHAN	KONYA	ŞIRNAK	ÇANKIRI
Dikili	Başiskele	Aliağa	Alaca	Damal	Akören	Beytüşşebap	Bayramören
Kocaeli	İZMİR	AMASYA	SAMSUN	MARDİN	KIRŞEHİR	BİNGÖL	YOZGAT
Dilovası	Bayraklı	Merkez	Alaçam	Dargeçit	Akpınar	Merkez	Boğazlıyan

³ 8 düzeye kümelenmiş 957 ilçeye ait listeler ekte verilen "kümelenmiş ilçelerin listesi.doc" isimli dosyadadır.

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : SAYIN, Gökhan
Uyruđu : T.C.
Dođum tarihi ve yeri : 22.10.1984 AFYON
Medeni hali : BEKAR
e-mail : gokhansayin35@mynet.com

Eđitim

Derece	Eđitim Birimi	Mezuniyet Tarihi
Lisans	Ege Üniversitesi/ İstatistik Bölümü	2007
Lise	Kemer Anadolu Lisesi	2002

Yabancı Dil

İngilizce

Hobiler

Müzik, kitap okumak, dođa yürüyüşü