

## ORGANİZE SANAYİ BÖLGELERİNİN HİYERARŞİK-K-ORTALAMALAR YÖNTEMİ İLE ANALİZİ

Arş.Gör. Ayşenur USLU  
Başkent Üniversitesi  
aysenur@baskent.edu.tr

Yrd.Doç.Dr. Cihan ÇETİNKAYA  
Gaziantep Üniversitesi  
cihancetinkaya@gantep.edu.tr

Yrd.Doç.Dr. Eren ÖZCEYLAN  
Gaziantep Üniversitesi  
erenozceylan@gmail.com

Doç.Dr. Selçuk Kürşat İŞLEYEN  
Gazi Üniversitesi  
isleyens@gazi.edu.tr

**ÖZ:** Organize Sanayi Bölgeleri (OSB), sanayinin etkinliğinin artırılması ve kentlerde düzenli yerleşimin sağlanması amacıyla sanayi tesislerinin bir araya toplanarak tesislerin ulaşım, enerji, su, Ar-Ge merkezi vb. gereksinimlerine ilişkin kolaylıklar sağlayan ve sanayinin çevreye olan olumsuz etkilerini en aza indirmek amacıyla çevre yönetim politikalarını uygulayan sanayi merkezleridir. OSB'ler sanayinin geliştirilmesinde ve kalkınmada önemli bir rol oynamaktadır. Bu çalışmada, ülkemizde bulunan 298 adet OSB içerisinde faal durumda olan ve 50'den fazla işletme bulunduran 89 OSB dikkate alınarak hiyerarşik-k-ortalamar yöntemi ile bir kümeleme analizi yapılmıştır. Analizler OSB'leri farklı özelliklerine göre sınıflandırabilmek amacıyla "genel bilgiler", "çevre yönetim bilgileri", "enerji altyapı ve kullanım bilgileri" ile "konum bilgileri" olmak üzere toplam 4 başlıkta 28 değişken dikkate alınarak gerçekleştirilmiştir. Analizler sonucunda benzer ve farklı özellikler gösteren OSB'ler belirlenerek sonuçlar tartışılmıştır.

**Anahtar kelimeler:** Organize sanayi bölgesi, Hiyerarşik-k-ortalamar, Kümeleme analizi.

## ANALYSIS OF ORGANIZED INDUSTRIAL ZONES USING HIERARCHICAL-K-MEANS METHOD

**ABSTRACT:** Organized Industrial Zones (OIZ) are the industrial centers in which industrial facilities are gathered together to increase the efficiency of the industry and to provide regular settlement in the cities. In addition they provide the facilities with transportation capabilities, energy & water support and R&D centers for their needs while implementing environmental management policies in order to minimize the negative effects of the companies on the environment. OIZs play an important role in the development of the industry. In this study, a cluster analysis by hierarchical-k-means method is performed considering 89 OIZs which are active (operational) and that has more than 50 enterprises among 298 OIZs of our country. In order to classify the OIZs according to their different characteristics, the analyzes are carried out taking into account 28 variables in total 4 topics; "general information", "environmental management information", "energy infrastructure and usage information" and "location information". Finally, the OIZs showing similar and different characteristics are determined and the results are discussed.

**Keywords:** Organized industrial zones, Hierarchical-k-means, Cluster analysis.

## 1. GİRİŞ

Türk sanayisinin gelişiminde organize sanayi bölgeleri (OSB) 1960'lı yıllardan itibaren önemli bir role sahip olmuştur. OSB'ler genel olarak; yeni yatırımların yapılması, istihdam sağlanması, bölgesel kalkınma gibi ekonomik önemi yanında, planlı ve düzenli sanayileşme ile kentleşme ve çevre politikaları açısından da etkin bir politika uygulama aracı durumundadır (Çetin ve Kara, 2008).

Önceki yıllarda “sanayileşme” temel amaçken, artık günümüzde “rekabet gücünü” artırmak uygulanan politikaların temel amacı haline gelmiştir. Söz konusu rekabet gücünün artırılabilmesinin yollarından biri de benzer özellik gösteren OSB'lerin ortaklaşa hareket etmelerinden geçmektedir (Çağlar, 2006). İşte bu noktada ortak özellikler gösteren OSB'lerin tespit edilmesinde – özellikle de politika geliştiriciler için – bilimsel tekniklerin kullanılması en sağlıklı sonucu verecektir. Bu tekniklerden birisi de son yıllarda çok fazla kullanım alanına sahip olan kümeleme analizidir.

Kümeleme analizi grupları kesin olarak bilinmeyen birimleri, değişkenleri birbiriyle benzer alt kümelere (grup, sınıf) ayırmaya yardımcı olan çok değişkenli istatistiksel analiz yöntemlerinden biridir (Ceylan, 2013). Kümeleme analizinin temel amacı birimleri sahip oldukları karakteristik özellikleri temel alarak gruplandırmaktır. Bu yöntem özellikle bilim ve iş alanında birçok durumda uygulanabilen, en kolay yorumlanabilen ve en etkili olan yöntem olma özelliğini taşır (Giray, 2016). Bu nedenle hemen hemen tüm bilim alanlarında bu yöntemden yararlanılmaktadır.

Ancak literatürde çok sayıda kümeleme analizi tekniğinin yer alması, bazı tekniklerin avantaj ve dezavantajlara sahip olmasına neden olmuştur. Bu sebeple, mevcut kümeleme yöntemlerinin dezavantajlarını ortadan kaldırarak kümeleme algoritmasının etkinliğini arttırmak amacıyla birden fazla kümeleme yönteminin birlikte kullanıldığı hibrid yaklaşımlar da bulunmaktadır (Ayдын ve Seven, 2015).

Bu çalışmada, ülkemizde bulunan 298 adet OSB içerisinde faal durumda olan ve 50'den fazla işletme bulunduran 89 OSB dikkate alınarak bir kümeleme analizi yapılmıştır. Söz konusu analiz için 56 farklı değişken belirlenmiş, bunlardan verisi temin edilen 28 değişken (OSB'deki firma sayısı, doğalgaz tüketimi, ulaşım yollarına yakınlık vb.) analize dahil edilmiştir. Değişkenler, OSB'leri farklı özelliklerine göre sınıflandırmak amacıyla “genel bilgiler”, “çevre yönetim bilgileri”, “enerji altyapı ve kullanım bilgileri” ile “konum bilgileri” olmak üzere toplam 4 ana başlıkta irdelenmiştir. Daha sonra, hiyerarşik kümeleme yöntemlerinden Ward's metodu ile hiyerarşik olmayan yöntemlerden k-ortalamlar metodu ardışık olarak kullanılmış ve hibrid bir yaklaşım sunulmuştur. Bu yaklaşımda hiyerarşik kümeleme ile elde edilen küme sayısı ve küme merkezleri, k-ortalamlar yöntemi için başlangıç değerlerini oluşturmaktadır.

Giriş bölümünün ardından, ikinci bölümde konu ve yönetime ilişkin literatür taramasına yer verilmiştir. Üçüncü bölümde, analizde kullanılan değişkenler ve uygulanan hibrid yaklaşım detayları ile beraber sunulmuştur. Dördüncü bölümde, elde edilen küme sonuçları karşılaştırmalı olarak verilirken, son bölümde ise sonuç ve öneriler ile çalışma sonlandırılmıştır.

## 2. LİTERATÜR İNCELEMESİ

Veri madenciliğinde sıklıkla kullanılan yöntemlerden olan “K-ortalamlar” algoritması ilk olarak 1955 yılında ortaya atılmıştır ve aynı alanda birçok algoritma olmasına rağmen hala efektif olarak kullanılmaktadır (Jain, 2010). Bu çalışmada, hiyerarşik k-ortalamlar tekniği kullanılmıştır. Fakat literatür incelemesinde, hem hiyerarşik hem de hiyerarşik olmayan k-ortalamlar tekniği kullanılan çalışmalara yer verilmiştir.

Mingoti ve Lima (2006), çalışmalarında “Kendiliğinden Örgütlenme Haritası, Yapay Sinir Ağları, Bulanık C-Ortalamlar ve K-Ortalamlar” gibi hiyerarşik ve hiyerarşik olmayan bir dizi kümeleme yöntemi arasında bir performans karşılaştırma yoluna gitmiştir. 2530 adet simüle edilmiş veri kullanılan bu çalışmada bulanık c-ortalamlar metodu çok daha dengeli ve güvenilir sonuçlar vermiş, kendiliğinden örgütlenme haritası kötü sonuçlar vermiştir. Hiyerarşik kümeleme ve k-ortalamlar yöntemleri ise birbirine yakın ve ortalamanın üzerinde sonuçlar vermiştir.

Arai ve Barakbah (2007), k-ortalamlar yönteminde kümelemeler genellikle yerel optimumda takıldığı için bu yöntemde özgü başlangıç merkezlerini optimize etmeye çalışan bir algoritma önermiştir. Hiyerarşik k-ortalamlar yöntemi; k-ortalamlar yönteminin hızından, hiyerarşik metodun ise hassasiyetinden faydalandığı için, rastgele normal dağılım verileri ve bir takım gerçek hayat verileri üzerinde test edildiğinde diğer metotlara göre (Merkezi bağlantı, bulanık c-ortalamlar vb.) daha sağlıklı sonuçlar vermiştir.

Çalışkan ve Soğukpınar (2008) nüfus tespiti için kullanılan kümeleme ve sınıflandırma yöntemlerinin eksik ve güçlü yönleri inceleyerek; kümelemeyi ve sınıflandırmayı, denetimli ve denimsiz öğrenimi, k-ortalamlar ve k en yakın komşu yöntemlerini bir arada kullanan hibrit bir yöntem geliştirmiştir. K-ortalamlar ve k en yakın komşu yöntemleri ile ayrı ayrı alınan sonuçların iyileştirilmesinin amaçlandığı çalışmada; karakteristik özelliklerine göre ayrılan her alt küme için ayrı k ve eşik değerler belirlenerek kümelere özgü değerler ile daha esnek bir yapı oluşturulmuştur. Geliştirilen yeni yöntemle yapılan testlerde, nüfus tespitinde doğruluk oranının arttığı görülmüştür.

Niknam ve Amiri (2010) doğrusal olmayan parçalı bölme (kümeleme) problemini ele almış ve bu problem için hibrit bir yaklaşım geliştirmiştir. Bu hibrit yaklaşımda; Bulanık Adaptif Parçacık Sürü Algoritması, Karınca Koloni Algoritması ve K-Ortalamlar Algoritmaları birlikte kullanılmıştır. Geliştirilen algoritmanın performansı, test veri setleri üzerinde defaten denenmiştir. Sonuçlar incelendiğinde, geliştirilen bu hibrit algoritmanın parçalı bölme problemi üzerinde diğer tekniklere nazaran daha iyi sonuçlar verdiği belirlenmiştir.

Fırat ve diğerleri (2012), k-ortalamlar yöntemi ile yıllık toplam yağışları sınıflandırmaya ve homojen bölgeleri belirlemeye çalışmıştır. Kümeleme analizi için Devlet Meteoroloji İşleri (DMİ) tarafından işletilen 188 yağış gözlem istasyonuna ait yıllık yağışlar ve enlem, boylam ile yükselti verileri kullanmışlardır. Ayrıca, kümeleme ile belirlenen bölgelerin homojenliğini test etmek için bölgesel homojenlik testi ile Ward yöntemi kullanılmıştır. Her iki kümeleme yöntemi ile belirlenen bölgeler için uygulanan bölgesel homojenlik testi sonuçlarına göre, yağış serilerinin sınıflandırılmasında K-Ortalamlar yöntemi kullandığı takdirde sadece 1 bölge “heterojen” olarak değerlendirilirken, Ward yönteminde ise 2 bölge “heterojen” olarak tespit edilmiştir. Bu değerlendirmelere göre, yıllık toplam yağışların sınıflandırılmasında K-Ortalamlar yöntemi ile elde edilen sonuçların kabul edilebilir seviyede olduğunu belirtmişlerdir.

Cengiz ve Öztürk (2012) Türkiye’deki illerin eğitim düzeyi benzerliklerini kümeleme analizi kullanarak belirlemeye çalışmıştır. Çalışmada; okuma-yazma bilmeyen, okuma-yazma bilen fakat bir okul bitirmeyen, ilköğretim mezunu, ortaöğretim mezunu, yükseköğretim mezunu, yüksek lisans mezunu vb. gibi birçok eğitim tabanlı oran kullanılmıştır.

Velmurugan (2014) çalışmasında veri madenciliğinde en sık kullanılan yöntemlerden olan k-ortalamlar ile bulanık c-ortalamlar tekniklerini telekomünikasyon verileri üzerinde kullanmak suretiyle bir karşılaştırma gerçekleştirmiştir. Analiz sonuçlarına göre bulanık c-ortalamlar yöntemi, k-ortalamlar yöntemine oranla biraz daha kısa zamanda sonuçlar üretmektedir.

Aydın ve Seven (2015), Türkiye'deki İl Nüfus ve Vatandaşlık Müdürlüklerini, iş yoğunluklarına göre hibrid hiyerarşik k-ortalamlar kümeleme analizi ile sınıflandırmıştır. Küme sayısına karar verirken silhouette endeksinden yararlanılmıştır. Analiz sonucunda, benzer iş yoğunluğuna sahip illerden oluşan altı farklı küme yapısı ortaya çıkarılmıştır. Elde edilen küme yapılarının geçerliliği ilgili istatistik testler yardımıyla da desteklenmiştir.

Literatürde K-Ortalamlar yönteminin kullanıldığı ek çalışmalar için; (Na vd., 2010), (Sarıman, 2011), (Celebi vd., 2013), (Bilen vd., 2014), (Silgu ve Çelikoğlu, 2014), (Yürük ve Erdoğan, 2015) çalışmaları incelenebilir.

Literatürde OSB'ler ile ilgili farklı amaç ve analiz yöntemlerini içeren bazı çalışmalar mevcuttur.

Zeng ve diğerleri (2008), Çin'de bulunan 25 imalat sektörünü hiyerarşik kümeleme yöntemi ile 4 kümeye ayırarak endüstriyel sürdürülebilirlik açısından değerlendirmiştir.

Çetin ve Kara (2008), OSB'lerin bölgesel kalkınmadaki etkilerini değerlendirmiş ve Isparta Süleyman Demirel OSB üzerinde bir uygulama gerçekleştirmiştir.

Bulut (2016), yedi çeyreklik dönem (2015 1.çeyrek - 2016 3. çeyrek) için OSB'lerin performansını doluluk oranları, istihdam sayısı, üretimdeki firma sayısı, sanayi parsel alanı ve üretime geçilen parsel alanı olmak üzere 5 kritere göre çok kriterli karar verme yöntemlerinden TOPSIS (ideal çözüme benzerliğine göre tercih sıralama tekniği) yöntemi ile analiz etmiştir.

Özden (2016), ülkemizdeki OSB'lerin mevcut durumunun saptanması amacıyla bir değerlendirme yapmış ve OSB'lerin kalkınma işlevini yerine getirmedeki etkinliğini tartışmıştır.

### 3. MATERYAL VE YÖNTEM

Bu bölümde araştırmanın amacı, kapsamı ve araştırmada kullanılan analiz yöntemi detaylı bir şekilde açıklanmıştır.

#### 3.1. Araştırmanın Amacı

Bu çalışma, OSB'lerin belirlenen 4 başlık (genel bilgiler, çevre yönetim bilgileri, enerji altyapı ve kullanım bilgileri, konum bilgileri) açısından benzerliklerini tespit etmek maksadıyla gerçekleştirilmiştir. OSB'lerin benzer yönlerinin tespiti, işletmelerin verimliliğinin artırılması ve sürdürülebilir bir yapı oluşturmaları açısından önemlidir. Özellikleri benzer firmalar arası ilişkilerin ve iletişimin oluşması yatırımların artmasına katkı sağlamakla birlikte, devlet yatırımları ile planlamaların da bu doğrultuda ilerlemesi daha planlı bir yapıya kavuşma açısından önemlidir. Kümelerin yapacakları güç birliktelikleri, ekonomik ve toplumsal gelişmeyi de sağlayacaktır. Bu bağlamda ele alınan bu çalışma hem ekonomik hem politik açıdan önem taşımaktadır.

#### 3.2. Araştırmanın Kapsamı

Çalışmada, ülkemizde bulunan 298 adet OSB arasından faal durumda olan ve 50'den fazla işletme bulduran 89 OSB dikkate alınmıştır (Tablo 1). Çalışma kapsamında incelenen 89 OSB'nin illere göre dağılımı Şekil 1'de verilmiştir. Çalışmada kullanılan veriler Bilim Sanayi ve Teknoloji Bakanlığı'nın OSB bilgi sitesi (URL1) ve OSB'lerin kendilerine ait web sitelerinden elde edilmiştir.



Şekil 1. İncelenen OSB'lerin illere göre dağılımı

Tablo 1. Çalışma kapsamında incelenen OSB'ler

No	OSB Adı	No	OSB Adı	No	OSB Adı
1	Adana Hacı Sabancı OSB	31	Aydın Astim OSB	61	Kırklareli OSB
2	Afyonkarahisar OSB	32	Bandırma OSB	62	Kırşehir OSB
3	Ostim OSB	33	Balıkesir OSB	63	Tosb Otomotiv YanSanayi İhtisas OSB
4	Ankara-İvedik OSB	34	Batman OSB	64	Kocaeli Gebze Plastikçiler OSB
5	Ankara Anadolu OSB	35	Burdur 1. OSB	65	Kocaeli Gebze Güzeller OSB
6	Antalya OSB	36	Uludağ OSB	66	Kocaeli Gebze Dilovası OSB
7	Bolu Karma ve Tekstil İhtisas OSB	37	Demirtaş OSB	67	Konya 1. OSB
8	İnegöl OSB	38	Kestel OSB	68	Konya Ereğli OSB
9	Nilüfer OSB	39	Hasanağa OSB	69	Malatya II. OSB
10	Bursa OSB	40	Diyarbakır OSB	70	Malatya Merkez 1 OSB
11	Denizli OSB	41	Düzce OSB	71	Manisa Turgutlu I. OSB
12	Eskişehir Sanayi Odası OSB	42	Erzincan OSB	72	Salihli OSB
13	Gaziantep OSB	43	Erzurum Merkez 1. OSB	73	Mardin OSB
14	Mersin Tarsus OSB	44	Antakya OSB	74	Niğde Bor Karma ve Deri İhtisas OSB
15	İzmir Atatürk OSB	45	Isparta Süleyman Demirel OSB	75	Ordu OSB
16	Kayseri OSB	46	İstanbul İkitelli OSB	76	Osmaniye OSB
17	Gebze OSB	47	İstanbul Anadolu Yakası OSB	77	Sakarya I.OSB
18	Konya OSB	48	İstanbul Deri OSB	78	Samsun Merkez OSB
19	Manisa OSB	49	Birlik OSB	79	Sivas Merkez I. OSB
20	Niğde OSB	50	İstanbul-Tuzla Kimya Sanayicileri OSB	80	Şanlıurfa OSB
21	Sakarya II. OSB	51	İstanbul Dudullu OSB	81	Çorlu Deri İhtisas ve Karma OSB
22	Çerkezköy OSB	52	İstanbul Tuzla OSB	82	Uşak Deri (Karma) OSB
23	Trabzon Arsin OSB	53	İstanbul Beylikdüzü OSB	83	Van OSB
24	Uşak OSB	54	İzmir Kemalpaşa OSB	84	Zonguldak Çaycuma OSB
25	Başkent OSB	55	Tire OSB	85	İtob OSB
26	Çorum OSB	56	Buca Ege OSB	86	Veliköy OSB
27	Adana Koza OSB	57	Kahramanmaraş OSB	87	Ergene-1 OSB
28	Adıyaman OSB	58	Karaman OSB	88	Ergene 2 OSB
29	Aksaray OSB	59	Kars OSB	89	Velimeşe OSB
30	Ankara Sanayi Odası 2. ve 3. OSB	60	Kayseri Mimarsinan OSB		

Çalışmada OSB'ler "genel bilgiler", "çevre yönetim bilgileri", "enerji alt yapı ve kullanım bilgileri" ile "konum bilgileri" başlıkları altında Tablo 2'de verilen 28 değişken dikkate alınarak analiz

edilmiştir. Daha sonra OSB'ler tüm değişkenler birlikte dikkate alınarak tekrar bir analize tabii tutulmuş ve sonuçlar verilmiştir.

Genel bilgiler başlığında OSB'ler, OSB'de bulunan firma sayısı (D1), OSB yönetiminde çalışan personel sayısı (D2), OSB'ye ait Ar-Ge merkezi (D3), Ar-Ge laboratuvarı (D4) ve teknokent (D5) bulunup bulunmaması ve endüstriyel tasarım tescili (D6) ile fiber internet altyapısının (D7) olup olmamasına göre değerlendirilmiştir. Çevre yönetim bilgileri başlığı OSB'lerin çevre yönetim uygulamalarını değerlendirmek amacıyla ele alınmış olup bu başlık altında OSB'ler atık su arıtma tesisinin (D8) bulunup bulunmaması, deşarj izin belgesinin (D9) olup olmaması, evsel atık (D10), proses atık (D11), tehlikeli atık (D12) ve ambalaj atığı (D13) yönetimini uygulayıp uygulamaması, çevre yönetim biriminin olup olmaması (D14), çevre yönetim biriminde çalışan personel sayısı (D15), çevre yönetim sistemi (D16) ve çevre laboratuvarı (D17) bulunup bulunmamasına göre değerlendirilmiştir. Enerji alt yapı ve kullanım bilgileri başlığı ise OSB'lerin kullandıkları enerji miktarları ve enerji altyapısının değerlendirilmesi amacıyla ele alınmış ve OSB'ler bu başlık altında elektrik şebekesinin özelliği (D18), su deposunun bulunup bulunmaması (D19) ve kapasitesi (D20), suyun bakteriyolojik analizinin yapıp yapılmaması (D21), klorlama tesisi olup olmaması (D22) ile elektrik (D23), doğalgaz (D24) ve su (D25) tüketim miktarları dikkate alınarak analiz edilmiştir. Son olarak konum bilgileri başlığı altında ise OSB'ler en yakın havalimanına (D26), demiryoluna (D27) ve limana (D28) uzaklıklarına göre değerlendirilmiştir.

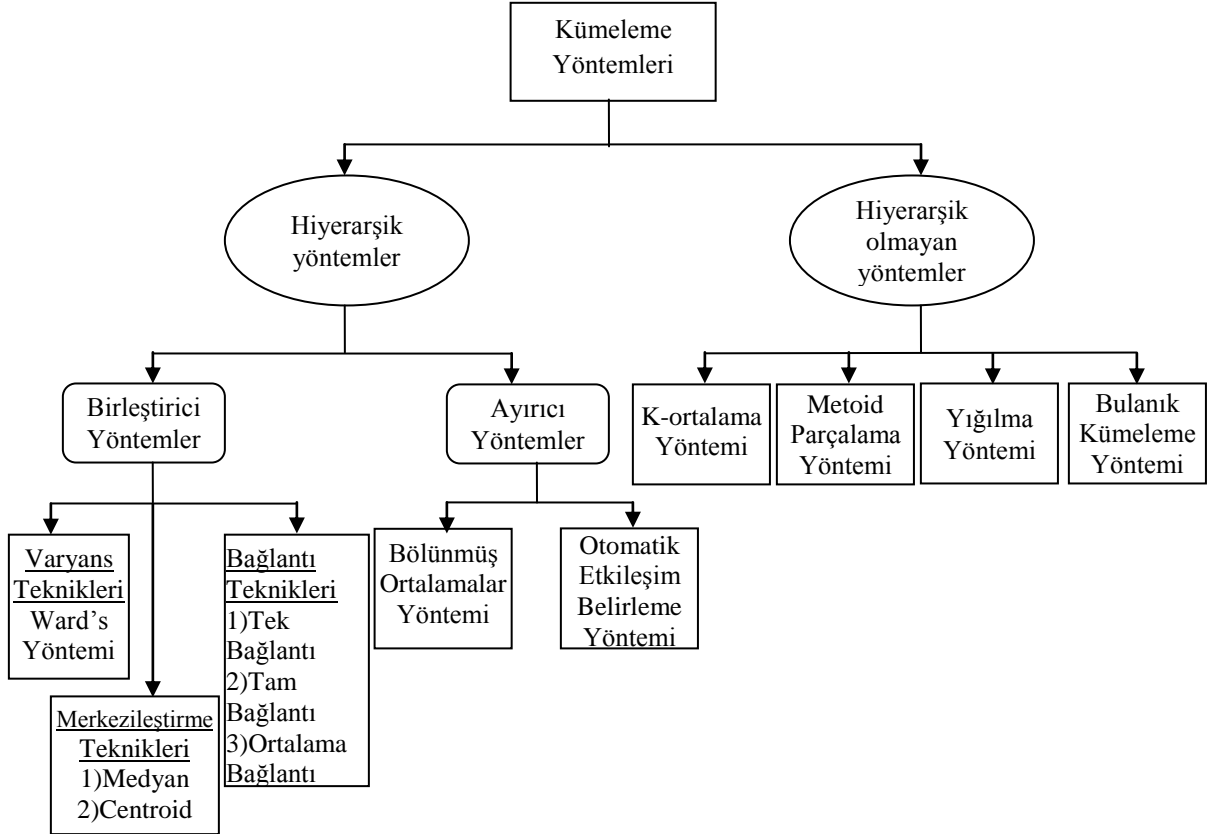
**Tablo2.** Çalışmada kullanılan değişkenler

<b>Genel Bilgiler</b>
D1: OSB'deki firma sayısı
D2: OSB yönetiminde çalışan personel sayısı
D3: Ar-Ge merkezi (var/yok)
D4: Ar-Ge laboratuvarı (var/yok)
D5: Teknokent (var/yok)
D6: Endüstriyel tasarım tescili (var/yok)
D7: Fiber internet altyapısı (var/yok)
<b>Çevre Yönetim Bilgileri</b>
D8: Atık su arıtma tesisi (var/yok)
D9: Deşarj izin belgesi (var/yok)
D10: Evsel atık yönetimi (var/yok)
D11: Proses atık yönetimi (var/yok)
D12: Tehlikeli atık yönetimi (var/yok)
D13: Ambalaj atığı yönetimi (var/yok)
D14: Çevre yönetim birimi (var/yok)
D15: Çevre yönetim birimi personel sayısı
D16: Çevre yönetim sistemi (var/yok)
D17: Çevre laboratuvarı (var/yok)
<b>Enerji Alt Yapı ve Kullanım Bilgileri</b>
D18: Elektrik şebekesinin özelliği (havai hat, yeraltı hattı, havai-yeraltı hat, diğer)
D19: Su deposu (var/yok)
D20: Su deposu kapasitesi (m <sup>3</sup> )
D21: Suyun bakteriyolojik analizi (var/yok)
D22: Klorlama tesisi (var/yok)
D23: Elektrik tüketimi (kWh)
D24: Doğalgaz tüketimi (m <sup>3</sup> )
D25: Su tüketimi (m <sup>3</sup> )
<b>Konum Bilgileri</b>
D26: En yakın havalimanına uzaklığı (km.)
D27: En yakın demiryoluna uzaklığı (km.)

D28: En yakın limana uzaklığı (km.)

### 3.3. Araştırma Yöntemi

Kümeleme analizi, veri matrisindeki gözlemleri benzerliklerine göre gruplamak amacıyla kullanılan çok değişkenli istatistiksel bir yöntemdir. Kümeleme analizi sonucunda elde edilen kümelerin kendi içerisinde homojen olması beklenirken, kümeler arasında ise heterojen bir yapı olmalıdır. Kümeleme analizi yöntemleri hiyerarşik kümeleme ve hiyerarşik olmayan kümeleme yöntemleri olmak üzere iki başlık altında incelenebilir (Şekil 2).



Şekil 2. Kümeleme yöntemlerinin sınıflandırılması (Çokluk vd.,2014).

Hiyerarşik kümeleme yöntemleri birleştirici yöntemler ve ayrııcı yöntemler olmak üzere ikiye ayrılmaktadır. Her birimin başlangıçta bir küme olarak kabul edildiği, daha sonra en yakın iki kümenin yeni bir kümede toplanarak birleştirilip, küme sayısının bu şekilde her seferinde bir azaltıldığı yöntemler birleştirici yöntemlerdir. Süreç dendogram adı verilen şekilde ifade edilebilir. Ayrııcı yöntemlerde ise başlangıçta tüm birimler bir kümede toplanmakta daha sonra benzer olmayan birimler başka bir kümeye atanmaktadır. Böylece küme sayısı her seferinde bir artırılmış olur. Herbir birim bir küme oluşturana kadar işlemlere devam edilir (Öztürk, 2012). Hiyerarşik kümelemenin en etkin ve yaygın kullanılan yöntemleri birleştirici yöntemler olup yöntem oluşturulması gereken küme sayısına kendisi karar verebilir (Kalaycı, 2014). Ayrıca okunuşunun ve yorumun kolay olması yönüne üstünlük sağlamaktadır (Kalaycı, 2014). Çalışmada küme içindeki varyansın en küçük olması amacıyla birleştirici yöntemlerden Ward's yöntemi kullanılmıştır. Ward's yöntemi, bir kümenin ortasına düşen gözlemin, aynı kümenin içinde bulunan gözlemlerden ortalama uzaklığını esas almakta olup toplam sapma karelerinden yararlanır (Kalaycı, 2014). Hiyerarşik kümeleme algoritmalarının çözümün hangi aşamasında durması gerektiğine

ilişkin kesin bir ölçüt bulunmaması ve bir gözlemin bir kümeye dahil edildikten sonra başka bir kümeye tekrar atanmaması ise yöntemin dezavantajlarıdır (Aydın ve Seven, 2015).

Hiyerarşik olmayan yöntemler ise küme sayısının önceden bilinmesi veya küme sayısına araştırmacının önceden karar verebilmesi durumunda tercih edilmekte olup en çok tercih edilen yöntem k-ortalama yöntemidir (Sarıman, 2011). K-ortalamlar yöntemi veri kümesini, başlangıçta verilen küme sayısına göre bölümlenmektedir. Algoritmanın ilk adımında başlangıçta verilen küme sayısı kadar merkez nokta rassal olarak seçilir. İkinci adımda, dikkate alınan uzaklık ölçüsüne göre veriler kendine en yakın olan kümeye atanır. Her eleman yalnızca bir kümeye atanabilmektedir. Daha sonra eklenen elemanlar ile birlikte küme merkezleri tekrar hesaplanır. Bu işlemler, küme merkezleri değişmeye kadar devam etmektedir. Bu yöntem ile sağlıklı bir sonuç elde edilmesinde küme sayısının belirlenmesi büyük önem taşımakta olup küme sayısının başlangıçta belirlenmesi gerekliliği yöntemin kullanımında bir dezavantaj oluşturmaktadır. Mevcut hiyerarşik ve hiyerarşik olmayan kümeleme yöntemlerinin dezavantajlarını azaltmak veya ortadan kaldırmak amacıyla birden fazla kümeleme yönteminin birlikte kullanılması hibrid kümeleme olarak adlandırılmaktadır. Bu çalışmada hibrid bir kümeleme yöntemi olan hiyerarşik-k-ortalamlar yöntemi kullanılmıştır. Yönteme ilişkin literatürde bulunan (Chen vd, 2005), (Aydın ve Seven, 2015) çalışmaları incelenebilir. Bu yaklaşımda hiyerarşik kümeleme ile elde edilen küme sayısı ve küme merkezleri, k-ortalamlar yöntemi için başlangıç değerlerini oluşturmaktadır. Çalışmada hiyerarşik-k-ortalamlar yönteminde hiyerarşik kümeleme yöntemlerinden Ward's metodu ile hiyerarşik olmayan yöntemlerden k-ortalamlar yöntemi kullanılmıştır. Yöntemin algoritması aşağıda verilmiştir.

#### *Adım 1: Veri matrisinin oluşturulması*

Kümeleme analizine tabii tutulmak istenen nesnelere ve değişkenlere ilişkin veriler elde edilerek veri matrisi oluşturulur. Çalışmada 89 adet OSB ve 28 adet değişken için veri matrisi Bilim Sanayi ve Teknoloji Bakanlığı'nın OSB bilgi sitesi (URL1) ve OSB'lerin kendilerine ait web sitelerinden elde edilen bilgiler doğrultusunda oluşturulmuştur.

#### *Adım 2: Verilerin normalize edilmesi*

K-means algoritması sayısal veriler ile çalıştığından veri matrisinde bulunan sözel ifadeler sayısal verilere dönüştürülmelidir (Sarıman, 2011). Daha sonra kullanılan değişkenlerin değer aralıkları birbirinden farklı ise verilerin normalize edilmesi gerekmektedir (Yılmaz ve Patır, 2011). Çalışmada "var-yok" şeklindeki sözel ifadeler sırasıyla 1-0 değerleri ile elektrik şebekesinin özelliği (D18) değişkeni ise havai hat:1, yeraltı hattı:2, havai-yeraltı hattı:3, diğer:4 değerleri ile sayısallaştırılmıştır. Değerler Eş.1 ile verilen min-max normalizasyon yöntemi kullanılarak 0-1 aralığında normalize edilmiştir (Han ve Kamber, 2006; Jain ve Bhandare, 2011).

$$V_i' = \frac{V_i - \min_A}{\max_A - \min_A} (new\_max_A - new\_min_A) + new\_min_A \quad (1)$$

Eş. 1'de  $\min_A$  değişkenin minimum değerini,  $\max_A$  değişkenin maksimum değerini ve  $V_i$   $i$ . verinin sahip olduğu değeri göstermektedir.  $new\_max_A$  ve  $new\_min_A$  ise yeni değer aralığımızı ifade etmektedir.

#### *Adım 3: Ward's hiyerarşik kümeleme yönteminin uygulanması*

Küme sayısının ve küme merkezlerinin belirlenmesi amacıyla ilk olarak hiyerarşik birleştirici yöntemlerden Ward's yöntemi uygulanır. Başlangıçta her gözlem kendi başına bir küme oluştururken, daha sonra birbirine en yakın iki gözlem birleştirilerek küme sayısı bir azaltılmış olur. Son aşamada bütün birimler büyük tek bir kümede toplanır. Birimlerin birbiriyle birleştirilmesinde, küme için hata kareler



toplamını minimize ederek homojen kümeler oluşturulması amacıyla Ward's yöntemi kullanılmıştır. Ward's metodunda p ve q kümelerini içeren bir t kümesi ile bir r kümesinin aralarındaki yakınlığın hesaplanmasında aşağıda verilen Eş.2 kullanılmaktadır (Jain ve Dubes, 1988).

$$d_{rt} = \frac{n_r+n_p}{n_r+n_t} d_{rp} + \frac{n_r+n_q}{n_r+n_t} d_{rq} - \frac{n_r}{n_r+n_t} d_{pq} \quad (2)$$

*Adım 4: Küme sayısı ve küme merkezlerinin belirlenmesi*

Adım 3'te uygulanan hiyerarşik kümeleme sonucu oluşan dendogram incelenerek küme sayısına karar verilir ve kümelerin merkez değerleri belirlenir.

*Adım 5: K-ortalamar yönteminin uygulanması*

Adım4'te belirlenen küme sayısı (k) ve küme merkezleri başlangıç değerleri olarak kullanılarak veriler k-ortalamar yöntemi ile yeniden kümelendir. İlk olarak gözlemler kullanılan uzaklık ölçüsüne göre en yakın kümeye atanır. Daha sonra küme merkezleri güncellenir. Bu işleme küme merkezleri değişmeye kadar devam edilir.

Çalışmada yapılan tüm istatistiksel analizlerde SPSS 20.0 paket programı kullanılmıştır.

## 4. BULGULAR VE TARTIŞMA

Çalışmanın bu bölümünde, kümeleme analizi sonucu elde edilen bilgilere yer verilmiştir. İlk olarak OSB'ler genel bilgiler, çevre yönetimi, enerji alt yapı ve kullanım bilgileri ile konum bilgilerine göre ayrı ayrı kümelendi, daha sonra tüm değişkenler birlikte dikkate alınarak bir kümeleme analizi yapılmıştır.

### 4.1. "Genel Bilgiler" Analiz Sonuçları

OSB'ler, genel bilgiler açısından [D1-D7] değişkenleri kullanılarak kümelendi. Veriler öncelikle hiyerarşik kümelemeye tabii tutulmuş, hiyerarşik kümeleme sonucunda elde edilen dendogram incelendiğinde küme sayısı 5 olarak belirlenmiştir. Küme sayısı 5 olduğunda k-ortalamar analizi sonuçları Tablo3'te, küme sayısının yanı sıra küme merkezlerinin de başlangıçta verildiği hiyerarşik-k-ortalamar yöntemine ilişkin sonuçlar ise Tablo 4'te verilmiştir.

**Tablo 3.** Genel bilgiler için k-ortalamar yöntemi sonuçları

Küme No	OSB No
1	1,2,5,8,10,11,18,19,20,21,22,23,24,28,29,30,31,33,34,37,40,41,42,43,45,47,49,50,52,53,54,56,58,60,64,65,66,67,69,70,73,75,76,77,78,80,83,88.
2	3
3	4,12,17,26,46.
4	6,7,9,14,25,27,32,35,36,38,39,44,48,55,57,59,61,62,68,71,72,74,79,81,82,84,85,86,87,89.
5	13,15,16,51,63.

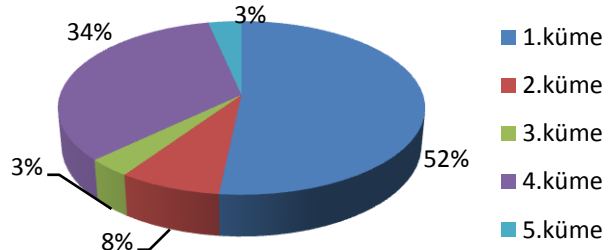
**Tablo 4.** Genel bilgiler için hiyerarşik-k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	1,2,5,8,10,11,18,19,20,21,22,23,24,28,29,30,31,37,40,41,42,43,45,47,49,50,52,53,54,56,58,60,64,65,66,67,69,70,73,75,76,77,78,80,83,88.
2	3,16,17,26,33,34,63.
3	4,12,46.
4	6,7,9,14,25,27,32,35,36,38,39,44,48,55,57,59,61,62,68,71,72,74,79,81,82,84,85,86,87,89.
5	13,15,51.

Sonuçlar incelendiğinde hiyerarşik-k-ortalamlar yönteminde k-ortalamlar yönteminden farklı olarak italik olarak belirtilmiş olan 6 adet OSB'nin (33, 34, 16, 17, 26, 63 numaralı OSB'ler) ait olduğu kümenin değiştiği görülmektedir. Her bir elemanın ait olduğu kümenin merkezine olan uzaklıklarının kareleri toplamının ortalamasını ifade eden hata kareler ortalamasına bakıldığında, k-ortalamlar yöntemine göre toplam hata kareler ortalaması 0.114 çıkarken hiyerarşik-k-ortalamlar yönteminde 0.110 çıkmıştır. Ayrıca, k-ortalamlar yöntemi ile 3 iterasyonda sonuca ulaşılrken hiyerarşik-k-ortalamlar ile 2 iterasyonda sonuca ulaşılmıştır. Hiyerarşik-k-ortalamlar yöntemine ilişkin ANOVA testi sonuçları Tablo 5'te verilmiştir.

**Tablo 5.** Genel bilgiler için hiyerarşik-k-ortalamlar yöntemi ANOVA sonuçları

Değişken No	F	Sig.
D1	23,992	0,000
D2	7,613	0,000
D3	87,736	0,000
D4	31,976	0,000
D5	29,301	0,000
D6	32,994	0,000
D7	13,052	0,000

**Şekil 3.** Genel bilgiler için OSB sayılarının kümelere göre dağılımı

Tablo 5'te tüm değişkenler için sig. değeri  $< 0.05$  olduğundan kümelerin değişkenler açısından birbirinden farklı özellikler gösterdiği söylenebilir.

#### 4.2. "Çevre Yönetimi" Analiz Sonuçları

OSB'ler çevre yönetimi açısından [D8-D17] değişkenleri kullanılarak kümelendiğinde hiyerarşik kümeleme sonucunda küme sayısı 6 olarak belirlenmiştir. K-ortalamlar yönteminin sonuçları Tablo 6'da, hiyerarşik-k-ortalamlar yöntemi sonuçları ise Tablo 7'de verilmiştir.

**Tablo 6.** Çevre bilgileri için k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	1,3,9,14,37,47,49,51,52,73,77.
2	2,4,7,35,62,71,78,86,88.
3	6,8,10,12,13,15,16,17,18,19,20,34,45,63,76.
4	53,64,65,70,81.
5	5,11,22,24,25,31,32,33,36,39,44,46,48,50,54,58,66,68,82,85.
6	21,23,26,27,28,29,30,38,40,41,42,43,55,56,57,59,60,61,67,69,72,74,75,79,80,83,84,87,89.

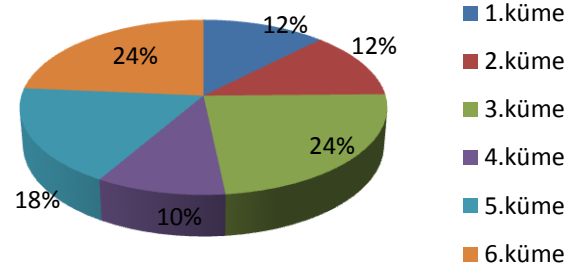
**Tablo 7.** Çevre bilgileri için hiyerarşik-k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	1,3,9,14,37,47,49,51,52,73,77.
2	2,4,7,35,36,46,62,71,78,86,88.
3	6,8,10,11,12,13,17,18,19,22,25,32,34,45,50,54,58,63,66,82,85.
4	15,16,20,53,64,65,70,76,81.
5	5,21,24,26,31,33,39,41,44,48,56,60,68,69,75,83.
6	23,27,28,29,30,38,40,42,43,55,57,59,61,67,72,74,79,80,84,87,89.

Hiyerarşik-k-ortalamlar yöntemi ile 24 adet OSB'nin (11, 15, 16, 20, 21, 22, 25, 26, 32, 36, 41, 46, 50, 54, 56, 58, 60, 66, 69, 75, 76, 82, 83, 85 numaralı OSB'ler) ait olduğu küme değişmiştir. Toplam hata kareler ortalaması ise k-ortalamlar yönteminde 0.888 çıkarken hiyerarşik-k-ortalamlar yönteminde 0.879 çıkmıştır. K-ortalamlar yöntemi ile 5 iterasyonda sonuca ulaşılrken hiyerarşik-k-ortalamlar ile 3 iterasyonda sonuca ulaşılmıştır. Hiyerarşik-k-ortalamlar yöntemine ilişkin ANOVA testi sonuçları Tablo 8 'de verilmiştir.

**Tablo 8.** Çevre bilgileri için hiyerarşik-k-ortalamlar yöntemi ANOVA sonuçları

Değişken No	F	Sig.
D8	11,789	0,000
D9	30,789	0,000
D10	8,538	0,000
D11	33,429	0,000
D12	55,977	0,000
D13	16,840	0,000
D14	99,098	0,000
D15	4,796	0,001
D16	15,767	0,000
D17	19,281	0,000

**Şekil 4.** Çevre bilgileri için OSB sayılarının kümelere göre dağılımı

Tablo 8'de tüm değişkenler için sig. değerinin  $< 0.05$  olduğu görülmektedir. Kümeler değişkenler açısından birbirinden farklı özellikler göstermektedir.

### 4.3. “Enerji Alt Yapı ve Kullanım Bilgileri” Analiz Sonuçları

OSB'ler enerji alt yapı ve kullanım bilgileri açısından 8 değişken [D18-D25] kullanılarak kümelendirilmiş olup hiyerarşik kümeleme sonucunda oluşan dendrogram incelendiğinde küme sayısı 6 olarak

belirlenmiştir. K-ortalamlar yönteminin sonuçları Tablo 9’da, hiyerarşik-k-ortalamlar yöntemi sonuçları ise Tablo 10’da verilmiştir.

**Tablo 9.** Enerji alt yapı ve kullanım bilgileri için k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	2,5,6,7,8,18,25,26,28,29,30,32,43,44,52,55,58,60,63,69,76,80,83,85.
2	15,53,78.
3	3,4,10,21,35,36,38,41,46,48,49,51,54,56,57,61,66,67,68,75,77,79,86,87,88,89.
4	1,11,12,16,19,37.
5	9,13,14,17,22,24,34,40,47,50,59,62,64,65,71,73,81,82.
6	20,23,27,31,33,39,42,45,70,72,74,84.

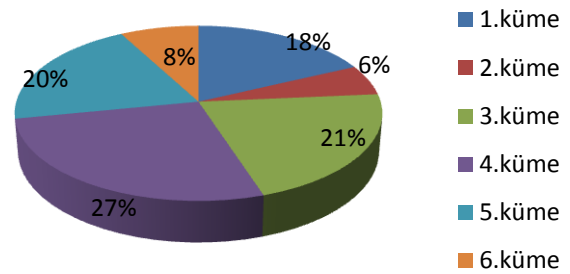
**Tablo 10.** Enerji alt yapı ve kullanım bilgileri için hiyerarşik-k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	1,6,7,8,18,25,26,28,29,32,44,58,76,80,83,85.
2	2,15,53,55,78.
3	3,4,21,35,36,41,46,51,54,66,67,68,75,77,79,86,87,88,89.
4	5,11,12,16,19,20,23,27,30,31,33,37,39,42,43,45,52,60,63,69,70,72,74,84.
5	9,13,14,17,22,24,34,40,47,50,59,62,64,65,71,73,81,82.
6	10,38,48,49,56,57,61.

Hiyerarşik-k-ortalamlar yöntemi ile 29 adet OSB’nin (1, 2, 5, 10, 20, 23, 27, 30, 31, 33, 38, 39, 42, 43, 45, 48, 49, 52, 55, 56, 57, 60, 61, 63, 69, 70, 72, 74, 84 numaralı OSB’ler) ait olduğu küme değişmiştir. K-ortalamlar yöntemine göre toplam hata kareler ortalaması 0.207, hiyerarşik-k-ortalamlar yönteminde ise 0.156 çıkmıştır. Her iki yöntem ile de 3 iterasyonda sonuca ulaşılmıştır. Hiyerarşik-k-ortalamlar yöntemine ilişkin ANOVA testi sonuçları Tablo 11 ‘de verilmiştir.

**Tablo 11.** Enerji bilgileri için hiyerarşik-k-ortalamlar yöntemi ANOVA sonuçları

Değişken No	F	Sig.
D18	9,708	0,000
D19	25,273	0,000
D20	4,155	0,002
D21	324,014	0,000
D22	15,129	0,000
D23	3,315	0,009
D24	9,466	0,000
D25	7,397	0,000



**Şekil 5.** Enerji bilgileri için OSB sayılarının kümelere göre dağılımı

Tablo 11’de tüm değişkenler için sig. değerinin  $< 0.05$  olduğu görülmektedir. Kümeler değişkenler açısından birbirinden farklı özellikler göstermektedir.

#### 4.4. “Konum Bilgileri” Analiz Sonuçları

OSB’ler konum bilgileri açısından 3 değişken [D26-D28] kullanılarak kümelendi ve hiyerarşik kümeleme sonucunda küme sayısı 6 olarak belirlenmiştir. K-ortalamlar yöntemi analiz sonuçları Tablo 12’de, hiyerarşik-k-ortalamlar yöntemi sonuçları ise Tablo 13’de verilmiştir.

**Tablo12.** Konum bilgileri için k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	1,5,11,12,13,14,15,17,19,20,21,22,23,24,31,32,33,35,42,44,45,46,47,48,49,50,51,52,53,54,55,56,58,61,63,64,65,66,68,71,72,74,76,77,78,81,82,84,85,86,87,88,89.
2	8,27,28,29,57,62.
3	7,41.
4	2,3,4,16,18,25,30,34,40,43,59,60,67,69,70,73,79,80,83.
5	9
6	6,10,26,36,37,38,39,75.

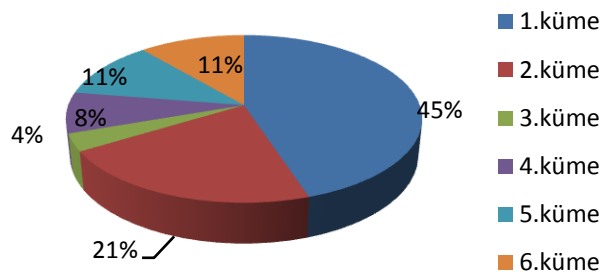
**Tablo13.** Konum bilgileri için hiyerarşik-k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	1,8,12,14,15,17,19,22,23,31,33,35,44,45,46,47,48,49,50,51,52,53,54,55,56,63,64,65,66,71,76,77,78,81,84,85,86,87,88,89.
2	2,4,5,11,13,16,18,25,28,42,43,57,59,60,67,70,79,80,82.
3	7,9,41.
4	3,30,34,40,69,73,83.
5	20,21,24,27,32,58,61,68,72,74.
6	6,10,26,29,36,37,38,39,62,75.

Hiyerarşik-k-ortalamlar yöntemi ile 31 adet OSB’nin (2, 4, 5, 8, 9, 11, 13, 16, 18, 20, 21, 24, 25, 27, 29, 32, 42, 43, 58, 59, 60, 61, 62, 67, 68, 70, 72, 74, 79, 80, 82 numaralı OSB’ler ) ait olduğu küme değişmiştir. K-ortalamlar yöntemine göre toplam hata kareler ortalaması 0.032, hiyerarşik-k-ortalamlar yönteminde ise 0.026 çıkmıştır. K-ortalamlar yöntemi ile 5 iterasyonda sonuca ulaşılmışken hiyerarşik-k-ortalamlar ile 2 iterasyonda sonuca ulaşılmıştır. Hiyerarşik-k-ortalamlar yöntemine ilişkin ANOVA testi sonuçları Tablo 14 ‘te verilmiştir.

**Tablo 14.** Konum bilgileri için hiyerarşik-k-ortalamlar yöntemi ANOVA sonuçları

Değişken No	F	Sig.
D26	53,477	0,000
D27	90,884	0,000
D28	87,189	0,000



**Şekil 6.** Konum bilgileri için OSB sayılarının kümelere göre dağılımı

Tablo 14’te tüm değişkenler için sig. değeri < 0.05 olup kümeler değişkenler açısından birbirinden farklı özellikler göstermektedir.

#### 4.5. Tüm Değişkenler İçin Analiz Sonuçları

OSB'ler tüm değişkenler [D1-D28] dikkate alınarak kümelendiğinde, hiyerarşik kümeleme sonucunda küme sayısı 9 olarak belirlenmiştir. K-ortalamlar yöntemi analiz sonuçları Tablo 15'te, hiyerarşik-k-ortalamlar yöntemi sonuçları ise Tablo 16'da verilmiştir.

**Tablo 15.** Tüm değişkenler için k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	1,14,33,37,52,73.
2	21,35,36,38,41,48,53,54,56,57,61,66,67,68,75,78,79,86,87,88,89.
3	3,51.
4	4,46.
5	13,15,16,17.
6	34,47,49,77.
7	2,5,23,24,26,27,28,29,30,31,40,42,43,55,59,60,69,70,72,74,80,81,83,84.
8	8,10,11,12,18,19,20,22,45,50,58,63,64,65,76.
9	6,7,9,25,32,39,44,62,71,82,85.

**Tablo 16.** Tüm değişkenler için hiyerarşik-k-ortalamlar yöntemi sonuçları

Küme No	OSB No
1	1,9,14,37,52,73.
2	6,7,10,18,19,20,25,32,33,39,44,45,58,62,64,71,76,85.
3	3,47,49,51,77.
4	2,4,35,46,67,78,79,86,87,88,89.
5	5,23,26,28,29,30,40,42,43,60,69,70,80,81,83
6	8,11,22,24,31,50,53,65,82
7	12,13,15,16,17,34,63
8	21,36,41,48,54,56,66,68,75
9	27,38,55,57,59,61,72,74,84

K-ortalamlar yöntemine göre toplam hata kareler ortalaması 2.019 çıkarken hiyerarşik-k-ortalamlar yönteminde ise 1.925 çıkmıştır. K-ortalamlar yöntemi ile sonuca 5 iterasyonda, hiyerarşik-k-ortalamlar yöntemi ile 3 iterasyonda ulaşılmıştır.

Hiyerarşik-k-ortalamlar yöntemine ilişkin ANOVA testi sonuçları Tablo 17 'de verilmiştir. Burada D1,D5,D18,D20,D25,D26,D27 değişkenleri için sig. değerinin  $< 0.05$  olmadığı görülmektedir. Kümeler bu değişkenler açısından birbirinden farklı özellikler göstermemektedir. Farklı küme sayıları ile yapılan denemeler sonucunda sig. değerinin tüm değişkenler için ancak 27 kümede  $< 0.05$  olduğu gözlemlenmiştir. 27 küme için ANOVA sonuçları Tablo 18'de, bu sonuçlara ilişkin kümeler ise Tablo 19'da verilmiştir.

**Tablo 17.** Tüm değişkenler için hiyerarşik-k-ortalamalar yöntemi ANOVA sonuçları (9 küme)

Değişken No	F	Sig.
D1*	1.693	.113
D2	3.157	.004
D3	11.000	.000
D4	12.558	.000
D5*	1.622	.132
D6	4.533	.000
D7	7.075	.000
D8	9.218	.000
D9	8.936	.000
D10	38.924	.000
D11	20.908	.000
D12	40.908	.000
D13	10.965	.000
D14	10.796	.000
D15	2.538	.016
D16	7.154	.000
D17	6.284	.000
D18*	.738	.657
D19	22.517	.000
D20*	2.015	.055
D21	23.569	.000
D22	7.938	.000
D23	3.187	.003
D24	3.992	.001
D25*	1.935	.066
D26*	1.687	.114
D27*	1.414	.203
D28	4.804	.000

**Tablo 18.** Tüm değişkenler için hiyerarşik-k-ortalamalar yöntemi ANOVA sonuçları (27 küme)

Değişken No	F	Sig.
D1	335.014	.000
D2	4.296	.000
D3	15.758	.000
D4	7.385	.000
D5	12.497	.000
D6	15.835	.000
D7	2.588	.001
D8	8.669	.000
D9	4.209	.000
D10	36.560	.000
D11	14.685	.000
D12	39.815	.000
D13	10.518	.000
D14	6.308	.000
D15	20.307	.000
D16	13.254	.000
D17	5.502	.000
D18	2.233	.005
D19	6.938	.000
D20	49.605	.000
D21	7.227	.000
D22	4.619	.000
D23	3.157	.000
D24	7.477	.000
D25	2.940	.000
D26	2.274	.004
D27	1.668	.050
D28	1.984	.014

**Tablo 19.** Tüm değişkenler için hiyerarşik-k-ortalamalar yöntemi sonuçları (27 küme)

Küme No	OSB No	Küme No	OSB No	Küme No	OSB No
1	1	10	10,19	19	34,45
2	2,78	11	33,39	20	51
3	3	12	8,12,63	21	21,35,56,67,75,79,86,87,88,89
4	4	13	13,17	22	11,22,24,50,54,82
5	14,37,52,73	14	48,68	23	70,81
6	47,49,77	15	15	24	46
7	7,62,71	16	16,20	25	6,18,25,32,44,58,76,85.
8	27,28,38,42,43,55,57,59,61,72,74,84.	17	5,23,29,30,31,40,60,69,80,83	26	26
9	9	18	53,64,65	27	36,41,66

## 5. SONUÇ VE ÖNERİLER

Organize Sanayi Bölgeleri, sanayi tesislerini bir araya toplayarak işletmelerin nitelikli ve ucuz altyapı hizmetlerine erişimini sağlamakla birlikte Ar-Ge merkezi, teknokent vb. ortak hizmet alanları oluşturarak işletmeler için çeşitli avantajlar sunmak suretiyle rekabet gücünün artırılmasında ve sanayinin

gelişiminde önemli bir rol oynamaktadır. Ekonomik kalkınmanın yanı sıra sanayi tesislerinin planlı bir şekilde yerleşmesi ve sanayinin çevreye olan olumsuz etkilerini azaltmak için uyguladıkları politikalar ile OSB'ler kentleşme ve çevre kirliliği açısından da büyük önem taşımaktadır. OSB'ler için alınacak olan stratejik kararlarda hangi OSB'lerin hangi açılardan benzer özellik gösterdiğinin bilinmesi karar vericiler açısından önem arz etmektedir. Elde edilen sonuçlar politikacılara da önemli çıkarımlar sunmaktadır. Analiz sonucunda elde edilen kümelenmeler, ihtiyaca göre bölgesel teşviklerle güçlendirilebilir.

Bu çalışmada, ülkemizde bulunan 89 OSB “genel bilgiler”, “çevre yönetim bilgileri”, “enerji altyapı ve kullanım bilgileri” ile “konum bilgileri” olmak üzere 4 başlıkta 28 değişken dikkate alınarak analiz edilmiş ve benzer özellik gösteren OSB'ler belirlenmiştir. Yapılan kümeleme analizinde, hiyerarşik-k-ortalamlar yöntemi kullanılarak, hiyerarşik kümeleme ve k-ortalamlar yöntemlerine ilişkin dezavantajların en aza indirilerek kümeleme algoritmasının etkinliğinin artırılması hedeflenmiştir. Sonuçlar incelendiğinde hiyerarşik-k-ortalamlar yöntemi ile yapılan kümelemenin k-ortalamlar yöntemine kıyasla hem hata kareler ortalaması hem de iterasyon sayısı açısından daha iyi sonuç verdiği görülmektedir.

Farklı değişkenler dikkate alınarak çalışma kapsamının genişletilmesi ve tüm OSB'leri kapsayacak bir analiz yapılması gelecekte yapılacak çalışmalar için potansiyel bir alan teşkil etmektedir. Kümeleme faaliyetleri sadece ülke içindeki coğrafya ile sınırlı kalmayıp uluslararası kümeler oluşturulmak suretiyle işbirlikleri ya da fikir alışverişleri artırılabilir. Ayrıca, ileriki çalışmalarda bulanık c-ortalamlar yönteminden yararlanılabilir.



**KAYNAKLAR**

- Aydın, N. ve Seven, A.N. (2015). İl nüfus ve vatandaşlık müdürlüklerinin iş yoğunluğuna göre hibrid kümeleme ile sınıflandırılması. *Yönetim ve Ekonomi Araştırmaları Dergisi*,13(2),181-201.
- Arai, K. ve Barakbah, A.R. (2007). Hierarchical K-means: an algorithm for centroids initialization for K-means, *Reports of the Faculty of Science and Engineering*, 36 (1), 25-31.
- Bilen, Ö., Hotaman, D., Aşkın, Ö.E. ve Büyüklü A.H. (2014). LYS başarılarına göre okul performanslarının eğitsel veri madenciliği teknikleriyle incelenmesi: 2011 İstanbul örneği. *Eğitim ve Bilim*, 39 (172), 78-94.
- Bulut, T. (2016). Türkiye'deki organize sanayi bölgelerinin performanslarının TOPSIS yöntemiyle değerlendirilmesi. *Kalkınmada Anahtar Verimlilik*, 335, 36-41.
- Celebi, M.E., Kingravi, H.A. ve Vela, P.A. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm, *Expert Systems with Applications*, 40 (1), 200–210.
- Cengiz, D. ve Öztürk, F. (2012). Türkiye' de illerin eğitim düzeylerine göre kümeleme analizi ile incelenmesi. *Trakya University Journal of Social Science*, 14 (1), 69-84.
- Ceylan, H.H. (2013). Perakende sektöründe konjoint ve kümeleme analizi ile fayda temelli pazar bölümlendirme. *Yönetim ve Ekonomi*, 20 (1), 141–154.
- Chen, B., Tai,P.C., Harrison,R ve Pan,Y. (2005). Novel hybrid hierarchical K-means clustering method (HK-means) for microarray analysis. *IEEE Computational Systems Bioinformatics Conference*, California-USA.
- Çağlar, E. (2006). Türkiye'de yerelleşme ve rekabet gücü: kümelenmeye dayalı politikalar ve organize sanayi bölgeleri. *I. Bölgesel Kalkınma ve Yönetişim Sempozyumu Bildiri Kitabı*, 305–316.
- Çalışkan, S.K. ve Soğukpınar, İ. (2008). KxKNN: K-means ve K en yakın komşu yöntemleri ile ağlarda nüfus tespiti. 2. Ağ ve Bilgi Güvenliği Sempozyumu, Girne, 120-124.
- Çetin, M. ve Kara, M. (2008). Bir kalkınma aracı olarak “Organize Sanayi Bölgeleri”: Isparta Süleyman Demirel Organize Sanayi Bölgesi üzerine bir araştırma. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 31, 49–68.
- Çokluk,Ö., Şekercioğlu,G. ve Büyüköztürk,Ş. (2014). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. PEGEM Akademi.
- Fırat, M., Dikbaş, F., Koç, A.C. ve Güngör M. (2012). K-ortalamar yöntemi ile yıllık yağışların sınıflandırılması ve homojen bölgelerin belirlenmesi. *İMO Teknik Dergi*, 6037-6050, Yazı 383.
- Giray, S. (2016). İki aşamalı kümeleme analizi ile hükümlü verilerinin incelenmesi. *İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik Dergisi*, 25, 1–31.
- Han,J. ve Kamber,M. (2006). *Data mining concepts and techniques*. Elsevier.
- Jain,Y.K. ve Bhandare,S.K. (2011). Min max normalization based data perturbation method for privacy protection. *International Journal of Computer & Communication Technology*, 2(8), 45-50.
- Jain, A.K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666.
- Jain, A.K. ve Dubes,R.C. (1988). *Algorithms for clustering data*. Prentice Hall Advanced Reference Series, Computer Science, Chapter 3.
- Kalaycı, Ş. (2014). *SPSS uygulamalı çok değişkenli istatistik teknikleri*, Asil Yayın Dağıtım, 5.Baskı.
- Mingoti, S.A. ve Lima, J.O. (2006). Comparing SOM neural network with fuzzy c-means, K-means and traditional hierarchical clustering algorithms. *European Journal of Operational Research*, 174 (3), 1742–1759.

- Na, S., Xumin, L. ve Yong, G. (2010). Research on k-means clustering algorithm: An improved k-means clustering algorithm. *Third International Symposium on Intelligent Information Technology and Security Informatics*, 63-67.
- Niknam, T. ve Amiri, B. (2010). An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis. *Applied Soft Computing*, 10, 183-197.
- Öztürk,F. (2012). *Kümeleme analizi ve uygulaması*. Yayımlanmış Yüksek Lisans Tezi, İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü.
- Özden Örnek, E. (2016). Kalkınma aracı olarak organize sanayi bölgelerini yeniden kurgulamak. *MEGARON*, 11(1), 106-124.
- Sarıman, G. (2011). Veri madenciliğinde kümeleme teknikleri üzerine bir çalışma: K-means ve K-metoids kümeleme algoritmalarının karşılaştırılması, *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü*, 15(3),192-202.
- Silgu, M.A. ve Çelikoğlu, H.B. (2014). K-means clustering method to classify freeway traffic flow patterns. *Pamukkale University Journal of Engineering Sciences*, 20 (6), 232-239.
- Velmurugan, T. (2014). Performance based analysis between k-Means and fuzzy C-Means clustering algorithms for connection oriented telecommunication data. *Applied Soft Computing*, 19, 134-146.
- URL1: <https://osbbs.sanayi.gov.tr/default.aspx> Erişim Tarihi: 01.03.2017.
- Yılmaz Koltan,Ş. ve Patır,S. (2011). Kümeleme analizi ve pazarlamada kullanımı. *Journal of Academic Approaches*, 2(1),91-113.
- Yürük, F. ve Erdoğan, P. (2015). Düzce ilinin hayvansal atıklardan üretilebilecek biyogaz potansiyeli ve K-means kümeleme ile optimum tesis konumunun belirlenmesi. *İleri Teknoloji Bilimleri Dergisi*, 4(1), 47-56.
- Zeng, S. X., Liu,H.C., Tam,C.M. ve Shao,Y.K. (2008). Cluster analysis for studying industrial sustainability: an empirical study in Shanghai. *Journal of Cleaner Production*, 16(10),1090-1097.

### **Citation Information/Kaynakça Bilgisi**

- Uslu, A., Çetinkaya, C., Özceylan, E. ve İşleyen, S. K. (2017). Organize sanayi bölgelerinin hiyerarşik -K-ortalamlar yöntemi ile analizi, *Türk Sosyal Bilimler Araştırmaları Dergisi*, 2(1), 20-37.