

İki Aşamalı Kümeleme Analizi ile Bireysel Emeklilik Sektöründe Müşteri Profiline Değerlendirilmesi

Zeynep CEYLAN^{1,2*}, Samet GÜRSEV², Serol BULKAN²

¹Endüstri Mühendisliği, Mühendislik Fakültesi, Ondokuz Mayıs Üniversitesi, Samsun, Türkiye

²Endüstri Mühendisliği, Mühendislik Fakültesi, Marmara Üniversitesi, İstanbul, Türkiye
zeynep.dokumaci@omu.edu.tr, sametgursev@gmail.com, sbulkan@marmara.edu.tr

(Geliş/Received:29.06.2017; Kabul/Accepted:08.10.2017)

DOI: 10.17671/gazibtd.323951

Özet— Türkiye'de Bireysel Emeklilik Sistemi (BES) 27 Ekim 2003 tarihinde yürürlüğe girmiştir. Bu tarihten itibaren katılımcı sayısı ve fon büyüklüğü açısından BES hızlı bir gelişme göstermiştir. Bu nedenle bireysel emeklilik şirketlerinin katılımcı sayılarını arttırmaları için ürün ve hizmetlerde farklılaşmaları ve doğru müşteriye doğru ürün ve hizmeti sunmaları gerekmektedir. Bunların yanı sıra, şirketler müşterilerini elde tutmaya yönelik, veri madenciliği teknikleriyle mevcut müşterilerini iyi tanımaları ve müşteri kaybını önlemek için doğru stratejileri uygulamaları gerekmektedir. Çünkü farklı müşteri ihtiyaçlarını doğru şekilde karşılamak, rekabet ortamında şirketlerin pazar payı ve karlılıklarını arttırmaları açısından önemli rol oynamaktadır. Bu nedenle, veri madenciliği büyük veri yığınlarına sahip olan işletme ve kurumların, rakiplerine üstünlük sağlayabilmesi için kullanılan önemli bir tekniktir. Bu çalışmada, Türkiye'de faaliyet gösteren bir emeklilik şirketinin müşterileri hakkındaki bilgilere veri madenciliği kümeleme yöntemlerinden biri olan iki aşamalı kümeleme yöntemi uygulanmıştır. İki aşamalı kümeleme yöntemi, diğer kümeleme yöntemleri ile kıyaslandığında hem sürekli hem de kategorik verileri işleyebilme yeteneğine sahiptir. Ayrıca büyük örneklem büyüklüklerinde esnek olması nedeniyle tercih edilmektedir. Veriler üç farklı kümeye ayrılmış ve her küme ayrı ayrı incelenmiştir. Elde edilen 3 kümede de özellikle tasarruflarını aylık 0-250 TL katkı payı ile düşük birikimlerle yapan katılımcıların çoğunlukta olduğu görülmüştür. Ayrıca, her üç kümede de BES'e katılımın en çok olduğu yaş grubunun 35-44 yaş arası olması, bu yaş aralığındaki katılımcıların geleceği açısından emeklilik sistemine daha olumlu yaklaşımlarına işaret ettiği sonucuna varılmıştır. Bunların yanı sıra, katılımcıların yaşadıkları şehirlerin sosyoekonomik ve demografik özelliklerinin farklı olması, katılımcıların tasarruf yapma kararlarını doğrudan etkilediği gözlemlenmiştir. Sonuç olarak, elde edilen bulgular şirketlerin yeni satış stratejileri oluştururken sadece sözleşme adedi odaklı olarak değil müşterilerinin cinsiyet, medeni durum, yaş, meslek, öğrenim durumu, yaşadığı şehir bilgilerinin yanı sıra ödeme ve tasarruf bilgilerini göz önüne alarak tasarlanması gerektiğini göstermiştir.

Anahtar Kelimeler— İki aşamalı kümeleme, müşteri analizi, veri madenciliği, bireysel emeklilik sistemi

Evaluation of Customer Profile in Individual Pension Sector by Two-Step Cluster Analysis

Abstract— The Individual Pension System (IPS) in Turkey entered into force on October 27, 2003. From this date, IPS shows a rapid development in terms of number of participants and fund size. Thus, individual pension companies need to differentiate in products and services and must offer the right products and services to the right customers in order to increase the number of participants. In addition to these, data mining techniques aimed at retaining corporate customers should be applied by existing customers with good definitions and correct strategies to prevent customer loss. Correctly meeting the needs of different customers plays an important role in increasing the market and profitability of companies in a competitive environment. For this reason, data mining is an important technique used by enterprises and institutions that have large data volumes to provide superiority to their competitors. In this study, two-step cluster method, one of the data mining clustering methods, was applied to the information of the customers of a pension company operating in Turkey. The two-step cluster method has the ability to process both continuous and categorical data when compared to other clustering methods. It is also preferred because of its large sample size and flexibility. The data were divided into three different clusters and each cluster was examined separately. In the obtained three clusters, it was seen that customers that especially made their savings with low accumulation as monthly 0-250 TL contribution formed the majority. In addition, for all three clusters, group of participants' with an age range between 35-44 was the most, thus it is concluded that this indicate that participants in this age range are more likely to approach the pension system

considering their future. Additionally, it has been observed that differences in socioeconomic and demographic characteristics of the cities where participants live in directly effects the participants' decisions to make savings. As a result, the findings show that while companies create new sales strategies, they should be designed not only based on contract number but also considering customers' gender, marital status, age, occupation, learning situation, living city information but also payment and savings.

Keywords— Two-step cluster, customer analysis, data mining, individual pension system

1. GİRİŞ (INTRODUCTION)

Küresel ekonomi, günümüz dünyasında hızla değişmekte ve borç tasarruf oranları bu değişime tepki göstermektedir. Tüm bunlara ek olarak genç nüfusu fazla olan ülkelerde emeklilik sistemleri ilerleyen dönemlerde büyük bir potansiyel sorun olarak görülmektedir. Bu sebeple, yaşlanan nüfus ve borç krizlerini engelleyici bazı çalışmaların önceden planlanması şart olmaktadır. Hükümetler tasarruf fonlarına ve emeklilik sistemlerine büyük önem vermek zorundadır. Hızla azalan genç nüfus sayısı toplumdaki üretim tüketim dengesini değiştirmektedir. Yaşlanan nüfus beraberinde sağlık masrafları ve emeklilik ödemeleri getirmektedir ve bu durum sağlık sigortaları, emeklilik sistemi ve kredi konularında yeni ve önleyici kararlar alınmasını gündeme getirmektedir. Ayrıca, günümüz dünyasında konut, araç, eğitim konularında ihtiyaç duyulan kredi miktarlarının hızla artması ve banka sistemlerindeki kredi kartı işleyişlerinin getirdiği borçlanma süreci hükümetleri hızla büyütmekte olan bir cari açık krizine sürüklemektedir. Bu nedenle, ekonomilerin ayakta kalabilmesi, cari açıkların kapatılması toplumların refah düzeyinin korunması için büyük önem arz etmektedir.

Türkiye, sosyal güvenlik reformu çalışmaları içerisinde yaşanan borç sorunu ve ileride karşılabilecek emeklilik sistemi sorunlarına bir altyapı oluşturması için Ekim 2003 tarihinde Bireysel Emeklilik Sistemi'ni (BES) resmen başlatmış ve başlangıcından bu yana önemli bir yol katetmiştir. Bireysel emeklilik, katılımcıya yüzde 25 devlet desteği verilerek uzun vadede sistemde kalınması ve bu sistemden emekli olunmasına yardımcı olan, kanun ve yönetmelikler ile kontrol edilen bir yapıdır. Katılımcıların sistemden emekli olabilmesi için 56 yaşını doldurması ve tercih ettikleri katkı payı tutarını, 10 yıl boyunca düzenli olarak seçtiği bir emeklilik şirketine yatırması gerekmektedir. Böylelikle, BES' te 10 yıl geçiren ve 56 yaşını tamamlamış katılımcılar, birikimlerini topluca alabilme ya da belli bir plan çerçevesinde aylık veya yıllık olarak düzenli maaş alabilme hakkını elde etmektedirler.

Türkiye'de BES gönüllü katılıma ve fona dayalı bir yapıya sahip iken yeni alınan kararlar doğrultusunda 2017 yılı itibariyle çalışanlara zorunlu bir uygulama ile kapsam ve içeriği geliştirilmektedir. Ödenecek katkı payları, çalışanların prime esas kazançlarının veya emeklilik keseneğine esas aylığının %3'ü oranında olup, daha

yüksek bir oranda katkı payı kesintisi çalışanın isteğine bağlı olarak işverenden talep edilebilmektedir. Bunun yanı sıra, zorunlu olarak BES' e dahil edilen çalışanın ilerleyen dönem içinde sistemden çıkması da her zaman mümkün olabilmektedir. Geliştirilen bu uygulama sayesinde topluma ikinci emeklilik geliri ile refah düzeyi sağlaması amaçlanmaktadır. 31.12.2004 tarihinde 314.257 kişinin katıldığı bireysel emeklilik sisteminde 21.04.2017 tarihi itibariyle 6.750.559 katılımcı sayısına ulaşılmış ve bunlardan toplanan fon tutarı da 58 milyar 275.01 milyon TL' ye yaklaşmıştır [1]. Ekonomik Kalkınma ve İşbirliği Örgütü verilerine göre 2020 yılında 115 milyar liralık bir fon büyüklüğünün geçileceği öngörülmektedir. Bu sayıya zorunlu bireysel emeklilik sistemi etkileri de eklenirse ciddi anlamda bir artış olacağı bir gerçektir [2]. Müşteri sayısındaki bu büyük artış emeklilik şirketleri için büyük bir rekabet avantajı oluşturmaktadır. Şirketlerin başarıya ulaşmasında; sisteme katılan bireylerin özellikle katkı paylarını düzenli ödemeleri, sistemin uzun vadeli bir yapıya sahip olduğunu bilerek gereken sabrı ve inancı göstermeleri, topludurumsal gelişmelere uygun olarak yatırım kararlarını belirlemeleri ve bireysel emeklilik sisteminin kendisi ve işleyişi konusunda bilgilencmeleri gerekmektedir. Bu nedenle, ister otomatik katılıma ister gönüllü katılıma bağlı olsun sistemdeki mevcut müşterin bilgilendirilmesi ve sürecin doğru aktarılması ve etkin satış tekniklerinin oluşturulması müşterilerin sistemde devamlılığının sağlanması açısından büyük önem teşkil etmektedir. Böylelikle, zorunlu emeklilik sigortasından yeteri kadar gelir elde edemediğini düşünen çalışanların sistemden ayrılmamaları sağlanmış olacaktır. Ayrıca, gönüllü olarak da katılabilecekleri bu sisteme teşvik edilerek sisteme yönlendirdikleri katkı paylarını arttırmaları ve daha fazla tasarruf etmeleri sağlanabilir.

Literatürde BES ile ilgili pek çok çalışma yapılmıştır. Örneğin, Kaydu (2006) çalışmasında, Antalya ve Isparta illerinde BES'e dahil olan 96 katılımcının profilini cinsiyet, yaş, gelir, öğrenim durumları, sisteme yönlendirdikleri katkı payları açısından değerlendirmiştir [3]. İşbilen (2008) çalışmasında İstanbul'da sisteme dahil olan 211 katılımcının BES hakkındaki tutumlarını belirlemeye çalışmıştır [4]. Ayrıca, katılımcıların cinsiyet, yaş grupları, eğitim durumları, medeni durum, mesleki durum, gelir, bağlı olunan güvenlik kurumu gibi demografik bilgileri ile bireysel emeklilik sözleşme profiline göre oluşabilecek farklılıkları incelemiştir. Şener ve Akın (2010) çalışmalarında, sisteme dahil olan 400 katılımcının demografik bilgilerini kullanarak BES'e giriş kararlarını etkileyen faktörleri belirlemeyi amaçlamışlardır. Çalışma ile, katılımcıların eğitim seviyesi ve gelir düzeyi değişkenlerinin sisteme giriş

kararında önemli bir etken olduğu sonucuna ulaşmışlardır [5]. Özer ve Çınar (2012) çalışmalarında bir vakıf üniversitesinde çalışan 289 öğretim üyesinin demografik özelliklerine, toplam çalışmış oldukları yıl sayısı ve gelir düzeyleri gibi değişkenlere göre emeklilik sistemine olan bakış açılarını değerlendirmişlerdir [6]. Yazıcı (2015) çalışmasında katılımcıların emeklilik yatırım fonları ve yatırım alternatifleri hakkında bilgi ve algısını ölçmeyi hedeflemiştir. Araştırma kapsamında, sistemde düzenli katkı payı ödeyen 1014 katılımcının yaş, cinsiyet, çalışma durumu, eğitim durumu, birikim tutarı, sistemde geçirdiği süre (kıdem) ve ikamet ettiği yere ilişkin bilgilerine başvurulmuştur [7]. Bayrak Salantur (2015) çalışmasında, 8.580 katılımcıya ait demografik ve sosyo-ekonomik değişkenlerin (cinsiyet, ödeme oranı, medeni durum, katkı payı, çocuk sayısı ve yaş) sistemden erken ayrılmalarına olan etkisini incelemiştir [8]. Güneş (2015) çalışmasında bireylerin sisteme yönelik tutum ve davranışlarını belirlemek amacıyla Karabük ilinde yaşayan 700 kişiye ait yaş, cinsiyet, aylık ortalama gelir, öğrenim düzeyi, mesleki statü gibi demografik bilgileri ve BES'e katılımı etkileyen diğer faktörleri detaylı incelemiştir [9]. Çömlekçi ve Gökmen (2017), BES'e katılmada etkili olan faktörlerin belirlenmesi ve bu faktörlerin katılımcıların demografik özelliklerine göre farklılık gösterip göstermediğini tespit etmek amacıyla TR42 Bölgesi'nde sisteme dahil olan 384 katılımcı üzerinde çalışmalarını yapmışlardır [10].

Günümüzde giderek yaygınlaşan ve gelişen bir sektör haline gelen bireysel emeklilik sisteminde veri tabanlarında saklı tutulan müşteri verilerinin keşfedilerek anlamlı bilgilerin çıkarılması büyük önem arz etmektedir. Bu nedenle, veri madenciliği yöntemi, bireysel emeklilik sistemi gibi büyük müşteri havuzunda yol gösterici olması açısından ihtiyaç haline gelmiştir. Bu sayede, müşteri davranışlarının ve isteklerinin daha net bir şekilde anlaşılması sağlanarak stratejik ve önemli kararların alınması ve rekabette kalınması açısından şirketlere yarar sağlanmış olacaktır.

Bu çalışmada ise, ülkemizde faaliyet gösteren bir bireysel emeklilik şirketinin 171.625 katılımcısına ait veriler kullanılmıştır. Katılımcıların yaş, cinsiyet, medeni durum, meslek, öğrenim durumu, katkı payı ve ödeme aracı, yaşadığı şehir gibi verilerine göre kümelenmesi ve bu kümelerin tanımlanması ele alınmıştır. Uygulamada, veri madenciliğinin özellikle pazarlama araştırmalarında geniş bir uygulama alanı bulan karma kümeleme analizi yöntemi olan iki aşamalı kümeleme analizi kullanılmıştır.

Bu çalışmanın temel amacı, BES'e dahil olan katılımcıların ödeme ve tasarruf kararlarının çeşitli sosyo-ekonomik ve demografik göstergelerle incelenerek, gruplara ayrılması ve belirlenmiş gruplar arasındaki farklılıkların değerlendirmesini yapmaktır. Literatürde, BES' te kümeleme analizi ile katılımcıların ödeme ve tasarruf bilgilerinin analiz edildiği ve pazarlama stratejisi geliştirildiği çok az çalışma bulunmaktadır. Bu yönüyle çalışmanın literatüre katkı sağlayacağı düşünülmektedir.

2. VERİ MADENCİLİĞİ (DATA MINING)

Günümüzde veri toplama ve depolama teknolojisindeki hızlı gelişmeler kuruluşların büyük miktarda veri biriktirmesine imkân sağlamıştır. Ancak, bu verilerden faydalı bilgilerin elde edilmesi oldukça zordur. Geleneksel veri analizi araçları ve teknikleri böyle büyük boyuttaki bir veri kümesinin analizinde yetersiz kalmaktadır. Bu noktada veri madenciliği önemli bir boşluğu doldurmaktadır. Veri madenciliği, geleneksel veri analiz yöntemleri ile büyük hacimli verileri işlemek için geliştirilmiş algoritmaları harmanlayan bir teknolojidir [11-15]. Chen ve arkadaşları (2006) veri madenciliğini "büyük veri depolarındaki ilginç (önemsiz olmayan, gizli önceden bilinmeyen ve potansiyel olarak yararlı) bilgi veya örüntüleri çıkarma işlemi olarak tanımlamışlardır" [16].

2.1. Kümeleme Analizi (Clustering Analysis)

Veri madenciliğinde kullanılan ilk tekniklerden birisi kümeleme analizidir. Küme analizi veya kümeleme, verilerde bulunan özelliklere göre veriler arasındaki benzerlikleri bulma ve benzer veri nesnelerini ayrık ve homojen gruplar oluşturacak şekilde gruplandırma işlemidir. Diğer bir deyişle, benzer nesnelerin aynı kümede ve benzer olmayan nesnelerin ise başka bir kümede gruplanmasıdır. Veri madenciliğinin temel yöntemlerinden biri olan kümeleme analizi pazar araştırması, makine öğrenimi, biyoinformatik, görüntü işleme, kalıp tanıma, veri sıkıştırma, bilgisayar grafikleri gibi pekçok alanda kullanılan istatistiksel bir tekniktir [17].

Veri Madenciliğinde kullanılmakta olan çok sayıda kümeleme algoritması bulunmaktadır. Kümeleme algoritmasının seçimi hem analiz edilecek verinin yapısına hem de amaç ve uygulamaya bağlıdır [18-19]. Literatürde yaygın kullanılan kümeleme algoritmaları, hiyerarşik (aşamalı) ve hiyerarşik olmayan kümeleme tekniği olmak üzere iki başlık altında toplanmıştır [20-21]. Hiyerarşik kümeleme yönteminde, gruplayıcı (agglomerative) ve bölücü (divisive) olmak üzere iki yöntem vardır. Gruplayıcı hiyerarşik yönteminde her birim başlangıçta bir küme olarak kabul edilir. Daha sonra en yakın iki küme yeni bir kümede birleştirilir. Bu şekilde, her adımda küme sayısı bir azaltılır. Oluşan küme yapıları, "dendogram" veya ağaç grafiği ile gösterilebilir [22]. Hiyerarşik olmayan kümeleme tekniklerinde ise küme sayısı hakkında ön bilgiler vardır veya araştırmacı küme sayısına karar vermektedir.

Çalışmada kullanılan İki Aşamalı Kümeleme (Two step cluster) yöntemi ise hiyerarşik olmayan kümeleme tekniklerinden "k Ortalamalar" ve hiyerarşik tekniklerinden ise "Ward'ın En Küçük Varyans" tekniklerinin birleştirilmesi ile oluşan hibrid kümeleme tekniğidir. Klasik kümeleme algoritmaları ile kıyaslandığında iki aşamalı kümeleme yöntemi daha öznetelikli kategoriler sağladığı için farklı disiplinlerden birçok araştırmacı tarafından uygulama alanı bulmuştur [23-24]. Örneğin, Choi ve arkadaşları (2004)

çalışmalarında multihop kablosuz algılayıcı ağları için iki aşamalı kümeleme yöntemini çalışmışlardır. İki aşamalı kümeleme yönteminin sağladığı yeni kümeleme yapılandırması yoluyla önemli ölçüde enerji tasarrufu sağlamışlardır [25]. Şchiopu (2010) çalışmasında bir bankanın mevcut ve olası müşterilerini etkin bir şekilde yönetmek için, iki aşamalı kümeleme analizi ile müşterileri 3 kümeye ayırmıştır [26]. Namvar ve arkadaşları (2010) iki aşamalı kümeleme yöntemini kullanarak banka müşterilerini eğitim seviyesi, yaş, meslek gibi demografik bilgilerine göre kümeler ayırmıştır [27]. Savaş ve Topaloğlu (2011) ise 3 farklı GSM şebekelerinin çekim gücü verilerini cep telefonlarından alarak veri tabanı oluşturan bir program geliştirmişlerdir. Programdan elde edilen verileri iki aşamalı kümeleme yöntemi ile analiz ederek 3 farklı kümeye ayırmış ve her 3 şebeke için sonuçları yorumlamışlardır [28]. Yılmaz (2012) çalışmasında üniversite öğrencilerinin eğlence ya da iletişim amacıyla, internet kullanımına göre üniversite öğrencilerinin profillerini belirlemek ve internetteki ilgisine bağlı olarak profillerinin farklı olup olmadığını anlamak için iki aşamalı kümeleme analizini çalışmıştır. Çalışmanın sonuçları, üniversite öğrencilerinin, eğlence ve iletişim amacıyla internet kullanımı anlamında iki profile bölündüğünü ve İnternet'e olan ilginin bu bölünme üzerinde büyük etkiye sahip olduğunu göstermektedir [29]. Çetin ve Irmak (2014) çalışmalarında iki aşamalı kümeleme analizini 40 farklı üniversitede yer alan akademisyenlerin elektronik alışverişte güvenlik ve risk algılarını belirlemek için kullanmışlardır. Kümeleme analizi sonucunda, katılımcıların risk algılarına göre 4 farklı kümede toplandığı sonucuna varmışlardır [30]. Arı ve arkadaşları (2016) çalışmalarında iki aşamalı kümeleme yöntemini, Borsa İstanbul (BİST)'da işlem gören 90 adet firmanın kamuya açık finansal tablolarından belirlenmiş olan finansal oranlarının performanslarını incelemek için kullanmışlardır [31]. Giray (2016) çalışmasında bir cezaevinin hükümlü kütüğü verilerinden faydalanarak hükümlülerin profillerini belirlemek amacıyla iki aşamalı kümeleme yöntemini kullanmıştır. Çalışma kapsamında yer alan *yaş*, *suç türü*, ve *suçun işlendiği bölge* gibi değişkenlere göre 4 farklı küme belirlenmiş, analiz sonucu oluşan kümelerin birbirlerine benzer ve birbirlerinden farklı özellikleri detaylı irdelenmiştir [32]. Kayri ve Günüş (2016) çalışmalarında iki aşamalı kümeleme analizi kullanarak bireylerin internet bağımlılık düzeylerini yaşadığı coğrafya, kültür ve sosyoekonomik koşullara göre 3 grupta incelemiştir [33].

2.2. İki Aşamalı Kümeleme Analizi (Two-Step Clustering Analysis)

İki aşamalı kümeleme yöntemi, çok büyük veri kümelerini işlemek üzere tasarlanmış ölçeklenebilir bir kümeleme analizi algoritmasıdır [34-35]. Bu algoritmanın en önemli özellikleri; büyük veri setlerinde uygulanabilmesi, kategorik ve sürekli değişkenleri işleyebilmesi, otomatik olarak en uygun küme sayısını belirleyebilmesi ve elde edilen kümelerle uyum

sağlamayan gözlemlerin istendiğinde veriden ayıklanabilmesidir. İki aşamalı kümeleme analizinde, süreç *ön kümeleme* ve *kümeleme* olmak üzere iki ana aşamadan oluşmaktadır:

Ön kümeleme aşamasında, gözlemler küçük alt kümeler olarak ön kümeleme yapılır ve daha sonra bu alt kümeler ayrı gözlemler olarak ele alınır. Gözlemin önceden oluşturulmuş bir kümeden içerisinden mi yoksa yeni bir kümeden mi katılacağı kararı uzaklık kriterlerine dayanarak hiyerarşik küme yöntemi ile yapılmaktadır.

Kümeleme aşamasında ise, ön kümelemeden kaynaklanan alt kümeler, gerekli küme sayısına göre gruplanır. Alt kümelerin sayısı gözlem sayısından daha küçük olduğundan, geleneksel gruplama yöntemlerinin kullanımı kolaydır. İki aşamalı kümeleme yöntemi kümelerin sayısını otomatik belirleyen aglomeratif hiyerarşik bir yöntem kullanır. Hiyerarşik kümeleme yöntemi, işlemin sonunda yalnızca tüm gözlemleri içeren bir küme kalana kadar kümelerin ardışık olarak birleştirildiği bir işlemi ifade etmektedir. İşlem, ön kümeleme aşamasında üretilen alt kümelerin her biri için bir başlangıç kümesi tanımlayarak başlamaktadır. Daha sonra tüm kümeler karşılaştırılır ve aralarındaki uzaklığı en küçük olan küme çifti seçilir ve tek bir kümeye birleştirilir. Birleşmeden sonra, yeni küme setleri karşılaştırılır; en yakın küme çifti birleştirilir ve süreç, tüm kümeler birleşene kadar tekrarlanır. Kümeler bu şekilde tekrar tekrar bir araya getirildiğinden, farklı küme sayılarıyla çözümleri karşılaştırmak daha kolaylaşmaktadır. İki aşamalı kümeleme yönteminin aşamalarında kullanılan işlemler aşağıda ifade edilmiştir:

2.2.1. Uzaklık ölçüsü (Distance Measure)

İki aşamalı kümeleme analizinin her iki aşamasında uzaklık ölçüsüne ihtiyaç duyulmaktadır. Eğer, verisinde bir veya daha fazla değişken kategorik ise, olasılığa dayalı mesafeyi temsil eden log-olabilirlik uzaklık ölçüsü kullanılmaktadır. Log-olabilirlik uzaklık hesaplamasında, sürekli değişkenler için normal dağılım ve kategorik değişkenler için ise multinomial dağılım varsayımı yapılmaktadır. Ayrıca, değişkenlerin birbirinden bağımsız olduğu varsayılmaktadır. Tüm değişkenlerin sürekli olduğu durumda ise, Öklid uzaklığı kullanılır, böylece gözlemler en küçük Öklid uzaklığına sahip kümede gruplanır. SPSS paket programı, kategorik ve sürekli değişkenlerle uyumlu olduğundan, kümeleri birleştirmek üzere mesafe ölçüsü olarak log-olabilirlik uzaklık ölçüsü kullanılmaktadır [36-37]. Kümeler arasındaki log-olabilirlik uzaklığı (1), (2) ve (3) nolu eşitlik ile hesaplanmaktadır:

$$d(i, j) = \xi_i + \xi_j - \xi_{\langle i, j \rangle} \quad (1)$$

burada, i ve j kümeler, ξ_i ise i kümesi içinde bir çeşit varyans (dağılım) olmak üzere, $d(i, j)$ i ve j kümeleri arasındaki log-olabilirlik uzaklığını, $\langle i, j \rangle$ ise i ve j kümelerini birleştirerek oluşturulan kümeyi temsil eder.

$$\xi_j = -N_j \left(\sum_{k=1}^{K^A} \frac{1}{2} \log(\sigma_k^2 + \sigma_{jk}^2) + \sum_{k=1}^{K^B} E_{jk} \right) \quad (2)$$

burada, j küme ve k değişken (kategorik veya sürekli) olmak üzere, N_j j kümesindeki gözlem sayısını, K^A toplam sürekli değişken sayısını, K^B toplam kategorik değişken sayısını, σ_k^2 k sürekli değişkeninin (tüm veri seti bazında) tahminlenmiş varyansını (yayılımı), σ_{jk}^2 k sürekli değişkeninin j kümesindeki tahminlenmiş varyansını temsil etmektedir.

$$E_{jk} = - \sum_{l=1}^{L^K} \frac{N_{jkl}}{N_j} \log \frac{N_{jkl}}{N_j} \quad (3)$$

burada, l kategori olmak üzere L^K k . kategorik değişkeninin kategori sayısını, N_j j kümesindeki gözlem sayısını, N_{jkl} l kategorili k değişkeninin bulunduğu j kümesindeki gözlem sayısını ifade etmektedir.

2.2.2. Küme sayısı: otomatik küme (Number of clusters: auto-cluster)

İki Aşamalı Kümeleme Analizi, küme sayısı konusunda bir ön bilginin olmadığı durumlarda kullanılabilir. Bu analiz tekniğinde optimum küme sayısı yöntem tarafından belirlenir. Küme sayısının otomatik olarak belirlenmesinde Bayesçi Bilgi Ölçütü (Bayesian Information Criterion-BIC) veya Akaike Bilgi Ölçütü (Akaike Information Criterion-AIC) kullanılmaktadır. BIC ve AIC denklemleri eşitlik (4), (5) ve (6) ile hesaplanmaktadır.

$$BIC(j) = -2 \sum_{j=1}^J \xi_j + m_j \log(N) \quad (4)$$

$$AIC(j) = -2 \sum_{j=1}^J \xi_j + 2m_j \quad (5)$$

$$m_j = J \left\{ 2K^A + \sum_{k=1}^{K^B} L^K - 1 \right\} \quad (6)$$

burada j küme k değişken olmak üzere, K^A toplam sürekli değişken sayısını, K^B toplam kategorik değişken sayısını, L^K ise k kategorik değişkenindeki kategori sayısını temsil etmektedir.

2.2.3. Önem ölçütü (Importance measure)

Kümeleri oluşturan değişkenlerin göreceli katkısı (önemi), her iki değişken türü (sürekli ve kategorik) için ayrı ayrı hesaplanmaktadır. Önem değerleri 0-1 arasında derecelendirilmektedir. 0 kümeleri belirlemede en önemsiz değişkeni ve 1 ise son derece önemli değişkeni ifade etmektedir. Önem ölçütü, sürekli değişkenler için t

testine, kategorik değişkenler için ise chi-square (χ^2) anlamlılık testine dayanmaktadır. Sürekli değişkenler için önem ölçütünün hesaplanması eşitlik (7)' de gösterilmiştir:

$$t = \frac{\mu_k - \mu_{jk}}{\sigma_{jk}} \sqrt{N_k} \quad (7)$$

burada, j küme ve k sürekli değişken olmak üzere, N_k tüm veri seti içindeki sürekli değişken sayısını, μ_k k sürekli değişkeninin ortalamasının tahmincisi, μ_{jk} k sürekli değişkeninin j kümesindeki tahminlenmiş ortalamasını, σ_{jk} k sürekli değişkeninin j kümesindeki tahminlenmiş varyansını ifade etmektedir. Kategorik değişkenler için önem ölçütünün hesaplanması eşitlik (8)' de gösterilmiştir:

$$x^2 = \sum_{l=1}^{L^K} \left(\frac{N_{jkl}}{N_{kl}} - 1 \right)^2 \quad (8)$$

burada, l kategori, k kategorik değişken, ve j küme olmak üzere, L^K k . kategorik değişkeninin kategori sayısını, N_{jkl} l kategorili k kategorik değişkeninin bulunduğu j kümesindeki gözlem sayısını, N_{kl} l kategorili k kategorik değişkeninin sayısını göstermektedir.

3. UYGULAMA (APPLICATION)

Çalışmada, Türkiye'de faaliyet gösteren bir emeklilik şirketinin 2014 yılı içerisinde bireysel emeklilik planı açmış 2017 yılına kadar ödemelerine devam eden aktif müşterileri ele alınmıştır. Kullanılan veri bireysel emeklilik sisteminde özel müşteri olarak bahsedilen (yüz bin lira ve üstü fon sahibi olan veriler hariç) tüm plan türlerini içeren bir yapıda hazırlanmıştır. Söz konusu şirketin veritabanından, sözleşme ve müşteri bilgileri gizliliği ihlal edilmeksizin emeklilik başlangıç tarihi göz önüne alınarak 171.625 katılımcının demografik, sosyo-kültürel bilgilerinin yanısıra ödeme ve tasarruf bilgileri alınmıştır. Çalışmada, katılımcı verilerinin öncelikle frekans tabloları ve betimleyici istatistikleri incelenmiş, ardından katılımcılar arasındaki homojen alt grupların araştırılması için verilere İki Aşamalı Kümeleme Analizi uygulanmıştır. Tablo 1' de analizde kullanılan değişkenlere ait frekans ve yüzde dağılımları oluşturulmuştur. Bu dağılımlara göre; katılımcıların %64.35'ini erkekler oluştururken, %35.65'ini kadınlar oluşturmaktadır.

Medeni durum değişkenine bakıldığında evli katılımcıların çoğunlukta (%60.49) olduğu görülmektedir. Verisinde yer alan katılımcıların yaş grubuna göre dağılımları incelendiğinde BES katılımcıları için en düşük yaşın 18, en büyük yaşın 76 ve yaş ortalamasının ise 40.61 olduğu görülmektedir. Ancak, analizi açıklama ve yorumlama açısından yaş değişkeni, Emeklilik Gözetim Merkezinin resmi sitesinde yer alan 2015 yılı BES gelişim raporuna paralel olarak kategorik hale getirilmiştir [38].

Tablo 1. Değişkenlere ait frekans ve yüzde dağılım değerleri
(Frequency and percentage distribution values of variables)

Katılımcı Özellikleri		Frekans	Yüzde (%)
Cinsiyet	Erkek	110434	64.35
	Kadın	61191	35.65
Medeni Durum	Evli	103818	60.49
	Bekar	67807	39.51
Öğrenim Durumu	Ortaöğretim ve Lise	13237	7.71
	Önlisans ve Lisans	145186	84.59
	Lisansüstü	13202	7.69
Aylık Katkı Payı	250 TL ve altı	141812	82.63
	251-500 TL	21573	12.57
	501-750 TL	4421	2.58
	751- 1000 TL	836	0.49
	1000 TL ve üzeri	2983	1.74
Ödeme Aracı	Havale	12590	7.34
	Kredi kartı	93597	54.54
	Otomatik ödeme	65438	38.13
Meslek	Mühendis	22416	13.06
	İşçi	22403	13.05
	Serbest Meslek	22337	13.02
	Avukat	14907	8.69
	Doktor	14894	8.68
	Devlet Memuru	14904	8.68
	Öğretmen	14965	8.72
	Asker	7472	4.35
	Sanatçı	7451	4.34
	Mimar	7462	4.35
	Esnaf	7473	4.35
	Özel sektör çalışanı	7462	4.35
	Akademisyen	7479	4.36
Kalkınmışlık	1. Kademe gelişmiş iller	107226	62.48
	2. Kademe gelişmiş iller	22778	13.27
	3. Kademe gelişmiş iller	18435	10.74
	4. Kademe gelişmiş iller	10876	6.34
	5. Kademe gelişmiş iller	7391	4.31
	6. Kademe gelişmiş iller	4919	2.87
Yaş	25 yaş altı	5736	3.34
	25 ve 34 yaş	43084	25.1
	35 ve 44 yaş	66486	38.74
	45 ve 55 yaş	43434	25.31
	56 yaş ve üstü	12885	7.51

Böylece, belirli bir ortalama değerdense, farklı yaş grupları arasındaki farklı tasarruf tutumları hakkında yorum yapmak daha anlamlı olacaktır. Tablo 1'e göre, katılımcıların büyük çoğunluğu (%38.74) 35-44 yaş aralığında bulunmaktadır. 25 yaş ve altı, %3.34 oranı ile en az katılımcının bulunduğu kategori durumundadır. Ayrıca, mühendis (%13.06), işçi (%13.05) ve serbest meslek (%13.02) gruplarına sahip katılımcıların çoğunlukta olduğu görülmektedir.

Öğrenim durumu değişkeni veri setinde, ortaöğretim ve lise, Önlisans ve lisans ve lisansüstü mezunu olmak üzere 3 kategori halinde bulunmaktadır. Katılımcıların eğitim seviyesi incelendiğinde Önlisans ve Lisans mezunu olanların büyük bir çoğunlukta olduğu görülmektedir (%84.59). Bu orana göre, eğitim seviyesinin yükseldikçe

uzun vadeli tasarruf eğiliminin de arttığını görmekteyiz. *Aylık katkı payı değişkeni*; katılımcıların tasarrufları için yaptıkları düzenli aylık ödeme miktarını göstermektedir.

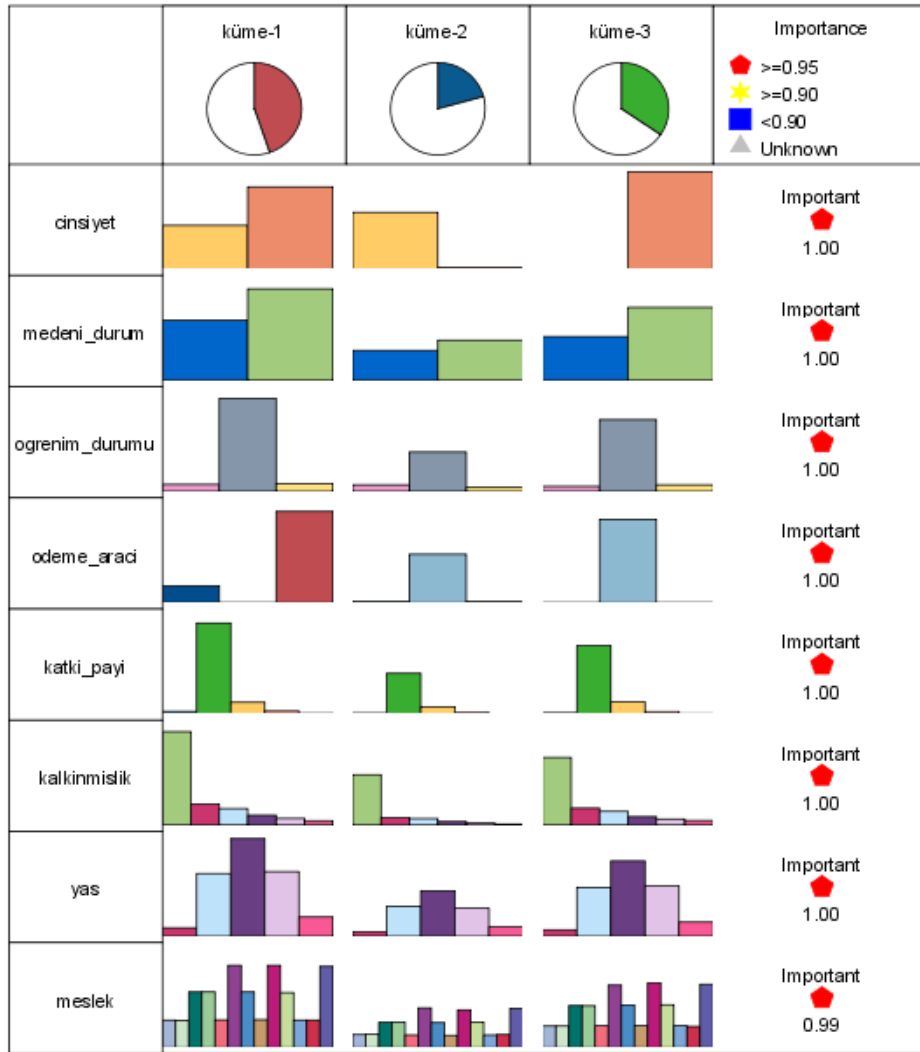
Veriseti içinde aylık 250 TL ve altı katkı payı ödeyenlerin oranı %82.63 iken, 1000 TL ve üstü ödeme planı olanların ise %1.74' tür. Bu durum, katılımcıların her ay küçük miktarda birikim yaparak ilerlemeyi tercih ettiğini göstermektedir. Ayrıca, katkı payları sözleşmelerde belirtilen havale, kredi kartı ve banka şubesine verilen otomatik ödeme talimatları gibi ödeme araçları ile yapılabilmektedir. Söz konusu verisetinde, katılımcıların, aylık ödemelerinde kredi kartını sıklıkla (%54.54) tercih ettiği görülmektedir. Katılımcıların en yoğun olduğu ilk dört il sırasıyla İstanbul (%30.61), Ankara (%9.77) ve İzmir (%6.63) ve Antalya'dır (%5.13).

Kalkınmışlık değişkeni ise, katılımcıların yaşadığı şehir bilgisi baz alınarak oluşturulmuştur. İkamet ili yurtdışı olan 101 adet katılımcı verisetinden çıkarılmıştır. Verisetinde yer alan şehir sütunu, Kalkınma Bakanlığı'nın resmi sitesinde yayınlanan "İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması" isimli raporunda yer alan illerin gelişmişlik seviyeleri sınıflamasına göre kategorilere ayrılmıştır.

İllerin gelişmişlik düzeyi belirlenirken "ilin ülke içerisindeki ekonomik ağırlığı, bireysel zenginlik ve refah düzeyi, sosyal gelişmişlik seviyesi, il ölçeğinde ekonomik ve sosyal gelişmişlik ile bireysel refah arasındaki denge" gibi değişkenlerle göz önüne alınmıştır [39]. Bu yöntemle oluşturulan kalkınmışlık değişkeni toplam altı kademeli bir gelişmişlik dağılımı göstermektedir. Buna göre, 1. kademe gelişmiş 8 il, 2. kademe gelişmiş 13 il, 3. kademe gelişmiş 12 il, 4. kademe gelişmiş 17 il, 5. kademe gelişmiş 16 il ve 6. kademe gelişmiş 15 il bulunmaktadır. Tablo 1'de de görülebileceği gibi BES'e dahil olan katılımcıların sayısı illerin gelişmişlik seviyesiyle ters orantılı şekilde azalma eğilimi göstermektedir. Buna göre, katılımcıların büyük çoğunluğu %62.48' i oranı ile 1. kademe gelişmiş illerde yaşamaktadır.

İki aşamalı kümeleme analizi, SPSS'in veri madenciliği modülü olan SPSS Clementine 12.0 programı kullanılarak yapılmıştır [40]. İki aşamalı kümeleme yöntemi, küme sayısı hakkında önbilginin olmadığı ve özellikle büyük veri setinde sürekli ve kategorik değişkenlerin bir arada bulunduğu durumlarda kolaylıkla kullanılabilen bir kümeleme yöntemidir.

Bu çalışmada, iki aşamalı kümeleme analizinin kullanılmasının temel amacı ise; ele alınan emeklilik şirketinin veritabanında yer alan büyük hacimli müşteri verisetini benzer özelliklere (değişkenlere) göre aynı kümede toplayarak farklı grupları belirlemek ve bu gruplar üzerindeki değişkenlerin önemini görebilmektir. Böylece, mevcut müşterileri farklı gruplara ayırarak her bir gruba özel müşteriler hakkında stratejik kararlar alınmasında şirkete yardımcı olması beklenmektedir. Sahip olunan verinin büyüklüğü de göz önüne alındığında kümelerin anlamlı ve güvenilir sonuçlar vermesi beklenmektedir. İki aşamalı kümeleme analizinin önemli avantajlarından biri, küme sayısının otomatik olarak seçilmesini sağlamasıdır.



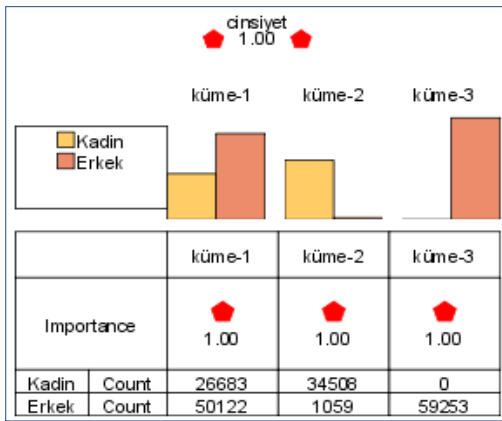
Şekil 1. İki aşamalı kümeleme yöntemi ile elde edilen kümeler
(Obtained clusters by two-step cluster analysis)

Program, 2 ile 15 küme arasında geniş bir aralık sunarak küme sayısını otomatik olarak bulmaktadır. Şekil 1'de görüldüğü gibi, bireysel emeklilik sistemine başvuruda bulunan 171.625 katılımcı üç kümeye bölünmüştür. Bu üç küme sırasıyla, 76805 (%44.75), 35567 (%20.72), ve 59253 (%34.52) katılımcıdan oluşmaktadır. Çalışmaya her ay düzenli olarak katkı payı ödeyen katılımcılar dâhil edilmiş, homojenliğin sağlanması açısından diğer katılımcılar veri setinden çıkarılmıştır. SPSS Clementine, her değişkenin (kategorik veya sürekli) küme üzerindeki önem seviyesi hakkında bilgi vermektedir. Programda, önem seviyesi 0.95 ve üzerinde olan değişkenlerin küme üzerinde önemli etkisinin olduğu kabul edilmektedir.

3.1. Bireysel Emeklilik Müşterilerinin İki Aşamalı Kümeleme Analizi ile Değerlendirilmesi (Evaluation of Individual Pension Customers by Two-Step Cluster Analysis)

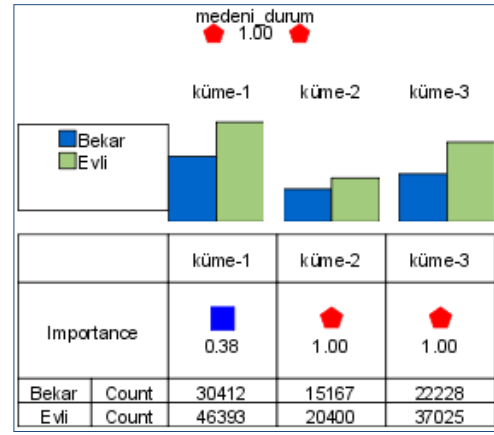
Bu aşamada, analiz sonucu oluşan üç kümede yer alan katılımcı profilleri genel hatlarıyla açıklanmıştır. Kümelerin yorumlanmasında katılımcıların cinsiyet, medeni durum, yaş, meslek, öğrenim durumu, yaşadığı şehir bilgilerinin yanısıra ödeme ve tasarruf bilgileri analiz için bize yönlendirme yapacaktır.

Cinsiyet değişkeni, Şekil 2'de görüldüğü gibi her üç küme için önemli çıkmıştır. Birinci kümede erkek (%65.26), ikinci kümede ise kadın katılımcıların (%97.02) çoğunlukta olduğu görülmektedir. Üçüncü kümenin tamamı ise sadece erkek katılımcılardan oluşmaktadır.



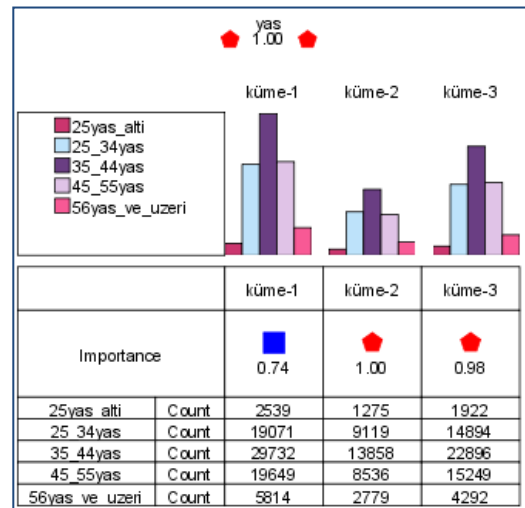
Şekil 2. Cinsiyet değişkeninin kümeler dağılımı
(Distribution of gender variable into clusters)

Medeni durum değişkenini ikinci ve üçüncü kümeler için anlamlı iken, birinci küme için önemsiz olduğu görülmektedir. Şekil 3' te de görüldüğü gibi her üç kümede evli katılımcılar çoğunlukta. Buna göre, evli olmak bireysel emeklilik ve tasarruf kararları için önemli bir değişkendir.



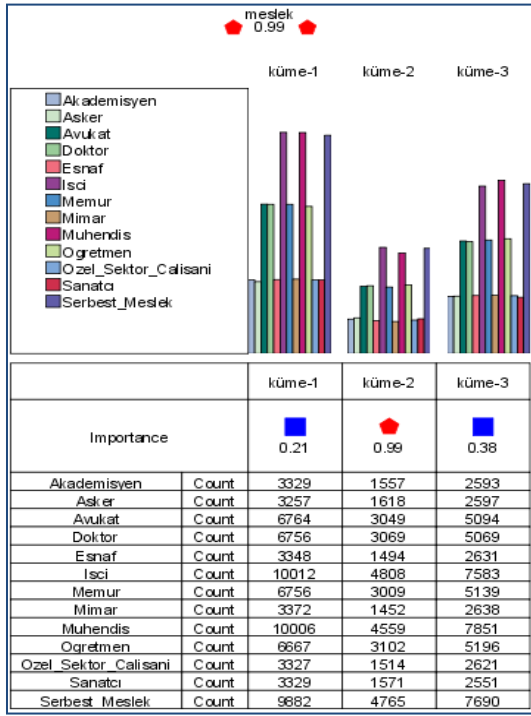
Şekil 3. Medeni durum değişkeninin kümeler dağılımı
(Distribution of marital status variable into clusters)

Şekil 4' te yer alan yaş değişkeni ikinci ve üçüncü kümeler için anlamlı iken, birinci küme üzerinde önemsiz olduğu görülmektedir. Sırasıyla, %38.71, %38.96 ve %38.64 oranı ile her üç kümede, BES'e katılımın en çok olduğu yaş grubunun 35-44 yaş aralığı olduğunu görmekteyiz. Bu grubu, birinci kümede %25.58 ile 45-55 yaş grubu, ikinci kümede %25.64 ile 25-34 yaş grubu ve üçüncü kümede %25.74 ile 45-55 yaş grubu izlemektedir. 56 yaş ve sonrası katılımcı sayısı oldukça düşüktür. Buna göre, ortaya çıkan kümelerin yaş değişkeni açısından dengeli olduğu söylenebilir. Ayrıca, bilindiği üzere bir katılımcının BES'ten emekli olabilmesi için 56 yaşını doldurmuş ve sistemde 10 yıl düzenli ödeme yapmış olması gerekmektedir.



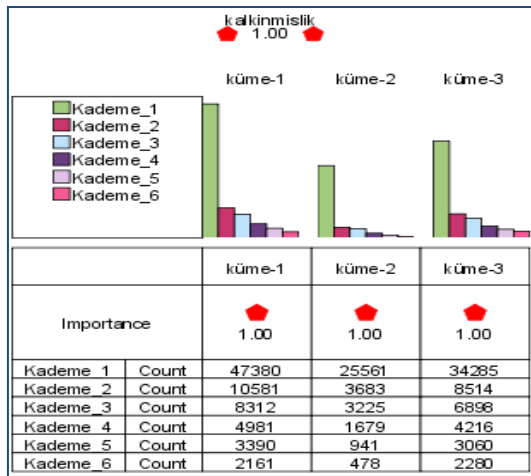
Şekil 4. Yaş değişkeninin kümeler dağılımı
(Distribution of age variable into clusters)

Meslek değişkeni, sadece ikinci küme için önemli iken, 1 ve 3. kümeler üzerinde etkili değildir. Şekil 5'te de görüldüğü üzere, birinci ve ikinci kümelerde sırasıyla %13.04 ve %13.52 oranıyla işçilerin çoğunlukta olduğu görülmektedir. Üçüncü kümede ise, %13.25 oranıyla mühendis ve %12.98 oranıyla serbest meslek grubundaki katılımcıların diğer meslek grubundaki katılımcılara kıyasla fazla olduğu görülmüştür.



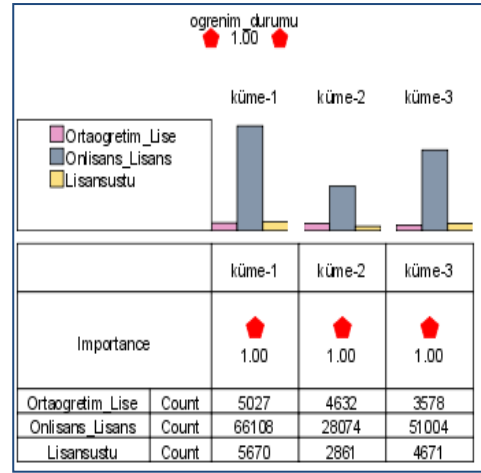
Şekil 5. Meslek değişkeninin kümelere dağılımı
(Distribution of occupation variable into clusters)

Katılımcı kümelerinin kalkınmışlık değişkenine göre kıyaslanması Şekil 6'da yer almaktadır. Şekil 6 incelendiğinde kalkınmış değişkeni her üç küme üzerinde önemli bir değere sahiptir. Her üç kümede yer alan katılımcıların büyük bir kısmı en çok gelişmiş 8 şehri ifade eden 1. kademe kategorisinde yer almaktadır. Bu kategorinin üç küme içerisindeki oranı sırasıyla % 61.69, %71.87 ve %57.86 şeklindedir. Diğer kademe gelişmiş illerde yaşayan katılımcıların sayısı ise, katılımcıların yaşadıkları şehirlerin gelişmişlik seviyesi düştükçe azalmaktadır. Buna göre, katılımcıların yaşadıkları şehirlerin sosyoekonomik ve demografik özelliklerinin farklı olması, katılımcıların tasarruf yapma kararlarını doğrudan etkilediği görülmektedir. Katılımcıların birinci kademe iller arasında yer almasının önemli sebeplerinden biri eğitim alanında gelişmişlik seviyesinin yüksek olmasıdır.



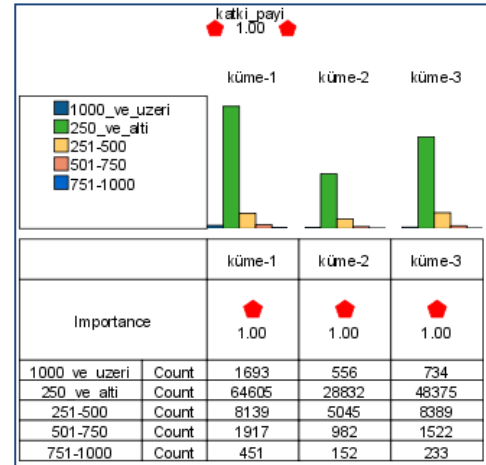
Şekil 6. Kalkınmışlık değişkeninin göre kümelere dağılımı
(Distribution of development variable into clusters)

Şekil 7' de görüldüğü gibi *öğrenim durumu* değişkeni, üç kümeyi etkileyen önemli bir değişkendir. Her üç kümede, önlisans ve lisans mezunu katılımcıların sayısı, diğer kategorilere kıyasla çok fazladır.



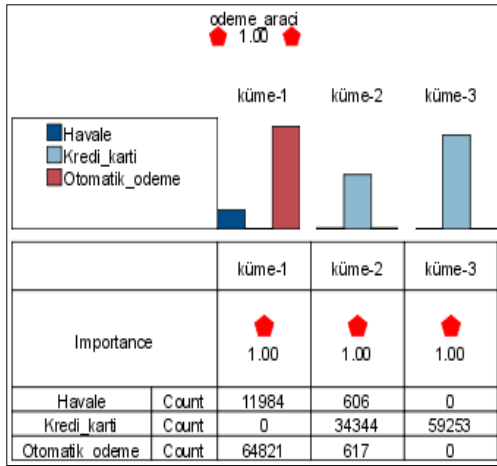
Şekil 7. Öğrenim durumu değişkeninin kümelere dağılımı
(Distribution of education status variable into clusters)

Her üç küme üzerinde önemli olan değişkenlerden biri de beş kategoriden oluşan *katkı payı* değişkenidir. Şekil 8'de de görüldüğü gibi her üç kümede yer alana katılımcıların %80' inden fazlası sözleşmelerini imzalarken düşük düzeylerde katkı payı ödemeyi tercih etmektedir. Bu duruma sebeplerden biri, ülkemizdeki bireylerin BES'i tasarruf aracı olarak yeterince anlayamaması ve tasarruflarını altın, döviz, hazine bonosu/devlet tahvili ve gayrimenkul gibi genel yatırım araçlarına yatırmalarından kaynaklandığı şeklinde yorumlanabilir.



Şekil 8. Katkı payı değişkeninin kümelere dağılımı
(Distribution of contribution rate variable into clusters)

Şekil 9'da belirtildiği gibi *ödeme aracı* değişkeni tüm kümeler için önemli bir değişkendir. 3. kümede yer alan katılımcıların tamamının ödemelerini kredi kartı ile yaptığını görmekteyiz. 1. kümede yer alan katılımcıların büyük çoğunluğu ödemelerini %84.4 oranı ile banka şubesine verilen hesaptan otomatik ödeme talimatı ve %15.6 oranı ile havale yoluyla yapmaktadırlar.



Şekil 9. Ödeme aracı değişkeninin kümelere dağılımı
(Distribution of payment type variable into clusters)

4. SONUÇ (RESULT)

Katılımcı hacmi ve fon büyüklüğü açısından istikrarlı bir şekilde büyümeye devam eden bireysel emeklilik sektöründe, müşteriye sunulan ürün ve hizmetin artırılması müşteri sürekliliğinin sağlanması ve müşteri kayıplarının önüne geçilmesi açısından önem arz etmektedir. Ayrıca, sistem içerisinde bulunan mevcut müşterilerin katkı paylarını arttırması ve ek sözleşmeler açması bu sektörde faaliyet gösteren şirketlerin pazar payı, karlılık ve fon büyüklüğü arttırmaları açısından oldukça önemlidir. Sektörün hızla büyümesi yeni şirketlerin pazara girmesine, istihdam artışına ve buna bağlı olarak yeni bir rekabet pazarına yol açacaktır. Bu nedenle, bireysel emeklilik şirketleri öncelikli olarak veri madenciliği teknikleriyle geliştirilen müşteri odaklı analiz ve yaklaşımlar geliştirmelidirler. Böylelikle, şirketlerin daha karlı ve müşteri memnuniyeti yüksek bir şekilde devam edebilmeleri için sistemi iyi tasarlanmış ve doğru tahminler içeren bir yapıda olması gerekmektedir.

Bu çalışmada, veri madenciliğinin önemli yöntemlerinden biri olan kümeleme analizi ile bireysel emeklilik sektöründe faaliyet gösteren bir şirkete dahil olan müşteriler demografik, sosyo-kültürel bilgilerinin yanı sıra ödeme ve tasarruf bilgilerine göre 3 farklı kümeye ayrılarak analiz edilmiştir. Analiz sonuçlarına göre, katılımcıların tasarruf yapma kararlarını doğrudan etkileyen önemli değişkenler kalkınmışlık, öğrenim durumu, ödeme aracı ve cinsiyet olup, her üç küme için de önemli değişkenlerdir. Ancak, katkı payı, yaş, medeni durum ve meslek değişkenleri her üç küme için eşit öneme sahip değildir. Buna göre, hedef müşteri grubunun daha net bir şekilde anlaşılması için katılımcının yaşadığı şehir, eğitim seviyesi, katkı payı ödeme şekli ve cinsiyet bilgisi önemli değişkenler olup stratejik ve önemli kararların alınması açısından karar vericilere (şirketlere) faydalı olması beklenmektedir.

Kümelerdeki katılımcı profilleri incelendiğinde; 1. ve 2. kümelerde katılımcıların büyük çoğunluğunu erkekler oluştururken, 3. kümeyi ise sadece kadın katılımcılar oluşturmaktadır. Ayrıca, her üç kümede de sistemi

çoğunlukla çalışan, evli ve eğitim düzeyi yüksek kişilerce tercih edildiği görülmüştür. Sistemdeki 35-44 yaş grubu aralığının toplam katılımcıların %38.74' ünü oluşturması bu yaş aralığındaki katılımcıların geleceği açısından emeklilik sistemine daha olumlu yaklaşımlarına işaret ettiği sonucuna varılmıştır. Bunların yanı sıra, katılımcıların yaşadıkları şehirlerin kalkınmışlık seviyesi arttıkça, katılımcıların daha çok tasarruf yapabilme eğilimi gösterdikleri görülmüştür. Ödeme türleri konusunda çoğunluğun kredi kartı üzerinden ödeme yapması müşterilerin büyük bir kısmının banka satış kanalı ve anlaşmalı bankalar ile sisteme dahil olduğunun sonucudur. Bu nedenle, sistemin bankacılık sistemi ile entegre gitmesi sürecin devamlılığı ve yasal takibi açısından oldukça önemlidir. Ayrıca, günümüzde kredi kartları sadece tüketim faaliyeti sürecinde yer alırken bu sistem ile tasarruf edinme sürecinde de yer alabildiği sonucuna varılabilir. Her üç kümede de sisteme dahil olan katılımcıların çoğunluğu tasarruflarını aylık 0-250 tl katkı payı ile düşük birikimlerle yaptığını görmekteyiz. Bu durum, toplumda henüz tasarruf kültürü ve emeklilik bakış açısına sahip olunmadığının ve sisteme duyulan güvenin yetersiz olduğunun kanıtıdır.

Sonuç olarak, BES' in katılımcılara daha etkin ifade edilmesi, kamunun bu sisteme güveninin artırılması ve tanıtım çalışmalarına daha fazla ağırlık verilmesi gerekmektedir. Özellikle, araştırma neticesinde de görüldüğü gibi uygun satış pazarlama stratejilerinin hazırlanmasında şirketlerin katılımcı tabanını nispeten düşük eğitilmiş ve düşük kalkınmışlık düzeyine sahip bölgelerde yaşayan katılımcıların lehine genişletmesi gerektiği sonucuna ulaşılabilir. Bu nedenle, bireysel emeklilik sisteminin kapsamına ilişkin daha detaylı reformların yapılması gerekmektedir.

KAYNAKLAR (REFERENCES)

- [1] Emeklilik Gözetim Merkezi. "BES Temel Göstergeleri", <http://www.egm.org.tr/webink/BESgostergeler.asp>, 01.05.2017.
- [2] The Organization for Economic Co-operation and Development (OECD). "Pension Markets in Focus, 2016". <http://www.oecd.org/daf/fin/private-pensions/Pension-Markets-in-Focus-2016.pdf>, 01.05.2017.
- [3] Z. Kaydu, **Bireysel Emeklilik Sisteminde Tüketici Profillerinin Farklılaşması: Antalya ve Isparta İllerinin Karşılaştırılması Örneği**, Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü, Isparta, 2006.
- [4] E. İşbilen, **Bireysel Emeklilik Sistemi ve Türkiye Uygulaması**, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul, 2008.
- [5] O. Şener, F. Akın, **Özel Emeklilik Fonları ve Türkiye'de Bireylerin Bireysel Emeklilik Sistemine Giriş Kararlarını Etkileyen Faktörlerin Belirlenmesi Üzerine Bir Araştırma**, *Marmara Üniversitesi İ.İ.B.F. Dergisi*, 28 (1), 291-312, 2010.
- [6] Ö. Özer, E. Çınar, "Bir Vakıf Üniversitesi Akademik Personelinin Bireysel Emeklilik Sistemine Bakış Açısının Değerlendirilmesi", *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9 (19), 75-88, 2012.
- [7] S. Yazıcı, **Bireysel Emeklilik Sistemi ve Türkiye Uygulaması**, Doktora Tezi, Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara, 2015.
- [8] Ş. B. Salantur, **Bireysel Emeklilik Katılımcılarının Sistemden Erken Çıkma Riskinin Değerlendirilmesi**, Doktora Tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara, Türkiye, 2015.

- [9] H. Güneş, **Bireylerin Bireysel Emeklilik Sistemine Yönelik Tutum ve Davranışları: Karabük İlinde Bir Alan Araştırması**, Yüksek Lisans Tezi, Karabük Üniversitesi Sosyal Bilimler Enstitüsü, Karabük, 2015.
- [10] İ. Çömlekçi, O. Gökmen, "Bireysel Emeklilik Sistemine Katılmada Etkili Olan Faktörler: TR42 Bölgesinde Bir Araştırma". *Uluslararası Sosyal Araştırmalar Dergisi*, 10(49), 2017.
- [11] Y. Yin, I. Kaku, J. Tang, J.M. Zhu, **Data Mining: Concepts, Methods and Applications in Management and Engineering Design**, 2011.
- [12] P.N. Tan, M. Steinbach, V. Kumar, **Introduction to Data Mining**, International Edition, Pearson Education Inc., Boston, USA, 2006.
- [13] G.S. Linoff, M.J. Berry, **Data Mining Techniques: for Marketing, Sales and Customer Relationship Management**, Indianapolis, Wiley Publishing Inc., 2011.
- [14] S. Savaş, N. Topaloğlu, M. Yılmaz, "Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri". *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 11(21), 1-23, 2012.
- [15] T. Özcan, **Perakende endüstrisinde raf alanı yönetimine veri madenciliği eşanlı analitik bir yaklaşım**. Doktora tezi, İstanbul Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, Türkiye, 2011.
- [16] Y.L. Chen, J.M. Chen, C.W. Tung, "Data mining approach for retail knowledge discovery with consideration of the effect of shelf-space adjacency on sales". *Decision support systems*, 42, 1503-1520, 2006.
- [17] J. Han, M. Kamber, J. Pei, **Data Mining Concepts and Techniques**, Morgan Kaufmann Publishers Inc., Waltham, USA, 2012.
- [18] İ. Ertuğrul, A. Organ, A. Şavlı, "Veri madenciliği uygulamasına ilişkin PAÜ hastanesinde hasta profilinin belirlenmesi". *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 19 (2), 97-103, 2013.
- [19] S. Giray, F.E. Gülel, "Avrupa Ülkelerinin İntihar Oranlarına Göre Sınıflandırılması". *Süleyman Demirel Üniversitesi Fen-Edebiyat Fakültesi Sosyal Bilimler Dergisi*, 31, 235-247, 2014.
- [20] R.K. Blashfield, M.S. Aldenderfer, "The literature on cluster analysis". *Multivariate Behavioral Research*, 13, 271-295, 1978.
- [21] A. Uslu, C. Çetinkaya, E. Özceylan, S.K. İşleyen, "Organize Sanayi Bölgelerinin Hiyerarşik-K-Ortalamalar Yöntemi ile Analizi", *Türk Sosyal Bilimler Araştırmaları Dergisi*, 2(1), 20-37, 2017.
- [22] R. Johnson, D. Wichern, **Applied Multivariate Statistical Analysis**, New Jersey: Prentice-Hall, 2002.
- [23] M. Kayri, "Two-Step Cluster analysis in researches: A case study". *Eurasian Journal of Educational Research (EJER)*, 28, 89-99, 2007.
- [24] Ç. Taşkın, G.G. Emel, "Veri madenciliğinde kümeleme yaklaşımları ve kohonen ağırları ile perakendecilik sektöründe bir uygulama". *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 15(3), 395-409, 2010.
- [25] Choi W, Shah P, Das SK. "A framework for energy-saving data gathering using two-phase clustering in wireless sensor networks". **Proceedings of the First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'04)**, Boston, USA, 2004.
- [26] D. Schiopu, "Applying Two Step Cluster Analysis for Identifying Bank Customers' Profile". *Universităţii Petrol – Gaze din Ploieşti, Ştiinţe Economice*, LXII (3), 66-75, 2010.
- [27] M. Namvar, M.R. Gholamian, S.K. Abi, "A Two Phase Clustering Method for Intelligent Customer Segmentation". *International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, Liverpool, United Kingdom, 2010.
- [28] S. Savaş, N. Topaloğlu, "Veri Madenciliği Yöntemi İle GSM Şebekelerinin Performans Analizi". *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 26(4), 741-751, 2011.
- [29] M.B. Yılmaz, "Profiles of University Students According to Internet Usage with the Aim of Entertainment and Communication and their Affinity to Internet". *International Online Journal Education Science*, 4(1), 225-242, 2012.
- [30] H. Çetin, S. İrmak, "Elektronik Alışverişte Akademisyenlerin Güvenlik ve Risk Algılarının Belirlenmesi". *Yönetim Bilimleri Dergisi*, 12(24), 275-294, 2014.
- [31] E.S. Arı, H. Özköse, A. Doğan, M.H. Calp, "İstanbul Borsası'nda İşlem Gören Firmaların Performanslarının Kümeleme Analizi İle Değerlendirilmesi". *Bilişim Teknolojileri Dergisi*, 9(1), 33-39, 2016.
- [32] S. Giray, "İki Aşamalı Kümeleme Analizi ile Hükümlü Verilerinin İncelenmesi". *İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik Dergisi*, 25, 1-31, 2016.
- [33] M. Kayri, S. Günüş, "Yüksek ve Düşük Sosyoekonomik Koşullara Sahip Öğrencilerin İnternet Bağımlılığı Açısından Karşılaştırmalı Olarak İncelenmesi". *The Turkish Journal on Addictions*, 3(2), 165-183, 2016.
- [34] S. Rundle-Thiele, K. Kubacki, A. Tkaczynski, J. Parkinson, "Using Two-Step Cluster Analysis to Identify Homogeneous Physical Activity Groups". *Marketing Intelligence & Planning*, 33(4), 522-537, 2015.
- [35] A. Tkaczynski, "Segmentation using two-step cluster analysis", **Segmentation in Social Marketing**, Editörler: Dietrich T, Rundle-Thiele S, Kubacki K. Springer, Singapore, 109-125, 2017.
- [36] A. Özdemir, K. Orçanlı, "İki Aşamalı Kümeleme Algoritması İle Pazar Bölümlemesi, Müşteri Profillerinin Belirlenmesi ve Niş Pazarların Tespiti". *Uşak Üniversitesi Sosyal Bilimler Dergisi*, 5 (3), 1-27, 2012.
- [37] J. Bacher, K. Wenzig, M. Vogler, "SPSS Twostep Cluster - A First Evaluation". In **RC33 Sixth International Conference on Social Science Methodology**, 2004.
- [38] Emeklilik Gözetim Merkezi. "Bireysel Emeklilik Sistemi Gelişim Raporu, 2015". <http://www.egm.org.tr/bes2015gr.asp>, 17.04.2017.
- [39] T.C. Kalkınma Bakanlığı. "İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması (SEGE-2011)". <http://www.kalkinma.gov.tr/Pages/content.aspx?l=7a0b8e4a-dd0f-43b1-880c-e682b9d15cc3&i=548>, 04.05.2017.
- [40] SPSS Inc. (2007). "Clementine 11.1 User's Guide". https://web.kku.ac.th/wichuda/DMining/ClementineUsersGuide_11.1.pdf, 03.03.2017.