

# KİŞİ BAŞINA DÜŞEN GSYİH DEĞERLERİNE GÖRE TÜRKİYE'DEKİ COĞRAFI BÖLGELERİN VE GSYİH'YI OLUŞTURAN SEKTÖRLERİN KÜMELENMESİ

Muammer YAYLALI\*  
Erkan OKTAY\*\*  
Yusuf AKAN\*\*\*

**Özet:** Bu çalışmanın teorik kısmında hiyerarşik kümeleme metotları ve iki-yönlü kümeleme metodu tanıtılmıştır. Çalışmanın uygulama bölümünde Türkiye'deki coğrafi bölgeler, Kişi Başına Düşen (KBD) Gayri Safi Yurtiçi Hasıla (GSYİH) değerleri itibariyle kümeleneştir. Ayrıca bölgeler itibariyle KBD GSYİH değerlerini oluşturan sektörler de kümeleneştir. Çalışmanın son bölümünde ise KBD GSYİH değerleri yardımıyla hem coğrafi bölgelere göre ve hem de GSYİH değerlerini oluşturan sektörlere göre kümeleme yapılmıştır. Bazı kümeleme tekniklerinde birbirinden farklı kümeler elde edilmiştir. Ancak kümeleme tekniklerinden elde edilen sonuçlar genelleştirildiğinde makul bir sonuca ulaşılmıştır.

**Anahtar Kelimeler:** Kümeleme analizi, hiyerarşik kümeleme analizleri, iki-yönlü kümeleme, kareli euclidean uzaklıklar, gayrisafı yurtiçi hasıla.

## I. Giriş

Bölgeler arasında dengeli bir gelişme sağlanması hedefi, ülke ekonomisi için büyük önem arz etmektedir. Dengeli gelişme amacı doğrultusunda alınması gereken tedbirlerin ve uygulanacak politikaların beşeri ve fiziki kaynakların dağılım deseni ile tutarlı olması ve kamu yatırımlarının dağılımında ekonomik coğrafyanın, bölgesel gelişmenin ve sosyal faydanın 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 önem taşımaktadır. Zira, Avrupa Birliği, üye ülkelerdeki bölgesel gelişmişlik farklılıklarını gidermek, mekansal 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 bağlamda ülkemiz gündeminde önemli bir yeri olan ve bölgesel gelişmişlik farklarının giderilmesi ve sürdürülebilir kalkınma ilkesi ile uyumlu politikaların gerçekleştirilmesinde gelişmişliğin belirlenmesi büyük önem arz etmektedir (DPT 2003:2-3). Bu amaçla bu çalışmada bölgelerin gelişmişlik düzeyinin belirlenmesinde Gay-

---

\* Prof. Dr., Atatürk Üniversitesi, IIBF, İktisat ABD

\*\* Doç. Dr., Atatürk Üniversitesi, IIBF, İşletme ABD

\*\*\* Doç. Dr., Atatürk Üniversitesi, IIBF, İktisat ABD

Gayrisafi Yurtiçi Hasıla (GSYİH) değişkeni kullanılmıştır. GSYİH değişkeni, ekonomik performansı ve gelir düzeyini yansıtan temel bir gösterge niteliğinde olması nedeniyle dikkate alınmıştır (Dornbusch ve Fischer 1998: 29).

Makro ekonomik hesaplar, ekonominin performansının değerlendirilmesine imkân tanıyan ve bu hesaplar aracılığıyla belirlenen ilişkileri kullanarak ileriye yönelik tahmin ve planlamaya yardımcı olan bütüncül bir sistemdir. Makro ekonomik hesaplar, ekonomide çok sayıda kişi ve kuruluş arasında geçen değişimleri sistemli ve ekonomik analize imkân verecek şekilde sunmayı amaçlamaktadır. Ülkedeki ekonomik faaliyetlerin genel düzeyini göstermesi bakımından, Gayrisafi Yurtiçi Hasıla bu hesap sisteminde en çok kullanılan değişkendir (Yükseler: 1998: 1). Bir ekonomide, belirli bir dönemde, ülkede yerleşik üretim birimleri tarafından üretilen nihai mal ve hizmetlerin değerlerinin toplamı Gayrisafi Yurtiçi Hasılayı vermektedir.

Gayri Safi Yurtiçi Hasıla'nın alternatif ölçme yöntemlerinden biri olan toplam üretim yaklaşımında, Gayrisafi Yurtiçi Hasılanın katma değerler toplamına eşit olduğu noktasından hareket edilmekte ve GSYİH, ekonomiyi oluşturan çeşitli faaliyet alanları arasındaki tüm firmaların katma değerleri hesaplanarak ölçülmektedir. Toplam üretim yaklaşımı, bir ekonomideki çeşitli üretim dallarının Gayri Safi Yurtiçi Hasıla'ya katkılarını ve dolayısıyla da üretimin faaliyet alanları itibariyle yapısını göstermektedir (Ünsal 2004:43-44). Faaliyet kollarına göre Gayri Safi Milli Hasıla, cari fiyatlarla ve sabit fiyatlarla olmak üzere iki yolla hesaplanmaktadır. Cari fiyatlarla hesaplanan Gayrisafi Milli Hasıladaki gelişmeler, gerçek gelişmeleri yansıtmayabilir. Zira, çeşitli yılların GSMH rakamlarının reel olarak karşılaştırması gerektiğinde, fiyat değişikliklerinin yol açtığı etkinin ortadan kaldırılması gerekmektedir. Sabit fiyatlarla GSMH bu etkiyi gidererek belli bir dönemde yer alan gerçek değişimleri ölçmektedir. Bu nedenle bu çalışmada, Devlet İstatistik Enstitüsü tarafından hazırlanan sabit fiyatlarla 1987 yılı temel alınarak hesaplanan GSYİH değerleri kullanılmıştır. Faaliyet Kollarına göre GSMH hesaplamalarında kullanılan iktisadi faaliyet kolları şunlardır: 1-Tarım (Çiftçilik, Ormancılık, Balıkçılık, Hayvancılık), 1-Sanayi (Madencilik-Taşocakçılığı, İmalat Sanayi, Elektrik-Gaz-Su), 3-İnşaat Sanayi, 4- Ticaret (Toptan-Perakende Ticaret, Otel-Lokanta Hizmetleri), 5-Ulaştırma ve Haberleşme, 6-Mali Kuruluşlar, 7-Konut Sahipliği, 8-Serbest Meslek ve Hizmetleri, 9-Diğer Faaliyetler (Devlet Hizmetleri + Kâr Amacı Olmayan Kuruluşlar + İthalat Vergisi – İzafi Banka Hizmetleri) (DİE 2002:XI).

#### **A. Hiyerarşik Kümeleme Metotları**

Kümeleme analizi kavramı, 1939 yılında Tryon tarafından literatüre tanıtılmıştır (Lorr, 1993: 2). Kümeleme analizi, ölçülen objelerin birbirine benzerliklerine göre sınıflamayı sağlayan çok değişkenli bir tekniktir. Objeler önceden tespit edilen kriterlere göre kümelendiğinde aynı kümedeki objelerin ölçülen karakteristikleri birbirine çok benzerken kümeler arası karşılaştırmalarda ölçülen

karakteristikler birbirine benzemezler (Hair, 1998: 473). Bugün literatüre kazandırılmış çok sayıda kümeleme tekniği mevcuttur. Bu tekniklerin en çok bilinenleri, hiyerarşik kümeleme analizleri, iki-yönlü kümeleme (blok kümeleme) analizi ve k-ortalama kümeleme analizidir.

Objeler, aralarındaki benzerliğe, ilişkiye veya uzaklığa göre kümelenebilir. Oktay, 2001'de benzerlik ve ilişki ölçülerini tanıtan monografik bir çalışma yapmıştır. Hiyerarşik kümelemelerin oluşturulmasında Euclidean Uzaklıklar, Kareli Euclidean Uzaklıklar, Manhattan Uzaklıklar, Chebychev Uzaklık Ölçüsü, Yüzde Uyumsuzluk, 1 – Pearson r ve  $\left[ \sum |x - y|^p \right]^{1/r}$  değerleri kullanılabilir. Bu çalışmada hiyerarşik kümelerin oluşturulmasında Kareli Euclidean Uzaklıklar kullanılacaktır. n boyutlu bir anakütlede iki nokta arasındaki farkların kareleri toplamı,

$$d_{ij}^2 = \sum_{k=1}^n (X_{ij} - X_{jk})^2 \quad (1)$$

formülü ile elde edilir.  $X_{ij}$  ve  $X_{jk}$  değerleri k adet değişkenin i ve j koordinatındaki değerlerini ifade eder. Objeler, birbirinden farklı ölçüm birimleriyle ölçülmüş değişkenlere göre kümelenecekse, değişkenler önce standart değerlere dönüştürülmeli ve sonra kareli Euclidean Uzaklıklar hesaplanmalıdır.

Hiyerarşik kümeleme metotları uygulanırken kaç küme oluşacağı önceden bilinmez. Önceden tespit edilen kriterlere göre (bu çalışma için Kareli Euclidean Uzaklıklar) birbiriyle en fazla benzeyen objeler aynı kümenin içinde yer alır. Veri setindeki en son obje kendisine en çok benzeyen kümenin içinde yer almasına kadar kümeleme işlemine devam edilir.

Başlıca yedi tür hiyerarşik kümeleme metodu vardır. Bunlar tek bağlantı metodu, tam bağlantı metodu, tartısız eşlenik grup ortalaması metodu, tartılı eşlenik grup ortalaması metodu, tartısız eşlenik grup merkezi metodu, tartılı eşlenik grup merkezi (medyan) metodu ve Ward metodudur.

#### 1. Tek Bağlantı (En Yakın Komşu) Metodu

Tek bağlantı metodu, ilk olarak Florek (1951) tarafından tarif edilmiştir. Bu metot daha sonra, ilk araştırmacıdan ve birbirlerinden bağımsız olarak Sneath (1957) ve McQuitty (1957) tarafından da ileri sürülmüştür (Lorr, 1993:62). Tek bağlantı kümeleme analizi hiyerarşik kümeleme analizleri içerisinde en basit olanıdır. Her aşamada p kümesi ile q kümesi birleştirildikten sonra oluşan yeni küme (t kümesi) ile başka bir küme (r kümesi) arasındaki benzerlik şöyle tarif edilir: Uzaklık ölçüsü  $d_{ij}$  ile gösterildiğinde, t kümesi ile r kümesinin birbirine en yakın olan iki değeri arasındaki uzaklık

$$d_{tr} = \min(d_{pr}, d_{qr}) \quad (2)$$

şeklindedir. t kümesi ile r kümesi birleştirilmişse, elde edilen kümedeki bir de-  
ğerin en yakın komşusuna olan uzaklığı en fazla  $d_{tr}$  kadar olur (Anderberg,  
1973:137).

Bu metodun sağladığı en önemli avantaj benzerlik matrisinin tekdüze  
transformasyonlara karşı değişken olmaması ve veri setindeki bağlı değerlerden  
etkilenmemesidir (Aldenderfer-Blashfield, 1984:38). Bu metot birbirinden yete-  
rince ayrıık olan kümeleri tespitite oldukça iyidir. Tek bağlantı metoduyla oluşturu-  
lulan kümedeki değerler, diğer kümedeki değerlere göre birbirine daha çok ben-  
zer.

Tek bağlantı metodu birbirinden çok az farklı olan kümeleri ayırmada yete-  
ersizdir. Tek bağlantı metodu elips şeklinde dağılmayan değerleri aynı kümede  
toplayabilen az sayıdaki kümeleme metodundan biridir. Mesela koordinat sis-  
teminde U şeklinde dağılan noktalar bu metoda göre aynı kümenin içerisinde  
yer alabilir. Bu sebeple kümenin zıt taraflarındaki değerler birbirinde oldukça  
farklı olabilir. Zincirleme etkiyle ortaya çıkan bu durum sebebiyle çoğu kez bir-  
birinden oldukça farklı üyeler aynı kümede yer alabilir (Anderberg, 1973:138).

## 2. Tam Bağlantı (En Uzak Komşu) Metodu

Tam bağlantı metodu birbirinden bağımsız çalışan birkaç araştırmacı ta-  
rafından farklı zamanlarda ileri sürülmüştür. Horn (1943), bu metodu test deęiş-  
kenlerini kümelemek için kullanmıştır. Sorensen (1948), ekoloji ile ilgili çalış-  
malarda kullanmak için bu metodu geliştirmiştir. McQuitty (1961), tür analizle-  
rini ve sıralı tür analizlerini (1963) önermiştir (Lorr, 1993:63).

Tam bağlantı metodu tek bağlantı metodunun tam tersi bir mantıkla kü-  
meleme yapmayı sağlar. Bu metodun uygulanması tek bağlantı metodu kadar  
kolaydır. Her aşamada p kümesi ile q kümesi birleştirildikten sonra oluşan yeni  
küme (t kümesi) ile başka bir küme (r kümesi) arasındaki benzerlik şöyle tarif  
edilir: Uzaklık ölçüsü  $d_{ij}$  ile gösterildiğinde, t kümesi ile r kümesinin birbirine  
en uzak olan iki değeri arasındaki uzaklık

$$d_{tr} = \max(d_{pr}, d_{qr}) \quad (3)$$

şeklindedir. Eğer t kümesi ile r kümesi birleştirilmişse, elde edilen kümedeki  
bir deęerin en uzak komşusuna olan uzaklığı en fazla  $d_{tr}$  kadar olur.  $d_{tr}$  değeri, t  
kümesi ve r kümesi birleştirildikten sonra elde edilen kümeyi kavrayan en kü-  
çük dairenin çapıdır (Johnson – Wichern, 1999:744-745).

Bu metoda göre de benzerlik matrisi tekdüze transformasyonlara karşı  
değişken değildir (Johnson, 1967:241-254). Tam bağlantı metodu tek bağlantı  
metodundan daha katı kurallara sahiptir. Bu sebeple tam bağlantı metodu X-Y  
koordinat sisteminde birbirine yakın noktaların yumurta şeklinde dağılım gös-  
termesi durumunda bu değerleri kümelemede tek bağlantı metoduna göre daha  
iyi sonuçlar sağlar (Aldenderfer – Blashfield, 1984:38).

### 3. Eşlenik Grup Ortalaması Metodu

Tartısız eşlenik grup ortalaması metodu ilk olarak Sokal ve Michener (1958) tarafından ortaya atılmıştır. Ancak Sokal ve Sneath (1963), literatüre tanıtımını sağlamıştır (Lorr, 1993:65). Tek bağlantı metodunda birbirine en yakın komşular ve tam bağlantı metodunda birbirine en uzak komşulardan başlanarak kümeleme yapılmaktadır. Her iki metot da aşırı değerlerin direk etkisi altındadır. Ortalama bağlantı metodunda bu dezavantaj giderilmiştir. Bu metoda göre bağlantı kurulacak objelerin ortalaması alınarak kümeleme yapılır. Ancak uygulamada tartısız eşlenik grup ortalaması metoduna göre bulunan sonuçlar tam bağlantı metoduna göre bulunan sonuçlara çok benzer.  $i$  kümesinden bir değer ve  $j$  kümesinden bir değer alınarak bulunan bütün eşlenik kombinasyonlarla ilişkili benzerlikler toplamı alınarak ilk benzerlik matrisindeki  $d_{ij}$  değerleri hesaplanabilir. Herhangi bir birleştirme yapılmadan önce her küme yalnız böyle bir değer ihtiva eder ve her küme çiftinde böyle bir değer çifti vardır.  $p$  kümesi ile  $q$  kümesi birleştirildiğinde oluşan yeni küme ( $t$  kümesi) ile başka bir  $r$  kümesi arasındaki eşlenik benzerlikler toplamı, benzerlik matrisi kullanılarak

$$d_{tr} = d_{pr} + d_{qr} \quad (4)$$

şeklinde yazılır.  $i$  kümesinde yer alan bütün eşlenik çiftler arasındaki benzerlikler toplamını  $top_i$  ve  $i$  kümesindeki değer sayısını  $n_i$  ile gösterildiğinde, ilk birleştirme işleminden önce bütün kümeler için  $top_i = 0$  ve  $n_i = 1$  olur.  $p$  kümesi ile  $q$  kümesi birleştirildiğinde elde edilen küme ( $t$  kümesi),

$$top_t = top_p + top_q + d_{pq} \quad (5)$$

$$n_t = n_p + n_q$$

olarak tarif edilir. Birbirine en çok benzeyen çifti araştırırken  $i$  kümesi ile  $j$  kümesinin aday çiftinin birleştirilmesinden sonra elde edilen kümedeki grup içi benzerlik ortalaması

$$\frac{top_i + top_j + d_{ij}}{(n_i + n_j)(n_i + n_j - 1) / 2} \quad (6)$$

formülüyle elde edilir. Formül 7'deki payda,  $i$  kümesi ile  $j$  kümesinin birleştirilmesinden sonra şekillenen birbirinden ayrık eşlenik çift sayısını göstermektedir (Anderberg, 1973:139).

### 4. Eşlenik Grup Merkezi Metodu (Medyan Metodu)

Bu metoda göre, kümelerde birbirine benzeyen ortalama vektörlere göre veya benzer küme merkezlerine göre kümeler birleştirilir. Benzerlik matrisinin kareli Euclidean Uzaklıklara göre bulunması durumunda ancak bu metot mantıklı sonuçlar verir.  $p$  kümesi ile  $q$  kümesi birleştirildikten sonra oluşan yeni küme ( $t$  kümesi) ile başka bir küme ( $r$  kümesi) arasındaki uzaklık

$$d_{tr} = \frac{N_p}{N_p + N_q} d_{pr} + \frac{N_q}{N_p + N_q} d_{qr} - \frac{N_p N_q}{N_p + N_q} d_{pq} \quad (7)$$

şeklinde hesaplanır (Lance – Williams, 1967). Bu eşitlik, hem değişkenlere hem de veri ünitelerine rahatlıkla uygulanabilir. Daha önce ifade edildiği gibi bu metodun tutarlı sonuçlar verebilmesi için benzerlik matrisi, i kümesi ile j kümesi merkezleri arasındaki kareli Euclidean Uzaklıklara göre hazırlanmalıdır.

Medyan metodu ilk olarak Gower tarafından ortaya atılmıştır (Gower, 1967:626). Değişken değerleri en az sıralama ölçeği ile ölçüldüğünde bu metod uygulanabilmektedir. p kümesi ile q kümesi birleştirildikten sonra oluşan yeni küme (t kümesi) ile başka bir küme (r kümesi) arasındaki uzaklık,

$$d_{tr} = \frac{1}{2}(d_{pr} + d_{qr}) - \frac{1}{4}d_{pq} \quad (8)$$

formülü yardımı ile elde edilir (Lance – Williams, 1967:375). Medyan metodu, kümeleme ölçüsü olarak p, q ve r küme merkezleri arasındaki kareli Euclidean Uzaklıkların Medyanını kullanır. Bu metoda göre kümelerdeki birim sayısı tartı olarak kullanılmaktadır (Sneath – Sokal, 1973:234-235).

##### 5. Ward Metodu

Ward metodu küme içindeki en küçük varyansı optimum kılmak için tasarlanmıştır. (Ward, 1963) Bu metod grup içi kareler toplamı veya hata kareler toplamı olarak da bilinir. Wishart (1969), küme merkezleri arasındaki kareli Euclidean Uzaklıklar matrisini kullanarak Ward metodunun nasıl kullanılacağını göstermiştir.

h kümelerinin k incisindeki  $m_k$  veri ünitelerinin j incisi için n değişkenlerinin i incisinden elde edilen ölçüm  $x_{ijk}$  ile gösterilecektir. Buna göre k inci kümedeki veri üniteleri için i inci değişkenin ortalaması

$$\bar{x} = \sum_{j=1}^{j=m_k} x_{ijk} / m_k \quad (9)$$

formülüyle elde edilir. k kümesi için hata kareler toplamı, k kümesindeki her bir veri noktasından k kümesinin ortalama vektörüne olan Euclidean Uzaklıklar toplamı ve k kümesi ortalaması etrafındaki grup içi kareli sapmalar

$$E_k = \sum_{i=1}^{i=n} \sum_{j=1}^{j=m_k} (x_{ijk} - \bar{x}_{ik})^2 = \sum_{i=1}^{i=n} \sum_{j=1}^{j=m_k} x_{ijk}^2 - m_k \sum_{i=1}^{i=n} \bar{x}_{ik}^2 \quad (10)$$

şeklinde gösterilebilir. Kümeler topluluğu için grup içi hata kareler toplamının toplamı

$$E = \sum_{k=1}^{k=h} E_k \quad (11)$$

formülüyle bulunur. Başlangıçta m adet veri ünitesi varsa, işe başlarken her veri ünitesinin kendi içinde bir küme oluşturduğu varsayılır. Bu sebeple, başlangıçta her küme için  $E_k = 0$  kabul edilir. Ward metoduna göre, her aşamada toplam grup içi kareler toplamı E'deki minimum artışı doğurabilecek kümeler bulunarak birleştirilir. Birleştirilecek kümeler p kümesi ile q kümesi ise elde edilecek t kümesinin E'de meydana getirdiği artış,

$$\Delta E_{pq} = m_p \sum_{i=1}^{i=n} \bar{x}_{ip}^2 + m_q \sum_{i=1}^{i=n} \bar{x}_{iq}^2 - m_t \sum_{i=1}^{i=n} \bar{x}_{it}^2 \quad (12)$$

formülü yardımıyla bulunur. Yeni küme için i inci değişkenin ortalaması

$$m_t \bar{x}_{it} = m_p \bar{x}_{ip} + m_q \bar{x}_{iq} \quad (13)$$

ilişkisinden bulunur. Her iki tarafın karesi alındığında

$$m_t^2 \bar{x}_{it}^2 = m_p^2 \bar{x}_{ip}^2 + m_q^2 \bar{x}_{iq}^2 + 2m_p m_q \bar{x}_{ip} \bar{x}_{iq} \quad (14)$$

bulunur. Ortalamaların çarpımı

$$2\bar{x}_{ip} \bar{x}_{iq} = \bar{x}_{ip}^2 + \bar{x}_{iq}^2 - (\bar{x}_{ip} - \bar{x}_{iq})^2 \quad (15)$$

şeklinde yazılabilir. Böylece

$$m_t^2 \bar{x}_{it}^2 = m_p (m_p + m_q) \bar{x}_{ip}^2 + m_q (m_p + m_q) \bar{x}_{iq}^2 + m_p m_q (\bar{x}_{ip} - \bar{x}_{iq})^2 \quad (16)$$

elde edilir.  $m_t = m_p + m_q$  'dur. Formül 15'deki eşitliğin her iki tarafı  $m_t^2$  'ye bölündüğünde

$$\bar{x}_{it}^2 = \frac{m_p}{m_t} \bar{x}_{ip}^2 + \frac{m_q}{m_t} \bar{x}_{iq}^2 - \frac{m_p m_q}{m_t^2} (\bar{x}_{ip} - \bar{x}_{iq})^2 \quad (17)$$

bulunur. Formül 9'daki ifade Formül 8'deki yerine yazıldığında

$$\Delta E_{pq} = \frac{m_p m_q}{m_p + m_q} \sum_{i=1}^{i=n} (\bar{x}_{ip} - \bar{x}_{iq})^2 \quad (18)$$

elde edilir. Görüldüğü üzere hata kareler toplamındaki minimum artış, birleştirilen kümelerin merkezleri arasındaki kareli Euclidean uzaklıkla orantılıdır. Bu sonuç tartılı eşlenik grup merkezi metodundan farklıdır. Hata kareler toplamı fonksiyonu azalmayan bir fonksiyondur ve bu metodun tersi bir fonksiyonun kullanılabilmesi mümkün değildir. p kümesi ile q kümesinin birleşmesiyle elde edilen kümeye t kümesi ve başka bir kümeye de r kümesi denirse, potansiyel olarak r ve t kümesinin birleştirilmesinden kaynaklanan E'deki artış

$$\Delta E_{rt} = \frac{m_r m_t}{m_r + m_t} \sum_{i=1}^{i=n} (\bar{x}_{ir} - \bar{x}_{it})^2 \quad (19)$$

eşitliği ile hesaplanır. Eşitlik 18'in sağ tarafındaki  $(\bar{x}_{ir} - \bar{x}_{it})^2$  ifadesinde  $\bar{x}_{it}$  yerine  $\bar{x}_{it} = (m_p \bar{x}_{ip} + m_q \bar{x}_{iq}) / m_t$  yazılır ve  $m_t = m_p + m_q$  olduğu dikkate alınır

$$(\bar{x}_{ir} - \bar{x}_{it})^2 = \frac{m_p}{m_t} (\bar{x}_{ir} - \bar{x}_{ip})^2 + \frac{m_q}{m_t} (\bar{x}_{ir} - \bar{x}_{iq})^2 - \frac{m_p m_q}{m_t^2} (\bar{x}_{ip} - \bar{x}_{iq})^2 \quad (20)$$

şeklinde yeniden yazılabilir. Bu ifade Formül 19'daki yerine yazılırsa

$$\Delta E_{rt} = \frac{1}{m_r + m_t} \sum_{i=1}^{i=n} \left[ m_r m_p (\bar{x}_{ir} - \bar{x}_{ip})^2 + m_r m_q (\bar{x}_{ir} - \bar{x}_{iq})^2 - \frac{m_r m_p m_q}{m_p + m_q} (\bar{x}_{ip} - \bar{x}_{iq})^2 \right] \quad (21)$$

bulunur. Böylece Formül 18'i kullanarak

$$\Delta E_{rt} = \frac{1}{m_r + m_t} \left[ (m_r + m_p) \Delta E_{rp} + (m_r + m_q) \Delta E_{rq} - m_r \Delta E_{pq} \right] \quad (22)$$

yazılabilir. Özetle Ward metodunun uygulanışı dört safhada gerçekleşir:

1. Benzerlik matrisi ( $d_{ij} = E_{ij}$  olmak üzere) oluşturulur.
2. Her aşamada hata kareler toplamını en az artıracak iki küme (p kümesi ve q kümesi) birleştirilir (t kümesi).
3. Benzerlik matrisindeki değerler Formül 22'ye göre düzenlenir. Formül 23'teki ilişkiyle hareketle hata kareler toplamındaki artış hesaplanabilir.

$$E = E + \frac{1}{2} \Delta E_{pq} \quad (23)$$

4. İkinci ve üçüncü maddedeki işlemler m adet veri ünitesi için m - 1 defa tekrar edilerek kümeleme işlemine son verilir.

### **B. İki-Yönlü Kümeleme**

Yukarıda izah edilen kümeleme metotları ya satır değişkenlerini ya da sütun değişkenlerini kümelemeyi sağlar. Hem satır değişkenine hem sütun değişkenine göre manalı kümelerin oluşturulabilmesi durumunda, iki-yönlü kümeleme metodu kullanılabilir. İki yönlü kümeleme grafiği oluşturulurken satır değişkeni için oluşturulan kümeler X ekseninde sütun değişkeni için oluşturulan kümeler Y ekseninde bantlarla gösterilir. Böylece X-Y koordinat sisteminde çeşitli renklerle tarif edilen alt kümeler belirlenir. Hem satır hem sütun değişkenine göre aynı alt kümede yer alma durumu aynı renklerle gösterilir (Hartigan, 1975; Zupan, 1982; Fisher, 1969).



**C. Türkiye’deki Coğrafi Bölgelerin KBD GSYİH Değerlerine Göre Hiyerarşik Kümeleme Metotlarıyla Gruplaması**

Gayri safi yurtiçi hasılının coğrafi dağılımı, o ülkedeki bölgeler ve illerin gelişmişlik düzeyini ortaya koyan göstergelerden biridir. Bu düşünceden hareketle, bir bölgedeki iller arasındaki gelişmişlik farkları ortaya konabilir (Güler, 1999: 68). Yapılan bir çalışmada, Doğu Anadolu Projesi (DAP) kapsamındaki illerin sektörlere göre GSYİH değerleri itibariyle kümelenmiştir (Oktay, 2001: 329-345). Bu çalışmada, Türkiye’deki coğrafi bölgelerin kişi başına düşen GSYİH değerleri itibariyle hiyerarşik kümeleme metotları ve iki-yönlü gruplama metoduna göre kümelenmesi hedeflenmiştir.

Tablo 1’de, 2000 yılı itibariyle Türkiye’deki coğrafi bölgelerin Kişi Başına Düşen GSYİH değerleri (1987 yılı fiyatlarıyla) verilmiştir.

Tablo 1: Türkiye’deki coğrafi bölgelere göre KBD GSYİH değerleri (1987 yılı fiyatlarıyla 2000 yılı değerleri)

Bölgeler	Tarım	Sanayi	İnşaat	Ticaret	Ulaş. Haber	Mali Kur.	Konut Sahip.	S. Meslek	Diğer Faal.
<b>Akdeniz</b>	323575	328804	69614	416946	214004	20526	75702	32103	118971
<b>Doğu A.</b>	170455	102515	35203	110237	76136	7989	40248	5897	85070
<b>Ege</b>	357832	616150	90675	522110	258568	39172	107510	46125	188569
<b>Güneydoğu</b>	260584	177079	40906	171682	113580	9642	59841	9368	75515
<b>İç Anadolu</b>	210606	306828	143349	379540	264012	23618	85188	37277	193803
<b>Karadeniz</b>	277620	288756	60297	254121	203490	17389	60117	17743	99099
<b>Marmara</b>	137640	1011774	110297	572923	315612	109156	108846	75741	156193
<b>Türkiye</b>	<b>235411</b>	<b>497580</b>	<b>88361</b>	<b>392419</b>	<b>230887</b>	<b>43626</b>	<b>83313</b>	<b>39638</b>	<b>140714</b>

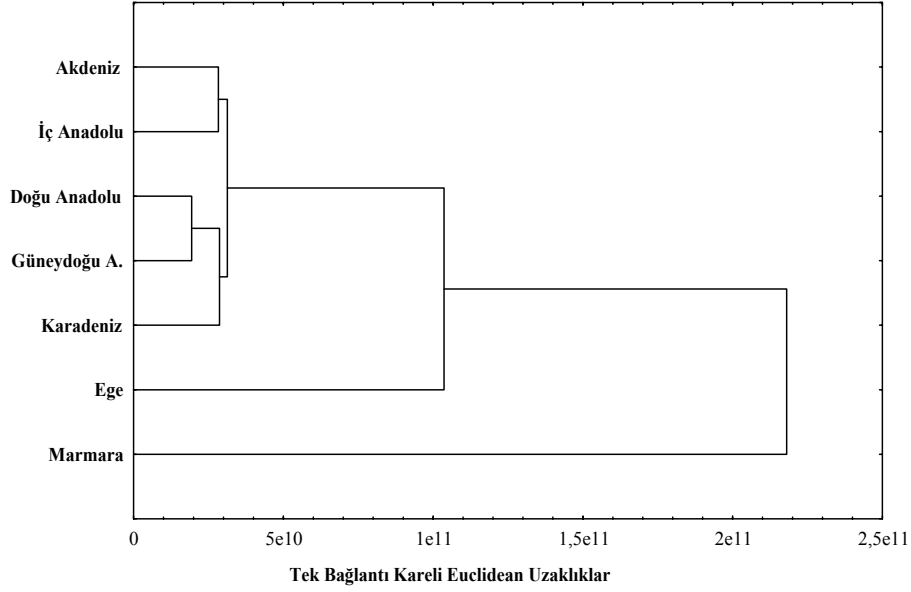
**İllere Göre Gayri Safi Yurtiçi Hasıla, 2000**, T.C. Başbakanlık Devlet İstatistik Enstitüsü, Ankara, Mayıs 2002, s.179-186.

Formül 1 yardımıyla KBD GSYİH değerleri kullanılarak bölgeler arasındaki Kareli Euclidean Uzaklıklar elde edilmiş ve Tablo 2’de verilmiştir.

Tablo 2: Türkiye’deki coğrafi bölgelere göre KBD GSYİH değerlerinden hesaplanan benzerlik matrisi (Kareli Euclidean Uzaklıklar)

Bölgeler	Akdeniz	Doğu A.	Ege	Güney A.	İç A.	Karadeniz
<b>Doğu A.</b>	192E9	...				
<b>Ege</b>	104E9	523E9	...			
<b>Güneydoğu</b>	101E9	19E9	366E9	...		
<b>İç Anadolu</b>	28E9	178E9	141E9	111E9	...	
<b>Karadeniz</b>	31E9	85E9	201E9	29E9	41E9	...
<b>Marmara</b>	550E9	1130E9	218E9	942E9	554E9	676E9

Şekil 1’deki dendogram, Kareli Euclidean Uzaklıkları kullanarak tek bağlantı metoduna göre coğrafi bölgelerin sektörel kişi başına düşen GSYİH değerleri itibariyle kümelenişini göstermektedir.

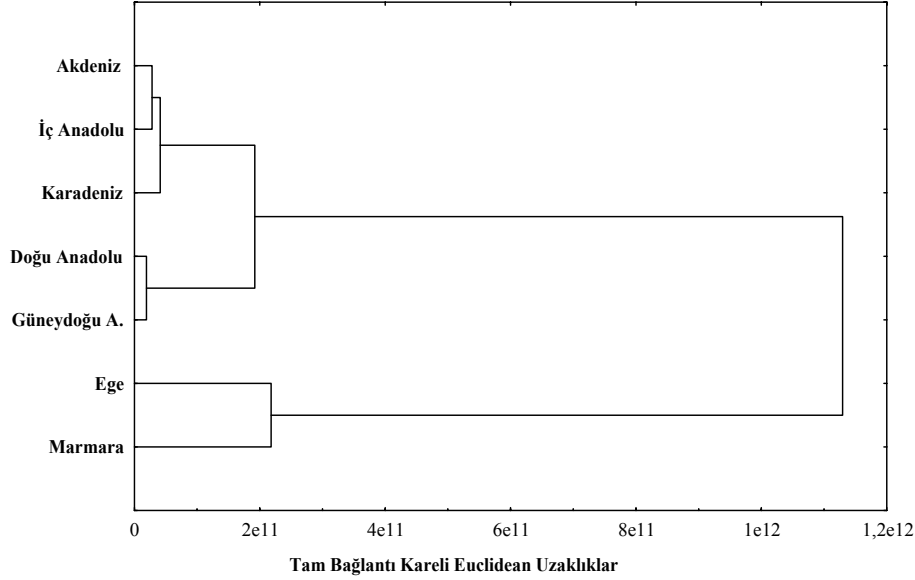


Şekil 1: Tek bağlantı Kareli Euclidean Uzaklıklar kullanılarak Türkiye’deki coğrafi bölgelerin 2000 yılı itibariyle kişi başına düşen GSYİH değerlerine göre (1987 yılı fiyatlarıyla) kümelenmesi

Tek bağlantı metoduna göre, ilk önce birleştirilen bölgeler Doğu Anadolu Bölgesi ve Güneydoğu Anadolu Bölgesidir (A Kümesi). A Kümesinde, KBD GSYİH değerlerine göre birbirine en çok benzeyen en zayıf bölgeler vardır. İkinci mertebede A Kümesi Karadeniz Bölgesi katılarak B kümesi elde edilmiştir. Üçüncü mertebede Akdeniz ve İç Anadolu Bölgeleri birleştirilerek C kümesi elde edilmiştir. Dördüncü mertebede B ve C kümesi birleştirilerek D kümesi elde edilmiştir. Beşinci mertebede D kümesi ile Ege Bölgesi birleştirilerek E Kümesi elde edilmiştir. Altıncı mertebede E kümesi ile Marmara Bölgesi birleştirilmiştir. 2000 yılı itibariyle kişi başına düşen GSYİH değeri bakımından zayıf bölgeler Güneydoğu Anadolu Bölgesi, Doğu Anadolu Bölgesi ve Karadeniz Bölgesidir. Akdeniz Bölgesi ve İç Anadolu Bölgesi Türkiye ortalamasına yakın bölgelerdir. Ege bölgesi Türkiye ortalamasının üzerindeki bölgedir. Nihayet, Marmara bölgesi kişi başına düşen GSYİH değeri bakımından en iyi bölgedir.

Şekil 2’deki dendogram tam bağlantı metoduna göre coğrafi bölgelerin sektörel kişi başına düşen GSYİH değerleri itibariyle kümelenişini göstermektedir.

Tam bağlantı metoduna göre, ilk önce birleştirilen bölgeler Doğu Anadolu Bölgesi ve Güneydoğu Anadolu Bölgesidir (A Kümesi). İkinci mertebede Akdeniz Bölgesi ve İç Anadolu Bölgesi birleştirilmiştir (B Kümesi). Üçüncü mertebede B kümesi ile Karadeniz Bölgesi birleştirilmiştir (C Kümesi). Dördüncü mertebede A Kümesi ile C Kümesi birleştirilmiştir (D Kümesi). Beşinci mertebede Ege Bölgesi ile Marmara Bölgesi birleştirilmiştir (E Kümesi). Altıncı mertebede D kümesi ile E kümesi birleştirilmiştir (F Kümesi). 2000 yılı itibariyle kişi başına düşen GSYİH değeri bakımından zayıf bölgeler Güneydoğu Anadolu Bölgesi ve Doğu Anadolu Bölgesidir. Akdeniz, İç Anadolu ve Karadeniz Bölgeleri Türkiye ortalamasına yakın bölgelerdir. Ege ve Marmara Bölgeleri ise Türkiye ortalamasının üzerindeki bölgelerdir.

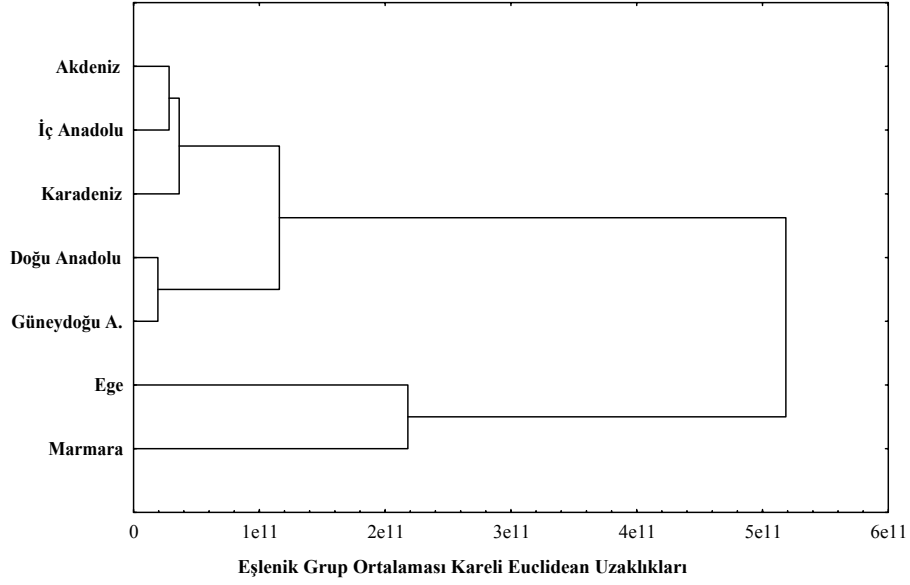


Şekil 2: Tam bağlantı Kareli Euclidean Uzaklıklar kullanılarak Türkiye'deki coğrafi bölgelerin 2000 yılı itibariyle kişi başına düşen GSYİH değerlerine göre (1987 yılı fiyatlarıyla) kümelenmesi

Şekil 3'deki dendrogram eşlenik grup ortalaması metoduna göre coğrafi bölgelerin sektörel kişi başına düşen GSYİH değerleri itibariyle kümelenişini göstermektedir.

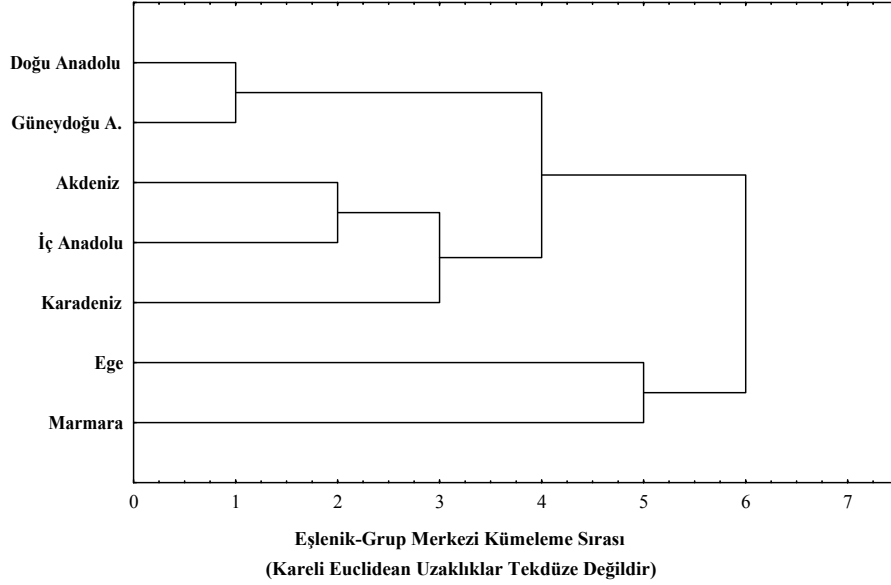
Eşlenik-grup ortalaması metoduna göre, ilk önce birleştirilen bölgeler Doğu Anadolu Bölgesi ve Güneydoğu Anadolu Bölgesidir (A Kümesi). İkinci mertebede Akdeniz Bölgesi ve İç Anadolu Bölgesi birleştirilmiştir (B Kümesi). Üçüncü mertebede B kümesi ile Karadeniz Bölgesi birleştirilmiştir (C Kümesi). Dördüncü mertebede A Kümesi ile C Kümesi birleştirilmiştir (D Kümesi). Be-

şinci mertebede Ege Bölgesi ile Marmara Bölgesi birleştirilmiştir (E Kümesi). Altıncı mertebede D kümesi ile E kümesi birleştirilmiştir (F Kümesi). 2000 yılı itibariyle kişi başına düşen GSYİH değeri bakımından zayıf bölgeler Güneydoğu Anadolu Bölgesi ve Doğu Anadolu Bölgesidir. Akdeniz, İç Anadolu ve Karadeniz Bölgeleri Türkiye ortalamasına yakın bölgelerdir. Ege ve Marmara Bölgeleri ise Türkiye ortalamasının üzerindeki bölgelerdir.



Şekil 3: Eşlenik-grup ortalaması Kareli Euclidean Uzaklıkları kullanılarak Türkiye'deki coğrafi bölgelerin 2000 yılı itibariyle kişi başına düşen GSYİH değerlerine göre (1987 yılı fiyatlarıyla) kümelenmesi

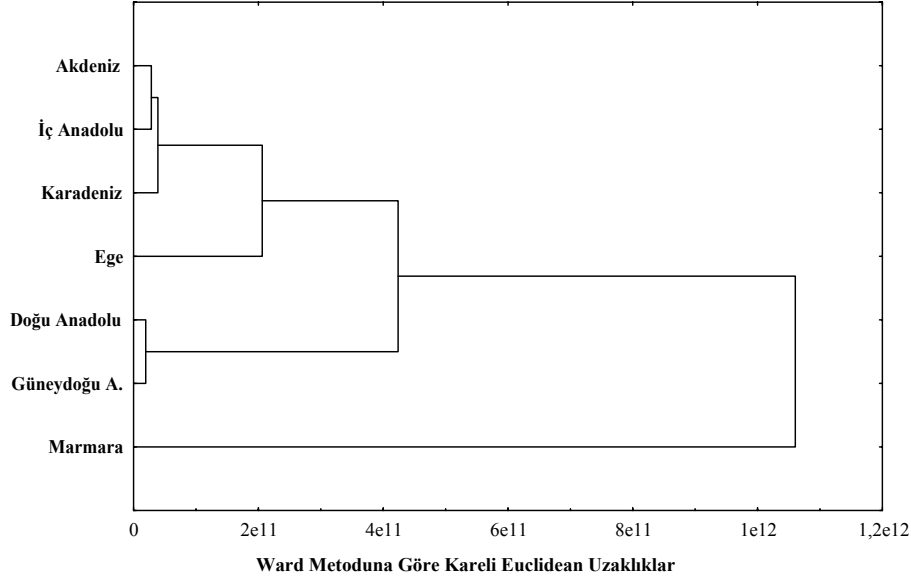
Şekil 4'teki dendogram eşlenik grup merkezi metoduna göre coğrafi bölgelerin sektörel kişi başına düşen GSYİH değerleri itibariyle kümelenişini göstermektedir.



Şekil 4: Eşlenik-grup merkezi metoduna göre Kareli Euclidean Uzaklıklar kullanılarak Türkiye'deki coğrafi bölgelerin 2000 yılı itibariyle kişi başına düşen GSYİH değerlerine göre (1987 yılı fiyatlarıyla) kümelenmesi

Eşlenik-grup merkezi metoduna göre, ilk önce birleştirilen bölgeler Doğu Anadolu Bölgesi ve Güneydoğu Anadolu Bölgesidir (A Kümesi). İkinci mertebede Akdeniz Bölgesi ve İç Anadolu Bölgesi birleştirilmiştir (B Kümesi). Üçüncü mertebede B kümesi ile Karadeniz Bölgesi birleştirilmiştir (C Kümesi). Dördüncü mertebede A Kümesi ile C Kümesi birleştirilmiştir (D Kümesi). Beşinci mertebede Ege Bölgesi ile Marmara Bölgesi birleştirilmiştir (E Kümesi). Altıncı mertebede D kümesi ile E kümesi birleştirilmiştir (F Kümesi). 2000 yılı itibariyle kişi başına düşen GSYİH değeri bakımından zayıf bölgeler Güneydoğu Anadolu Bölgesi ve Doğu Anadolu Bölgesidir. Akdeniz, İç Anadolu ve Karadeniz Bölgeleri Türkiye ortalamasına yakın bölgelerdir. Ege ve Marmara Bölgeleri ise Türkiye ortalamasının üzerindeki bölgelerdir.

Şekil 5'teki dendrogram Ward metoduna göre coğrafi bölgelerin sektörel kişi başına düşen GSYİH değerleri itibariyle kümelenişini göstermektedir.



Şekil 5: Ward metoduna göre Kareli Euclidean Bağlantı Uzaklıkları kullanılarak Türkiye'deki coğrafi bölgelerin 2000 yılı itibariyle kişi başına düşen GSYİH değerlerine göre (1987 yılı fiyatlarıyla) kümelmesi

Ward metoduna göre, ilk önce birleştirilen bölgeler Doğu Anadolu Bölgesi ve Güneydoğu Anadolu Bölgesidir (A Kümesi). İkinci mertebede Akdeniz Bölgesi ve İç Anadolu Bölgesi birleştirilmiştir (B Kümesi). Üçüncü mertebede B kümesi ile Karadeniz Bölgesi birleştirilmiştir (C Kümesi). Dördüncü mertebede C Kümesi ile Ege Bölgesi birleştirilmiştir (D Kümesi). Beşinci mertebede A ve D kümesi birleştirilmiştir (E Kümesi). Altıncı mertebede E kümesi ile Marmara Bölgesi birleştirilmiştir (F Kümesi). 2000 yılı itibariyle kişi başına düşen GSYİH değeri bakımından zayıf bölgeler Güneydoğu Anadolu Bölgesi ve Doğu Anadolu Bölgesidir. Akdeniz, İç Anadolu ve Karadeniz Bölgeleri Türkiye ortalamasına yakın bölgelerdir. Ege ve Marmara Bölgeleri ise Türkiye ortalamasının üzerindeki bölgelerdir.

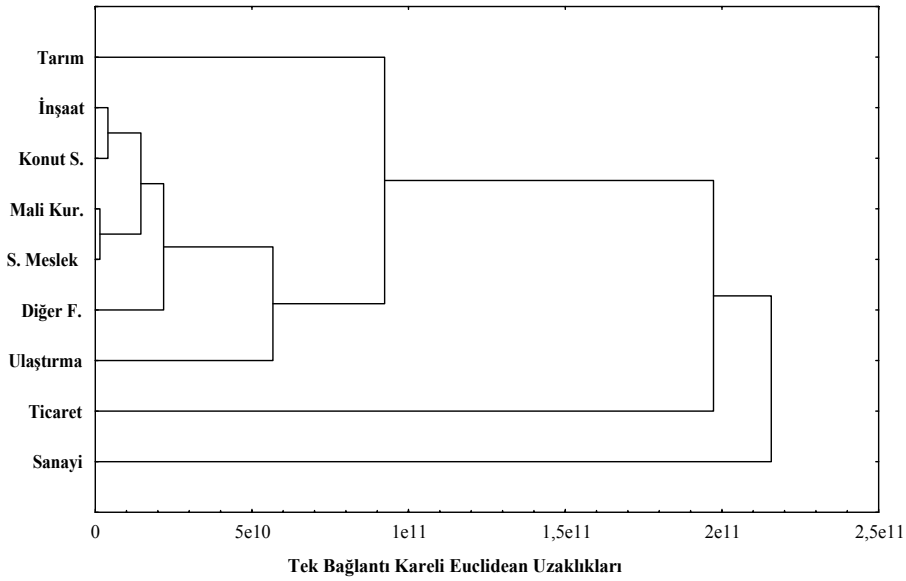
#### ***D. Sektörlere İtibariyle KBD GSYİH Değerlerinin Coğrafi Bölgelere Göre Hiyerarşik Kümeleme Metotlarıyla Gruplaması***

Formül 1 yardımıyla Türkiye'deki coğrafi bölgeler itibariyle KBD GSYİH değerleri kullanılarak GSYİH'yı oluşturan sektörlere göre Kareli Euclidean Uzaklıklar elde edilmiş ve Tablo 3'te takdim edilmiştir.

Tablo 3: Coğrafi bölgeler itibariyle sektörlere göre KBD GSYİH değerlerinden hesaplanan benzerlik matrisi (Kareli Euclidean Uzaklıklar)

Bölgeler	Tarım	Sanayi	İnşaat	Ticaret	Ulaştır.	Mali K.	Kon. S.	S.Mesl.
Sanayi	852E9	...						
İnşaat	255E9	1258E9	...					
Ticaret	266E9	216E9	637E9	...				
Ulaştırma	92E9	640E9	133E9	197E9	...			
Mali Kur.	386E9	1433E9	23E9	825E9	236E9	...		
Konut S.	245E9	1257E9	4E9	645E9	141E9	17E9	...	
S. Meslek	374E9	1473E9	20E9	832E9	237E9	1E9	15E9	...
Diğer F.	145E9	1018E9	22E9	442E9	57E9	80E9	26E	76E9

Şekil 6'daki dendogram tek bağlantı metoduna göre Türkiye'deki coğrafi bölgeler itibariyle GSYİH oluşturan sektörlerinin kümelenişini göstermektedir.

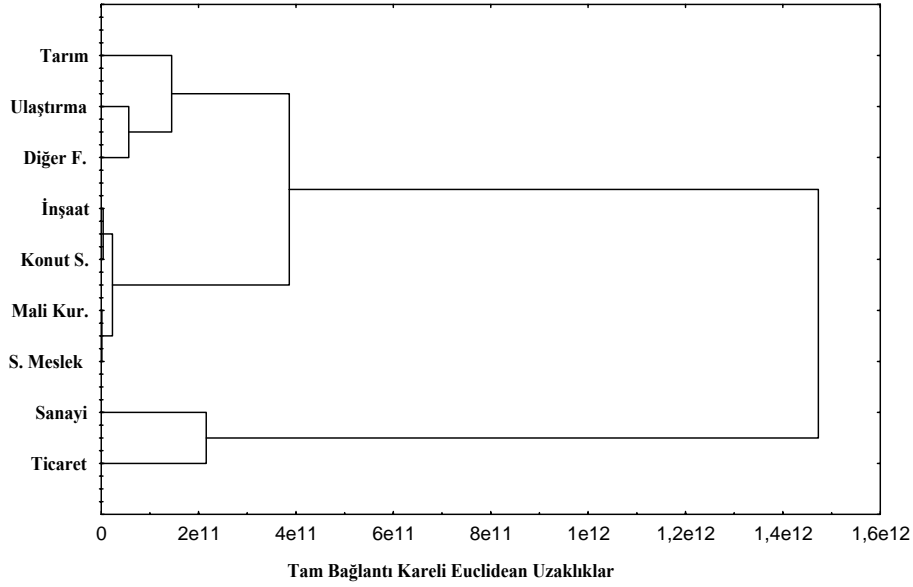


Şekil 6: Tek bağlantı Kareli Euclidean Uzaklıklar kullanılarak Türkiye'deki coğrafi bölgeler itibariyle KBD GSYİH sektörlerinin kümelenişi (1987 yılı fiyatlarıyla 2000 yılı değerleri)

Tek bağlantı metoduna göre ilk önce birleştirilen sektörler Mali Kuruluşlar ve Serbest Meslek Sektörleridir (A Kümesi). İkinci mertebede birleştirilen sektörler İnşaat ve Konut Sahipliği Sektörleridir (B Kümesi). Üçüncü mertebede A ve B Kümeleri birleştirilmiştir (C Kümesi). C Kümesindeki sektörler, KBD GSYİH değerlerine göre zayıf sektörlerdir. Dördüncü mertebede C Kümesi ile

Diğer Faaliyetler birleştirilmiştir (D Kümesi). Beşinci mertebede D Kümesi ile Ulaştırma Sektörü birleştirilmiştir (E Kümesi). Altıncı mertebede E kümesi ile Tarım Sektörü birleştirilmiştir (F kümesi). C kümesi hariç, F kümesindeki diğer sektörler KBD GSYİH değerlerine göre Türkiye ortalaması civarındaki sektörlerdir. Yedinci mertebede F kümesi ile Ticaret Sektörü birleştirilmiştir (G Kümesi). Nihayet sekizinci mertebede G kümesi ile Sanayi Sektörü birleştirilmiştir (H Kümesi). C ve F Kümeleri hariç, G Kümesindeki Sektörler KBD GSYİH değerlerine göre gelişmiş sektörlerdir.

Şekil 7’deki dendogram tam bağlantı metoduna göre Türkiye’deki coğrafi bölgeler itibariyle GSYİH oluşturan sektörlerin kümelenişini göstermektedir.



Şekil 7: Tam bağlantı Euclidean farklar kullanılarak Türkiye’deki coğrafi bölgeler itibariyle KBD GSYİH sektörlerinin kümelenişi (1987 yılı fiyatlarıyla 2000 yılı değerleri)

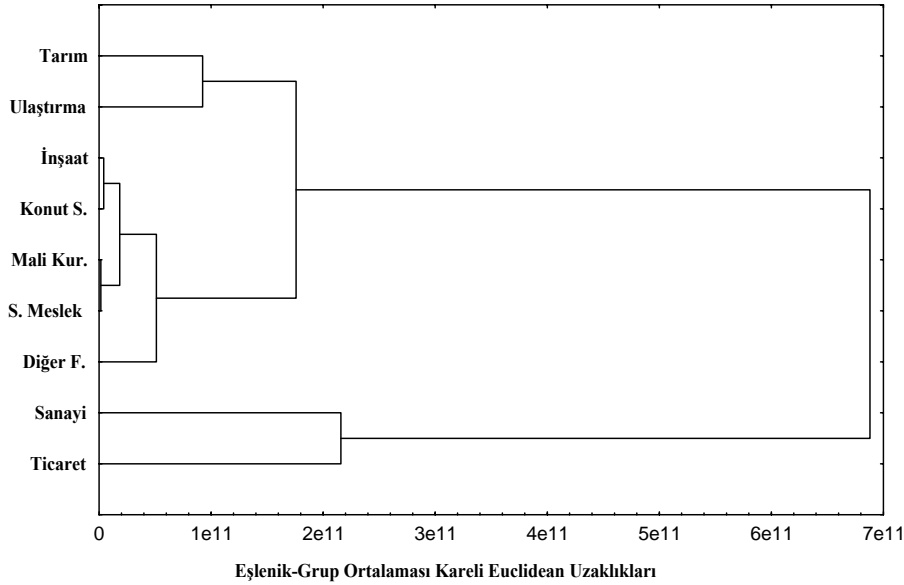
Tam bağlantı metoduna göre ilk önce birleştirilen sektörler Mali Kuruluşlar ve Serbest Meslek Sektörleridir (A Kümesi). İkinci mertebede birleştirilen sektörler İnşaat ve Konut Sahipliği Sektörleridir (B Kümesi). Üçüncü mertebede A ve B Kümeleri birleştirilmiştir (C Kümesi). C Kümesindeki sektörler, KBD GSYİH değerlerine göre zayıf sektörlerdir. Dördüncü mertebede Ulaştırma ve Diğer Faaliyetler Sektörü birleştirilmiştir (D Kümesi). Beşinci mertebede D Kümesi ile Tarım Sektörü birleştirilmiştir (E Kümesi). C Kümesi hariç, E Kümesindeki diğer sektörler KBD GSYİH değerlerine göre Türkiye ortalaması civarındaki sektörlerdir. Altıncı mertebede C kümesi ile E Kümesi birleştirilmiştir



(F Kümesi). Yedinci mertebede Sanayi ve Ticaret Sektörleri birleştirilmiştir (G Kümesi). Bu kümedeki sektörler KBD GSYİH değerlerine göre en gelişmiş sektörlerdir. Nihayet sekizinci mertebede F Kümesi ile G Kümesi birleştirilmiştir (H Kümesi).

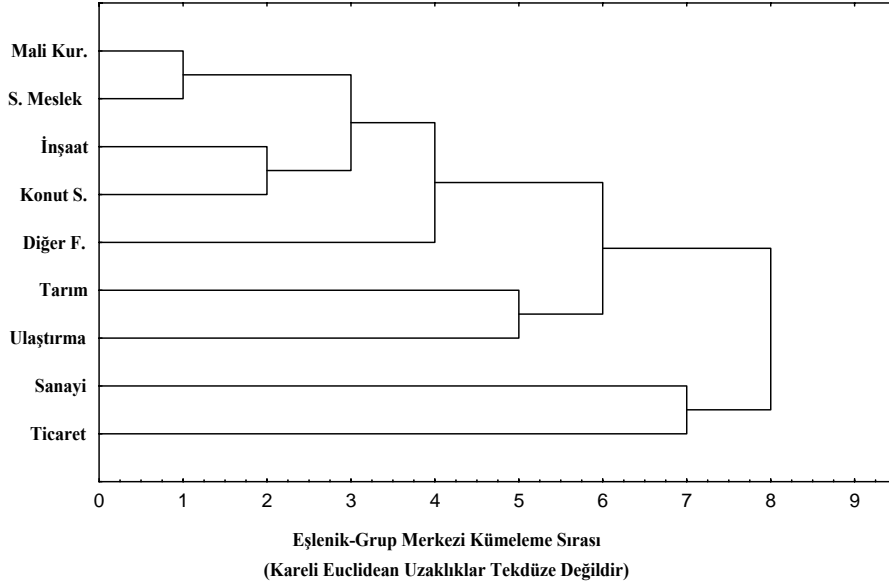
Şekil 8'deki dendogram eşlenik grup ortalaması metoduna göre Türkiye'deki coğrafi bölgeler itibariyle GSYİH oluşturan sektörlerin kümelenişini göstermektedir.

Eşlenik-grup ortalaması metoduna göre ilk önce birleştirilen sektörler Mali Kuruluşlar ve Serbest Meslek Sektörleridir (A Kümesi). İkinci mertebede birleştirilen sektörler İnşaat ve Konut Sahipliği Sektörleridir (B Kümesi). Üçüncü mertebede A ve B kümeleri birleştirilmiştir (C Kümesi). Dördüncü mertebede C Kümesi ile Diğer Faaliyetler birleştirilmiştir (D Kümesi). D kümesindeki sektörler KBD GSYİH değerlerine göre en zayıf sektörlerdir. Beşinci mertebede Tarım ile Ulaştırma Sektörü birleştirilmiştir (E Kümesi). E Kümesindeki sektörler KBD GSYİH değerlerine göre ortalama etrafındaki sektörlerdir. Altıncı mertebede D kümesi ile E Kümesi birleştirilmiştir (F kümesi). Yedinci mertebede Ticaret ve Sanayi Sektörü birleştirilmiştir (G Kümesi). G kümesindeki sektörler KBD GSYİH değerlerine göre Türkiye ortalamasının üzerindeki sektörlerdir. Nihayet sekizinci mertebede F Kümesi ile G Kümesi birleştirilmiştir (H Kümesi).



Şekil 8: Eşlenik-grup ortalaması Euclidean farkları kullanılarak Türkiye'deki coğrafi bölgeler itibariyle KBD GSYİH'yı oluşturan sektörlerin kümelenişi (1987 yılı fiyatlarıyla 2000 yılı değerleri)

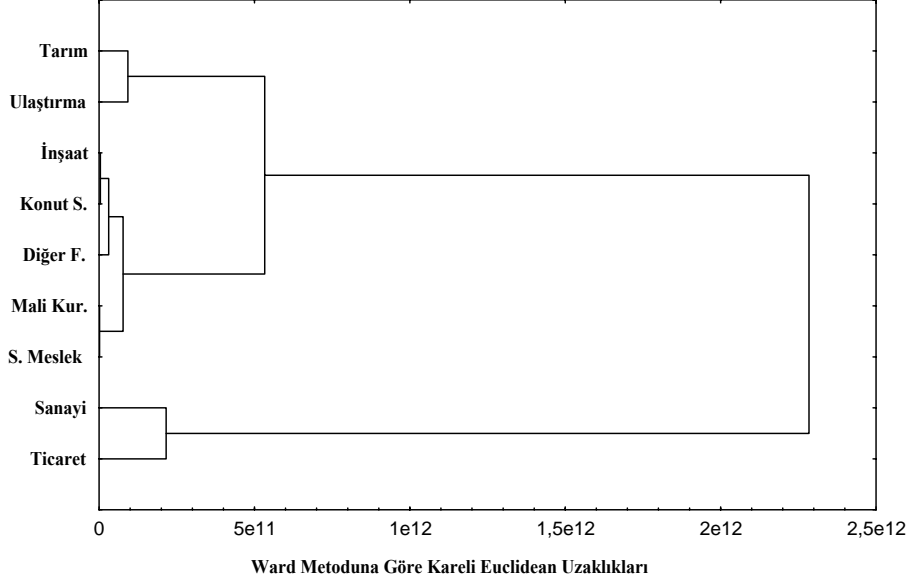
Şekil 9'daki dendogram eşlenik grup merkezi metoduna göre Türkiye'deki coğrafi bölgeler itibariyle GSYİH oluşturan sektörlerinin kümelenişini göstermektedir.



Şekil 9: Eşlenik-grup merkezi metoduna göre Kareli Euclidean Uzaklıklar kullanılarak Türkiye'deki coğrafi bölgeler itibariyle KBD GSYİH'yı oluşturan sektörlerin kümelenişi (1987 yılı fiyatlarıyla 2000 yılı değerleri)

Eşlenik-grup merkezi metoduna göre ilk önce birleştirilen sektörler Mali Kuruluşlar ve Serbest Meslek Sektörleridir (A Kümesi). İkinci mertebede birleştirilen sektörler İnşaat ve Konut Sahipliği Sektörleridir (B Kümesi). Üçüncü mertebede A ve B kümeleri birleştirilmiştir (C Kümesi). Dördüncü mertebede C Kümesi ile Diğer Faaliyetler birleştirilmiştir (D Kümesi). D kümesindeki sektörler, KBD GSYİH değerlerine göre en zayıf sektörlerdir. Beşinci mertebede Tarım ve Ulaştırma Sektörleri birleştirilmiştir (E Kümesi). E kümesindeki sektörler, KBD GSYİH değerlerine göre Türkiye ortalaması civarındaki sektörlerdir. Altıncı mertebede D kümesi ile E Kümesi birleştirilmiştir (F kümesi). Yedinci mertebede Ticaret ve Sanayi Sektörleri birleştirilmiştir (G Kümesi). G Kümesindeki sektörler, KBD GSYİH değerlerine göre Türkiye ortalamasının üzerindeki sektörlerdir. Nihayet sekizinci mertebede F kümesi ile G Kümesi birleştirilmiştir (H Kümesi).

Şekil 10'daki dendogram Ward metoduna göre Türkiye'deki coğrafi bölgeler itibariyle GSYİH oluşturan sektörlerinin kümelenişini göstermektedir.



Şekil 10: Ward metoduna göre Kareli Euclidean Bağlantı Uzaklıkları kullanılarak Türkiye'deki coğrafi bölgeler itibariyle KBD GSYİH'yı oluşturan sektörlerin kümelenişi (1987 yılı fiyatlarıyla 2000 yılı değerleri)

Ward metoduna göre ilk önce birleştirilen sektörler Mali Kuruluşlar ve Serbest Meslek Sektörleridir (A Kümesi). İkinci mertebeye birleştirilen sektörler İnşaat ve Konut Sahipliği Sektörleridir (B Kümesi). Üçüncü mertebeye B Kümesi ile Diğer Faaliyetler sektörü birleştirilmiştir (C Kümesi). Dördüncü mertebeye A Kümesi ile C Kümesi birleştirilmiştir (D Kümesi). D kümesindeki sektörler, KBD GSYİH değerlerine göre en zayıf sektörlerdir. Beşinci mertebeye Tarım ve Ulaştırma Sektörü birleştirilmiştir (E Kümesi). Altıncı mertebeye D Kümesi ile E kümesi birleştirilmiştir (F kümesi). Yedinci mertebeye Ticaret ve Sanayi Sektörleri birleştirilmiştir (G Kümesi). G Kümesindeki sektörler, KBD GSYİH değerlerine göre Türkiye ortalamasının üzerindeki sektörlerdir. Nihayet sekizinci mertebeye F kümesi ile G Kümesi birleştirilmiştir (H Kümesi).

Bulunan sonuçlar arasında önemli farklılıklar vardır. Mesela bir metoda göre zayıf sektörler arasında görülen bir sektör başka bir kümeleme metoduna göre Türkiye ortalaması civarında görülebilmektedir. Hiçbir metot tek başına tutarlı bir sonuç vermemiştir. Bu sebeple beş metottan üç veya daha fazlasının uzlaştıkları sonuçların genelleştirilmesi gereklidir.

### E. Sektörlerin ve Coğrafi Bölgelerin KBD GSYİH Değerlerine Göre İki-Yönlü Kümeleme Metoduyla Gruplanması

Şekil 11'deki iki-yönlü grafik, KBD GSYİH değerlerine göre Türkiye'deki coğrafi bölgelerin ve sektörlerin kümelenişini göstermektedir.

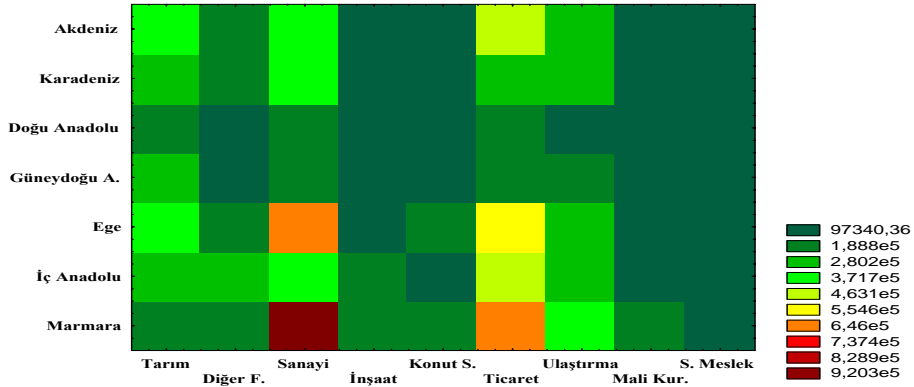
Akdeniz Bölgesinde zayıf sektörler İnşaat, Konut Sahipliği, Mali Kuruluşlar ve Serbest Meslek Sahipliğidir. Bunu Sırasıyla Diğer Faaliyetler ve Ulaştırma Sektörü izlemektedir. Dördüncü kümede Tarım ve Sanayi Sektörleri vardır. Bu bölgede KBD GSYİH'da en büyük pay Ticaret Sektöründen gelir.

Karadeniz Bölgesinde de zayıf sektörler İnşaat, Konut Sahipliği, Mali Kuruluşlar ve Serbest Meslek Sahipliğidir. Bu kümeyi Diğer Faaliyetler izlemektedir. Üçüncü kümede Tarım, Ticaret ve Ulaştırma Sektörleri yer almaktadır. Bu bölgede KBD GSYİH'da en büyük pay Sanayi Sektöründen gelir.

Doğu Anadolu Bölgesinde zayıf sektörler Diğer Faaliyetler, İnşaat, Konut Sahipliği, Ulaştırma, Mali Kuruluşlar ve Serbest Meslek Sahipliğidir. Bunu Tarım, Sanayi ve Ticaret Sektörlerinin oluşturduğu küme izlemektedir.

Güneydoğu Anadolu Bölgesinde zayıf sektörler Diğer Faaliyetler, İnşaat, Konut Sahipliği, Mali Kuruluşlar ve Serbest Meslek Sahipliğidir. Bunu Sanayi, Ticaret ve Ulaştırma Sektörlerinin oluşturduğu küme izlemektedir. Bu bölgede KBD GSYİH'da en büyük pay Tarım Sektöründen gelir.

Ege Bölgesinde zayıf sektörler İnşaat, Mali Kuruluşlar ve Serbest Meslek Sahipliğidir. Bu kümeyi Diğer Faaliyetler ve Konut Sahipliği Sektörlerinin oluşturduğu küme izlemektedir. Bu kümeyi küçükten büyüğe doğru sırasıyla Ulaştırma, Tarım, Ticaret ve Sanayi Sektörleri izlemektedir.



Şekil 11: Sektörlerin ve Türkiye'deki Coğrafi Bölgelerin KBD GSYİH Değerlerine Göre İki-Yönlü Kümeleme Metoduyla Gruplanması (1987 yılı fiyatlarıyla 2000 yılı değerleri)

İç Anadolu Bölgesinde zayıf sektörler Konut Sahipliği, Mali Kuruluşlar ve Serbest Meslek Sahipliğidir. Bu kümeyi İnşaat Sektörü izlemektedir. Üçüncü kümede Tarım, Diğer Faaliyetler ve Ulaştırma Sektörleri vardır. Bu kümenin üzerinde Sanayi Sektörü vardır. Nihayet bu bölgede KBD GSYİH’da en büyük pay Ticaret Sektöründen gelir.

Marmara Bölgesinde en zayıf sektör Serbest Meslek Sahipliğidir. Bunu Tarım, Diğer Faaliyet, İnşaat, Konut Sahipliği, Mali Kuruluşlar Sektörlerinin oluşturduğu küme izlemektedir. Bu kümenin üzerinde küçükten büyüğe doğru sırasıyla Ulaştırma, Ticaret ve Sanayi Sektörleri vardır.

Türkiye’de Tarım Sektörüne en düşük katkı Marmara ve Doğu Anadolu Bölgelerinden gelmektedir. Bir üst kümede Karadeniz, Güneydoğu Anadolu ve İç Anadolu Bölgeleri vardır. Tarım Sektörüne en yüksek katkı ise Akdeniz ve Ege Bölgelerinden gelir.

Diğer Faaliyetlere en düşük katkı Doğu ve Güneydoğu Anadolu Bölgelerinden gelmektedir. Bir üst kümede Akdeniz, Karadeniz, Ege ve Marmara Bölgeleri vardır. Diğer Faaliyetlere en yüksek katkı İç Anadolu Bölgesinden gelir.

Sanayi Sektörüne en düşük katkı Doğu ve Güneydoğu Anadolu Bölgelerinden gelmektedir. Bir üst kümede Akdeniz, Karadeniz ve İç Anadolu Bölgeleri vardır. Sanayi Sektörüne en yüksek katkı ise küçükten büyüğe doğru sırasıyla Ege ve Marmara Bölgelerinden gelir.

İnşaat Sektörüne en düşük katkı Akdeniz, Karadeniz, Doğu Anadolu, Güneydoğu Anadolu ve Ege Bölgelerinden gelmektedir. Bir üst kümede ise İç Anadolu ve Marmara Bölgesi vardır.

Konut Sahipliği Sektörüne en düşük katkı Akdeniz, Karadeniz, Doğu Anadolu, Güneydoğu Anadolu ve İç Anadolu Bölgelerinden gelmektedir. Bir üst kümede ise Ege ve Marmara Bölgesi vardır.

Ticaret Sektörüne en düşük katkı Doğu ve Güneydoğu Anadolu Bölgelerinden gelmektedir. Bu kümenin üzerinde ise Karadeniz Bölgesi vardır. Bir üst kümede Akdeniz ve İç Anadolu Bölgeleri vardır. Ticaret Sektörüne en yüksek katkı ise küçükten büyüğe doğru sırasıyla Ege ve Marmara Bölgelerinden gelir.

Ulaştırma Sektörüne en düşük katkı Doğu Anadolu Bölgesinden gelmektedir. Güneydoğu Anadolu Bölgesi bir üst basamaktadır. Bu bölgenin üstündeki basamakta ise Akdeniz, Karadeniz, Ege ve İç Anadolu Bölgeleri vardır. Ulaştırma Sektörüne en yüksek katkı ise Marmara Bölgesinden gelir.

Mali Kuruluşlar Sektörüne en düşük katkı Akdeniz, Karadeniz, Doğu Anadolu, Güneydoğu Anadolu, Ege ve İç Anadolu Bölgelerinden gelmektedir. Bir üst kümede ise Marmara Bölgesi vardır.

Serbest Meslek Sektörüne bütün bölgelerin katkısı benzer olmakla beraber çok düşük seviyededir.

## II. Sonuç

KBD GSYİH değerlerine göre bölgeler hiyerarşik olarak kümelenirken, tek bağlantı metodu hariç diğer kümeleme yöntemleri benzer sonuçlar vermiştir. Beş kümeleme metodu ile elde edilen sonuçlar genelleştirildiğinde KBD GSYİH değerlerine göre zayıf bölgeler Doğu Anadolu Bölgesi ve Güneydoğu Anadolu Bölgesidir. Türkiye ortalaması civarında olan bölgeler Karadeniz Bölgesi, Akdeniz Bölgesi ve İç Anadolu Bölgesidir. Türkiye ortalamasının üzerinde olan bölgeler ise, Ege Bölgesi ve Marmara Bölgesidir.

KBD GSYİH değerlerine göre sektörler hiyerarşik olarak kümelenirken, tam bağlantı metodu hariç diğer kümeleme yöntemleri benzer sonuçlar vermiştir. Beş kümeleme metodu ile elde edilen sonuçlar genelleştirildiğinde, KBD GSYİH değerlerine göre zayıf sektörler, Mali Kuruluşlar, Serbest Meslek Sahipliği, İnşaat, Konut Sahipliği ve Diğer Faaliyet Sektörleridir. Türkiye ortalaması civarında olan sektörler Tarım ve Ulaştırma Sektörleridir. Nihayet Türkiye ortalamasının üzerinde olan sektörler Ticaret ve Sanayi Sektörleridir.

İki yönlü kümeleme tekniğinde ise yukarıdakilere ilave olarak Hangi bölgeden hangi sektörün KBD GSYİH'da ne ölçüde pay sahibi olduğu da görülmektedir. KBD GSYİH'nın oluşmasında en yüksek pay Marmara Bölgesinin Sanayi Sektöründen gelir (Birincil Bölgesel Sektör). Daha sonra Ege bölgesinin Sanayi Sektörü ve Marmara Bölgesinin Ticaret Sektörü gelir (İkincil Bölgesel Sektörler). Bu kümeden sonra Ege Bölgesinin Ticaret Sektörü gelir (Üçüncül Bölgesel Sektör). Bunu izleyen küme Akdeniz ve İç Anadolu Bölgelerinin Ticaret Sektörleridir (Dördüncül Bölgesel Sektörler). Sonraki küme Akdeniz Bölgesinin Tarım ve Sanayi Sektörü, Karadeniz Bölgesinin Sanayi Sektörü, Ege Bölgesinin Tarım Sektörü, İç Anadolu Bölgesinin Sanayi Sektörü ve Marmara Bölgesinin Ulaştırma Sektörünün oluşturduğu kümedir (Beşincil Bölgesel Sektörler). Altıncı kümedeki alt sektörler, Akdeniz Bölgesinin Ulaştırma Sektörü, Karadeniz Bölgesinin Tarım, Ticaret ve Ulaştırma Sektörleri, Güneydoğu Anadolu Bölgesinin Tarım Sektörü, Ege Bölgesinin Ulaştırma Sektörü, İç Anadolu Bölgesinin Tarım, Diğer Faaliyetler ve Ulaştırma Sektörleridir (Altıncıl Bölgesel Sektörler). Yedinci kümede Akdeniz Bölgesinin Diğer Faaliyetler Sektörü, Karadeniz Bölgesinin Diğer Faaliyetler Sektörü, Doğu Anadolu Bölgesinin Tarım, Sanayi ve Ticaret Sektörleri, Güneydoğu Anadolu Bölgesinin Sanayi, Ticaret ve Ulaştırma Sektörleri, Ege Bölgesinin Diğer Faaliyetler ve Konut Sahipliği Sektörleri, İç Anadolu Bölgesinin İnşaat Sektörü, Marmara Bölgesinin Tarım, Diğer Faaliyetler, İnşaat, Konut Sahipliği ve Mali Kuruluşlar Sektörleri yer alır (Yedincil Bölgesel Sektörler). Diğer bölgesel sektörler KBD GSYİH'da en düşük paya sahip sektörlerdir (Sekizincil Bölgesel Sektörler).

**Abstract:** In the theoretical part of this study, it has been emphasized on the measure of Squared Euclidean Distances which are base for clustering. Then the hierarchical agglomerative methods and two-way joining method have been presented. In the practical part of the study, the geographic regions in Turkey have been agglomerated according to the Per Capita Gross Domestic Product for sectors. Furthermore the sectors in GDP have been agglomerated by regional GDP Per Capita. When hierarchical agglomerative methods applied to the existing data, the clusters which are different from each other have been obtained. When the clusters obtained from five hierarchical agglomerative methods were generalized, reliable results were reached. Finally the regional sectors have been agglomerated by means of two-way joining method.

**Key Words:** Cluster analysis, the hierarchical agglomerative methods, two-way joining , squared euclidean distances, gross domestic product.

### **Kaynakça**

- Aldenderfer, Mark S., Roger K. Blashfield (1986), **Cluster Analysis**, Sage Publications, Beverly Hills.
- Anderberg, Michael R. (1973), **Cluster Analysis for Applications**, Academic Press, New York.
- Dornbusch, R. ve S. Fischer (1998), **Makroekonomi**, Birinci Basım, Çeviri, McGraw-Hill –Akademi Ortak Yayını, İstanbul.
- Fisher, Walter Dummer (1969), **Clustering and Aggregation in Economics**, Johns Hopkins Pres, Baltimore.
- Güler, Fazıl (1999), **Hiyerarşik Kümeleme Analizi ve Bir Uygulama**, İstanbul.
- Gower, J. C. (1967), **A Comparison of Some Methods of Cluster Analysis**, Biometrics, Vol: 23, No: 4, 623-637.
- Hair, Jr. J. F., R. E. Anderson, R. L. Tatham, W. C. Black (1998), **Multivariate Data Analysis**, Prentice-Hall, Upper Saddle River, New Jersey.
- Hartigan, John A. (1975), **Clustering Algorithms**, Wiley, New York.
- Lance, G. N., W. T. Williams (1967), **A General Theory of Classificatory Sorting Strategies 1. Hierarchical Systems**, Comput. Journal, Vol: 9, No: 4, 373-380.
- Lorr, Maurice (1993), **Cluster Analysis for Social Scientists**, Jossey-Bass, San Francisco.
- Miyamoto, Sadaakio (1990), **Fuzzy Sets in Information Retrieval and Cluster Analysis**, Kluwer Academic Publishers, Dordrecht-Boston.
- Jain, Anil K. (1988), **Algorithms for Clustering Data**, Prentice-Hall, Englewood Cliff.
- Johnson, Richard A., Dean W. Wichern (1999), **Applied Multivariate Statistical Analysis**, Fourth Editon, Prentice Hall, Upper Saddle River, New Jersey.

- Öztürk, İrfan (1999), **N < P Boyutlu Biyolojik Verilerde Farklı Kümeleme Yöntemlerinin Karşılaştırmalı Olarak İncelenmesi**, Harran Üniversitesi Sosyal Bilimler Enstitüsü, Yayınlanmamış Doktora Tezi, Şanlıurfa.
- T.C. Başbakanlık Devlet İstatistik Enstitüsü (2002), **İllere Göre Gayri Safi Yurtiçi Hasıla, Türkiye, 2000**, DİE Yayınları, Yayın No: 2582, Ankara.
- T.C. Başbakanlık Devlet Planlama Teşkilatı (2003), **İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması**, DPT Yayını, Yayın No:2671, Ankara.
- T.C. Başbakanlık Devlet İstatistik Enstitüsü (2003), **İller ve Bölgeler İtibariyle Gayri Safi Yurtiçi Hasıladaki Gelişmeler (1987-2000)**, DİE Yayınları, Ankara.
- Sneath, Peter H. A., Robert R. Sokal (1973), **Numerical Taxonomy: The Principles and Practice of Numerical Classification**, W. H. Freeman and Company, San Francisco.
- Oktay, Erkan (2001), **Kontenjans Tablolarından Elde Edilen İlişki Ölçüleri**, Aktif Yayınevi, Erzurum.
- , **Doğu Anadolu Projesi Kapsamındaki İllerin GSYİH Değerlerine Göre Hiyerarşik Kümeleme Analizleriyle Gruplanması**, Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, Cilt: 15, Sayı: 3-4, 329-345.
- Ünsal, Erdal M. (2004), **Makroiktisat**, Genişletilmiş 5. Bası, Turhan Kitapevi, Ankara.
- Ward, Jr. J. H. (1963), **Hierarchical Grouping to Optimise an Objective Function**, Journal of The American Statistical Association, Vol: 58, 236-244.
- Wishart, D. (1969), **An Algorithm for Hierarchical Classification**, Biometrics, Vol: 22, No: 1, 165-170.
- Zupan, Jure (1982), **Clustering of Large Data Sets**, Research Studies Pres, New York.
- <http://ekutup.dpt.gov.tr/ekonomi/makro/yukselez/odemeler.pdf>YükselerMakro Ekonomik Hesaplar ve Ödemeler Dengesi.
- <http://sunsite.univie.ac.at/textbooks/statistics/stcluan.html#two>
- <http://singlerow.com/src/asp/guidetoadvanalytics.asp?msg=twj>
- [http://64.233.187.104/search?q=cache:IRtgiVvfJ\\_EJ:www.geocities.com/artha\\_indica/cluster.pdf+%22two-way+joining%22&hl=tr](http://64.233.187.104/search?q=cache:IRtgiVvfJ_EJ:www.geocities.com/artha_indica/cluster.pdf+%22two-way+joining%22&hl=tr)