

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KÜMELEME ANALİZİ İLE ELEKTRİK
TÜKETİMİNİN SINIFLANDIRILMASI

YÜKSEK LİSANS TEZİ
Huri ALKAN

Anabilim Dalı: İstatistik
Programı: İstatistiksel Bilgi Sistemleri

Danışman: Doç. Dr. Sinan ÇALIK

OCAK-2012

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KÜMELEME ANALİZİ İLE ELEKTRİK TÜKETİMİNİN SINIFLANDIRILMASI

YÜKSEK LİSANS TEZİ

Huri ALKAN

(091133104)

Anabilim Dalı: İstatistik

Programı: İstatistiksel Bilgi Sistemleri

Danışman: Doç. Dr. Sinan ÇALIK

Tezin Enstitüye Teslim Edildiği Tarih: 26 Aralık 2011

OCAK- 2012

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KÜMELEME ANALİZİ İLE ELEKTRİK TÜKETİMİNİN SINIFLANDIRILMASI

YÜKSEK LİSANS TEZİ

Huri ALKAN

(091133104)

Tezin Enstitüye Verildiği Tarih: 26. 12. 2011

Tezin Savunulduğu Tarih:

Tez Danışmanı: Doç. Dr. Sinan ÇALIK(F.Ü)

Diğer Jüri Üyeleri: Doç. Dr. Mustafa İNŞ (F.Ü.)

: Yrd. Doç. Dr. Mehmet GÜRCAN (F.Ü.)

OCAK- 2012

ÖNSÖZ

Tez konusunun belirlenmesi ve yürütülmesi aşamasında, her türlü yardımı ve desteği esirgemeyen kıymetli danışmanım Doç. Dr. Sinan ÇALIK'a, teşekkür eder, saygılarımı sunarım.

Huri ALKAN

ELAZIĞ- 2012

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ.....	II
İÇİNDEKİLER	III
ÖZET	VI
SUMMARY	VII
ŞEKİL LİSTESİ.....	VIII
TABLO LİSTESİ.....	IX
SEMBOLLER LİSTESİ.....	X
1. GİRİŞ.....	1
1.1. Kümeleme Analizi.....	1
2. MATERYAL VE METOD	3
2.1. Kümeleme Analizinin İşlevleri.....	3
2.2. Benzerlik.....	3
2.2.1. Benzerlik Kavramı	3
2.2.2. Değişken Seçimi ve Standardize İşlemleri	5
2.2.3. Z Değerleri Standardizasyonu.....	5
2.2.4. 0-1 ve -1,1 Standardizasyonu.....	5
2.2.5. $\mu = 1$, $\sigma = 1$ Standardizasyonu	6
2.2.6. Benzerlik Ölçüleri	6
2.2.7. Uzaklık Ölçüleri.....	7
2.2.8. Korelasyon Katsayısı.....	8
2.2.9. İlişki Katsayısı.....	9
2.2.10. Olasılıksal Benzerlik Katsayısı	10
2.3. Hiyerarşik Kümeleme Yöntemi	10
2.3.1. En Yakın Komşuluk Yöntemi.....	11
2.3.2. En Uzak Komşuluk Yöntemi.....	13
2.3.3. Ortalama Bağlantı Kümeleme Yöntemi	13
2.3.4. Ward's Bağlantı Kümeleme Yöntemi	14
2.3.5. Merkezi Bağlantı Kümeleme Yöntemi.....	14
2.4. Hiyerarşik Olmayan Kümeleme Yöntemleri	14
2.4.1. K-Ortalamalar Yöntemi.....	15

2.5.	Diskriminant Analizi	15
2.5.1.	İki Grup için Diskriminant Analizi	15
2.5.2.	Çoklu Grup İçin Diskriminant Analizi	17
3.	UYGULAMA.....	19
3.1.	Uygulamanın Amacı.....	19
3.2.	Uygulama Verileri ve Düzeni	19
3.3.	Elektrik Tüketimi	20
3.4.	Örnekleme Seçimi.....	21
3.5.	Elektrik Tüketim Verileri	21
3.6.	Kümeleme Yöntemlerinin Belirlenmesi.....	22
3.7.	Gruplar Arası Ortalama Bağlantı yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi	23
3.8.	Grup içi Ortalama Bağlantı yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi	26
3.9.	En Yakın Komşu yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi.....	29
3.10.	En Yakın Komşu yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi.....	31
3.11.	Merkezi Kümeleme Yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi.....	33
3.12.	Ward's Yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi.....	35
4.	BULGULAR.....	37
4.1.	Gruplar Arası Ortalama Bağlantı yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri.....	37
4.2.	Grup içi Ortalama Bağlantı yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri	37
4.3.	En Yakın Komşu Yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri	38
4.4.	En uzak Komşu Yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri	39
4.5.	Ward's Kümeleme Yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri	40

4.6.	Merkezi Bağlantı Yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri	41
4.7.	Kümeleme Analizi Yöntemlerinin Etkinlikleri İncelenmesi	41
5.	SONUÇLAR VE TARTIŞMA	43
6.	ÖNERİLER	45
	KAYNAKLAR	46
	EKLER	47
	EK 1. Elektrik Tüketim Ortalamaları	47
	ÖZGEÇMİŞ.....	48

ÖZET

Kümeleme analizi teknikleri, verileri sınıflandırma ve gruplandırma işlemlerinde kullanışlı çok değişkenli analiz tekniğidir. Kümeleme analizinin genel amacı, sınıflandırılmamış verileri benzerliklerine göre tasnif ederek ve araştırmacıya özetleyici bilgiler elde etmede yardımcı olmaktır.

Bu çalışmada kümeleme analizi ve kümeleme tekniklerine ilişkin bilgiler sunulmuştur. Uygulamada farklı kümeleme yöntemleri ile Bingöl, Elazığ, Malatya ve Tunceli il ve ilçe merkezlerindeki hanelerin yıllık elektrik tüketim değerleri dikkate alınarak bu yerleşim merkezleri elektrik tüketimleri bakımından 5 kümeye ayrılmıştır. Elde edilen kümelerin diskriminant analizi ile doğru sınıflara atanma oranları elde edilmiştir.

Anahtar Kelimeler: Elektrik tüketimi, Diskriminant Analizi, Kümeleme Yöntemleri

SUMMARY

Classification of Electricity Consumption by Cluster Analysis

Cluster analysis techniques are multivariate analysis in data classification and grouping processes useful technique. The overall objective of clustering analysis are to assist for researcher about classified data and summarizing the classification of similar information.

In this study presents theoretical information on the cluster analysis and clustering techniques. A theoretical information were given for cluster analysis methods.

In application Bingol, Elazig, Malatya, Tunceli and county center of this city were classified by different cluster methods. Electricity consumptions of city centers were considered for classification. Five sets were obtained. Finally five sets were examined by discriminant analysis.

Key Words: Electricity Consumption, Discriminant Analysis, Cluster Methods.

ŞEKİL LİSTESİ

	<u>Sayfa No</u>
Şekil 2.1. Dendogram Grafiği	12
Şekil 2.2. En yakın komşuluk yöntemi ile en uzak komşuluk yöntemi.....	13
Şekil 3.1. Gruplar arası Ortalama Bağlantı Yöntemi ile Yerleşim Merkezlerinin Elektrik.....	25
Şekil 3.2. Grup içi Ortalama Bağlantı Yöntemi ile Yerleşim Merkezlerinin Elektrik	28
Şekil 3.3. En yakın komşu Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi.....	30
Şekil 3.4. En uzak komşu Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi Bakımından Dendogram Grafiği	32
Şekil 3.5. Merkezi Kümeleme Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi	34
Şekil 3.6. Ward's Kümeleme Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi.....	36

TABLO LİSTESİ

	<u>Sayfa No</u>
Tablo 3.1. Örneklemin Elde Edildiği Merkezler.....	21
Tablo 3.2. 2010 Yılına ait örneklemden elde edilen yıllık hane elektrik tüketimleri ortalamaları.....	22
Tablo 3.3. Gruplar arası ortalama bağlantı yöntemi ile birleştirme tablosu	24
Tablo 3.4. Grup içi Ortalama Bağlantı yöntemi ile elde edilen Birleştirme Tablosu	27
Tablo 3.5. En Yakın Komşu yöntemine ile elde edilen Birleştirme Tablosu	29
Tablo 3.6. En uzak Komşu Yöntemi ile elde edilen Birleştirme Tablosu.....	31
Tablo 3.7. Merkezi Kümeleme Yöntemi ile elde edilen Birleştirme Tablosu.....	33
Tablo 3.8. Ward's Kümeleme yöntemi ile elde edilen Birleştirme Tablosu	35
Tablo 4.1. Gruplar arası ortalama bağlantı yöntemi ile yerleşim merkezlerinin kümelenmesi.....	37
Tablo 4.2. Grup içi ortalama bağlantı yöntemi ile yerleşim merkezlerinin kümelenmesi ..	38
Tablo 4.3. En yakın Komşu Yöntemi ile yerleşim merkezlerinin kümelenmesi.....	39
Tablo 4.4. En uzak Komşu Yöntemi ile yerleşim merkezlerinin kümelenmesi	40
Tablo 4.5. Ward's Kümeleme Yöntemi ile yerleşim merkezlerinin kümelenmesi.....	40
Tablo 4.6. Merkezi Bağlantı Yöntemi ile yerleşim merkezlerinin kümelenmesi	41
Tablo 4.7. Doğru Sınıflandırma Oranları	42

SEMBOLLER LİSTESİ

Kw/h: Elektrik güç birimi

1. GİRİŞ

1.1. Kümeleme Analizi

Tanım 1.1.1. (*Kümeleme Analizi*)

Kümeleme analizi, bir sınıflandırma oluşturmak için kullanılan prosedürleri geniş bir yelpazede sunan sistemlerin genel adıdır. Bu prosedürleri ampirik kümeler veya benzer varlıklar oluşturmaktadır. Daha açık bir ifade ile kümeleme analizi varlık örneklerine ilişkin bilgi içeren veri setleri ile başlayarak bu varlık kümelerine nispeten homojen grupları oluşturmaya çalışan çok değişkenli istatistiksel yöntemlerdir [1].

Kümeleme analizinde amaç değişkenlerdeki verileri birbiri ile benzerlik gösteren alt gruplara veya kümelere tasnif etmektir. Bundan dolayı kümeleme analizi değişkenlerdeki benzerliklerden ve farklılıklardan yararlanır.

Kümeleme analizi için farklı tanımlar mevcuttur. Bu tanımlardan biride Hair'e aittir. Hair'e göre "Kümeleme analizi esas amacı birimlerde bulunan karakteristik özellikleri dikkate alarak gruplamak olan çok değişkenli istatistiksel teknikler grubudur. Kümeleme analizi birimler, küme içerisinde çok benzer biçimde, kümeler arasında farklı olacak şekilde kümeler. Kümeleme işlemi başarılı olursa bir geometrik çizim yapıldığında birimler küme içerisinde çok yakın, kümeler ise birbirinden uzak olacaktır [2].

Kümeleme analizi 20.yüzyılda fark edilmesine rağmen kümeleme analizi ile ilgili literatür son 20 yıl boyunca elde edilmiştir. Kümelene analizinin gelişimindeki en önemli ilk çalışmalardan biri Robert Sokal ve Peter Sneath adlı iki biyolog tarafından 1963 yılında yayınlanan "Principles of Numerical Taxonomy" adlı kitap olmuştur [3]. Sokal ve Sneath bu kitaplarında biyolojik sınıflandırmaların üretimi için etkili prosedürleri tartışmıştır. Böylece organizmalar arasındaki benzerlik derecesi ortaya konacak ve kümeleme analizinin kullanımı ile nispeten benzer organizmalar aynı gruplara yerleştirilecektir. Kümeleme analizine ilişkin çalışmalar Sokal ve Sneath'ın kitabından sonra artmaya başlamıştır. 1963 ile 1975 yılları arasında kümeleme analizine ilişkin yayınlanan çalışma sayısı katlanarak artmıştır. Bilgisayar teknolojisinin gelişmesi ile kümeleme analizi yöntemlerinde ve uygulanabilir alanlarda gelişmeler yaşanmıştır. Farklı kümeleme teknikleri ortaya sunulmuş ve farklı disiplinlerde kümeleme analizi yöntemleri uygulama

sahası bulmuştur. Kümeleme analizi sosyal bilimlerin birçok alanında ilgi konusu olmuştur [1].

Kümeleme analizi kendi içerisinde popüler bir alan olmasına rağmen, hala kümeleme analizinin faktör analizi, diskriminant analizi ve çok boyutlu ölçekleme gibi çok boyutlu istatistiksel yöntemlere göre kıyaslandığında çok bilinen bir yöntem olmadığı anlaşılmaktadır. Kümeleme analizi üzerine sosyal bilimler literatürü sıklıkla şaşırtıcı ve çelişkili bir terminoloji dizisi, yöntemleri ve tercih edilen yaklaşımları sunar [1].

2. MATERYAL VE METOD

2.1. Kümeleme Analizinin İşlevleri

Kümeleme analizi, birimleri benzerliklerine göre tasnif eden çok değişkenli istatistiksel yöntemler topluluğudur. Bu yöntemler belirli işlemleri bazı parametrelere göre gerçekleştirmektedir. Kümeleme tekniklerinin başlıca işlevleri;

- i. n sayıda birim ve p sayıda değişkenin olduğu varsayılın. n sayıdaki birimi p adet değişkenin belirlenen özelliklerine göre gruplar içerisinde benzer ve gruplar arasında farklı alt kümeler oluşturmak,
- ii. Uygun bir uzaklık ölçüsünün belirlenerek birimler arası uzaklıkları belirlemek,
- iii. Kümeleme yöntemleri ile birimleri alt gruplara ayırmak,
- iv. Grupları ve elde edilen grafikleri analiz ederek ve yorumlayarak hipotezleri test etmek.

2.2. Benzerlik

2.2.1. Benzerlik Kavramı

Birimleri benzerlik durumlarına göre tanımlamak sınıflama sürecinin temelidir. Görünüşte basit olmasına rağmen benzerlik ölçüleri temelden itibaren uzaklık ölçülerinin prosedürlerini kullanmaktadır. Benzerlik kavramı, bazı epistemolojik sorunların ortaya çıkmasına neden olmaktadır. Sadece birimler için benzemek ya da benzememek istatistiksel bir sınıflama yöntemi olamaz. Bu nedenle sınıflandırma yöntemleri için bilimsel araştırmalar gerçekleştirilerek bilimsel süreçler takip edilmiştir. İstatistiksel prosedürlerdeki gelişmeler benzerlik kavramı açısından doğal bir sonuç olarak gerekli ve güvenilir olmuştur. Ölçüm kavramı benzerliğin kantitatif tahminlerine hakim olmuştur. Benzerliğin bu yaklaşımı koordinat uzayında gözlenen benzerlik ile farklılık noktalarına karşılık aralarındaki metrik uzaklık ile temsil edilir [4]. Uzay boyutluluğu olayların tanımlanmasında kullanılan değişken sayısı tarafından belirlenmektedir. Benzerlik

ölçütlerinden doğru olup olmadığını kararını vermek için kullanılan dört standart mevcuttur [1].

1. Simetri, x ve y iki birim olmak üzere;

$$d(x, y) = d(y, x) \geq 0 \quad (2.1)$$

eşitliği ile ifade edilmektedir.

2. Üçgen eşitsizliği, x, y, z üç birim olmak üzere;

$$d(x, y) \leq d(x, z) + d(y, z) \quad (2.2)$$

eşitliği ile ifade edilmektedir.

3. Özdeş olmayan ayırt edilebilirlik, x ve y iki birim olmak üzere;

$$d(x, y) \neq 0, \quad x \neq y \quad (2.3)$$

ile ifade edilmektedir.

4. Özdeş olmayan ayırt edilemezlik, x ve x' için

$$d(x, x') = 0 \quad (2.4)$$

ile ifade edilmektedir. Bu durum iki birim arasındaki uzaklığın 0'a eşit olduğunu göstermektedir.

Bu standartlar önemli matematiksel özelliklerdir. Jardine ve Sibson (1971) ve Clifford ve Stephenson (1975) başta olmak üzere benzerlik kavramının rutin kullanımlarına karşı metrik niteliklere sahip bazı öneriler sunmuştur [5], [6]. Ölçüler önemli olmasına rağmen, nesneler arasındaki benzerliği temsil etmenin tek yolu anlamına gelmemektedir. Benzerlik kavramı benzerliğin temsil edilebilmesi için kalıtsal boyutluluğa sahip değildir [1].

2.2.2. Değişken Seçimi ve Standardize İşlemleri

Kümeleme analizinde değişken seçimi araştırma süreçlerinde en önemli adımlardan birini teşkil etmektedir. Fakat bu adımın önemi birçok uygulamada kullanılmamış olmasından dolayı yeterince bilinmemektedir [1].

Benzerlik kavramını temsil eden değişkenler kümesinin bulunması temel problemdir. Değişkenler sınıflandırmaya destek olma amacı ile teoriye uygun olarak seçilmelidir. Çalışmalarda kullanılmak üzere değişkenlerin rasyonel seçimi için teori temeli teşkil etmektedir. Bu durum pratikte uygulanamamaktadır. Teori uygulamada sadece kapalı araştırmalarda sınıflandırmayı destekleyici nitelikte kullanılmaktadır. Bu nedenle bu tür çalışmalardaki değişken uygunluğunu değerlendirmek güç olmaktadır [1].

Birçok istatistiksel analizde uygun yöntemler kullanılarak standardize işlemi uygulanmaktadır. Değişkenlerin normal dağılımı söz konusu olduğunda logaritmik veya diğer dönüşümler sıklıkla kullanılmaktadır. Verilerin aynı ölçekte olmadığı durumlarda ise değerler birim varyans ile standardize edilmektedir. Bazı araştırmacılar alt gruplara ayırmada kümeleme analizinde kullanılan standardize yönteminin uygunluğunu tartışmışlardır [1].

2.2.3. Z Değerleri Standardizasyonu

Bu yöntem ile verilere istatistiksel bir dağılımdan yararlanılarak Z değerlerine dönüşüm uygulanmaktadır. Veriler oransal ve aralık ölçek türü olduğunda bu yöntem kullanılabilir.

$\bar{x}$ aritmetik ortalama ve s standart sapma olmak üzere Z değerleri standardizasyonu;

$$Z_{ik} = \frac{x_{ij} - \bar{x}}{s_j} \quad (2.5)$$

eşitliği ile sağlanmaktadır.

2.2.4. 0-1 ve -1,1 Standardizasyonu

Verilerin homojen olmadığı durumlarda kullanılan standardize işlemlerindendir. En küçük x değerinin en büyük x değerinden küçük olduğu durumlar için elverişlidir. x_i serideki herhangi bir değer olmak üzere -1 ile 1 arası standardize işlemi;

$$x'_i = \frac{x_i}{x_{max}} \quad (2.6)$$

eşitliği ile sağlanmaktadır.

0 -1 standardizasyonu

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (2.7)$$

eşitliği ile sağlanmaktadır.

2.2.5. $\mu = 1, \sigma = 1$ Standardizasyonu

Değişkene ilişkin ortalamanın 1 olması için uygulanan standardize işlemidir. Standardize işlemi;

$$x'_i = \frac{x_i}{\bar{x}} \quad (2.8)$$

eşitliği ile sağlanmaktadır. Ortalamanın sifra eşit olduğu durumlarda standardize işlemi

$$x'_i = \frac{x_i + 1}{\bar{x} + 1} \quad (2.9)$$

eşitliği ile sağlanmaktadır. Standart sapmanın 1'e eşit olarak standardize işlemi gerçekleştirilmesi durumunda

$$x'_i = \frac{x_i}{s_x} \quad (2.10)$$

eşitliği ile standardize işlemi gerçekleştirilmektedir.

2.2.6. Benzerlik Ölçüleri

Yaygın olarak kullanılan dört benzerlik ölçüsü mevcuttur. Bunlar korelasyon katsayısı, uzaklık ölçüsü, ilişki katsayısı ve olasılıksal benzerlik katsayısıdır. Her yöntem kullanıldığı veri setine göre belirli avantajlara ve dezavantajlara sahiptir.

2.2.7. Uzaklık Ölçüleri

i. Öklit Uzaklık Ölçüsü;

x_i ve x_j iki birim olmak üzere, x_i ve x_j birimleri arasındaki uzaklık;

$$d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ip} - x_{jp})^2} \quad (2.11)$$

eşitliği ile sağlanmaktadır. Öklit uzaklık ölçüsü kümeleme analizinde sıklıkla kullanılmaktadır. Uzaklık ölçüsünün kullanımı değişkenlerin ağırlıklandırma durumuna göre farklılık göstermektedir. Değişkenlerin ağırlıklandırılmadığı takdirde x_i ve x_j birimleri arasındaki uzaklık;

$$d(i, j) = \sqrt{w_1(x_{i1} - x_{j1})^2 + w_2(x_{i2} - x_{j2})^2 + \dots + w_p(x_{ip} - x_{jp})^2} \quad (2.12)$$

eşitliği ile elde edilmektedir.

ii. Minkowski Uzaklık Ölçüsü;

x_i ve x_j iki birim olmak üzere, x_i ve x_j birimleri arasındaki uzaklık;

$$d(i, j) = [|x_{i1} - x_{j1}|^m + |x_{i2} - x_{j2}|^m + \dots + |x_{ip} - x_{jp}|^m]^{\frac{1}{m}} \quad (2.13)$$

eşitliği ile elde edilmektedir. Değişkenlerde ağırlıklandırma kullanılmadığı durumlarda x_i ve x_j birimleri arasındaki uzaklık;

$$d(i, j) = [w_1|x_{i1} - x_{j1}|^m + w_2|x_{i2} - x_{j2}|^m + \dots + w_p|x_{ip} - x_{jp}|^m]^{\frac{1}{m}} \quad (2.14)$$

eşitliği ile elde edilmektedir.

iii. Pearson Uzaklık Ölçüsü;

Pearson uzaklık ölçüsü x_i ve x_j iki birim arasındaki uzaklık hesaplamalarında değişkenin varyansında dikkate almaktadır. x_i ve x_j iki birim olmak üzere, x_i ve x_j birimleri arasındaki uzaklık;

$$d(i, j) = \sqrt{(x_{i1} - x_{j1})^2/s_1^2 + (x_{i2} - x_{j2})^2/s_2^2 + \dots + (x_{ip} - x_{jp})^2/s_p^2} \quad (2.16)$$

eşitliği ile elde edilir. x_i ve x_j iki birim olmak üzere değişkenlere ağırlıklandırma uygulanmadığı durumlarda pearson uzaklık ölçüsü ile iki birim arasındaki uzaklık;

$$d(i, j) = \sqrt{w_1(x_{i1} - x_{j1})^2/s_1^2 + w_2(x_{i2} - x_{j2})^2/s_2^2 + \dots + w_p(x_{ip} - x_{jp})^2/s_p^2} \quad (2.17)$$

eşitliği ile elde edilir.

iv. Manhattan Uzaklık Ölçüsü;

x_i ve x_j iki birim olmak üzere, Manhattan Uzaklık ölçüsü ile iki birim arasındaki uzaklık;

$$d(i, j) = (|x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}|) \quad (2.18)$$

eşitliği ile elde edilir. Manhattan uzaklık ölçüsünde mutlak değerler ile belirtilen mutlak uzaklıklar eşitlikte kullanılmaktadır. Değişkenlerin ağırlıklandırılmadığı durumlarda x_i ve x_j birimleri arasındaki uzaklık;

$$d(i, j) = (w_1|x_{i1} - x_{j1}| + w_2|x_{i2} - x_{j2}| + \dots + w_p|x_{ip} - x_{jp}|) \quad (2.19)$$

eşitliği ile sağlanır.

2.2.8. Korelasyon Katsayısı

Geometrik olarak yorumlanmasından dolayı açısal ölçü olarakta nitelendirilen korelasyon katsayısı benzerlik ölçüleri arasında en çok kullanılanlardan biridir. Genelde iki değişken arasındaki ilişkiyi ortaya koymak için kullanılan bu katsayı;

$$r_{jk} = \frac{\sum(x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum(x_{ij} - \bar{x}_j)^2 \sum(x_{ik} - \bar{x}_k)^2}} \quad (2.20)$$

eşitliği ile sağlanmaktadır. x_{ij} i. değişkeninin j. durumdaki değerini temsil etmektedir. $\bar{x}_j$ j. durumdaki tüm değerlerin ortalamasını temsil etmektedir. Korelasyon katsayısı -1 ile 1 aralığında değerler almaktadır. Korelasyon katsayısının 0 olması iki değişken arasında ilişkinin olmadığı korelasyon katsayısının +1 yönünde olması iki değişken arasında paralel

yönlü kuvvetli bir ilişki olduğu, korelasyon katsayısının -1 yönünde olması iki değişken arasında ters yönlü bir ilişkinin olduğunu göstermektedir [1].

2.2.9. İlişki Katsayısı

İlişki katsayısı ikili değişkenlerde tanımlanmış birimler arasında benzerlik kurmak için kullanılmaktadır. 2x2'lik bir kontejans tablosunda 1 varlığı 0 ise yokluğu temsil etmek üzere;

	1	0
1	a	b
0	c	d

Bu yöntem ile temsil edilebilecek üç katsayı bulunmaktadır. Bu katsayılar Matching katsayısı, Jaccard's katsayısı ve Gower's katsayısıdır [1]. Matching katsayısı;

$$S = \frac{(a+d)}{(a+b+c+d)} \quad (2.21)$$

şeklindedir. S, 0 ile 1 arasında olan iki değişken arasındaki benzerliktir.

Jaccard's katsayısı;

$$S = \frac{a}{(a+b+c)} \quad (2.22)$$

eşitliği ile ifade edilmektedir. Gower's katsayısı benzerlik tahminlerinde ölçülerin farklı ölçeklerin değişkenlerinin kullanımından dolayı kendiliğinden ortaya çıkan birleşik bir katsayıdır [1].

$$s_{ij} = \frac{\sum_{k=1}^p s_{ijk}}{\sum_{k=1}^p w_{ijk}} \quad (2.23)$$

Eşitliği ile Gower's katsayısı elde edilir. w_{ijk} ağırlıklandırılmış değişken s_{ijk} ise benzerlik skorudur [1].

2.2.10. Olasılıksal Benzerlik Katsayısı

Bu katsayılar teknik olarak iki değişken arasındaki benzerliği hesaplamaz. Bu ölçü tipi bunun yerine doğrudan ham veriler üzerinde çalışmaktadır. Elde edilen bilgiler sonucu olasılıksal sonuçlar elde edilmektedir [1].

2.3. Hiyerarşik Kümeleme Yöntemi

Belirli özelliklere göre birimlerin benzerliklerini ortaya koyarak ve bu benzerlikleri esas kabul ederek birimleri doğru kategorilere sınıflandırmak kümeleme analizi yöntemlerinin temelini teşkil etmektedir. Kümeleme analizi bu açıdan çok değişkenli istatistiksel tekniklerden biri olan diskriminant analizi ile örtüşmektedir. Fakat diskriminant analizi ile kümeleme analizi tekniklerinde uygulama açısından farklı durumlar söz konusu olmaktadır. Kümeleme analizi teknikleri ile birimler için mevcut andaki durumlar göz önünde bulundurulmaktadır. Diskriminant analizinde geleceğe yönelik tahminler için öngörülerde bulunmaktadır [7] - [9].

Kümeleme analizi teknikleri önemli iki başlık altında sınıflanmaktadır. Bunlar hiyerarşik kümeleme yöntemleri ve hiyerarşik olmayan kümeleme yöntemleridir. Hiyerarşik kümeleme yöntemleri birleştirici ve parçalayıcı hiyerarşik kümeleme yöntemleri olmak üzere iki temel grupta incelenmektedir. Literatürde birleştirici hiyerarşik kümeleme yöntemleri parçalayıcı kümeleme yöntemlerine nazaran daha yaygın olarak kullanılmaktadır. Parçalayıcı hiyerarşik kümeleme yönteminde mevcut gözlemlerin tamamı tek bir küme olarak kabul edilmektedir. Küme sayısı bir eksiltiyle benzerlik matrisi yeniden düzenlenir, benzerlik ve uzaklık matrisine göre benzer birimleri bir araya getirilmektedir. n birim aşamalı olarak sırası ile $1, 2, 3, \dots, (n-r), \dots, (n-3), (n-2), (n-1)$ ve n kümeye yerleştirilir [7], [10]. Birleştirici hiyerarşik kümeleme yönteminde her birim bir kümeyi temsil etmektedir. Daha sonra parçalayıcı hiyerarşik kümeleme yönteminin tersi işlemler gerçekleştirilir. n birim sırası ile $n, (n-1), (n-2), \dots, (n-r), \dots, 3, 2, 1$ kümeye yerleştirilir. Her işlem sonunda küme sayısı azalmaktadır ve son aşamada tek bir küme bütün birimleri ihtiva etmektedir. Birleştirici hiyerarşik kümeleme yöntemi farklı algoritmalar ile uygulanmaktadır [7].

2.3.1. En Yakın Komşuluk Yöntemi

Tek bağlantılı kümeleme yöntemi olarak da nitelendirilen en yakın komşuluk yönteminde uzaklıklar matrisi kullanılarak birbirine en yakın birimler birleştirilmekte ve birleştirme işlemi tekrarlanarak yürütülmektedir.

Bu yöntemde ilk adım iki noktanın bulunmasıdır. İki nokta bir kümeyi temsil etmektedir. Üçüncü bir noktanın bu kümeye olan uzaklığı araştırılmaktadır. Üçüncü noktanın kümedeki herhangi bir elemana olan uzaklığı ile küme dışındaki elemanlara uzaklığı ile kıyaslandığında, kümedeki elemanlara uzaklığı daha küçük olduğu durumlarda üçüncü eleman bu kümeye atanmaktadır. En yakın komşuluk yöntemi ile yapılan hiyerarşik kümeleme elips şeklinde dağılım göstermeyen birimleri aynı kümede toplayabilme yeteneği bakımından diğer kümeleme yöntemlerine karşı üstünlük sağlamaktadır. Bu yöntemin avantajlı olduğu durumlar bulunmasına karşın birimler arasındaki farkın az olduğu durumlarda ayırım işlemini yapmada yeterli değildirler [7]. En yakın kümeleme yöntemi bir uygulama ile şu şekilde ifade edilebilmektedir;

$$D_1 = \begin{matrix} & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 0 & & & & \\ 2 & 0 & & & \\ 6 & 5 & 0 & & \\ 10 & 9 & 4 & 0 & \\ 9 & 8 & 5 & 3 & 0 \end{pmatrix} \end{matrix}$$

D_1 matrisinde 5 birim bulunmaktadır ve her birimin birbirine olan uzaklığı matris içerisinde verilmiştir. 1 ve 2 birimlerinde oluşan iki birimlik bir küme oluşturulsun ve bu kümelerin diğer birimlere olan uzaklığı;

$$d_{(12)3} = \min(d_{13}, d_{23}) = d_{23} = 5$$

$$d_{(12)4} = \min(d_{14}, d_{24}) = d_{24} = 9$$

$$d_{(12)5} = \min(d_{15}, d_{25}) = d_{25} = 8$$

1 ve 2 birimin tek bir küme oluşturulmasından sonra yeni bir matris elde edilmektedir.

$$D_2 = \begin{matrix} (1,2) \\ 3 \\ 4 \\ 5 \end{matrix} \begin{pmatrix} 0 & & & \\ 5 & 0 & & \\ 9 & 4 & 0 & \\ 8 & 5 & 3 & 0 \end{pmatrix}$$

D_2 matrisinde birimler arasındaki en küçük uzaklık 4 ve 5 birim arasındadır. 2. Küme 4. ve 5 birimlerden oluşmaktadır.

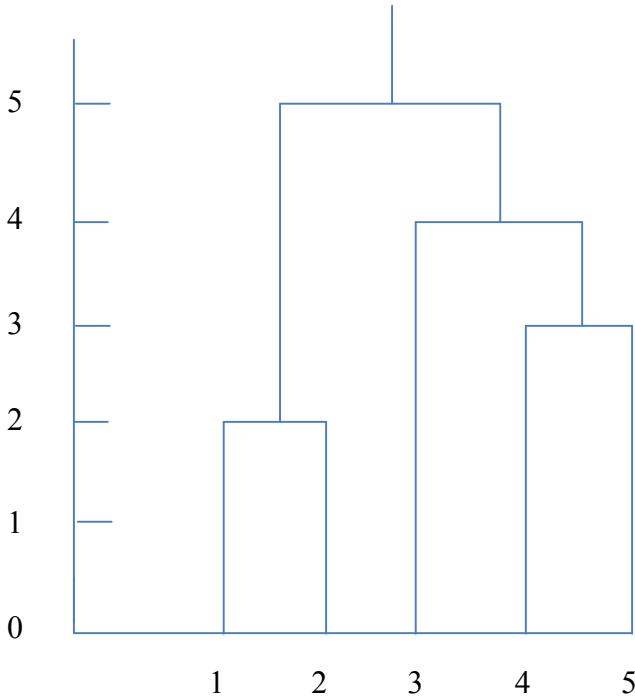
$$d_{(12)(45)} = \min(d_{14}, d_{15}, d_{24}, d_{25}) = d_{25} = 8$$

$$d_{(45)3} = \min(d_{34}, d_{35}) = d_{34} = 4$$

Bu bilgilere göre matris yeniden düzenlendiğinde;

$$D_3 = \begin{matrix} (1,2) \\ 3 \\ (4,5) \end{matrix} \begin{pmatrix} 0 & & \\ 5 & 0 & \\ 8 & 4 & 0 \end{pmatrix}$$

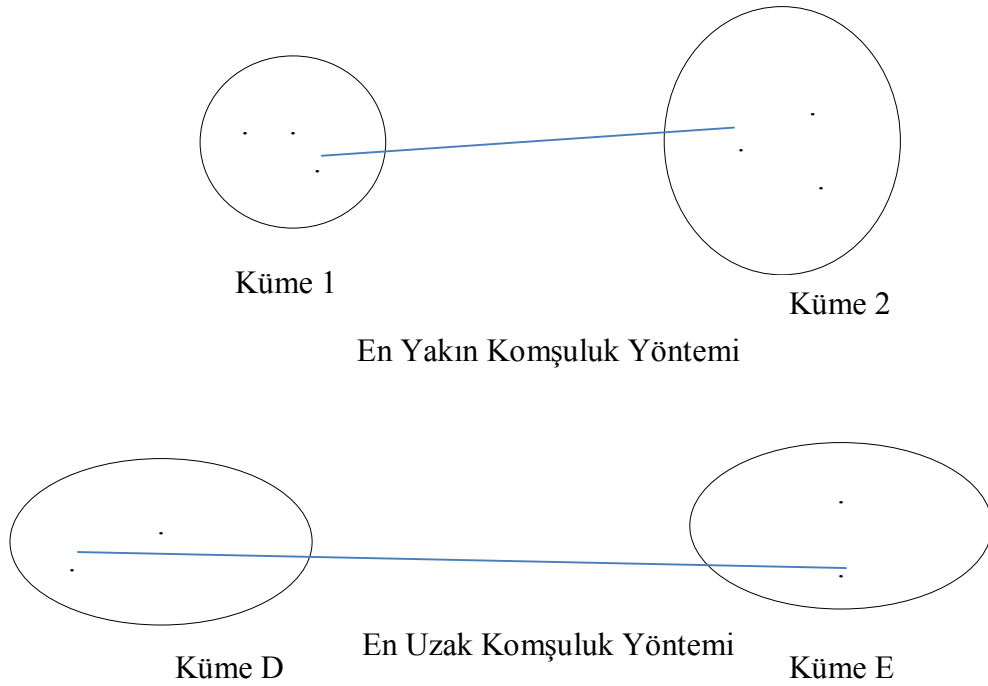
Son olarak 3.birimde kendisine en küçük uzaklıktaki (4,5) kümesine üye olacaktır. Son olarak birleştirici hiyerarşik kümeleme yöntemine göre tüm kümelerden bir küme altında birleşecektir [12].



Şekil 2.1. Dendrogram Grafiği

2.3.2. En Uzak Komşuluk Yöntemi

Bu yöntem prensipte en yakın komşuluk yöntemine benzemektedir. Fakat en yakın komşu birimlerin kümелendiği en yakın komşuluk yönteminden farklı olarak bu yöntemde birbirine en uzak iki birim bir küme oluşturmaktadır. Diğer kümelerin de birleştirilmesi ile yine tek bir küme elde edilmektedir. Birimlerin veya kümelerin birleştirilmesindeki temel ölçü birimler arasındaki maksimum uzaklığın dikkate alınmasıdır. Bu yöntem en yakın komşuluk yöntemine göre birimlerin elips şeklinde dağılması durumunda avantaj sağlamaktadır [7].



Şekil 2.2. En yakın komşuluk yöntemi ile en uzak komşuluk yöntemi

2.3.3. Ortalama Bağlantı Kümeleme Yöntemi

Bu yöntemde birimler arasındaki benzerlikler önem kazanmaktadır. İki küme arasındaki uzaklığın hesaplanmasında, ilk kümedeki birimlerin ikinci kümedeki birimlere olan uzaklıklarının ortalaması elde edilmektedir. Diğer yöntemlere karşı avantajlı olduğu durumlar söz konusudur. Ortalama bağlantı kümeleme yöntemi uç değerlere karşı hassasiyeti oldukça azaltmaktadır [7].

2.3.4. Ward's Bağlantı Kümeleme Yöntemi

Ward's bağlantı kümeleme yöntemi diğer kümeleme yöntemlerine göre farklılık göstermektedir. Diğer kümeleme yöntemlerinde kümeler arasındaki uzaklıklar elde edilirken bu yöntemde küme içerisindeki birimlerin hata kareleri toplamını minimize edilerek küme içerisinde homojen kümeler oluşturulması amaçlanmaktadır. Aşama aşama ilerlenerek hata kareler toplamı en küçük olan kümeler birleştirilir. Bu sayede küme içerisinde homojen kümeler arası heterojenlik sağlanmaktadır [7].

2.3.5. Merkezi Bağlantı Kümeleme Yöntemi

İki küme merkezine olan uzaklıkları, kümelerin benzerliklerini belirlemede kullanılmaktadır. Küme merkezleri olarak belirtilen değerler ortalama değerleri teşkil etmektedir. Kümeye yeni birimler dahil edildikçe kümenin merkezi değişmektedir. Ortalamaya dayalı olarak küme merkezini belirlediği için uç birim hassasiyetine karşı dirençlidir [7].

2.4. Hiyerarşik Olmayan Kümeleme Yöntemleri

Birimlerin bulunduğu küme sayısı, kümeleme analizi için elde edilmek istenen temel bilgilerden biridir. Küme sayısının bilinmediği durumlarda hiyerarşik kümeleme yöntemleri uygulanmaktadır. Böylelikle küme sayısı önceden bilinmediği durumlarda veya küme sayısı hakkında bir kanı bulunmadığı durumlarda hiyerarşik kümeleme yöntemleri uygun bir alternatif olmaktadır. Küme sayısı hakkında bilginin bulunduğu veya kümeleme analizinde elde edilecek küme sayısı için öngörü oluştuğunda hiyerarşik olmayan kümeleme yöntemleri tercih edilmektedir. Hiyerarşik olmayan yöntemler diğer kümeleme yöntemlerine göre daha sağlam teorik temellere dayanmaktadır [11].

Hiyerarşik olmayan kümeleme yöntemleri, hiyerarşik kümeleme yöntemlerine göre farklılıklar göstermektedir. Hiyerarşik olmayan yöntemlerde birimler öncelikle belirlenmiş olan kümelere yerleştirilmektedir. Hiyerarşik yöntemlerde ise birimler kümelere belirli bir sıra takip edilerek alınmaktadır. Hiyerarşik olmayan yöntemler teorik yapısına göre büyük sayıdaki örneklerde daha uygun bir çalışma alanı bulmuştur [11].

2.4.1. K-Ortalamalar Yöntemi

Hiyerarşik olmayan kümeleme yöntemleri içerisinde en yaygın olarak kullanılan yöntem K-ortalamalar yöntemidir. Bu yöntem, küme içerisindeki kareler toplamını en küçük biçimde k sayıda kümeye bölmektedir.

$x_1, x_2, \dots, x_n$ p değişkenli gözlem vektörleri ve $a_{1n}, a_{2n}, \dots, a_{kn}$ küme merkezleri olmak üzere;

$$W_n = \frac{1}{n} \sum_{i=1}^n \min_{1 \leq j \leq k} \|x_i - a_{jn}\| \quad (2.24)$$

bağıntısı K-ortalama tekniğinin algoritmasını temsil etmektedir [11].

2.5. Diskriminant Analizi

İstatistiksel analizlerde birimleri bilinen özelliklerine göre tasnif etmek önemli bir yer tutmaktadır. Birçok fen ve sosyal bilimi için geniş uygulama sahası bulan diskriminant analizi, bilinen özelliklere göre birimleri sınıflara ayırmaktadır. Fakat bu yönüyle kümeleme analizi ile aynı işlevi gerçekleştirdiği algısı yanıltıcı olacaktır. Diskriminant analizi istatistiksel temellere dayanan güçlü bir karar biçimidir. Kümeleme analizinde değişkenlerin özelliklerine göre sınıflandırma işlemi uygulanırken diskriminant analizinde amaç hatalı sınıflandırma yapılmasını önlemek için birimleri gerçek gruplarına ayırmak ve birimlerin gelmiş olduğu yığınları tespit etmektir [11]. Diskriminant analizi tam olarak birimleri sınıflandırma işlemini en az hata ile gerçekleştirmek için mevcut istatistiksel yöntemlerdir. Bu yöntemde esasında gruplar arasındaki farklılığın ortalamar cinsinden en yüksek değerlere ulaşmasını hedeflemektedir. Bunun için diskriminant analizi birimleri tasnif edebilecek bir fonksiyona ihtiyaç duymaktadır [11].

2.5.1. İki Grup için Diskriminant Analizi

Birimlere ilişkin özelliğin çok sayıda olması kümelerin daha belirgin şekilde ayrılmasına neden olsa da, her bir özellik için birimleri sınıflara ayırmak bazı durumlar için mümkün olmamaktadır. Değişken sayısının indirgenmesi sınıflama işlemini kolaylaştırmaktadır. Bu nedenle diskriminant analizi çok değişkenli sistemi tek değişkene dönüştürmektedir. Bu işlemi gerçekleştirirken tüm değişkenleri ağırlıklandırarak fonksiyona katılmasını sağlar [11].

$x_1, x_2, \dots, x_p$ değişkenler ve $a_1, a_2, \dots, a_p$ değişkenlere verilen ağırlıklar olmak üzere p değişkenli bir birim için;

$$y_i = a_1 x_{i1} + a_2 x_{i2} + \dots + a_p x_{ip} \quad (2.25)$$

Bağıntısı elde edilmektedir. Bu eşitlik için gruplar arası varyansın grup içi varyansa göre en büyük değer ulaştırılması istenmektedir. Bu durum;

$$F = \frac{\text{Gruplar arası varyans}}{\text{Grup içi varyans}} \quad (2.26)$$

eşitliği ile gösterilmektedir [11].

y_i eşitliğinde kullanılan a değişken ağırlık katsayılarının bulunmasında Fisher;

$$f(a_1, a_2, \dots, a_p) = \frac{a'Ba}{a'Wa} \quad (2.27)$$

eşitliğini öne sürmüştür.

a, $p \times 1$ boyutlu katsayılar vektörü B, $p \times p$ boyutlu gruplar arası, W $p \times p$ boyutlu grup içi varyans matrisidir.

$$C = \frac{1}{(n_1+n_2-2)} W \quad (2.28)$$

ise;

$\bar{x}_j^{(1)}, \bar{x}_j^{(2)}$ değişken ortalamaları olmak üzere

$$d_j = (\bar{x}_j^{(1)} - \bar{x}_j^{(2)}) \text{ ve } d' = (d_1, d_2, \dots, d_p) \quad (2.29)$$

$$f(a_1, a_2, \dots, a_p) = \frac{n_1 n_2}{n_1 + n_2} \frac{a' d d' a}{a' C a} \quad (2.30)$$

eşitliği elde edilir. İki grubun grup içi varyansları eşitliğinden [11];

$$a = C^{-1} d \quad (2.31)$$

elde edilir. C matrisi grup içi kovaryans matrisinden elde edilmiştir. C matrisinin inversi ile grup ortalamaları vektörlerinden elde edilen fark değerleri vektörlerinin çarpımı sonucu sağlanabilmektedir. Bu durum [11];

$$\bar{y}^{(1)} = a_1 \bar{x}_1^{(1)} + a_2 \bar{x}_2^{(1)} + \dots + a_p \bar{x}_p^{(1)} \quad (2.32)$$

$$\bar{y}^{(2)} = a_1 \bar{x}_1^{(2)} + a_2 \bar{x}_2^{(2)} + \dots + a_p \bar{x}_p^{(2)} \quad (2.33)$$

eşitlikleri ile gösterilebilir.

Elde edilecek diskriminant fonksiyonu ;

$$Var(y) = \sum_{i=1}^p \sum_{j=1}^p c_{ij} a_i a_j = a' C a \quad (2.34)$$

eşitliği ile ifade edilebilir [11].

2.5.2. Çoklu Grup İçin Diskriminant Analizi

Çok sayıda grup için diskriminant analizi yine doğrusal bağlantılar kurarak p değişkenli çok sayıda kitledeki birimler için ayırma gücü en yüksek sınıflandırmayı gerçekleştirmektedir. Diskriminant fonksiyonunun elde edilmesindeki yapı iki grup için diskriminant analizi için yapılan fonksiyonun elde edilmesi ile paralellik göstermektedir. Fisher tarafından elde edilen eşitlik;

$$\lambda = \frac{a' B a}{a' W a} \Big|_{max} \quad (2.35)$$

λ_i özdeğerlerine karşılık gelen özvektörler ayırıcı fonksiyonlardır [11].

a' ya göre türev alındığında;

$$\frac{\partial \lambda}{\partial a} = \frac{2[(Ba)(a'Wa) - (a'Ba)(Wa)]}{(a'Wa)(a'Wa)} \quad (2.36)$$

$$\frac{2Ba}{a'Wa} - \frac{2\lambda Wa}{a'Wa} = 0 \quad (2.37)$$

$$\frac{2(Ba - \lambda Wa)}{a'Wa} = 0 \quad (2.38)$$

elde edilir. Buradan;

$$2(Ba - \lambda Wa) = 0 \quad (2.39)$$

$$Ba - \lambda Wa = 0 \quad (2.40)$$

$$(B - \lambda W)a = 0 \quad (2.41)$$

$$(W^{-1}B - \lambda I) = 0 \quad (2.42)$$

λ_i özdeğerleri $|W^{-1}B - \lambda I| = 0$ determinantının çözümü r tane özvektör elde edilmesini sağlayacaktır. Bu şekilde elde edilen bir diskriminant fonksiyonu;

$$y = a_1x_1 + a_2x_2 + \dots + a_px_p \quad (2.43)$$

eşitliği ile ifade edilir [11].

3. UYGULAMA

3.1. Uygulamanın Amacı

Kümeleme Analizi yöntemleri tasnif etme ve değişkenleri sınıflandırma işlemleri için kullanılabilir. Bu yöntemi uygularken benzerlik ve uzaklık ölçülerinden yararlanılmaktadır. Çalışmanın amacı, kümeleme analizi ile Doğu Anadolu Bölgesi'nde bulunan Bingöl, Elazığ, Malatya ve Tunceli illeri ve bu illere bağlı ilçelerin hanelerindeki elektrik tüketiminin sınıflandırılması ve yerleşim merkezlerinin kümelerinin belirlenmesidir. Ayrıca elektrik tüketiminin sınıflandırılmasında kullanılan kümeleme yöntemlerinin diskriminant analizi ile inceleyerek etkinlikleri karşılaştırılmıştır.

3.2. Uygulama Verileri ve Düzeni

Elektrik tüketimine ilişkin veriler 2010 yılına ait Doğu Anadolu Bölgesi'nde Bingöl, Elazığ, Malatya ve Tunceli illerinden elde edilmiştir. Belirtilen dört il ve ilçe merkezlerindeki hanelere ait 12 aylık kw cinsinden elektrik tüketim değerleri elde edilmiştir. Bazı ilçe merkezlerinde elektrik tüketim tespitinin aylık yerine 2 aylık yapılması nedeniyle bu hanelerdeki 2 aylık tüketim değerleri ikiye bölünerek ilgili aylar için ortalama tüketim değerleri elde edilmiştir. Verilerin sağlanması için öncelikle bölgedeki il ve ilçe merkezlerindeki hane sayısı ve nüfus dikkate alınarak örnekleme yöntemleri ile verilerin çekileceği haneler belirlenmiştir. Belirlenen hanelere ilişkin aylık tüketim değerleri elde edilmiştir. 12 aylık elde edilen tüketim değerleri toplanarak hanelerin yıllık elektrik tüketim değerleri elde edilmiştir. İl ve ilçe merkezlerindeki ortalama bir hanenin yıllık elektrik tüketim değerleri belirlenmiştir. İl ve ilçe merkezlerindeki ortalama yıllık elektrik tüketimine göre bu merkezler “1.derece” , “2.derece” , “3.derece” , “4.derece” ve “5.derece” olmak üzere 5 kümeye tasnif edilmiştir. Farklı kümeleme yöntemleri kullanılarak en uygun kümeleme yönteminin tespiti için diskriminant analizi ile elde edilen kümeleme sonuçları incelenmiştir. Verilerin analizi ve incelenmesi SPSS paket programında gerçekleştirilmiştir. Sonuçlar tablo ve dendogram grafikleri ile sunularak bölgedeki yerleşim merkezlerinin elektrik tüketimi sınıflandırılarak tablo üzerinde

gösterilmiştir. Elektrik tüketiminin önemi ve elektrik tüketimindeki farklılıklar incelenerek nedenleri üzerine tartışılmıştır.

3.3. Elektrik Tüketimi

İnsanlar yaşamını devam ettirebilmek adına sürekli bir gelişim süreci içerisinde bulunmuştur. Gelişim süreci içerisinde kendisini ve teknolojisini yenilemiştir. Teknoloji ve teknikler geliştikçe insanların ihtiyaç duyduğu enerji miktarı da artmıştır. Yaşamın devamı için ihtiyaç duyulan enerjiyi döneme göre farklı nesnelerden sağlanması belirli bir süre sonra farklı enerji kaynaklarına yönelmesine neden olmuştur. İlk çağlarda çalışabilmek ve hayatını sürdürebilmek için enerji kaynağı olarak kendi vücut gücünü kullanan insanoğlu daha sonraları doğadan yararlanmayı öğrenmiştir. Öncelikle rüzgardan ve yanma olayından enerji sağlanmış daha sonra fosil yakıt kaynakları keşfedilmiştir. Daha sonra ise elektrik keşfedilmiştir. Elektriğin enerji için kullanımı günden güne hızla büyüyerek artmıştır. Elektrik doğrudan enerji kaynağı olarak kullanılmaması elektriği farklı bir enerji kaynağı statüsüne yerleştirmiştir. Çünkü elektrik enerjisi doğrudan elde edilememektedir ve farklı enerji kaynaklarının bir takım mekanik teknoloji yardımı ile dönüştürülmesi sonucu sağlanabilen bir enerji kaynağıdır. Elektrik insanın yaşamındaki her alanda geniş çalışma sahası bulmuştur. Gerek insanın iş yaşamında gerekse de sosyal alanında elektrik tüketimi günden güne artmaktadır. İnsanın yoğun olduğu ve sanayi kullanımının bulunduğu bölgelerde elektrik tüketimi fazla olmakta iken insan yoğunluğunun ve iş gücünün az olduğu bölgelerde enerjiye ihtiyacın az olması nedeniyle elektrik tüketimi azalmaktadır. Fakat elektrik tüketiminin bölgeden bölgeye farklılık göstermesinden ziyade hanelerde bile farklılık göstermesi sadece sanayi ve dış çevre faktörünün etkili olmadığını göstermektedir. Zira ailedeki birey sayısı, sahip olunan elektrikli aletler, yaşam tarzı, iklimsel faktörler gibi birçok neden hanelerdeki elektrik tüketimini de etkileyebilmektedir. Elektrik tüketimindeki artış ve azalışlar enerji talebini karşılanabilmesi açısından incelenmektedir. Kentlerdeki farklı elektrik tüketimi, yine bu kentlerin elektriğe olan ihtiyacının karşılanması açısından önem taşımaktadır. Bu nedenle elektrik tüketiminin önemi ve tüketimdeki farklılıkların nedenleri büyük bir önem taşımaktadır.

3.4. Örneklem Seçimi

Çalışmada Bingöl, Elazığ, Malatya ve Tunceli illerine ait kent ve ilçe merkezlerindeki haneler evreni oluşturmaktadır. Evrenin yüksek boyutu çalışmaya tüm evrenin dahil edilmesini kısıtlamıştır. Bu nedenle evrendeki istatistiksel yöntemler ile örneklem çekimi yapılmıştır. Örneklem yöntemi olarak sistematik örneklem kullanılmıştır. Ayrıca örneklem yapılırken il merkezlerinin ve ilçe merkezlerinin nüfus büyüklükleri dikkate alınarak ağırlıklandırma yapılmıştır. Toplamda 4 il merkezi ve 35 ilçe merkezi için 345 hane örneklem dahil edilmiştir.

Tablo 3.1. Örneklem Elde Edildiği Merkezler

İl Merkezleri				
	Bingöl	Elazığ	Malatya	Tunceli
1	Bingöl Merkez	Elazığ Merkez	Malatya Merkez	Tunceli Merkez
2	Adaklı	Alacakaya	Akçadağ	Çemişgezek
3	Genç	Arıcak	Arapgir	Hozat
4	Karlıova	Ağın	Arguvan	Mazgirt
5	Kiğı	Baskil	Battalgazi	Nazımiye
6	Solhan	Karakoçan	Darende	Ovacık
7	Yedisu	Keban	Doğanyol	Pertek
8		Kovancılar	Doğanşehir	Pülümür
9		Maden	Hekimhan	
10		Palu	Kale	
11		Sivrice	Pötürge	
12			Yazıhan	
13			Yeşilyurt	

3.5. Elektrik Tüketim Verileri

Elektrik tüketim verileri örneklem dahilindeki hanelerin aylık veya 2 aylık tüketim değerleri elde edilerek 2010 yılındaki tüketim değerleri hesaplanarak hanelerin yıllık tüketim değerleri elde edilmiştir. Her il ve ilçe merkezi için ortalama yıllık elektrik tüketim ortalaması elde edilmiştir. İl ve ilçelere göre ortalama yıllık tüketim değerleri tablo 3.2.' de

belirtilmiştir. Bu verilere göre en yüksek elektrik tüketimi Pötürge ilçesinde en düşük elektrik tüketimi ise Çemişgezek İlçesinde gerçekleşmektedir.

Tablo 3.2. 2010 Yılına ait örneklemden elde edilen yıllık hane elektrik tüketimleri ortalamaları

Sıra	Merkez	Tüketim (kw)	Merkez	Tüketim (kw)	Merkez	Tüketim (kw)	Merkez	Tüketim (kw)
1	Bingöl	2010	Elazığ	1925,6	Malatya	2117	Tunceli	1291,4
2	Adaklı	1380,7	Alacakaya	1444,5	Akçadağ	1594,5	Çemişgezek	573,3
3	Genç	1196,3	Arıcak	1290,2	Arapgir	1625	Hozat	1374,3
4	Karlıova	1727,4	Ağın	1198,7	Arguvan	1454,7	Mazgirt	1317,6
5	Kiğı	933,5	Baskil	1120,3	Battalgazi	1529,6	Nazımiye	1388,6
6	Solhan	1584,3	Karakoçan	1734,7	Darende	1868,6	Ovacık	1070,3
7	Yedisu	1064,5	Keban	1595,6	Doğanyol	1630	Pertek	802
8			Kovancılar	1427,9	Doğanşehir	1554	Pülümür	1441,3
9			Maden	1519,2	Hekimhan	1340		
10			Palu	1825,8	Kale	1154,7		
11			Sivrice	1782,7	Pötürge	2181,9		
12					Yazıhan	1760		
13					Yeşilyurt	1860		

3.6. Kümeleme Yöntemlerinin Belirlenmesi

Çalışma da il ve ilçe merkezlerinin kümeleme analizi yöntemleri ile yıllık elektrik tüketim değerleri açısından sınıflandırılması için veriler, kümeleme işlemlerine uygun hale getirilmiştir. Kümeleme işlemi ve il ve ilçe merkezlerinin sınıflandırılması için 5 küme oluşturulmuştur. Değişkenlerin tümü aynı ölçek düzeyinde seçilmesi nedeniyle herhangi bir standardize işlemi kullanılmamıştır. Farklı kümeleme yöntemleri uygulanarak kümeleme yöntemlerinin etkinliği araştırılmıştır.

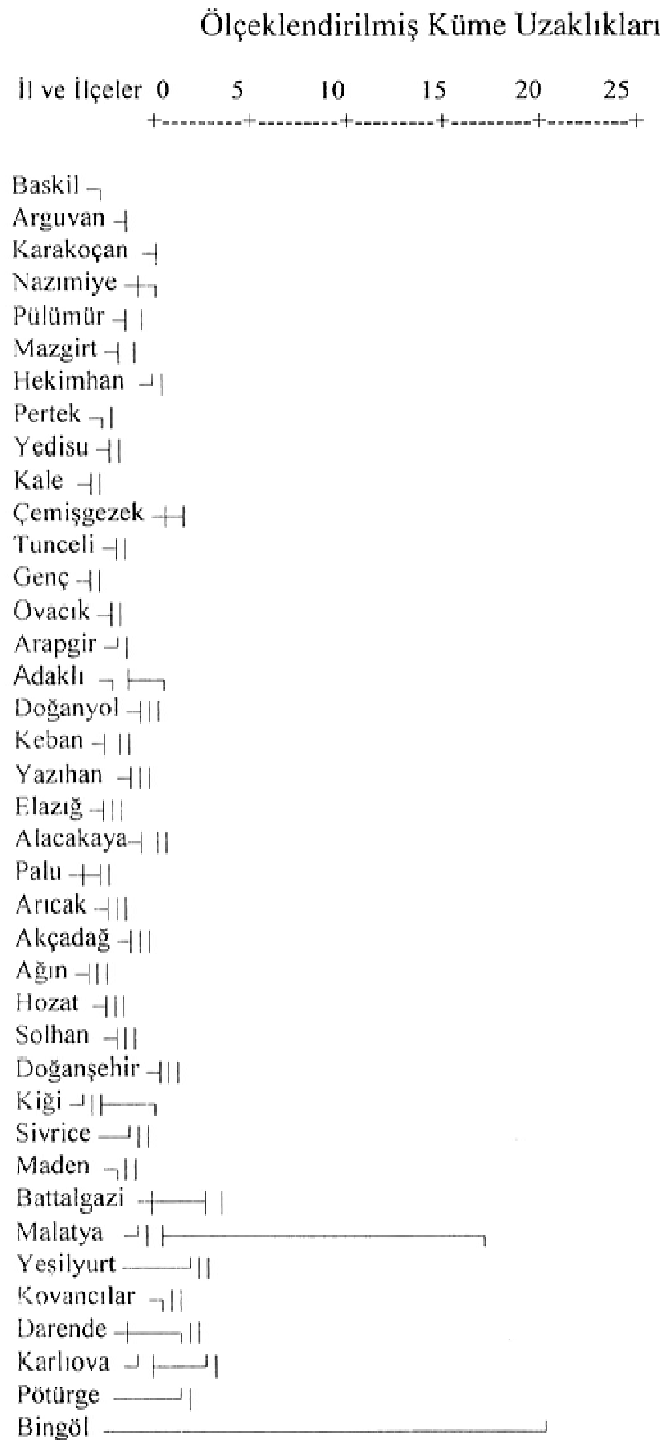
3.7. Gruplar Arası Ortalama Bağlantı yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi

39 yerleşim merkezi 2010 yılına ait yıllık hane elektrik tüketim ortalamalarına göre gruplar arası ortalama bağlantı yöntemi ile 5 kümeye ayırma işlemi gerçekleştirilmiştir. Kümeleme işlemi ikili kümelerden tek bir küme elde edilinceye kadar birleştirme işlemi sürdürülerek gerçekleştirilmiştir. Verilerin tamamı tek bir ölçek türü ile elde edilmesi nedeniyle verilerde standardize işlemi gerçekleştirilmemiştir. Kümelerin elde edilmesi tablo 3.3.'de verilmiştir. Tablo 3.3'deki elektrik tüketimi ortalamaları bakımından birbirine en yakın birimler 4 numaralı yerleşim merkezi olan Baskil ile 29 numaralı yerleşim merkezi olan Arıvan ilçeleri olduğu tespit edilmiştir. Bu iki birime ortalama benzerliği en fazla sağlayan birim 5 numaralı yerleşim merkezi olan Karakoçan ilçesi dahil edilerek ilk küme elde edilmiştir. Elektrik tüketimi verilere göre Ortalama bağlantı yöntemi ile elde edilen kümeler ve kümeler arasındaki bağlantılar Şekil 3.1' de dendogram grafiği ile belirtilmiştir.

Tablo 3.3. Gruplar arası ortalama bağlantı yöntemi ile birleştirme tablosu

Aşama	Birleştirilmiş Küme		Katsayılar	İlk Kümeleşmenin görüldüğü aşamalar		Sonraki Aşama
	Küme 1	Küme 2		Küme 1	Küme 2	
1	4	29	5981,000	0	0	19
2	15	18	9734,000	0	0	21
3	17	25	20602,000	0	0	8
4	19	20	48510,000	0	0	20
5	2	9	56586,000	0	0	9
6	1	13	61933,000	0	0	10
7	16	28	71133,000	0	0	20
8	17	35	97887,000	3	0	13
9	2	3	109299,000	5	0	14
10	1	24	159158,500	6	0	16
11	22	33	159760,000	0	0	18
12	6	37	174946,000	0	0	18
13	12	17	188158,333	0	8	24
14	2	27	220657,000	9	0	23
15	7	31	257430,000	0	0	27
16	1	32	262963,000	10	0	23
17	8	30	274110,000	0	0	26
18	6	22	308467,000	12	11	25
19	4	5	314147,500	1	0	30
20	16	19	353068,000	7	4	24
21	14	15	375870,000	0	2	22
22	14	34	507906,333	21	0	30
23	1	2	511029,750	16	14	28
24	12	16	567215,500	13	20	32
25	6	11	598921,500	18	0	29
26	8	39	642603,000	17	0	35
27	7	23	700820,000	15	0	34
28	1	21	784380,750	23	0	29
29	1	6	1074937,800	28	25	31
30	4	14	1115760,583	19	22	32
31	1	10	1860665,643	29	0	33
32	4	12	2001058,321	30	24	33
33	1	4	2122521,320	31	32	35
34	7	36	4319600,667	27	0	37
35	1	8	4784075,967	33	26	36
36	1	38	5495307,636	35	0	37
37	1	7	8561372,294	36	34	38
38	1	26	34726817,053	37	0	0

Dendrogram Grafiği



Şekil 3.1. Gruplar arası Ortalama Bağlantı Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi Bakımından Dendogram Grafiği

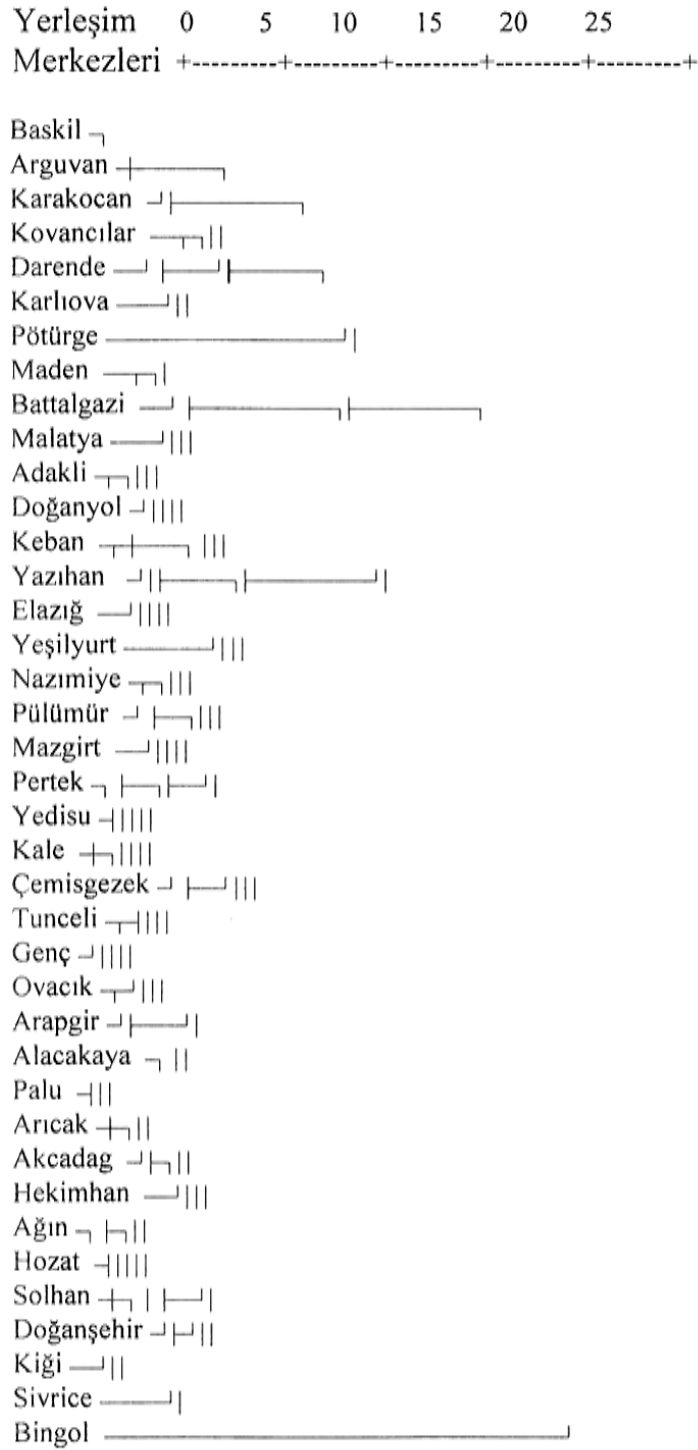
3.8. Grup içi Ortalama Bağlantı yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi

Yıllık hane elektrik tüketim ortalamalarına göre grup içi ortalama bağlantı yöntemi ile 5 kümeye ayırma işlemi gerçekleştirilmiştir. Kümeleme işlemi ikili kümelerden tek bir küme elde edilinceye kadar birleştirme işlemi sürdürülerek gerçekleştirilmiştir. Standardize işlemi veriler tek bir ölçek türü ile elde edilmesi nedeniyle uygulanmamıştır. Birimlerin birbirlerine olan uzaklığı karesel öklit uzaklığı ile elde edilmiştir. Kümeleme işlemine ilişkin bilgiler ve işlem aşamaları tablo 3.4’de verilmiştir. Tablo 3.4’deki elektrik tüketimi ortalamaları bakımından birbirine en yakın birimler 4 numaralı yerleşim merkezi olan Baskil ile 29 numaralı yerleşim merkezi olan Arguvan ilçeleri olduğu tespit edilmiştir. Bu iki birime ortalama benzerliği en fazla sağlayan birim 5 numaralı yerleşim merkezi olan Karakoçan ilçesi dahil edilerek ilk küme elde edilmiştir. Grup için ortalama bağlantı yöntemi ile elde edilen kümeler ve kümeler arasındaki bağlantılar Şekil 3.2’de dendogram grafiği ile belirtilmiştir.

Tablo 3.4. Grup içi Ortalama Bağlantı yöntemi ile elde edilen Birleştirme Tablosu

Aşama	Birleştirilmiş Kümeler		Katsayılar	İlk Kümeleşmenin Görüldüğü		
				Aşamalar		Sonraki Aşamalar
	Küme 1	Küme 2	Küme 1	Küme 2	Küme 1	Küme 2
1	4	29	5981,000	0	0	16
2	15	18	9734,000	0	0	17
3	17	25	20602,000	0	0	8
4	19	20	48510,000	0	0	18
5	2	9	56586,000	0	0	9
6	1	13	61933,000	0	0	10
7	16	28	71133,000	0	0	18
8	17	35	72125,333	3	0	11
9	2	3	91728,000	5	0	12
10	1	24	126750,000	6	0	15
11	12	17	130141,833	0	8	25
12	2	27	156192,500	9	0	22
13	22	33	159760,000	0	0	20
14	6	37	174946,000	0	0	20
15	1	32	194856,500	10	0	23
16	4	5	211425,333	1	0	33
17	14	15	253824,667	0	2	29
18	16	19	255319,167	7	4	25
19	7	31	257430,000	0	0	27
20	6	22	261429,000	14	13	24
21	8	30	274110,000	0	0	26
22	2	34	340411,600	12	0	28
23	1	21	380063,400	15	0	28
24	6	11	396426,000	20	0	31
25	12	16	406721,929	11	18	29
26	8	39	519772,000	21	0	35
27	7	23	553023,333	19	0	33
28	1	2	564159,689	23	22	30
29	12	14	736343,236	25	17	32
30	1	10	755739,818	28	0	32
31	6	38	846817,867	24	0	34
32	1	12	1110893,312	30	29	34
33	4	7	1236513,133	16	27	36
34	1	6	1689983,071	32	31	35
35	1	8	2235882,869	34	26	37
36	4	36	2560243,333	33	0	37
37	1	4	3562877,351	35	36	38
38	1	26	5161028,105	37	0	0

Birleştirilmiş Kümelerin ölçeklendirilmiş Uzaklıkları



Şekil 3.2. Grup içi Ortalama Bağlantı Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi Bakımından Dendrogram Grafiği

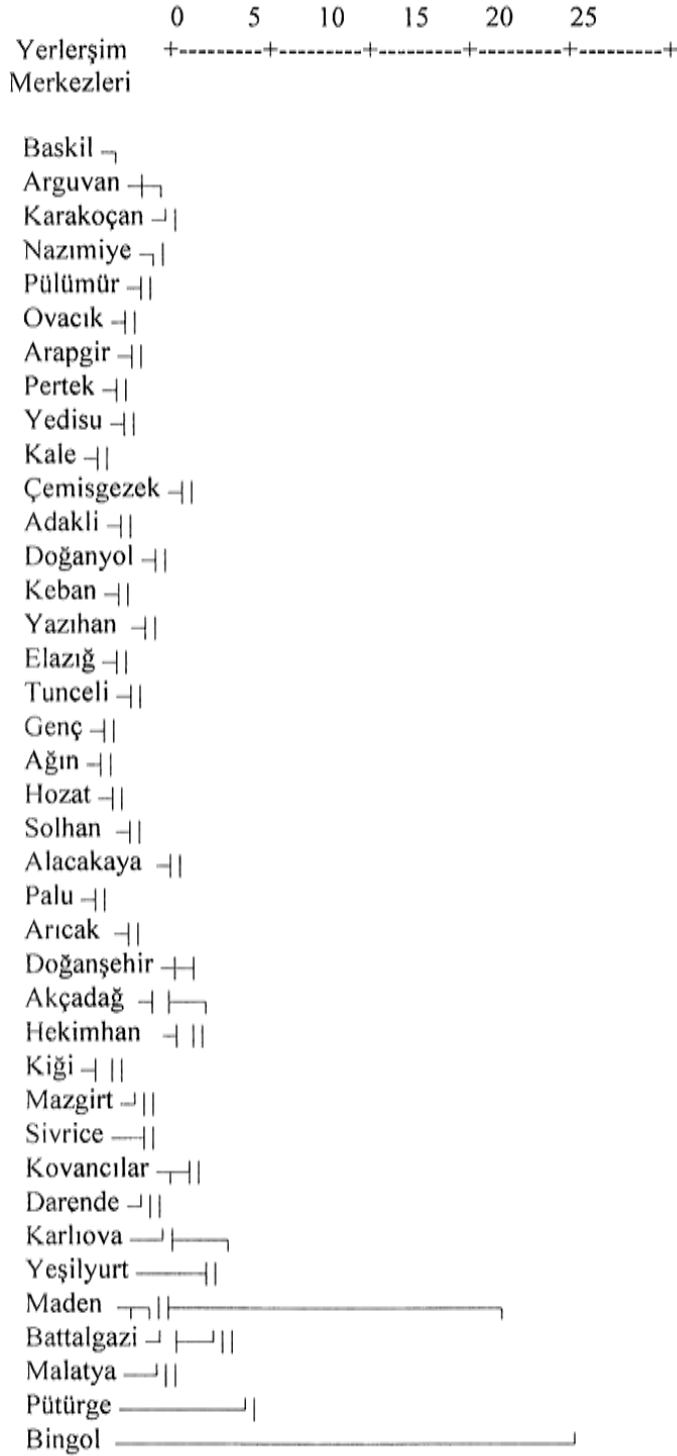
3.9. En Yakın Komşu yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi

Yıllık hane elektrik tüketim ortalamalarına göre grup en yakın komşu yöntemi ile 5 kümeye ayırma işlemi gerçekleştirilmiştir. Standardize işlemi veriler tek bir ölçek türü ile elde edilmesi nedeniyle uygulanmamıştır. Birimler arası uzaklık ölçümleri karesel Öklit uzaklığı ile belirlenmiştir. Kümeleme işlemine ilişkin bilgiler ve işlem aşamaları tablo 3.5'te verilmiştir. En yakın bağlantı yöntemi ile elde edilen kümeler ve kümeler arasındaki bağlantılar Şekil 3.3'de dendogram grafiği ile belirtilmiştir.

Tablo 3.5. En Yakın Komşu yöntemine ile elde edilen Birleştirme Tablosu

Aşama	Birleştirilmiş Kümeler		Katsayılar	İlk Kümeleşmenin Görüldüğü Aşamalar		Yeni aşama
	Küme 1	Küme 2		Küme 1	Küme 2	
1	4	29	5981,000	0	0	25
2	15	18	9734,000	0	0	26
3	17	25	20602,000	0	0	10
4	19	20	48510,000	0	0	20
5	2	9	56586,000	0	0	9
6	1	13	61933,000	0	0	8
7	16	28	71133,000	0	0	19
8	1	24	74765,000	6	0	11
9	2	3	91285,000	5	0	11
10	17	35	96621,000	3	0	14
11	1	2	120494,000	8	9	12
12	1	32	129394,000	11	0	17
13	22	33	159760,000	0	0	15
14	12	17	160930,000	0	10	19
15	6	22	164025,000	0	13	16
16	6	37	174946,000	15	0	18
17	1	27	193398,000	12	0	20
18	6	11	193821,000	16	0	22
19	12	16	201126,000	14	7	23
20	1	19	245496,000	17	4	22
21	7	31	257430,000	0	0	32
22	1	6	261520,000	20	18	23
23	1	12	267361,000	22	19	26
24	8	30	274110,000	0	0	30
25	4	5	292954,000	1	0	31
26	1	15	336381,000	23	2	27
27	1	34	339098,000	26	0	28
28	1	21	340058,000	27	0	29
29	1	14	360980,000	28	0	31
30	8	39	440630,000	24	0	36
31	1	4	592835,000	29	25	33
32	7	23	615837,000	21	0	34
33	1	10	709074,000	31	0	34
34	1	7	837378,000	33	32	35
35	1	38	1288038,000	34	0	36
36	1	8	1320746,000	35	30	37
37	1	36	2669013,000	36	0	38
38	1	26	10476158,000	37	0	0

Birleştirilmiş Kümelerin Ölçeklendirilmiş Uzaklıkları



Şekil 3.3. En yakın komşu Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi Bakımından Dendrogram Grafiği

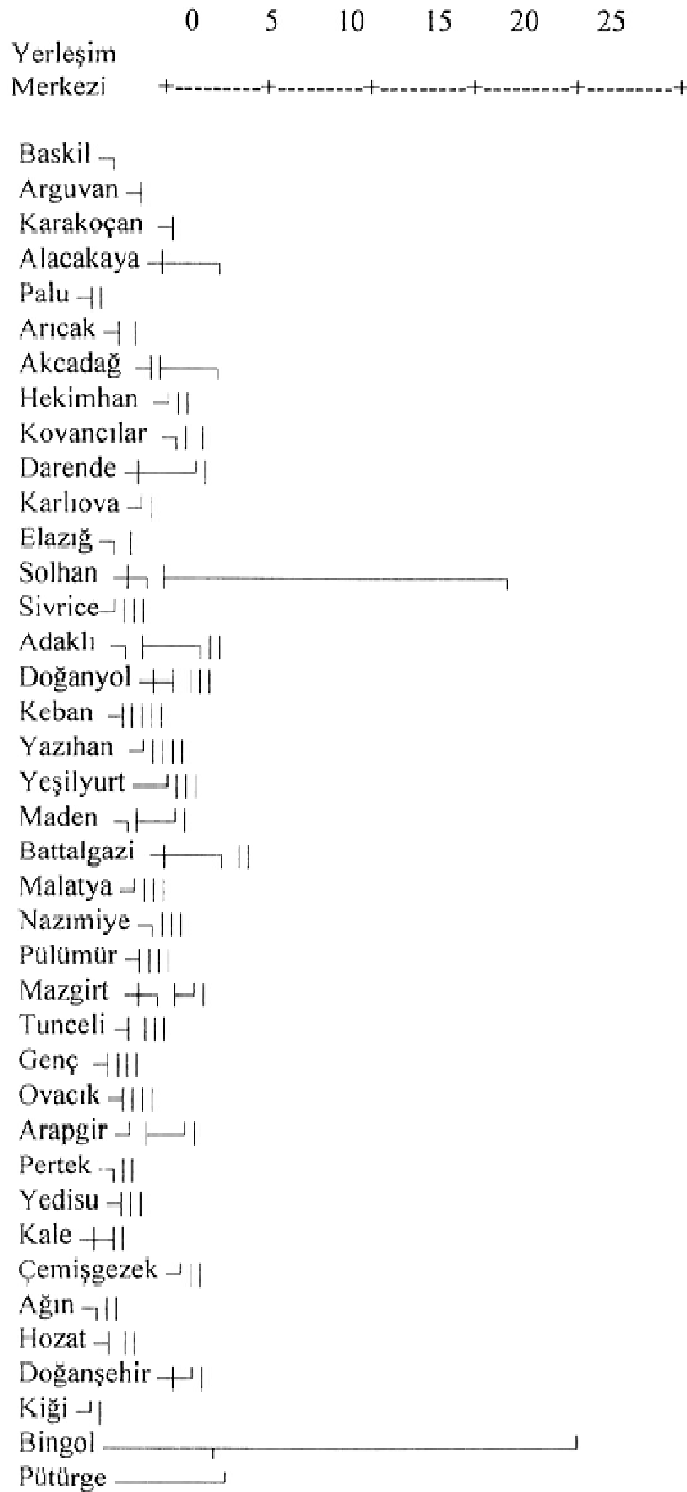
3.10. En Yakın Komşu yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi

Yıllık hane elektrik tüketim ortalamalarına göre grup en uzak komşu yöntemi kullanılarak kümeleme işlemi gerçekleştirilmiştir. Bu yöntemde kümeleme işlemi birim değerleri en yakın olan birimler birleşerek bir küme oluşturulmaktadır. Bu işlem adım adım tek bir küme elde edilinceye kadar devam etmektedir. Standardize işlemi verilerin tek bir değişken ile elde edilmesi nedeniyle uygulanmamıştır. Birimler arası uzaklık ölçümleri karesel Öklit uzaklığı ile belirlenmiştir. Kümeleme işlemine ilişkin bilgiler ve işlem aşamaları tablo 3.6’te verilmiştir. En uzak bağlantı yöntemi ile elde edilen kümeler ve kümeler arasındaki bağlantılar Şekil 3.4’ de dendogram grafiği ile belirtilmiştir.

Tablo 3.6. En uzak Komşu Yöntemi ile elde edilen Birleştirme Tablosu

Aşama	Birleştirilmiş küme		Katsayılar	İlk Kümelemenin Görüldüğü Aşama		Sonraki Aşama
	Küme 1	Küme 2		Küme 1	Küme 2	
1	4	29	5981,000	0	0	18
2	15	18	9734,000	0	0	19
3	17	25	20602,000	0	0	8
4	19	20	48510,000	0	0	20
5	2	9	56586,000	0	0	9
6	1	13	61933,000	0	0	12
7	16	28	71133,000	0	0	20
8	17	35	99153,000	3	0	13
9	2	3	127313,000	5	0	15
10	22	33	159760,000	0	0	21
11	6	37	174946,000	0	0	21
12	1	32	228653,000	6	0	22
13	12	17	236721,000	0	8	29
14	7	31	257430,000	0	0	24
15	2	27	266667,000	9	0	23
16	8	30	274110,000	0	0	25
17	11	24	314125,000	0	0	27
18	4	5	335341,000	1	0	28
19	14	15	376481,000	0	2	26
20	16	19	445515,000	7	4	26
21	6	22	495113,000	11	10	30
22	1	21	714834,000	12	0	29
23	2	34	764922,000	15	0	28
24	7	23	785803,000	14	0	33
25	8	39	844576,000	16	0	34
26	14	16	1071899,000	19	20	31
27	10	11	1109699,000	0	17	32
28	2	4	1982787,000	23	18	33
29	1	12	2041245,000	22	13	31
30	6	38	2472090,000	21	0	32
31	1	14	3000989,000	29	26	34
32	6	10	3356941,000	30	27	35
33	2	7	6016730,000	28	24	37
34	1	8	7958426,000	31	25	35
35	1	6	9889117,000	34	32	37
36	26	36	10476158,000	0	0	38
37	1	2	13830701,000	35	33	38
38	1	26	50017864,000	37	36	0

Birleştirilmiş kümelerin ölçeklendirilmiş uzaklıkları



Şekil 3.4. En uzak komşu Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi Bakımından Dendogram Grafiği

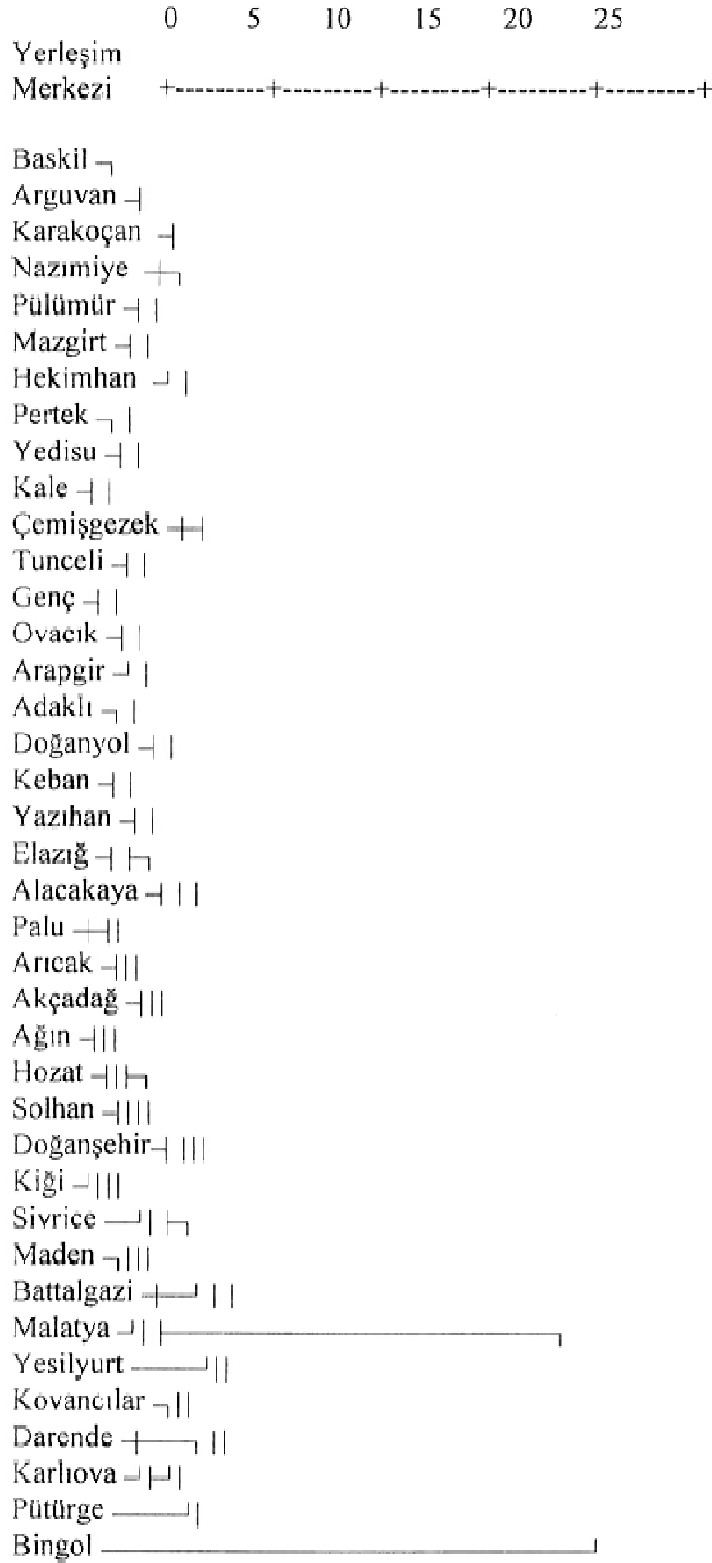
3.11. Merkezi Kümeleme Yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi

Yıllık hane elektrik tüketim ortalamalarına göre grup merkezi kümeleme yöntemi kullanılarak kümeleme işlemi gerçekleştirilmiştir. Standardize işlemi verilerin tek bir değişken ile elde edilmesi nedeniyle uygulanmamıştır. Birimler arası uzaklık ölçümleri karesel Öklit uzaklığı ile belirlenmiştir. Kümeleme işlemine ilişkin bilgiler ve işlem aşamaları tablo 3.7’ de verilmiştir. Merkezi kümeleme yöntemi ile elde edilen kümeler ve kümeler arasındaki bağlantılar Şekil 3.5’ de dendogram grafiği ile belirtilmiştir.

Tablo 3.7. Merkezi Kümeleme Yöntemi ile elde edilen Birleştirme Tablosu

Aşama	Birleştirilmiş küme		Katsayılar	İlk Kümelemenin Görüldüğü Aşama		
	Küme 1	Küme 2		Küme 1	Küme 2	Sonraki Aşama
1	4	29	5981,000	0	0	19
2	15	18	9734,000	0	0	21
3	17	25	20602,000	0	0	8
4	19	20	48510,000	0	0	20
5	2	9	56586,000	0	0	9
6	1	13	61933,000	0	0	10
7	16	28	71133,000	0	0	20
8	17	35	92736,500	3	0	12
9	2	3	95152,500	5	0	14
10	1	24	143675,250	6	0	15
11	22	33	159760,000	0	0	16
12	12	17	164116,556	0	8	23
13	6	37	174946,000	0	0	16
14	2	27	190081,000	9	0	22
15	1	32	220713,000	10	0	22
16	6	22	224790,500	13	11	25
17	7	31	257430,000	0	0	28
18	8	30	274110,000	0	0	26
19	4	5	312652,250	1	0	30
20	16	19	323157,250	7	4	23
21	14	15	373436,500	0	2	24
22	1	2	379386,375	15	14	27
23	12	16	422667,625	12	20	31
24	14	34	423298,111	21	0	30
25	6	11	500885,625	16	0	29
26	8	39	574075,500	18	0	34
27	1	21	623712,469	22	0	29
28	7	23	636462,500	17	0	35
29	1	6	711949,919	27	25	31
30	4	14	902460,910	19	24	32
31	1	12	1379880,904	29	23	32
32	1	4	1206857,923	31	30	33
33	1	10	1690558,313	32	0	34
34	1	8	3819410,890	33	26	36
35	7	36	4135259,556	28	0	37
36	1	38	4444441,592	34	0	37
37	1	7	6500922,877	36	35	38
38	1	26	32992258,342	37	0	0

Birleştirilmiş Kümelerin Ölçeklendirilmiş Uzaklıkları



Şekil 3.5. Merkezi Kümeleme Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi Bakımından Dendrogram Grafiği

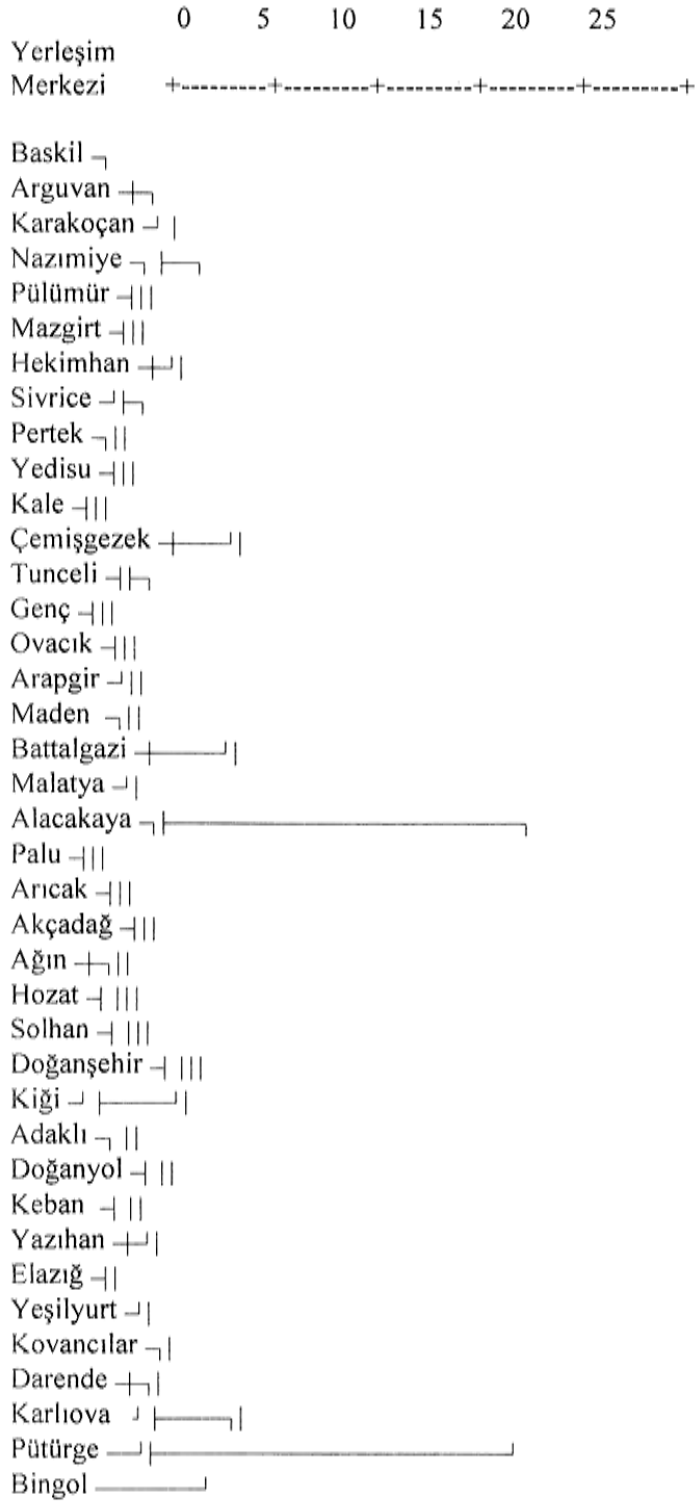
3.12. Ward's Yöntemi ile Elektrik Tüketimi Bakımından İl ve İlçe Merkezlerinin Kümelenmesi

Yıllık hane elektrik tüketim ortalamalarına göre grup Ward's yöntemi kullanılarak kümeleme işlemi gerçekleştirilmiştir. Standardize işlemi verilerin tek bir değişken ile elde edilmesi nedeniyle uygulanmamıştır. Birimler arası uzaklık ölçümleri karesel Öklit uzaklığı ile belirlenmiştir. Kümeleme işlemine ilişkin bilgiler ve işlem aşamaları tablo 3.8' de verilmiştir. Ward's kümeleme yöntemi ile elde edilen kümeler ve kümeler arasındaki bağlantılar Şekil 3.6' da dendogram grafiği ile belirtilmiştir.

Tablo 3.8. Ward's Kümeleme yöntemi ile elde edilen Birleştirme Tablosu

Aşama	Birleştirilmiş küme			İlk Kümelemenin Görüldüğü Aşama		Sonraki Aşama
	Küme 1	Küme 2	Katsayılar	Küme 1	Küme 2	
1	4	29	2990,500	0	0	18
2	15	18	7857,500	0	0	20
3	17	25	18158,500	0	0	8
4	19	20	42413,500	0	0	22
5	2	9	70706,500	0	0	9
6	1	13	101673,000	0	0	12
7	16	28	137239,500	0	0	22
8	17	35	199063,833	3	0	13
9	2	3	262498,833	5	0	16
10	22	33	342378,833	0	0	19
11	6	37	429851,833	0	0	19
12	1	24	525635,333	6	0	17
13	12	17	648722,750	0	8	27
14	7	31	777437,750	0	0	25
15	8	30	914492,750	0	0	23
16	2	27	1057053,500	9	0	28
17	1	32	1222588,250	12	0	26
18	4	5	1431023,083	1	0	31
19	6	22	1655813,583	11	10	24
20	14	15	1904771,250	0	2	21
21	14	34	2222244,833	20	0	29
22	16	19	2545402,083	7	4	27
23	8	39	2928119,083	15	0	35
24	6	11	3328827,583	19	0	30
25	7	23	3753135,917	14	0	32
26	1	21	4220977,967	17	0	28
27	12	16	5066313,217	13	22	34
28	1	2	5911655,000	26	16	33
29	10	14	6978159,950	0	21	31
30	6	38	8302352,617	24	0	33
31	4	10	10438649,333	18	29	34
32	7	36	13540094,000	25	0	36
33	1	6	17110796,667	28	30	37
34	4	12	23254305,542	31	27	35
35	4	8	31951109,930	34	23	37
36	7	26	41692431,930	32	0	38
37	1	4	52388092,176	33	35	38
38	1	7	98059534,000	37	36	0

Birleştirilmiş Kümelerin Ölçeklendirilmiş Tablosu



Şekil 3.6. Ward's Kümeleme Yöntemi ile Yerleşim Merkezlerinin Elektrik Tüketimi Bakımından Dendogram Grafiği

4. BULGULAR

4.1. Gruplar Arası Ortalama Bağlantı yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri

Çalışmada gruplar arası ortalama bağlantı yöntemi ile yerleşim merkezleri elektrik tüketimlerine göre 5 farklı kümeye ayrılmıştır. Kümeleme sonuçları tablo 4.1’de verilmiştir. Tablo 4.1’ e göre yerleşim merkezlerinin % 76,9 u küme 1’e atanmıştır. Yerleşim merkezlerinin sadece %2,7 küme 4’e ve yine %2,7’si küme 5’e atanmıştır. Yeşilyurt yerleşim merkezleri arasında en yüksek tüketime sahip olmamasına rağmen küme 5’e atanmıştır.

Tablo 4.1. Gruplar arası ortalama bağlantı yöntemi ile yerleşim merkezlerinin kümeleneşmesi

	Küme 1	Küme 2	Küme 3	Küme 4	Küme 5
Ağın	Pülümür	Kovancılar	Maden	Bingöl	Yeşilyurt
Alacakaya	Tunceli	Karlıova	Battalgazi		
Arıcak	Genç	Darende	Malatya		
Baskil	Kiğı	Pütürge			
Karakoçan	Adaklı				
Keban	Solhan				
Palu	Yedisu				
Sivrice	Akçadağ				
Elazığ	Arapgir				
Çemişgezek	Arguvan				
Hozat	Doğınşehir				
Mazgirt	Doğanyol				
Nazımiye	Hekimhan				
Ovacık	Kale				
Pertek	Yazihan				

4.2. Grup içi Ortalama Bağlantı yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri

Grup içi ortalama bağlantı yöntemi ile yerleşim merkezleri elektrik tüketimlerine göre 5 farklı kümeye ayrılmıştır. Kümeleme sonuçları tablo 4.2’ de verilmiştir. Tablo 4.2’ ye göre yerleşim merkezlerinin % 71,7’si küme 1’e atanmıştır. Yerleşim merkezlerinin sadece

%2,7 küme 4'e ve yine %2,7'si küme 5'e atanmıştır. Gruplar arası ortalama bağlantı yönteminde Küme 5' e atanan Yeşilyurt, grup içi ortalama kümeleme yönteminde küme 1'e atanmıştır. En yüksek elektrik tüketimine sahip yerleşim merkezi olan Pötürge küme 5'e atanmıştır.

Tablo 4.2. Grup içi ortalama bağlantı yöntemi ile yerleşim merkezlerinin kümelenmesi

	Küme 1	Küme 2	Küme 3	Küme 4	Küme 5
Ağın	Pülümür	Baskil	Malatya	Bingöl	Pötürge
Alacakaya	Tunceli	Karakoçan	Battalgazi		
Arıcak	Genç	Kovancılar	Maden		
Yeşilyurt	Kiğı	Karlıova			
Yazıhan	Adaklı	Arguvan			
Keban	Solhan	Darende			
Palu	Yedisu				
Sivrice	Akçadağ				
Elazığ	Arapgir				
Çemişgezek	Kale				
Hozat	Doğanşehir				
Mazgirt	Doğanyol				
Nazımiye	Hekimhan				
Ovacık					
Pertek					

4.3. En Yakın Komşu Yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri

En yakın komşu yöntemi ile yerleşim merkezleri elektrik tüketimlerine göre 5 farklı kümeye ayrılmıştır. Kümeleme sonuçları tablo 4.3' da verilmiştir. Tablo 4.3'a göre yerleşim merkezlerinin % 84,6'sı küme 1'e atanmıştır. Yerleşim merkezlerinin sadece %2,7 küme 4'e ve yine %2,7'si küme 5'e atanmıştır.

Tablo 4.3. En yakın Komşu Yöntemi ile yerleşim merkezlerinin kümelenmesi

	Küme 1	Küme 2	Küme 3	Küme 4	Küme 5
Ağın	Pülümür	Kovancılar	Maden	Bingöl	Yeşilyurt
Alacakaya	Tunceli	Karlıova	Battalgazi		
Arıcak	Genç	Arguvan	Malatya		
Darende	Kiğı				
Yazıhan	Adaklı				
Keban	Solhan				
Palu	Yedisu				
Sivrice	Akçadağ				
Elazığ	Arapgir				
Çemişgezek	Kale				
Hozat	Doğanşehir				
Mazgirt	Doğanyol				
Nazımiye	Hekimhan				
Ovacık	Baskil				
Pertek	Karakoçan				

4.4. En uzak Komşu Yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri

En uzak komşu yöntemi ile yerleşim merkezleri elektrik tüketimlerine göre 5 farklı kümeye ayrılmıştır. Kümeleme sonuçları tablo 4.4’ de verilmiştir. Tablo 4.4’ e göre yerleşim merkezlerinin % 46,2’si küme 1’e atanmıştır. Diğer kümeleme yöntemlerine göre daha dağınık bir atama yapan en uzak komşu yönteminde küme 2’ye atanan yerleşim merkezi oranı %28,2 olarak elde edilmiştir. Yerleşim merkezlerinin sadece %2,7 küme 4’e ve yine %2,7’si küme 5’e atanmıştır.

Tablo 4.4. En uzak Komşu Yöntemi ile yerleşim merkezlerinin kümelenmesi

	Küme 1	Küme 2	Küme 3	Küme 4	Küme 5
Ağın	Arapgir	Alacakaya	Keban	Bingöl	Pötürge
Malatya	Kale	Arıcak	Sivrice		
Maden	Doğanşehir	Baskil	Elazığ		
Battalgazi		Karakoçan	Adaklı		
Pülümür		Kovancılar	Solhan		
Tunceli		Palu	Doğanyol		
Genç		Karlıova	Yeşilyurt		
Kiğı		Akçadağ	Yazıhan		
Yedisu		Arguvan			
Çemişgezek		Darende			
Hozat		Hekimhan			
Mazgirt					
Nazımiye					
Ovacık					
Pertek					

4.5. Ward's Kümeleme Yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri

Ward's kümeleme yöntemi ile yerleşim merkezleri elektrik tüketimlerine göre 5 farklı kümeye ayrılmıştır. Kümeleme sonuçları tablo 4.5.'de verilmiştir. Tablo 4.5' e göre yerleşim merkezlerinin % 43,5'ü küme 2'ye atanmıştır. Küme 1'e atanan yerleşim merkezlerinin oranı %33,3 olarak belirlenmiştir. Yerleşim merkezlerinin sadece %7,7 küme 4'e ,%2,6'si küme 5'e atanmıştır.

Tablo 4.5. Ward's Kümeleme Yöntemi ile yerleşim merkezlerinin kümelenmesi

	Küme 1	Küme 2	Küme 3	Küme 4	Küme 5
Ağın	Pülümür	Alacakaya	Kovancılar	Malatya	Bingöl
Akçadağ	Tunceli	Arıcak	Karlıova	Maden	
Doğanyol	Genç	Baskil	Darende	Battalgazi	
Solhan	Karakoçan		Pötürge		
Kiğı	Çemişgezek				
Adaklı	Yedisu				
Hozat	Arapgir				
Yeşilyurt	Arguvan				
Yazıhan	Kale				
Doğanşehir	Hekimhan				
Keban	Sivrice				
Palu	Mazgirt				
Elazığ	Nazımiye				
	Ovacık				
	Pertek				

4.6. Merkezi Bağlantı Yöntemi ile Elektrik Tüketimi Bakımından Yerleşim Merkezleri Kümeleri

En yakın komşu yöntemi ile yerleşim merkezleri elektrik tüketimlerine göre 5 farklı kümeye ayrılmıştır. Kümeleme sonuçları tablo 4.6' da verilmiştir. Tablo 4.6' ya göre yerleşim merkezlerinin % 84,6'sı küme 1'e atanmıştır. Yerleşim merkezlerinin sadece %2,7 küme 4'e ve yine %2,7'si küme 5'e atanmıştır.

Tablo 4.6. Merkezi Bağlantı Yöntemi ile yerleşim merkezlerinin kümelenmesi

	Küme 1	Küme 2	Küme 3	Küme 4	Küme 5
Ağın	Pülümür	Malatya	Kovancılar	Bingöl	Pötürge
Akçadağ	Tunceli	Maden	Karlıova		Yeşilyurt
Doğanyol	Genç	Battalgazi	Darende		
Solhan	Karakoçan				
Kiğı	Çemişgezek				
Adaklı	Yedisu				
Hozat	Arapgir				
Alacakaya	Arguvan				
Yazıhan	Kale				
Doğanşehir	Hekimhan				
Keban	Sivrice				
Palu	Mazgirt				
Elazığ	Nazımiye				
Arıcak	Ovacık				
Baskil	Pertek				

4.7. Kümeleme Analizi Yöntemlerinin Etkinlikleri İncelenmesi

2010 yılına ait Bingöl, Elazığ, Malatya ve Tunceli il merkezleri ve bu illere ait bazı ilçe merkezlerinin hanelerden rasgele elde edilen yıllık elektrik tüketim değerleri ortalamaları kümeleme yöntemleri ile sınıflandırılmıştır. Uygulanan kümeleme yöntemlerinin doğru sınıflama başarısı ise diskriminant analizi ile test edilmiştir. Diskriminant analizi doğru sınıflandırma oranlarını vermektedir. Bu şekilde kümeleme yöntemlerinin doğru sınıflandırma oranlarını elde etmek mümkün olmaktadır. Tablo 4,7'te diskriminant analizine ilişkin kümeleme yöntemlerinin doğru sınıflandırma oranları verilmiştir. Tablo 4.7' deki sonuçlara göre en başarılı kümeleme yöntemi bu çalışma için % 58,6 doğru sınıflama oranı ile en yakın komşu yöntemi olarak elde edilmiştir.

Tablo 4.7. Doğru Sınıflandırma Oranları

Uygulanan Kümeleme Yöntemi	Doğru Sınıflandırma Oranı
Gruplar arası ortalama bağlantı yöntemi	% 49,9
Grup içi ortalama bağlantı yöntemi	% 42,0
En yakın komşu yöntemi	% 58,6
En uzak komşu yöntemi	% 32,5
Merkezi bağlantı kümeleme yöntemi	% 55,4
Ward's Kümeleme yöntemi	% 28,4

5. SONUÇLAR VE TARTIŞMA

Yapılan çalışma da kümeleme yöntemleri ve diskriminant analizi üzerine durulmuş elektrik tüketimine ilişkin veriler kullanılarak uygulama yapılmıştır.

Uygulamada Bingöl, Elazığ, Malatya ve Tunceli il merkezleri ve bu illere bağlı ilçe merkezlerindeki rasgele seçilen hanelere ilişkin 2010 yılına ait yıllık ortalama elektrik tüketim değerleri incelenmiştir. Örneklemi il ve ilçe merkezlerinden seçilen 345 hane oluşturmaktadır.

Gruplar arası ortalama bağlantı yöntemi ile elektrik tüketimi ortalamaları bakımından yerleşim merkezleri 5 farklı sınıfa ayrılmıştır. Kümeleme sonuçları ve kümelenmenin elde edilmesine ilişkin dendogram grafiği elde edilmiştir.

Grupiçi ortalama bağlantı yöntemi ile elektrik tüketimi ortalamaları bakımından yerleşim merkezleri 5 farklı sınıfa ayrılmıştır. Kümeleme sonuçları ve kümelenmenin elde edilmesine ilişkin dendogram grafiği elde edilmiştir.

En yakın komşu yöntemi ile elektrik tüketimi ortalamaları bakımından yerleşim merkezleri 5 farklı sınıfa ayrılmıştır. Kümeleme sonuçları ve kümelenmenin elde edilmesine ilişkin dendogram grafiği elde edilmiştir.

En uzak komşu yöntemi ile elektrik tüketimi ortalamaları bakımından yerleşim merkezleri 5 farklı sınıfa ayrılmıştır. Kümeleme sonuçları ve kümelenmenin elde edilmesine ilişkin dendogram grafiği elde edilmiştir.

Merkezi bağlantı yöntemi ile elektrik tüketimi ortalamaları bakımından yerleşim merkezleri 5 farklı sınıfa ayrılmıştır. Kümeleme sonuçları ve kümelenmenin elde edilmesine ilişkin dendogram grafiği elde edilmiştir.

Ward's kümeleme yöntemi ile elektrik tüketimi ortalamaları bakımından yerleşim merkezleri 5 farklı sınıfa ayrılmıştır. Kümeleme sonuçları ve kümelenmenin elde edilmesine ilişkin dendogram grafiği elde edilmiştir.

Kümeleme yöntemlerinin başarısını test etmek amacıyla kümeleme yöntemleri ile elde edilen kümeleme sonuçlarına ilişkin diskriminant analizi ile doğru sınıflandırma oranları elde edilmiştir. En başarılı kümeleme yöntemi %58,6 ile en yakın komşu yöntemi olarak elde edilmiştir.

En yakın komşu yöntemine göre yerleşim merkezlerinin %84,6'sının küme 1'e atandığı küme 3, küme 4 ve küme 5'e atanma oranlarının sadece %2.6 da kaldığı gözlemlenmiştir.

Küme 5'e atanan yerleşim merkezi Malatya ilinin Yeşilyurt ilçesi olmuştur. Bu ilçe yüksek elektrik tüketimine sahip ilçelerden biridir. En yüksek elektrik tüketimi ortalamasına sahip ilçe olan Pötürge küme 4'e atanmıştır.

Kümeleme analizi yöntemlerinin elektrik tüketimini sınıflamada başarılı bir yöntem olduğu gözlemlenmektedir.

Elektrik tüketim ortalamaları incelendiğinde il merkezlerinin ilçelere nazaran gelen olarak daha yüksek tüketime sahip olduğu gözlemlenmiştir. Bu farklılığın sebebi yerleşim merkezlerindeki elektrik tüketimi alışkanlıkları sebep gösterilebilir. Örneğin il merkezlerindeki aileler ısınma veya soğutma gibi iklimsel faktörler karşısında doğrudan teknolojik aletlerden faydalanırken ilçe merkezlerinde teknolojik aletlerin kullanım alışkanlığı daha az olduğu düşünülebilir.

6. ÖNERİLER

Kümeleme yöntemleri ile elektrik tüketiminde başarı ile uygulanabildiği ve uygun sonuçlar alınabildiği görülmüştür. Kümeleme analizinin sınıflandırma işlemleri için farklı çalışma sahalarında da uygulanabileceği düşünülmüştür.

Yıllık ortalama elektrik tüketimi değerleri bakımından kümeleme işlemi gerçekleştirilmiştir. Bu tüketim değerleri ile ilişkileri ortaya çıkarmak adına farklı değişkenlerde eklenerek tüketime neden olan faktörler incelenebilir.

İl merkezlerindeki elektrik tüketimi ile ilçe merkezlerindeki elektrik tüketimi ortalamaları birbirinden farklılık göstermektedir. Bu farklılığın nedenleri araştırılabilir.

Araştırmacılar daha fazla örnekleme ulaşarak ve çalışma alanını genişletebilir böylece daha sağlıklı ve hassas sonuçlar elde edilebilir.

Araştırmacılar farklı yöntemler ile yerleşim merkezlerini elektrik tüketimine göre sınıflandırabilir ve sonuçların tutarlılığını ve güvenilirliğini inceleyerek sonuçları karşılaştırabilir.

KAYNAKLAR

- [1] **Aldenderfer, M.S., Blashfield R.K.**, 1984, Cluster Analysis, Sage Publications
- [2] **Hair, J. F., Anderson, R. E., Tatham, R. L., Tatham, R. L. & Black, W. C.**, 1995, Multivariate Data Analysis with Readings. Prentice-Hall.
- [3] **Sokal, R., Sneath, P.** (1963) Principles of Numerical Taxonomy, San Francisco W.H. Freeman
- [4] **Tversky, A.**, 1977, Features of similarity, Psychological Review 84: 327-352.
- [5] **Jardine, N., Sibson R.**, 1971, Mathematical Taxonomy. New York, John Wiley.
- [6] **Clifford, H., Stephenson W.**, 1975 An Introduction to Numerical Taxonomy. New York, Academic Press.
- [7] **Çokluk, Ö., Şekercioğlu G., Büyüköztürk, Ş.**, 2010, Sosyal Bilimler için çok Değişkenli İstatistik Spss ve Lisrel Uygulamaları, Pegem Akademi, Ankara.
- [8] **Jensen, R.E.**, 1971, A Cluster analysis study of financial performance of selected, Business firms, Accounting rev. 46(1):36-56
- [9] **Selanik, M.**, 2007, “Türk Tarımının Avrupa Birliği İçindeki Yerinin Kümeleme Analizi İle Belirlenmesi”, Yüksek Lisans Tezi, Ankara
- [10] **Everitt, B.** 1974. Cluster Analysis. Heinmann.London,
- [11] **Tatlıdil, H.**,2002, Uygulamalı Çok Değişkenli İstatistiksel Analiz, Akademi Matbaası, Ankara.
- [12] **Everitt B.S., Landau S.,Leese M., Stahl D.**, 2011. Cluster Analysis 5th Edition, Wiley
- [13] **Fielding A.H.**, 2007, Cluster and Classssification Techniques fort he biosciences, Cambrigde universty press.

EKLER

EK 1. Elektrik Tüketim Ortalamaları

Sıra	Merkez	Tüketim (kw)	Merkez	Tüketim (kw)	Merkez	Tüketim (kw)	Merkez	Tüketim (kw)
1	Bingöl	2010	Elazığ	1925,6	Malatya	2117	Tunceli	1291,4
2	Adaklı	1380,7	Alacakaya	1444,5	Akçadağ	1594,5	Çemişgezek	573,3
3	Genç	1196,3	Arıcak	1290,2	Arapgir	1625	Hozat	1374,3
4	Karlıova	1727,4	Ağın	1198,7	Arguvan	1454,7	Mazgirt	1317,6
5	Kiğı	933,5	Baskil	1120,3	Battalgazi	1529,6	Nazımiye	1388,6
6	Solhan	1584,3	Karakoçan	1734,7	Darende	1868,6	Ovacık	1070,3
7	Yedisu	1064,5	Keban	1595,6	Doğanyol	1630	Pertek	802
8			Kovancılar	1427,9	Doğanşehir	1554	Pülümür	1441,3
9			Maden	1519,2	Hekimhan	1340		
10			Palu	1825,8	Kale	1154,7		
11			Sivrice	1782,7	Pötürge	2181,9		
12					Yazıhan	1760		
13					Yeşilyurt	1860		

ÖZGEÇMİŞ

1987 yılında Elazığ'da doğan Huri ALKAN, orta öğrenimini Elazığ Balakgazi Lisesi'nde tamamlamış, 2005-2009 yılları arasında Dumlupınar Üniversitesi İktisadi İdari Bilimler Fakültesi İşletme Bölümü'nde tamamlamıştır. 2009 yılında Fırat Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilimdalı İstatistiksel Bilgi Sistemleri Bilim Dalı'nda yüksek lisans eğitimine başlamış olup halen eğitimini sürdürmektedir.