

İki Aşamalı Kümeleme Algoritması ile Pazar Bölümlemesi, Müşteri Profillerinin Belirlenmesi ve Niş Pazarların Tespiti

Abdulkadir ÖZDEMİR*
Kenan ORÇANLI**

Özet

Günümüzde, bilgi teknolojilerinin hızlı bir şekilde gelişmesine paralel olarak elde edilen verilerin depolanması kabiliyeti de artmıştır. Bu sayede, büyük veri yığınları oluşmuş ve bu veri yığınlarından nasıl faydalanılacağı, nasıl anlamlı hale getirileceği sorununun ortaya çıkarmıştır. İşte bu sorunun çözümüne imkân sağlayan Veri Madenciliği, son yıllarda yoğun olarak kullanılmaya başlanmıştır. Veri Madenciliği tekniklerinin kullanılması, büyük veri yığınlarına sahip olan işletmeler ve kurumlar açısından rakiplerine üstünlük sağlayabilecek son derece faydalı yöntemlerdendir. Bu çalışma da, Veri Madenciliği kümeleme tekniklerinden İki Aşamalı Kümeleme Algoritması kullanılarak demografik ve sosyo-kültürel özelliklere ait değişkenlerin verileri ile pazarların bölümlere ayrılması, işletmeler açısından son derece kritik ve stratejik karar olan hedef pazar olabilecek pazar nişlerinin tespit edilmesine yönelik bir örnek uygulama oluşturulmaya çalışılmıştır. Uygulama da, Türkiye İstatistik Kurumu (TÜİK)'na ait veri tabanı ve bu veri tabanında 2010 yılı Hane halkı Bütçe Anketindeki veriler kullanılmıştır.

Anahtar Kelimeler: Veri Madenciliği, Kümeleme Analizi, Pazar, Bölümleme, Niş Pazar.

Market Segmentation with Two Step Clustering Algorithms, Determination of Customer's Profiles and Nish Markets

Abstract

Nowadays, the information technologies have being rapidly developed, in the same vain the storage capacity of the dataset having being increased. Thus, it creates the large masses of data and it have caused some problems about how to benefit these masses and how to make them meaningful. In order to solve these kinds of problems, Data Mining is extensively used in recent years. Using Data Mining Techniques in businesses and institutions which have huge masses of data are useful in terms of competing with others. In this study, we tried to create a

* Yrd. Doç. Dr., Atatürk Üniversitesi, İİBF, Yönetim Bilişim Sistemleri Bölümü

** Türk Silahlı Kuvvetleri

sample application about data segmentation, niche markets and target market which is highly strategic and critical decision for businesses by using one of the Data Mining Techniques, Two Step Clustering Algorithm, with the data acquired from variables of demographic and socio-cultural features. In this application, we employed Turkish Statistical Institute's database and the data obtained from The Households Budget Survey-2010.

KeyWords: Data Mining, Cluster Analysis, Marketing, Segmentation, Niches Marketing.

Giriş

Günümüzde tüketici gereksinim ve istekleri umulmadık bir şekilde giderek artmakta ve çeşitlenmektedir. Bu gelişime paralel olarak, iletişim ve bilgi teknolojilerindeki ilerleme, pazarlama alanında da birçok değişime neden olmaktadır. Özellikle, günümüzde işletmelerin müşterilerle ilgili olarak hem kendi faaliyetleri sonucunda sahip oldukları hem de dış ortamdaki elde ettikleri veri sayısı artmakta, bu verilerden elde edilen bilgileri yönetme biçimleri değişmekte ve ayrıca yeni bilgilere duydukları ihtiyaç da fazlalaşmaktadır. Bazı bilgi teknolojileri araçları sayesinde toplanabilen müşteriler ile ilgili büyük hacimli veriler, işletmelerin rakiplerine karşı avantajlı bir hale gelmesini sağlamak için fırsatlar yaratmaktadır. Ancak, bu verilerin stratejik kararları desteklemede ve etkin pazarlama stratejilerini oluşturmada kullanılabilmesi için çok boyutlu ve analitik olarak analiz edilmesi gerekir. Bu amacı gerçekleştirmek için Veri Madenciliği en önemli araçlardan biridir. Büyük veri yığınlarında Veri Madenciliği teknikleri kullanılarak pazarlama bölümleri için müşterilerle ilgili gerekli bilgiler sağlanmaktadır.

1990'lı yıllardan itibaren büyük miktarda veri içerisinde gizli kalmış ve kullanılabilir bilgileri açığa çıkarmak amacıyla kullanılan Veri Madenciliği, başta pazarlama alanında olmak üzere birçok alanda yeni bakış açıları yaratmış ve kullanım alanları hızla artmaya devam eden bir yöntem haline gelmiştir (Koyuncu ve Özgülbaş, 2009: 22).

Bu gelişmeler çerçevesinde; ülkemizde daha önce pazarları bölümlere ayırmak ve müşteri profillerinin belirlenmesiyle ilgili farklı teknikler kullanılarak çalışmalar yapılmıştır. Ancak, Türkiye'de ekonomik konjonktürün ve küreselleşmeyle talep pazarının yapısı hızlı bir biçimde değişmesi neticesinde sahip olunan bilgilerin etkinliği azalmakta ve çoğunlukla işe yaramaz hale gelmektedir. Bu çerçevede düşünüldüğünde;

en güncel bilgiler çerçevesinde, Türkiye genelindeki tüketiciler ve pazarlama ile ilgili olarak Veri Madenciliği tekniklerinin kullanılarak işletmelere genel bir bakış açısı vermek amacıyla hazırlanan bu makalede, pazarı bölümlere ayırmada önemli bir faktör olan sosyo-ekonomik ve demografik özellikler ele alınarak 2010 yılında TÜİK tarafından düzenlenmiş “Hane Halkı Bütçe Anketi” verilerini kullanarak “İki Aşamalı Kümeleme Algoritması” ile Türkiye’de genel bir pazar bölümlenmesi yapılmış, elde edilen bölümlenmelerdeki etkili tüketici profilleri belirlenmiş ve ayrıca mevcut bilgiler çerçevesinde olabilecek niş pazar bölümleri tespit edilmeye çalışılmıştır.

Çalışma da, Veri Madenciliği kümeleme tekniklerinden İki Aşamalı Kümeleme Algoritması kullanılmıştır. Kümeler oluştururken, SPSS 11,5 istatistiksel paket programı kullanılarak çözüme ulaşılmıştır. Bu kümeleme algoritmasının seçiminde birçok faktör rol oynarken, özellikle verilerin ölçeklenebilirliği, veri yığının büyüklüğü ve değişkenlerde kategorik ve nicel verilerin aynı anda kullanılması en çok dikkat edilen hususlar olmuştur.

Veri Madenciliği, Pazar Bölümlenmesi ve Niş Pazarlama

Veri Madenciliği

Karar vericiler, güncel işlemler, eğilimler ve değişiklikler hakkında özlü ve güvenilir bilgiye ihtiyaç duyarlar. Birçok firmada hemen elde edilebilecek şey sadece o anki verilerdir. Bu yüzden idareciler eksik bilgilerle karar vermek durumunda kalırlar. Böyle durumlarda kullanıcılar ve bilgi sistemleri uzmanları verileri toplamak ve yerli yerine koymak maksadıyla aşırı zaman harcamak zorunda kalabilmektedirler (Watson ve Haley, 1998: 34).

Günümüzde teknolojik gelişmeler sonucunda birçok veri elektronik ortamda tutulmaya başlanmıştır. Gelişen bilgi teknolojileri ile elektronik ortamda tutulan bu verilerde zamanla çok büyük bir artış yaşanmıştır. Bilgisayar ve iletişim teknolojileri, bu artan verileri hızlı bir şekilde depolanmasına, işlenmesine ve bilgiye dönüşüne imkân sağlamaktadır. Bu sayede bilgiye ulaşmak isteyenler hızlı ve kolay bir şekilde istenilen bilgiye ulaşmakta ve sağlanan bu bilgiler çerçevesinde karar almaktadırlar. Ancak doğru karara en hızlı bir şekilde ulaşmak sadece verileri toplamakla değil, aynı zamanda toplanan bu veri yığınlarını analiz ederek ve yorumlayarak anlamlı raporlar oluşturmakla mümkündür. Ayrıca, verilerin boyutlarının çok büyük oluşu, herhangi bir araç kullanmaksızın verilerin etkin bir

biçimde analiz edilmesini ve karar destek aşamasında kullanılmasını imkânsız hale getirmiştir. Bu noktalarda Veri Madenciliği kavramı ortaya çıkmaktadır.

Veri Madenciliği, önceden bilinmeyen, veri içinde gizli, anlamlı ve yararlı örüntülerin büyük ölçekli veri tabanlarından otomatik biçimde elde edilmesini sağlayan veri tabanlarında bilgi keşfi süreci içerisinde bir adımdır (Sever ve Oğuz, 2002: 174). Aslında, Veri Madenciliği, büyük veri yığınlarından anlamlı bilgiler elde etmek için, bilgisayar destekli bir bilgi çözümleme işlemidir. Kendiliğinden oluşan kümelenmeler, örüntüler, birliktelikler ve istisnalar veri tabanlarındaki bilgi kaynaklarıdır. Veri Madenciliği yöntem ve teknikleri bu kaynakları analiz ederek taşıdıkları bilgiyi keşfetmeye çalışırlar (Dinçer, 2006: 1).

Veri madenciliğinin günümüzde yaygın bir kullanım alanı bulunmaktadır. Örneğin pazarlama, bankacılık ve sigortacılık gibi alanlarda ve elektronik ticaret ile ilgili alanlarda yaygın bir şekilde kullanılmaktadır (Akpınar, 2000: 32).

Veri Madenciliğinde kullanılan veri çevrimiçi veya çevrimdışı işletimsel veridir. İşletimsel veri organizasyonel aktiviteler düşünülerek düzenlenir ve normalleştirilir. Bu bilgi keşfi süreci için gerekli verilerin ya bir arada bulunmamasına, ya hiç tutulmamasına ya da ilgili veri içeriğinin birden fazla yorumlanmasına yol acar. Bu yüzden bilgi keşfi açısından işletimsel veri ister çevrimiçi ister çevrimdışı olsun "işlenmemiş/ham veri" olarak kabul edilir.(Sever ve Oğuz, 2002: 177).

İşlenmemiş/ham verilerin Veri Madenciliğinde kullanılması için ilk önce bazı düzenlemelerin yapılması gerekmektedir. Veri ön işleme tekniği adı verilen bu düzenlemeler dört aşamada yapılmaktadır. Bunlardan biri olan veri temizleme verilerdeki gürültünün giderilmesi ve tutarsızlıkların düzeltilmesi için uygulanır. Bir diğer teknik olan veri birleştirme, farklı kaynaklı verileri uygun bir veri tabanında birleştirir. Veri indirgeme de ise fazla olan bazı değişkenlerin atılması ve birleştirilmesi veya kümeleme yolu ile veri büyüklüğünün azaltılması amaçlanır. En son yöntem ise veri dönüştürmedir. Bu sözü edilen veri ön işleme teknikleri, veri madenciliğinden önce uygulanması sonucu elde edilen sonuçların kalitesinin artmasına fayda sağlamaktadır (Han ve Kamber, 2001: 105).

Veri Madenciliği teknikleri kullanıldıkları veri yapılarına ve keşfedebildikleri örüntü biçimlerine göre kategorilere ayrılır. Birçok kaynak veri madenciliği teknikleri için farklı gruplandırmalar yapmıştır. Zhong ve Zhou (1999), Veri Madenciliğinde kullanılan modeller tahmin edici ve tanımlayıcı olmak üzere iki başlık altında incelenmektedir. Tahmin edici

modellerde, sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerlerinin tahmin edilmesi amaçlanmaktadır. Tanımlayıcı modellerde ise, karar vermeye rehberlik etmede kullanılacak mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır. Tanımlayıcı modeller, İlişki Analizi ve Kümeleme Analizi olmak üzere iki grupta incelenmektedir. Birliktelik Kuralları ile Ardışık Zamanlı Örüntüler İlişki Analizi kapsamında yer almaktadır. Tahmin edici modeller, Sınıflandırma ve İstatistiksel Tahmin Modelleri olmak üzere iki ana başlık altında incelenmektedir. Diğer bir ayırım ise Han ve Kamber (2001) tarafından yapılmıştır. Ancak bunlardan kabul göreni Han ve Kamber (2001)'in ortaya sürdüğü kategorilerdir. Bu kategoriler, Tanımlama ve Ayrımlama, Birliktelik Analizi, Sınıflandırma ve Öngörü, Kümeleme Analizi ve Sıradışılık Analizidir.

Veri madenciliği ile ilgili kullanılan pek çok yöntemin yanına hemen her geçen gün yeni yöntem ve algoritmalar eklenmektedir. Bu yöntemler genellikle istatistiği temel alan ama daha çok makine öğrenimi ve yapay zekâ destekli yeni nesil yöntemlerdir. Yeni nesil yöntemlerin başlıcaları; Yapay Sinir Ağları (Berson, Smith ve Thearling,1999: 23), Temel Bileşenler Analizi, Diskriminant Analizi, Faktör Analizi, Kohonen Ağları, Bulanık Mantığa Dayalı Yöntemler, Genetik Algoritmalar, Bayesci Ağlar, Pürüzlü Küme Teorisine Dayalı Yöntemler olarak sıralanabilir (Chen, 2001: 56).

Yukarıda sayılan yöntemlerin dışında birden fazla tekniği içine alan hibrid yöntemler ve zaman serilerine dayalı yöntemlerden de veri madenciliği yöntemi olarak faydalanılmaktadır. Bilgi keşfine yarayan her yöntem Veri Madenciliği yöntemi olarak kullanılabilir (Kovalerchuk ve Vityaev, 2001: 45).

Pazar Bölümlemesi ve Niş Pazarlama

Günümüzde pazarlama anlayışı, müşteri odaklı olma yönünde bir anlayışa doğru kaymış ve şirketlerin müşteri ile ilişkileri büyük ölçüde değişmiştir. İşletmelerde başarılı bir müşteri analiziyle, küresel bir dünyada ve giderek artan rekabetçi piyasalarda, işletmeler için yaşamsal önem taşıyan müşteri için değer yaratmak, müşteri sadakati sağlamak ve bu konularda kurumsallaşmayı gerçekleştirmek mümkün olabilecektir (Rud, 2001). Bu nedenle işletmeler tüketici davranışlarını anlamak ve buna göre pazarlama stratejileri uygulamak zorunluluğu ile karşı karşıya kalmıştır. Fakat pazarlarda birçok farklı ihtiyaç ve isteği olan tüketici ve tüketici grupları bulunmaktadır. Bir işletmenin bu kadar çeşitli tüketici

profilinibelirlenmesi, bu profillerin ihtiyaçlarını ve isteklerini tatmin etmesi çok zor ve maliyetli bir durumdur.

Pazarların giderek büyüdüğü hatta birleşerek küresel pazarları oluşturduğu günümüzde, işletmelerin bir pazara bütün olarak hitap etmesi verimlilik ve etkinlik açısından uygun değildir. Bu nedenle çağdaş pazarlama stratejilerinin ilk adımı olan pazar bölümlendirme kavramı ortaya çıkmıştır. Veri Madenciliği sonuçları pazar bölümlenmesi, kredi derecelendirme ve pazar sepeti analizi gibi pek çok tüketici odaklı uygulamaya girdi teşkil etmektedir (Rygielski, Wang ve Yen, 2002: 486; Aydoğan, Gencer ve Akbulut, 2008: 42).

Günümüzde işletmeler, ne tür pazarlara yönelik faaliyet gösterirlerse gösterebilirler, pazarlama yönetimleri hedef pazarlarını seçmek için pazarları bölümlere ayırmak zorundadır. Çünkü pazarın bölümlenmesi, hedef pazarların belirlenmesinde başlangıç olarak kabul edilmektedir (Kotler, 1994: 78). Bu çerçevede bu pazarları ayrıntılı olarak inceleyerek onları oluşturan tüketicilere ulaşmanın en uygun yolları bulunmaya çalışılmalıdır (Mucuk, 1997: 112).

Pazar bölümlenmesi kuruluşun amacını gerçekleştirebilmek amacıyla türdeş olmayan pazarı türdeş bölümlere ayırmak ve en etkili hizmeti bu bölümlere sunabilmektir. Bölümlendirmede temel, grupların aralarındaki farklılıklar ve grup içindeki benzerlikler üzerine oturtulmalıdır (Odabaşı, 1994: 39-40).

Pazar bölümlenme, hedef pazarın belirlenmesinde hedef pazar seçimi ve pazar konumlandırma aşamalarından önceki basamaktır. Hedefe giden ilk adım olması sebebiyle pazarlama stratejisinin oluşturulmasında ve pazarlama çabalarının başarıya ulaşmasındaki etkisi de büyüktür (Altunışık, Özdemir ve Torlak, 2001: 34). Pazarlama bölümlenmesiyle, işletmelerin kendileri açısından en uygun pazar fırsatlarını yakalayabilme şansına sahip olacaklardır. Çünkü pazarın tüketim alışkanlıklarına göre farklı bölümleri, farklı pazarlama karmalarına ihtiyaç duyarlar. Bu çerçevede, bunun başarılması aynı zamanda işletmenin kaynaklarını doğru planlamasını ve doğru kullanmasını da beraberinde getirir.

Bir pazarı bölümlere ayırmanın basit ve tek bir yöntemi yoktur. Pazarlamacının, pazar yapısını en iyi şekilde ortaya koyması için farklı bölümlendirme değişkenlerini tek başına veya bir arada kullanması gerekmektedir (Kotler ve diğerleri, 2001: 321).

Günümüzde başarılı olmak isteyen işletmeler için pazar bölümlenme çalışmaları çok önemlidir. Kaynaklara ve pazarlara göre farklılık

göstermesine rağmen, pazarlamanın büyük duayenlerinden sayılan Kotler (1994)'e göre, pazarlama bölümlenme süreci üç bölümden oluşmaktadır. İlk aşama olan pazar araştırma safhasında, araştırmacı tüketicilerin tüketim tercihleri ve davranışları hakkında bilgiler toplamak amacıyla araştırmalar yürütür ve tüketicinin demografik, sosyo-ekonomik, davranış kalıpları gibi bilgilere ulaşmaya çalışır. İkinci aşama olan analiz aşamasında ise, araştırmacı analizler uygulayarak birbirleriyle ilgili bağlı değişkenleri bulur ve ardından kümeleme analizi yaparak elde edilebilecek pazar bölümlerine ulaşmaya çalışır. Üçüncü ve son aşama olan profil oluşturma aşamasında ise her bir küme demografik, davranışsal gibi farklılıklara göre bölümlenir ve her bir bölümün profilleri çıkarılır. Ardından her bir pazar bölümüne en belirgin ayırıcı özelliği baz alınarak bir isim verilir.

Pazar bölümlenme ile pazarlamada istenen başarıya ulaşmak amacıyla pazarın başlıca karakteristikleri olan coğrafik, sosyoekonomik, demografik, psikografik ve davranışsal karakteristikler göz önüne alınarak birbirinden farklı ve homojen bölümlere ayrılması sağlanmaktadır. Pazar bölümlendirmesi yapmak isteyen işletmeler pazardaki hedef pazarı belirlemeye çalışmaktadırlar. Bu sayede işletmeler pazarda belirledikleri hedef pazarlara uygun olarak pozisyon almaktadır ve bunun sonucunda hedef pazar içerisindeki müşterilerin beklentilerini sağlamaktadırlar (Mykletun, Crofts ve Mykletun, 2001: 494).

Pazar bölümlenme faaliyetlerinde sosyo-ekonomik ve demografik özellikler ile sosyal sınıf kavramı çok yaygın olarak kullanılmaktadır. Bunun temel nedeni; sosyo-ekonomik, demografik (yaş, cinsiyet, gelir, meslek, eğitim, medeni durum v.b.) ve sosyal sınıf ölçütlerin kolay ölçülebilir, kolay belirlenebilir ve kolay analiz edilebilir olmasıdır. Maliyetlerin düşük olmasını sağlayan bu sebepler, sosyo-ekonomik, demografik ve sosyal sınıf ölçütlerin bölümlendirme çalışmalarında ilk akla gelen ölçüt olmalarına neden olmaktadır. İlk pazar bölümlenme çalışmalarını yapan araştırmacılar demografik özellikleri ile sosyal sınıf kavramından yararlanmışlardır.

Pazardaki çeşitlilik, gelişen teknoloji yeni pazarlama yaklaşımlarını ortaya çıkartmakta, büyük işletmeler ve onların klasik pazarlama anlayışları sorgulanmaktadır (Dalgıç ve Maarten, 1994: 34). Bu sorgulamalar doğrultusunda işletmeler kendilerini gerek rakiplerin gerekse pazarın yapısı içerisinde bir değişim süreci içerisinde sokmaları gerekmektedir. Ancak bu değişim sürecinde yerini koruyabilmek, rakiplerin yolunda ilerlemek mantıklı gibi görünse de, asıl olan yepyeni hedefler çizip yeteneklerin bu yönde geliştirme ile başarılacağını söylemek mümkündür (Semiz, 2008: 4).

İşte bu durumda niş pazarlama stratejisi karşımıza çıkmaktadır. "Niş pazarlama" göreceli olarak bir pazar bölümünde veya bölümlerinde benzer niteliklerde mal/hizmetlere gereksinim duyan, bir ya da daha çok benzer özelliği paylaşan, küçük bir tüketici kitlesinin istemlerini daha iyi karşılamak amacıyla geliştirilen pazarlama faaliyetlerine denir (Özkan, 2003). Buna neden olarak, aynı veya ayrı pazar bölümünde bulunsalar bile, bazı tüketicilerin benzerlerinden ayrılmak isteyecekleri gösterilmektedir.

Pazar bölümlenmesi ile niş pazar bölümlenmesi çoğu zaman birbirine karıştırılmaktadır. Pazar bölümlenmesi, büyük bir pazarı daha iyi yönetilebilir alt pazarlara bölme sürecidir. Amaç, her biri diğerinden belirli bir şekilde farklı olan homojen alt pazarlar belirlemektir. Böylece işletme bir veya birden fazla dilim seçer ve her birine ufak kitlesel pazar muamelesi yapar (Shani ve Chalasani, 1992: 45). Niş pazarlamada ise, işletme az sayıdaki müşterinin ihtiyaçlarından yola çıkar ve genellikle geniş müşteri tabanına ulaşır (Dalgıç ve Maarten, 1994: 42).

Pazar bölümlenmesi, niş pazarlamanın temelini oluşturmaktadır. Niş pazarlama; pazara odaklanma, veri tabanı oluşturma, pazar bölümlenmesi ve niş pazar belirleme, ürün ve hizmetleri niş pazara uygun olarak farklılaştırma aşamalarından oluşan bir süreci ifade eder. Bu sürecin her hangi bir aşamada kesilmesi, niş pazarlamanın uygulanamaması anlamına gelir ve bu aşamalardan birisi de pazar bölümlenmesidir (Magrath, 1988: 8).

Niş pazarlama, pazar bölümlenmesi sürecinin çeşitli düzeylerinden biridir. Pazar bölümlenmesi, tüketicileri benzer ihtiyaçlara sahip homojen gruplara ayırmaktır, ama süreç hiçbir zaman mükemmel değildir. Yani aynı bölüm içerisinde yer alan ve benzer ihtiyaçlara sahip olan kişiler bile hala demografik özellikleri veya kullanım alışkanlıkları açısından birbirinden farklılık gösterebilmektedir ve bunlara tek bir bölüm stratejisiyle ulaşılamamaktadır (Ercan, 2007: 3).

Pazar Bölümlenmesi, Niş Pazar Bölümlerinin Tespiti ve Veri Madenciliği ile İlgili Uygulama

Araştırmanın Kapsamı

Hane halkı Bütçe Anketleri, hanelerin sosyoekonomik yapıları, yaşam düzeyleri ve tüketim kalıpları hakkında bilgi veren, uygulanan sosyoekonomik politikaların geçerliliğinin test edilmesi amacıyla kullanılan en önemli kaynaklardan biridir (Bağdadioğlu ve diğerleri, 2009: 15).

Bu kapsamda; Hane halkı Bütçe Anketi ile;

- ✓ Tüketici fiyat indekslerinde kullanılacak maddelerin seçimi ve temel yıl ağırlıklarının elde edilmesi ve güncellenmesi,
- ✓ Hanelerin tüketim yapısı ve tüketim kalıplarında zaman içinde meydana gelen değişikliklerin izlenmesi,
- ✓ Milli gelir hesaplamalarında özel nihai tüketim harcamaları tahminlerine yardımcı olacak verilerin derlenmesi,
- ✓ Asgari ücret tespit çalışmaları için ihtiyaç duyulan verilerin elde edilmesi,
- ✓ Harcamaya dayalı yoksulluk sınırının belirlenmesi, hane halklarının yaşam seviyeleri, beslenme sorunları vb. diğer sosyoekonomik analizler için gerekli verilerin elde edilmesi amaçlanmıştır (TÜİK Yayın No: 3186, 2008: 2).

Ankette, bireylerin ve bunların oluşturduğu hane halklarının tüketim yapılarını, gelir düzeylerini, sosyoekonomik gruplara, kırsal, kent ve bölgelere göre ortaya çıkaran bu çalışmayla tüketim alışkanlıkları, tüketim harcaması türleri ile mal ve hizmet harcamalarının çeşitliliği, hane halkının sosyoekonomik özellikleri, hane halkı fertlerinin çalışma durumları, hane halkının toplam geliri, gelirin elde edildiği kaynaklar vb. konular hakkında bilgiler derlenmiştir (TÜİK Yayın No: 627, 2010: 2).

Anketin uygulama alanı olarak Türkiye Cumhuriyeti sınırları içinde bulunan tüm yerleşim yerleri kapsama dâhil edilmiştir. Bu yerleşim yerleri, kent-kırsal tanımı dikkate alınarak iki tabakaya ayrılmıştır:

- ✓ Kentsel Yerler: Nüfusu 20.001 ve daha fazla olan yerleşim yerleri,
- ✓ Kırsal Yerler: Nüfusu 20.000 ve daha az olan yerleşim yerleri olarak tanımlanmıştır (TÜİK Yayın No: 627, 2010: 2).

2010 Hane Halkı Bütçe Anketi, 1 Ocak–31 Aralık 2010 tarihleri arasında bir yıl süre ile her ay değişen 1.104, yıllık toplam 13.248 örnek hane halkına uygulanmıştır. Ancak kurumsal nüfus olarak tanımlanan yaşlılar evi, huzur evleri, hapishane, askeri kışla, özel nitelikli hastane, çocuk yuvalarında bulunanlar ile göçerler anket dışı tutulmuştur (TÜİK Yayın No: 627, 2010: 2).

2010 Hane Halkı Bütçe Anketi'nde birinci aşama örnekleme birimi olan blokların seçiminde kullanılan temel örnekleme çerçevesi Ulusal Adres Veri Tabanıdır (TÜİK Yayın No: 627, 2010: 3).

Yukarıda verilen çerçeve kullanılarak bloklar oluşturulmuş, kentsel yerlerden ve belediye teşkilatı olan kırsal yerlerden ve köylerden yerleşim

yerinin büyüklüğüne orantılı olasılıkla bloklar belirlenmiş ve her bloktan hane halkları sistematik olarak seçilmiştir. Nihai örnekleme birimi olarak örnek adreste bulunan hanehalkı tanımlanmış ve tabakalı iki aşamalı küme örnekleme yöntemi kullanılmıştır (TÜİK Yayın No: 627, 2010: 3).

Ankette örnekleme tekniklerine göre seçilen esas örnek hane halkı ile anket yapılamaması durumunda "cevaplamama formu" doldurulmuş, bu durum dikkate alınarak ağırlık katsayıları hesaplanmıştır. 2010 Hane halkı Bütçe Anketi'nde cevaplamama oranı, kırsal yerlerde % 21,4 ve kentsel yerlerde % 25 olmak üzere, Türkiye genelinde % 23,9 olarak gerçekleşmiştir (TÜİK Yayın No: 627, 2010: 4).

2010 Hane halkı Bütçe Anketi'nin örnekleme yapısı; "Türkiye, Kent, Kır" bazında tahmin verme amacına uygun olarak oluşturulmuştur (TÜİK Yayın No: 627, 2010: 4).

TÜİK tarafından yapılan 2010 Hane halkı Bütçe Anketi, her ay ortalama 1104 farklı örnek hanenin dönüşümlü olarak izlenmesi yoluyla yürütülmüştür. Hane halklarından bilgiler; görüşme, kayıt ve gözlem metodları kullanılarak derlenmiştir. Anket ayı öncesinde örnek hane halkına yapılan ilk ziyarette, hane halkı ile tanışma sağlandıktan sonra hane halkının sosyoekonomik durumuna ait bilgiler alınmakta, harcama kayıt defterinin nasıl doldurulacağı açıklanmaktadır. Anket ayı içinde yapılan ziyaretler sırasında, örnek hanenin anket ayında yaptığı gıda, giyim, sağlık, ulaştırma, haberleşme, eğitim, kültür, eğlence, konut, ev eşyası vb. gibi tüketim harcamaları hane halkına bırakılan kayıt defterinden kayıtlama ve görüşme yöntemleri ile alınmıştır (TÜİK Yayın No, 2010: 627: 4).

Anket ayı bitiminde yapılan son görüşmede ise, hane halkı fertlerinin anket ayı içindeki istihdam durumu, iktisadi faaliyeti, mesleği, işteki durumu, anket ayı ve son bir yıl içerisinde elde ettiği kullanılabilir net gelir bilgileri derlenmiştir (TÜİK Yayın No: 627, 2010: 4).

Araştırmanın Amacı ve Yöntemi

Bu çalışmanın temel amacı, işletmelerin hedef pazarları belirlemede ve belirli yönlerden performanslarını değerlendirmede dikkate alınması gereken en önemli ölçütlerin (değişkenlerin) sınıflandırılmasında kümeleme analizi tekniklerinden iki Aşamalı Kümeleme Algoritmasının uygun bir yöntem olarak kullanılabilmesini göstermek ve karar alıcılara bir örnek oluşturmaktır.

Kümeleme

Kendi başına bir disiplin olmaktan çok istatistik, veri madenciliği, vektör nicemleme, makine öğrenmesi gibi diğer disiplinlerin içinde veya onlarla birlikte kullanılan bir teknik olan kümeleme analizi, bir dizi örüntüyü ayrık ve homojen gruplar oluşturacak şekilde gruplandırma işlemidir (Han ve Kamber,2001). Kümeleme analizinin temel amacı, birey veya nesnelerin temel özelliklerini dikkate alarak onları gruplandırmaktır. Diğer bir deyişle kümeleme analizi, gruplanmamış verileri benzerliklerine göre gruplandırarak araştırmacıya özetleyici bilgiler sunmaktır (Kalaycı, 2010: 349).

Kümeleme işlemi örüntülerin benzerlik değerlerine göre sınıflara veya kümelere ayrılmasıyla gerçekleştirilir (Han ve Kamber, 2001: 10). Bu işlemin sonucunda elde edilen her başarılı küme içinde yer alan nesnelere arasında maksimum benzerlik ve kümeleme sonucunda elde edilen her kümedeki nesnelere arasında maksimum farklılık oluşması sağlanır (Demiralay, 2005: 10).

Veri kümeleme analizinde ele alınan problemlerin amacı, içeriği ve boyutu değişkenlik gösterdiğinden ve algoritmaların kümeleme yaparken izledikleri kriterlerden (Jain ve Dubes, 1988) dolayı her probleme uygulanacak genel geçerli bir yöntem bulunmamaktadır. Yaygın olarak kullanılan kümeleme yöntemleri birimler arasındaki uzaklıklara dayanan benzerlik ya da benzemezlik matrisine göre işlem yaptıklarından, farklı kümeleme yöntemleri farklı uzaklık ölçülerine göre farklı sonuçlar verebilmektedir (Fraley ve Raftery, 1998: 582).

Literatürde kümeleme metoduyla ilgili birçok çalışma bulunmaktadır. Ancak en kabul göreni Han ve Kamber(2001)'in yaptığı bölümleridir. Han ve Kamber kümeleme metodlarını altı bölümde incelemektedir. Bunlar Bölümleyici Kümeleme Metodları (hijerarşik olmayan), Hijerarşik Kümeleme Metodları, Yoğunluğa Dayalı Kümeleme Metodları, Grid Tabanlı Kümeleme Metodları, Model Tabanlı Kümeleme Metodları, Kısıtlamalı Kümeleme Metodları'dır.

İki Aşamalı Kümeleme Algoritması

Kümeleme algoritmalarında kümelemeye alınan değişkenlerin ölçümleri kategorik veya metrik olması önemli rol oynamaktadır. Bazı algoritmalar sadece metrik değişkenlere uygulanırken bazı algoritmalar ise kategorik değişkenlere uygulanmaktadır. Ancak bir grupta hem metrik ölçümler hem de kategorik değişkenlerin bulunması da mümkündür (Alpar, 2011: 349).

Punj ve Steward (1983) tarafından önerilen (Savaş ve Topaloğlu, 2011: 747) ve hiyerarşik olmayan bir kümeleme analiz yöntemi olan İki Aşamalı Kümeleme Analizi, veri yapısında hem kategorik hem de sürekli değişkenlerin olması durumunda kullanılabilen (Alpar, 2011: 349) ve Ward'ın Minimum Varyans Yöntemi ile "K-means" yönteminden oluşan hibrid bir kümeleme (Kuo, Hove Hu, 2002; 1476) yaklaşımıdır. Bu yöntemin en önemli özelliği, büyük veri setlerinde uygulanabilmesi, elde edilen kümelerle uyum sağlamayan gözlemlerin istendiğinde veriden ayıklanabilmesi (Alpar, 2011: 349) ve AIC ya da BIC kriterlerine (Savaş, 2011: 747) göre en uygun küme sayısını otomatik olarak belirleyebilmesidir.

Yöntem iki aşamadan oluşur. İlk aşamada gözlemler teker teker işleme alınarak ön kümelere gruplanır. İkinci aşamada bu ön kümelere standart aşamalı kümeleme yaklaşımları uygulanır. Her iki aşamada da kullanılacak iki uzaklık ölçüsü Log-Olabilirlik Uzaklığı ve Öklit Uzaklığıdır (Alpar, 2011: 349). Log-olabilirlik uzaklığı veride sürekli ve kategorik değişkenlerin olması durumunda kullanılabilir. Olasılık temelli bir yaklaşıma dayanmaktadır (Alpar, 2011: 349).

Araştırmanın Uygulanması

Uygulamada kullanılan 2010 yılı Hane halkı Bütçe Anketi'nde üç temel değişken grubu bulunmaktadır. Bunlar;

- ✓ Hane halkı ile ilgili değişkenler,
- ✓ Hane halkı fertleri ile ilgili değişkenler,
- ✓ Tüketim harcamaları ile ilgili değişkenlerdir.

Hane Halkı Bütçe Anketinde değişken gruplarında, birinci grupta 118 adet değişken, ikinci grupta 109 adet değişken ve üçüncü grupta ise 4 adet değişken bulunmaktadır. Pazar bölümlenmesi yaparken sosyo-ekonomik ve demografik değişkenlerden analiz birimi olarak "haneyi mi yoksa bireyi mi esas almalı?" sorusu değişkenlerin belirlenmesinde anahtar konumundadır. Edgell (1998)'e göre pazar bölümlenmesi kavramını geleneksel çerçevede inceleyen araştırmacılar hanenin baz alınması gerektiğini belirtmekte ve erkeği de hanedeki yapıyı belirleyen kişi olarak tanımlamaktadırlar. Goldthorpe (1983)'te yaptığı çalışmalarda bu düşüncüyü desteklemektedir. Goldthorpe, çalışmalarında hane üyelerinin aynı ortamı paylaştığını ve paylaşılan bu ortam hakkında bir yargıya aile reisine bakarak ulaşılabileceğini belirtmektedir. Çünkü bir hanede en fazla sorumluluk aile reisindedir. Bu görüşe karşıt olarak bireyin analiz nesnesi olması gerektiğini savunan görüşlerde mevcuttur. Bireyin analiz birimi olması gerektiğini savunan görüşlere göre bireyin tavır ve davranışları mantık gereği, bireyin

dolaysız tecrübelerine göre incelenmelidir (Edgell, 1998: 58). Bireyin dolaysız tecrübelerini de ölçmenin yolu da onun konumunu haneden bağımsız analiz etmektir. Buna göre analiz biriminin “birey mi yoksa hane mi olacağı” araştırmanın amacına ve bağlamına bağlıdır (Edgell ve Duke, 1991). Bireyin üretimle ilişkili davranış ve tavırlarına yönelik doğrudan bireyi ilgilendiren bir odağı olan araştırmalarda, analiz birimi birer birey olarak kadınlar ve erkekler olmalıdır. Bununla beraber, ailenin bir bütün olarak birlikte tüketim yapmasının söz konusu olduğu tüketim davranışlarına ve eğilimlerine ilişkin çalışmalar da uygun analiz birimi hanedir (Edgell, 1998: 59).

Değişkenler seçilirken hibrid bir yaklaşım sergilenmiştir. Bazı değişkenler hane ile ilgiliyken bazı değişkenlerde bireyle yani hanenin davranışlarını ve eğilimlerinin en iyi temsil edildiği olarak değerlendirilen hane halkı reisinin özellikleri alınarak kullanılmıştır. Çünkü Türkiye’de aile reisinin davranışı aileye yön vermektedir. Bu değişkenlerden başlangıçta amaca uygun ve birbirleriyle benzer olduğu değerlendirilen değişkenleri seçmek maksadıyla çok boyutlu ölçeklendirme yönteminden yararlanılmıştır. Çok boyutlu ölçeklendirme yönteminde verilerin dağılımları ile ilgili herhangi bir varsayımda bulunulmamakla beraber uzaklık matrisinin grafik koordinatlarına dönüştürülmesi ile haritalar elde edilmektedir. Dolayısıyla çok boyutlu ölçeklendirme, boyut indirgeyerek nesnelere/değişkenler arasındaki ilişkileri belirlemeyi amaçlayan grafik tabanlı bir yöntemdir. Bu gösterimde, birbirine benzer nesnelere/değişkenleri uzayda temsil eden noktaların birbirine yakın, aynı mantıkla benzemeyenlerin ise birbirinden uzak olması beklenir. Dolayısıyla başlangıçta 58 adet değişken seçilmiş ve çok boyutlu ölçeklendirme yönteminden sonra birbirine benzer olduğu tespit edilen 26 değişkene indirgenmiştir.

Değişkenler kullanılmadan önce ön işleme tabi tutulmuştur. Bu ön işlemede, eksik veriler bazı değişkenlerde verilerin ortalamaları alınarak tamamlanmış, çok fazla eksik verisi olan değişkenler analizden çıkartılmıştır (%50’den eksiği olanlar) (Alpar, 2011: 154). Ayrıca, yapılan incelemeler neticesinde altı adet çiftte yüksek korelasyon (Pearson Korelasyon Katsayısı:0,75-0,86 arası) çıkması neticesinde bu altı tane çiftten birer değişken gruptan çıkartılmıştır. Geriye kalan ve bizim kullanacağımız hane halkı tipi, hane halkı büyüklüğü, konut tipi, konut mülkiyeti, ısıtma sistemi, ikinci konut sahipliği, otomobil sahipliği, kır-kent yerleşim durumu, hane halkı reisinin eğitim durumu, hane halkı reisinin hangi sektörde çalıştığı, hane halkı reisinin medeni durumu, hane halkı reisinin mesleği, iş yeri

statüsü ve bu iş yerinde çalışan kişi sayısı, hane halkı reisinin cinsiyeti ve yaşı, yıllık kullanılabilir gelir, konut büyüklüğü, oturduğu konutun değeri, harcama durumu değişkenleri belirlenmiştir. Bu değişkenlerden yıllık kullanılabilir gelir, oturduğu konutun değeri, oturduğu konutun büyüklüğü ve aylık harcama değişkenleri sürekli değişken geriye kalan 16 değişken ise kategorik değişkendir.

Uygulaması anlatılan İki Aşamalı Kümeleme Algoritması küme sayısını belirleme konusunda bize bir olanak tanımaktadır. Küme sayısının bulunması amacıyla çeşitli yöntemler geliştirildiyse de çoğunlukla bunlar küçük örneklemelerde uygulanmakta ve büyük örneklemelerde sağlıklı sonuçlar vermemektedir.(Akın, 2008: 26). Bu çerçevede, en ideal küme sayısını bulmaya yarayan herhangi bir istatistik yöntem bulunmamaktadır.(Leisen, 2001: 57). Dolayısıyla tek kümeden başlayıp yanıtlayıcı sayısına kadar küme belirlenebilir. Ancak yönetsel açıdan 2-5 arasında seçim yapmakta fayda vardır. Çok fazla sayıda küme belirleyecek olursak kümelerin her birini tanımlama zorlaşacak ve ayrıca işletmeler için her birine uygun pazarlama karması geliştirmek çok maliyetli olacaktır. Doğru küme sayısını belirlemek birazda yargıya bağlıdır. Elde edilen kümeler tanımlandığında belirli özellikler ön plana çıkarılarak her bir kümeye farklı isimler verilebiliyorsa belirlenen kümelerin doğru sayıda olduğu söylenebilir.(Lilien ve Rangaswamy, 1998: 85). Burada önemli olan kümelerin stratejik planlama yapılırken kullanılabilmesidir. Dolayısıyla kümelerin birbirinden farklı ama belirgin özellikler taşıyor olması oluşturulacak pazarlama karmalarına ışık tutacaktır.(Arslan, 2004: 259)

Araştırmadan Elde Edilen Bulgular

Çalışmamızda en uygun olan kümeleme sayısını tespit etmek amacıyla çeşitli denemelerde bulunulmuştur. Başlangıç olarak küme sayısı 2 olarak belirlenmiş ve paket program ile 2 adet küme oluşturulmuştur. Ancak, kümelere ait eleman sayısı 2 nolu kümede yoğunlaşmış ve değişkenlerde iki küme arasında heterojenlik kümeler içinde homojenlik sağlanamamıştır. Aynı denemeler küme sayısı 3, 4 ve 5 olacak şekilde denenmiş ve en iyi küme sayısı olarak en son denemede bulunmuş ve küme sayısının 5 olmasına karar verilmiştir.

Yapılan analiz sonucunda elde edilen sonuçlar Tablo 1 ile Tablo 12'de görülmektedir.

Tablo 1: Küme Dağılımı

Küme	N	%
1	972	12,32
2	1598	20,26
3	1684	21,35
4	1103	13,99
5	2527	32,05
	7884	100

Tablo 2: Sürekli Değişkenlere Ait Verilerin Kümelere Dağılımı

Küme	Konut değeri		Konut alanı		Yıllık kullanılabilir gelir		Aylık harcama	
	Frekans	S.S	Frekans	S.S	Frekans	S.S	Frekans	S.S
1	62198	44589	101	29	22053	14454	1700	1017
2	100454	96538	113	35	36851	39822	2644	2031
3	40730	29678	97	32	16039	10213	1255	734
4	63012	44771	103	28	22889	15694	1788	1222
5	61878	44951	101	29	22169	15185	1741	1146
	65378	60281	103	31	23922	22941	1822	1385

Tablo 3: Hane Tipi Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	Tek çocuklu çekirdek aile		İki çocuklu çekirdek aile		Üç veya daha fazla çocuklu çekirdek aile		Çocuk-suz çift		Ataerkil veya geniş aile		Tek yetişkinli aile		Bir arada yaşayan kişiler	
	f	%	f	%	f	%	f	%	f	%	f	%	f	%
1	181	12	202	13	127	11	168	12,56	179	12,32	80	10,38	35	12,86

2	351	25	620	40	134	11	222	16,60	168	11,57	54	7,01	49	18,01
3	229	16	11	1	360	31	320	23,93	432	29,75	265	34,41	67	27,63
4	198	14	214	13	167	14	200	14,95	179	12,32	107	13,89	38	13,97
5	435	31	482	31	342	30	427	31,3	494	34,02	264	34,28	83	30,51
	1394	100	1529	100	1130	100	1337	100	1452	100	770	100	272	100

Tablo 4: Oturulan Konut Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	Müstakil konut bağımsız		Müstakil konut tek duvarlı bitişik		Apartman bodrum/zemin		Apartman normal kat		Apartman dubleks daire		Apartman çatı katı		Diğer	
	f	%	f	%	f	%	f	%	f	%	f	%	f	%
1	403	11	94	12	37	10,9	426	13	6	6,9	6	37,5	0	0
2	216	6	86	11	112	33,2	1135	34,6	47	54,6	0	0	2	18
3	1214	35	221	29	39	11,5	207	6,3	1	1,16	0	0	2	18
4	487	14	89	12	53	15,7	463	14,1	9	10,4	1	6,25	1	9
5	1104	32	247	35	96	28,4	1042	31,8	23	26,7	9	56,2	6	54
	3424	100	737	100	337	100	3273	100	86	100	16	100	11	100

Tablo 5: Isıtma Sistemi Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	Soba		Merkezi ısıtma		Kombi		Klima		Diğer	
	f	%	f	%	f	%	f	%	f	%
1	639	12,07	102	13,09	188	12,78	40	12,61	3	11,11
2	523	9,88	342	43,90	605	41,15	121	38,17	7	25,92
3	1631	30,82	9	1,15	28	1,90	13	4,10	3	11,11
4	740	13,98	99	12,70	208	14,14	51	16,08	5	18,52
5	1758	33,22	227	29,13	441	30	92	29,02	9	33,33
	5291	100	779	100	1470	100	317	100	27	100

Tablo 6: Faaliyet Gösterilen Alan Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	Toptan ve parekendecileret, motorlu taşıtlar		Konaklama ve yiyecek hizmeti faaliyetleri		Ulaştırma ve depolama		Bilgi ve iletişim	
	f	%	f	%	f	%	f	%
1	4	0,48	32	1,10	10	4,25	9	28,12
2	375	45,12	191	6,60	109	46,38	12	37,5
3	441	53,06	154	5,32	110	46,80	11	34,37
4	8	0,96	0	0	1	0,46	0	0
5	3	0,36	2516	86,96	5	2,12	0	0
	831	100	2893	100	235	100	32	100

Tablo 7: Medeni Durum Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	Hiç evlenmedi		Evli		Eşi öldü		Boşandı	
	f	%	f	%	f	%	f	%
1	73	36,31	871	12,91	4	0,54	24	11,76
2	28	13,93	1526	22,62	20	2,72	24	11,76
3	21	10,44	1605	23,79	16	2,17	42	20,58
4	17	8,45	993	14,72	80	10,88	13	6,37
5	62	30,84	1749	25,93	615	83,67	101	49,50
	201	100	6744	100	735	100	204	100

Tablo 8: Sektör Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	Özel		Devlet		KİT	
	f	%	f	%	f	%
1	128	1,83	807	94,82	37	80,43
2	1564	22,38	31	3,64	3	6,52
3	1665	23,82	13	1,52	6	13,04
4	1103	15,78	0	0	0	0
5	2527	36,16	0	0	0	0
	6987	100	851	100	46	100

Tablo 9: Cinsiyet Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	Erkek		Kadın	
	f	%	t	%
1	883	13,11	89	7,74
2	1538	22,83	60	5,22
3	1620	24,05	64	5,57
4	997	14,80	106	9,22
5	1697	25,19	830	72,23
	6735	100	1149	100

Tablo 10: İş Yeri Büyüklük Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	1-9 kişi		10-24 kişi		25-49 kişi		50 ve daha çok	
	f	%	f	%	f	%	f	%
1	129	4,04	143	4,58	170	44,61	530	44,12
2	924	29,00	236	7,57	111	29,13	327	27,22
3	1039	32,61	201	6,45	100	26,24	344	28,64
4	1091	34,24	12	0,38	0	0	0	0
5	3	0,094	2524	81,00	0	0	0	0
	3186	100	3116	100	381	100	1201	100

Tablo 11: Meslek Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	Kanun yapımcılar, üst düzey yöneticiler ve müdürler		Profesyonel meslek mensupları		Yardımcı profesyonel meslek mensupları		Büro ve müşteri hizmetlerinde çalışan elemanlar		Hizmet ve satış elemanları	
	f	%	f	%	f	%	f	%	f	%
1	94	12,91	325	90,02	127	42,76	118	50	113	3,72
2	291	39,97	15	4,15	86	28,95	48	20,33	211	6,94

3	340	46,70	21	5,81	83	27,94	70	29,66	193	6,35
4	3	0,41	0	0	1	0,33	0	0	0	0
5	0	0	0	0	0	0	0	0	2520	82,97
	728	100	361	100	297	100	236	100	3037	100

Tablo 11'in devamı

Küme	Nitelikli tarım, hayvancılık, ormancılık ve su ürünleri		Sanatkârlar ve ilgili işlerde çalışan		Tesis ve makine operatörleri ve montajcıları		Nitelik gerektirmeyen işlerde çalışanlar	
	f	%	f	%	f	%	f	%
1	3	0,28	52	6,13	57	8,76	83	12,13
2	21	2,01	374	44,15	300	46,15	252	36,84
3	3	0,28	421	49,70	286	44	267	39,03
4	1017	97,41	0	0	3	0,46	79	11,54
5	0	0	0	0	4	0,61	3	0,43
	1044	100	847	100	650	100	684	100

Tablo 12: Yaş Değişkenine Ait Verilerin Kümelere Dağılımı

Küme	50-54 yaş		55-59 yaş		60-64 yaş		65-üzeri yaş		15-19 yaş	
	f	%	f	%	f	%	f	%	f	%
1	114	12,92	30	3,94	12	2,10	2	0,16	0	0
2	137	15,53	101	13,27	47	8,23	35	2,87	0	0
3	170	19,27	95	12,48	46	8,05	41	3,36	1	14,28
4	160	18,14	170	22,33	129	22,59	212	17,40	0	0
5	301	34,12	365	47,96	337	59,01	928	76,19	6	85,71
	882	100	761	100	571	100	1218	100	7	100

Tablo 12'nin devamı

Küme	20-24 yaş		25-29 yaş		30-34 yaş		35-39 yaş		39-49 yaş	
	f	%	f	%	f	%	f	%	f	%
1	16	19,51	96	20,29	147	17,41	177	17,80	173	18,19
2	23	28,04	149	31,50	310	36,72	299	30,08	253	26,60
3	22	26,82	161	34,03	273	31,99	321	32,29	312	32,80
4	1	1,21	28	5,91	50	6,27	92	9,25	106	11,14
5	20	24,39	39	8,24	64	7,58	105	10,56	107	11,25
	82	100	473	100	844	100	994	100	951	100

Çalışmamızda bulunan 7.884 hane(nesne) beş kümeye bölünmüştür. Oluşan kümelerimizde 1. kümede; 972 hane (%12,3), 2. kümede; 1.598 hane (%20,3), 3. kümede; 1684 hane (%21,4), 4. kümede; 1.103 hane (%14) ve 5. kümede ise 2.527 hane (%32,1) bulunmaktadır. Dağılım incelendiğinde, çok fazla aşırılıklar bulunmamaktadır. Bu kümelerin oluşmasında;

✓ 1. kümede, mülkiyet durumu, hane halkı reisinin eğitim durumu, mesleği, iş statüsü, çalışılan yerin büyüklüğü değişkenleri,

✓ 2. kümede, hane halkının büyüklüğü, hane halkı yapısı, oturlan konut tipi, hanenin mülkiyet durumu, oturlan konutun ısıtma sistemi, hanenin ikinci bir konuta sahip olup olmadığı, otomobili olup olmadığı, hanenin kırsal alanda veya kentte yaşayıp yaşamadığı, konut değeri, yaşanan konutun büyüklüğü, yıllık kullanılabilir gelir, yapılan harcama değişkenleri,

✓ 3. kümede, hane halkının büyüklüğü, hane halkı yapısı, oturlan konut tipi, mülkiyet durumu, oturlan konutun ısıtma sistemi, hanenin ikinci bir konuta sahip olup olmadığı, otomobili olup olmadığı, hanenin kırsal alanda veya kentte yaşayıp yaşamadığı, hane halkı reisinin mesleğinde faaliyet gösterdiği sektör, konut değeri, yaşanan konutun büyüklüğü, yıllık kullanılabilir gelir, yapılan harcama değişkenleri,

✓ 4. kümede, hane halkı reisinin eğitim durumu, faaliyet gösterilen sektör, meslek, çalışılan yerin büyüklüğü değişkenleri,

✓ 5. kümede, mülkiyet durumu, hanenin ikinci bir konuta sahip olup olmadığı, otomobili olup olmadığı, hane halkı reisinin eğitim durumu, faaliyet gösterilen sektör, medeni durum, meslek, iş statüsü, cinsiyet, yaş, çalışılan yerin büyüklüğü değişkenleri etkili olmuştur.

Kümelerin oluşmasında etkili olan değişkenler en genel haliyle incelendiğinde 2. ve 3. kümede yer alan haneler için daha çok maddi kaynaklar ve gelir gibi somut kavramlar, 1. ve 4. kümede yer kişiler için ise alınan eğitim, çalışılan sektör, meslek gibi daha çok soyut kavramlar ön planda olduğu görülmektedir. 5. kümede ise hem somut hem de soyut

değişkenler önemli iken aynı zamanda ayırıcı değişken olarak cinsiyet ve yaş değişkenleri de eklenmektedir.

Oluşan beş kümede yer alan tüketici profilleri genel hatlarıyla aşağıda açıklanmıştır.

1 no.lu kümede 972 hane bulunmaktadır. Bulunan hanelerin büyük çoğunluğu kentte yaşamaktadır(kent:346 hane, kırsal:626 hane). Hane halkı reisi büyük çoğunlukla erkek olmasına rağmen az sayıda olsa da kadın hane halkı reisi de bulunmaktadır (erkek:883 kişi, kadın:89 kişi). Hane halkı üyeleri soba ile ısınan (639/972) ya müstakil konutta (403/972) veya apartman katında (426/972) oturmaktadır. Hane halkı reisinin eğitim durumuna bakıldığında oldukça yüksek eğitim seviyesi ile karşı karşıya kalmaktayız (lisans:349/972, ön lisans:144/972, y.lisans:57/972). Bu kişiler daha çok devlet kurumlarında olmak üzere sosyal güvenlik ve eğitim (244/972) ile savunma (388/972) sektöründe çalışmaktadır. Kümede bulunan hane halkı reislerinin yaşları 40 ile 50 arasındadır.

2 no.lu kümede 1.598 hane bulunmaktadır. Yıllık kullanılabilir gelir ve servet (ikinci konut sahipliği, otomobil sahipliği ve konut değeri) açısından incelendiğinde bu küme diğer kümelere göre en iyi seviyede olduğu görülmektedir. Servet durumuna paralel olacak şekilde harcamada en yüksek seviyededir. Bu haneler çoğunlukla kentsel alanda yaşamaktadır. Grup gelirini imalat sektöründen sağlamaktadır.

3. no.lu kümede 1.684 hane bulunmaktadır. Bu grup 2 no.lu kümenin aksine yıllık kullanılabilir gelir ve servet açısından en son sırada olan kümedir. Küme elemanları genellikle kırsal alanda yaşamaktadır. Kümede eğitim seviyesi olarak genellikle ilkökul seviyesinde bulunanlar toplanmıştır. Hane halkının çoğunluğunun kırsal alanda ikamet etmesinden dolayı geçimlerini tarım ve hayvancılıktan sağlanması beklenmektedir. Ancak bu kümede tarım ve hayvancılık yok denecek kadar azdır. Bu kümede bulunan haneler geçimlerini daha çok sektörlerinde çalışarak sağlamaktadır. Bunun nedeni olarak ta toprak işleriyle uğraşmanın zorlukları, gelecekte daha farklı yaşam sürmek istemeleri ve çocukların daha iyi eğitim almasından kaynaklandığı değerlendirilmektedir. Ayrıca dikkat çekici diğer bir noktada bu kümede üst düzeylerde görev alanların diğer kümelere oranla daha fazla olmasıdır. Toprak işleriyle uğraşmak gibi zor koşullarda ve güç ekonomik şartlarda insanlar kendilerine daha iyi bir gelecek sağlamak amacıyla kendilerini eğitime vermekte ve sonuçta rahat bir yaşama sahip olmaktadır.

4 no.lu kümede 1.103 hane bulunmaktadır. Bu ve 5 no.lu kümeyi gelir ve servet durumlarına bağlı olarak orta gelir düzeyinde olan sınıf

olarak niteleyebiliriz. Bu küme kazançlarını daha çok nitelikli tarım ve hayvancılık işlerinden sağlamaktadırlar. Bu kümede bulunanlar daha çok kentsel alanda oturmaktadırlar. Buradan tarım ve hayvancılıkla ilgili olarak illa kırsal alanda oturmak zorunluluğun bulunmadığı sonucuna varabiliriz. Bu kesim daha çok teknik lise veya üniversitelerin ön lisans programlarından mezun olan kişiler bulunmaktadır.

Son olarak, 5 no.lu kümede ise 2527 hane yer almaktadır. Bu kümede bulunan haneler ataerkil aile yapısına sahip olup çoğunluğu kentsel alanda yaşamaktadır. Geçimlerini büyük çoğunlukla yiyecek ve konaklama hizmetlerinden sağlamaktadırlar. Hane halkı reisi olan kadınların büyük çoğunluğu bu kümede toplanmıştır. İş olarak çoğunlukla kendilerinin sahip olduğu yerlerde çalışmaktadırlar.

Sonuç

Günümüzde hızla değişen ve gelişen dünyamızda, işletmelerin rakiplerine karşı üstünlük elde edebilmeleri için potansiyel müşterilerine ve tüketicilerine ait verilerin düzenli bir şekilde işlemeleri gerekmektedir. Özellikle pazar bölümlenmesi, hedef pazar seçimi ve niş pazar seçimi gibi stratejik kararlar çok boyutlu analiz yapabilen karar destek sistemleri yardımı ile alınmalıdır. Veri madenciliği, veri tabanlarında gizli kalmış bazı örüntüleri ortaya çıkarmak için kullanılan çok yararlı yöntemleri kapsamaktadır. Bu çalışma ile demografik ve sosyo-kültürel değişkenler ile 5 adet pazar bölümlenmesi elde edilmiştir. Daha sonra bu bölümlerden örnek olacak şekilde dört adet pazar nişi tespit edilmiştir.

Niş pazar olabilecek gruplar ile Veri Madenciliği kapsamında tespit edilen bilgiler aşağıya çıkartılmıştır.

✓ 5 kümenin yıllık kullanılabilir gelir ortalaması 23922 TL ve aylık harcaması 1872 TL'dir. Sadece 2 no.lu küme bu ortalamanın üzerinde bulunurken diğer kümeler ise bu ortalamanın altında bulunmaktadır. 2 no.lu kümeye yönelik olarak pazarlamacılar daha çok lüks ürünlere ağırlık verecek olurlarsa oldukça kazançlı çıkacaklardır.

✓ Hane Halkı reisinin bayan olanlar daha çok 5 no.lu kümede toplanmaktadır. Bu kişiler daha çok kırsal alanda yaşamaktadırlar. Yaş ortalaması 40 ile 55 arasındadır. Hane reisinin bayan olması hanenin ihtiyaçları da bu duruma göre şekillenmektedir. Pazarlamacıların bu gruba yönelik olarak geliştirecekleri pazarlama karmalarıyla başarılı olacakları değerlendirilmektedir.

✓ 1 no.lu kümede bulunan hane halkı reislerinin çoğunluğu yüksek tahsilli olup devlette çalışmaktadır. Devlette çalışan bu kişiler

düzenli gelire sahip kişilerdir. Pazarlamacılar için oldukça uygun bir niş pazardır. Bu kişilere yönelik daha çok taksitli alışveriş imkânının tanınması uzun vadede pazar payının büyümesine olanak sağlayacaktır.

✓ 4 no.lu kümede bulunanları çoğunluğu nitelikli tarım ve hayvancılık ile uğraşmakta ve bu kişiler daha çok kentsel alanda yaşamaktadır. Kanımıza göre hayvancılık daha çok kırsal alanda yapılmaktadır. Fakat gelişen teknoloji nitelikli tarım ve hayvancılık günümüzde kentsel ve kırsal alan dinlenmektedir. Dolayısıyla bu kesim pazarlamacılar için bulunmaz bir alan olduğu düşünülmektedir.

✓ Elde edilen bulgulara göre Türkiye’de büyük çoğunluk olarak çekirdek aile yapısının hüküm sürmesine rağmen halen ataerkil aile yapısının da azımsanmayacak ölçüde bulunduğu olgusuna rastlanmıştır. Dolayısıyla tüketimde bu özelliğe göre şekillenecektir.

✓ Türkiye’de yaşayan insanlar güçleri ve kuvvetleri yerinde olduğu 30 ile 43 yaşları arasında iken daha çok inşaat, imalat ve nitelikli tarım/hayvancılık işleriyle uğraşırken yaşlarının ilerlemesiyle beraber daha çok daha az çaba gerektiren satış elemanlığı ve hizmet işleri alanlarına yöneldiği tespit edilmiştir. Aslında bu durum Türkiye’nin bir gerçeğini yansıtmaktadır. Türkiye bir taraftan sanayileşme yolunda ilerken diğer taraftan teknolojik gelişmelere rağmen halen insan gücüne dayalı işlerin yapıldığı gözler önüne serilmektedir.

Yapılan literatür çalışması neticesinde çeşitli teknikler kullanılarak pazar bölümlenmesi ile ilgili yapılmış çalışmalara rastlanılmıştır. Ancak daha ileriye gidilerek Pazar nişlerinin tespit edilmesi ve bölümlendirme esnasında İki Aşamalı Kümeleme Tekniği kullanılmasıyla ilgili başka çalışmaya rastlanılmamıştır. Yapılan bu çalışma ile işletmeler daha uygun ve stratejik kararlar alması konusunda yol göstereceği düşünülmektedir.

Kaynakça

Akpınar, H. (2000). Veri tabanlarında Bilgi Keşfi ve Veri Madenciliği, İstanbul Üniversitesi İşletme Fakültesi Dergisi, 29(1), 1-22.

Alpar, R. (2011). Uygulamalı Çok Değişkenli İstatistiksel Yöntemler, Üçüncü Baskı, Detay Yayıncılık, Ankara.

Altunışık, R., Özdemir, Ş. ve Torlak, Ö. (2001). Modern Pazarlama, Değişim Yayınları, Sakarya.

Arslan, M. (2004). Spor Ayakkabısı Satınalma ve Kullanım Amaçlarına İlişkin Pazar bölümlerinin Oluşturulması, Marmara Üniversitesi, İ.İ.B.F Dergisi, Cilt XIX, Sayı 1.

Aydoğan, E., Gencer, C. ve Akbulut, S. (2008). Veri Madenciliği Teknikleri ile Bir Kozmetik Markanın Ayrılan Müşteri Analizi ve Müşteri Bölümlenmesi, Mühendislik ve Fen Bilimleri Dergisi, 26(1).

Bağdadioğlu, N, Başaran, A., Kalaycıoğlu, S. ve Pınar, A. (2009). Kamu Kolaylıkları Yönetişiminde Yoksulluğun Dikkate Alınması, Hacettepe Üniversitesi, Piyasa Ekonomisini ve Girişimciliği Geliştirme Merkezi Yayınları, Ankara

Berson, A., Smith, S. and Thearling, K. (1999). Building Data Mining Applications for CRM, McGrawHill, 510, USA.

Chen, Z. (2001). Data Mining and Uncertain Reasoning: An Integrated Approach, John Wiley and Sons, Inc., 370, Canada.

Demiralay, M. (2005). Hiyerarşik Kümeleme Metotları ile Veri Madenciliği Uygulamaları, Yayınlanmamış Yüksek Lisans Tezi, Maltepe Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Diñer, E. (2006). Veri Madenciliğinde K-Means Algoritması ve Tıp Alanında Uygulanması, Yayınlanmamış Yüksek Lisans Tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Kocaeli.

Edgell, S. (1998). Sınıf, Dost Kitapevi Yayınları, Ankara.

Edgell, S and Duke, V. (1991). A Measure of Thatcherism: A Sociology of Britain, Londra, HarperCollins.

Ercan, G., B. (2007). Niş Pazarlama Stratejilerin Turizm Sektöründe A Grubu Seyahat Acentalarında Uygulanması: ETS Tur Örneği, Yayınlanmamış Yüksek Lisans Tezi, Trakya Üniversitesi, Trakya.

Fraley, C., Raftery, A.E. (1998). How Many Clusters? Which Clustering Method? Answers via Model-based Analysis, The Computer Journal, 4(8):578-588.

Goldthorpe, J.H. (1983). Women and Class Analysis: in Defence of the Conventional View, Sociology, 17, pp. 465-488.

Han, J. and Kamber, M. (2001). Data Mining: Concepts and Techniques. USA: Morgan Kaufmann Publishers.

Jain A. K., Dubes R. C. (1988). Algorithms for Clustering Data, Prentice Hall, Englewood Cliffs, NJ.

Kalaycı, Ş. (2010). SPSS Uygulamalı Çok Değişkenli İstatistiksel Teknikleri, 5. Baskı, Asil Yayın Dağıtım, Ankara.

Kotler, P. (1994). Marketing Management, 7. Edition, Prentice Hall Englewood Cliffs N.J., International.

Kotler, P., Armstrong, G., Saunders, J. and Wong, V. (2001). Principles of Marketing, Third European Edition, Pearson Education Limited Edinburg Gate Harlow.

Kovalerchuk, B. and Vityaev, E. (2001). Data Mining in Finance: Advances in Relational and Hybrid Methods, Kluwer Academic Publishers, 308, USA.

Koyuncuğil, A.S. ve Özgülbaş, N. (2009). Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları, Bilişim Teknolojileri Dergisi, CİLT: 2(2).

Kuo, R.J., Ho, L.M. and Hu, C.M. (2002). Integration of Self-Organizing Feature Map and K-means Algorithm for Market Segmentation, Computers & Operations Research, Vol: 29, Issue: 11.

Liesen, B. (2001). Image Segmentation: The Key of A Tourist Destination, The Journal of Services Marketing, Vol. 15, Issue: 1, Santa Barbara, USA, S. 49-70.

Lilien, G.L. and Rangaswamy, A. (1998). Marketing Engineering Computer-Assisted Marketing Analysis and Planning, Addison-Wesley, Educational Publishers, Inc, USA.

Magrath, A. J. (1988). Market Smart: Proven Strategies to Outfox and Outflank Your Competition John Wiley and Sons, New York.

Mucuk, İ. (1997). Pazarlama İlkeleri, 7. Basım, Türkmen Kitapevi, İstanbul.

Mykletun, J. R., Crotts, C. J., Mykletun, A. (2001). Positioning An Island Destination In The Peripheral Area of The BALTICS: A Flexible Approach to Market Segmentation, Tourism Management, 22, p.p: 493-500.

Odabaşı, Y. (1994). Sağlık Hizmetleri Pazarlaması, Eskişehir: Anadolu Üniversitesi Yayın No:799, Açıköğretim Fakültesi Yayın No:409.

Özkan, M. www.danismend.com/kategori/altkategori/nis-pazarlama, Erişim tarihi: 04.05.2012

Rud, O. (2001). Data Mining Cookbook: Modelling Data for Marketing, Risk and Customer Relationship Management, John Wiley and Sons, Canada, U.S.A.

Rygielski C., Wang J., and Yen D. (2002). Data Mining Techniques for Customer Relationship Management, Technology in Society, 24:483-502.

Savaş, S. (2011). Analysis of GSM Networks' Performance With Data Mining, M.Sc. Thesis, Gazi University, Institute of Informatics.

Savaş, S. ve Topaloğlu, N. (2011). Veri Madenciliği Yöntemi İle Gsm Şebekelerinin Performans Analizi, Gazi Üniv. Müh. Mim. Fak. Der. Gazi Univ. Cilt 26, No 4, 741-751.

Sever, H. ve Oğuz, B. (2002). Veri Tabanlarında Bilgi Keşfine Formel Bir Yaklaşım: Kısım 1- Eşleştirme Sorguları ve Algoritmalar. Bilgi Dünyası, 3(2): 173-204.

Semiz, D. (2008). Niş Pazarlama Stratejisi ve Organik Ürünler Pazarında Bir Araştırma, Yayınlanmamış Yüksek Lisans Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.

Shani D. and Chalasani S. (1992). Exploiting Niches Using Relationship Marketing, The Journal of Services Marketing, Vol: 6, No.4, p.45.

Dalgıç, T. ve Maarten, L. (1994). Niche Marketing Revisited: Concept Applications and Some European Cases, European Journal of Marketing, Vol:28, No:4, p.39.

TÜİK (2008). Tüketim Harcamaları, Yoksulluk ve Gelir Dağılımı. Sorularla Resmi İstatistikler Dizisi – 6, Ankara

TÜİK (2011). Türkiye 2010 Yılı Hane halkı Bütçe Anketi Micro Veri Seti, Yayın No: 627, Ankara.

Watson, H. and Haley, B. (1998). Managerial considerations. Communications of the ACM, 41(9), 32-37.

Zhong, N. and Zhou, L. (1999). Methodologies for Knowledge Discovery and Data Mining, The Third Pacific-Asia Conference, Pakdd-99, Beijing, China, 26-28.