



Hanehalkı Tüketim Harcamalarına Göre İibs Düzey 2 Bölgelerin K- Ortalamalar ve Kırpılmış K- Ortalamalar Yöntemleri İle Kümelenmesi

Clustering Nuts Level 2 Regions By K-Means and Trimmed K Means Methods According To Household Consumption Expenditures

Neslihan Akın Özdemir¹ ve Cem Gürler²

¹Öğr. Gör., Zonguldak Bülent Ecevit Üniversitesi, Alaplı Meslek Yüksekokulu, Yönetim ve Organizasyon Bölümü,
neslihan.aozdemir@beun.edu.tr, Orcid ID: 0000-0002-6577-2525

²Dr. Öğr. Üyesi, Yalova Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, cem.gurler@yalova.edu.tr, Orcid ID: 0000-0001-5127-6726

MAKALE BİLGİSİ

Anahtar Kelimeler

*Hanehalkı tüketim harcaması,
Kümeleme analizi,
K-ortalamlar,
K kırpılmış ortalamlar*

Makale Geçmişi:

*Geliş Tarihi: 18 Nisan 2022
Kabul Tarihi: 26 Ekim 2022*

ARTICLE INFO

Keywords

*Household consumption
expenditures,
Cluster analysis,
K-means,
Trimmed k-means*

Article History:

*Received: 18 April 2022
Accepted: 26 October 2022*

ÖZET

Mevcut çalışmada, Türkiye’de, tüketim harcamalarına göre benzer bölgelerin belirlenmesi amaçlanmıştır. Bu bağlamda, Türkiye İstatistik Kurumu’nun (TÜİK) yayınlamış olduğu Hanehalkı Tüketim Harcamasının verileri kullanılmıştır. İlgili veride gıda ve alkolsüz içecekler, alkollü içecekler, sigara ve tütün, giyim ve ayakkabı, konut ve kira, mobilya, ev aletleri ve bakım hizmetleri, sağlık, ulaştırma, haberleşme, eğlence ve kültür, eğitim hizmetleri, lokanta ve oteller, çeşitli mal ve hizmetler olmak üzere toplamda 12 harcama mevcuttur. Ayrıca veride, TÜİK’in belirlediği Türkiye İstatistik Bölge Birimleri Sınıflandırması Düzey 2 bölgeleri yer almaktadır. Benzer bölgelerin belirlenmesi için kümeleme yöntemlerinden k-ortalamlar ve k kırpılmış ortalamlar kullanılmıştır. K-ortalamlar yönteminde küme sayısının belirlenmesi için ise Silhouette İndeksinden yararlanılmıştır. Analiz sonucunda 3 küme olduğu tespit edilmiş ve sırasıyla kümelerde 4, 4 ve 18 bölge yer almıştır. K kırpılmış ortalamlar analiz sonucuna göre 3 küme olduğu tespit edilmiş ve sırasıyla kümelerde 4, 10 ve 11 bölge yer almıştır TR10 (İstanbul) aykırı gözlem olarak belirlenmiştir.

ABSTRACT

This paper attempts to determine similar regions in Turkey according to consumption expenditures. For this reason, the data used in the analysis was gathered from Turkish Statistical Institute, Household Consumption Expenditures research. In the data, there were 12 variables: food and non-alcoholic beverages, alcoholic beverages, cigarettes, clothing and foot wear, housing and rent, furniture, house appliances, health, transportation, communication, entertainment and culture, educational services, restaurants, food services and hotels, various goods and services. Moreover, NUTS Level 2 regions was analyzed. In order to determine similar regions, k-means and k-trimmed means from clustering methods were used. The Silhouette Index was used to determine the number of clusters in the k-means method. As a result of the analysis, it was determined that there were 3 clusters and there were 4, 4 and 18 regions in the clusters, respectively. According to the results of the trimmed k-means analysis, it was determined that there were 3 clusters and there were 4, 10 and 11 regions in the clusters, respectively. TR10 (Istanbul) was determined as outlier.

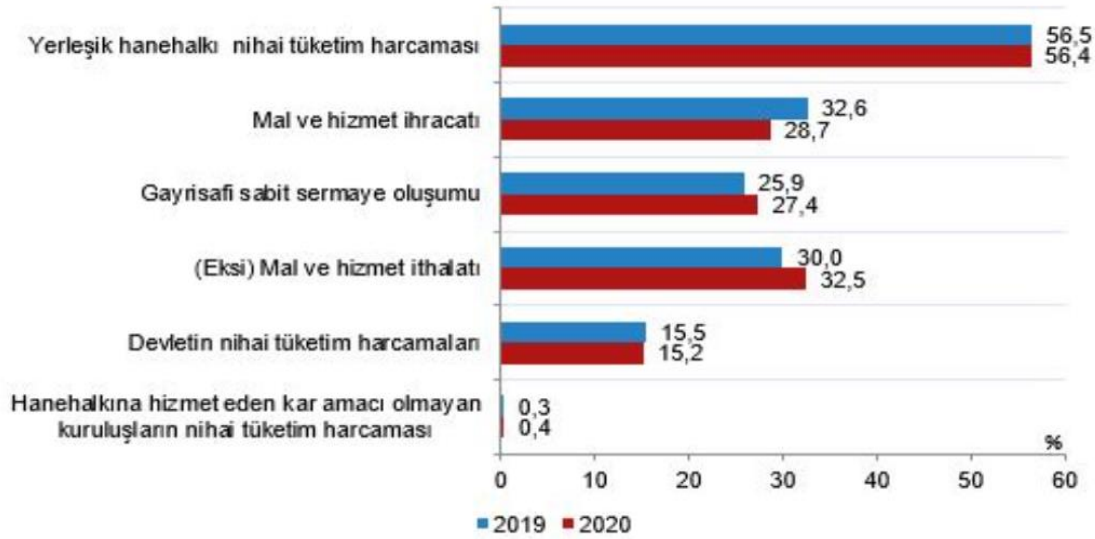
Ekonomide tüketim terimi, mal ve hizmetlerin birey veya toplum tarafından tatmin edilmek için kullanılması anlamına gelen anahtar bir kavramdır. Hanehalklarının davranışlarında tüketim harcamaları ve tasarrufları, gelirlerine bağlı olan çok önemli unsurlardır. Hanehalkları tüketim ve tasarruf arasındaki dengeyi bütçe kısıtlamaları ve ihtiyaçları ile sağlarlar. Hanehalkının gelirinde tüketim önemli bir yer tutmaktadır. Hanehalkı tüketim harcamaları, hanehalkı nihai tüketim harcamalarını ifade eder. Hanehalkı tarafından satın alınan dayanıklı ürünler (araba, çamaşır makinesi ve ev

bilgisayarı gibi) dahil tüm mal ve hizmetlerin piyasa değerini ifade eder. Konut satın alımlarını içermez, ancak sahibi tarafından kullanılan konutlar için emsal kirayı içerir. Ayrıca, izin ve lisans almak için hükümetlere yapılan ödemeleri ve ücretleri de içerir (Aslam, 2017:222; Obinna, 2020:104).

Mikro düzeydeki harcamalara (haneler, aileler) ilişkin veriler 19. yüzyılın sonları ve 20. yüzyılın başlarından beri toplanmış ve dünya çapındaki ülkelerdeki nüfusların ekonomik durumunun çeşitli yönlerine ışık tutmak için kullanılmıştır. Hanehalkı harcamalarına ilişkin mikro verilerin en yaygın kullanım alanları olarak tüketici fiyat endeksi, ulusal hesaplar ve refah analizleri olduğu görülmektedir (OECD, 2013).

Tüketici Fiyat Endeksi (TÜFE), iki dönem arasındaki fiyatlardaki bir değişiklik karşısında, bir hane halkının sabit bir mal ve hizmet sepeti satın alma maliyetindeki değişikliğin ölçülmesi esasına dayanır. Doğrudan yaşam maliyetindeki devinimleri gösteren tüketici fiyat endeksi aynı zamanda, yaşam standartlarının önemli bir göstergesi olan kişi başına düşen reel tüketimdeki değişikliklere ilişkin tahminlerin elde edilmesine de olanak sağlar. TÜFE'nin gıda, giyim, konut vb. gibi ana tüketim harcama kategorilerinin alt endekslerinin bir araya getirilmesiyle oluşturulduğu durumlarda, bu kategorilerin nispi fiyatlarındaki hareketler aynı zamanda hanehalkı harcama modellerindeki değişimlerin nedenlerine dair fikir verecektir (McLellan, 1999: 2). Ayrıca, ekonomik performansın önemli bir göstergesi olan tüketici fiyat endeksi ücretlerin, kiralarn, sözleşmelerin ve sosyal yardım ödemelerinin düzenlenmesinde ve para ve maliye politikaları için bir araç olarak yaygın bir biçimde kullanılmaktadır (Berry vd., 2019:6; Schultze, 2003: 3).

Temel ekonomik işlevi tüketim olan hanehalklarının ekonomik davranışları, ulusal ekonomi içinde özel bir rol oynamaktadır. Temel gelir kaynakları işgücü maliyetleri (veya ürettikleri ürün ve hizmetlerin satışından elde edilen kazançlardır) olan hanehalkları elde ettiğinden daha az harcayan bir özne olarak gelirin yeniden dağıtım sürecinde yer alırlar. Böylece harcanabilir gelir şeklinde, nihai tüketim harcamalarını karşılamaya yetecek kadar kaynak yaratırken, aynı zamanda tasarruf da sağlarlar. Böylece, ekonomide ulusal tasarruf için yeterli kaynakları sağlayan bir sektör oluştururlar (Hronova ve Hinds, 2013: 4). Hanehalkları ulusal tasarruf yanında tüketim harcamasında bulunarak GSYİH'ya katkı sağlamaktadırlar. Hanehalkı tüketim harcamaları, yatırım, kamu harcamaları ve net ihracat GSYİH'nın bileşenleridir.



Şekil 1. Harcama yöntemiyle GSYH ana bileşenlerinin payları, 2019, 2020

Kaynak: TÜİK (2021)

Şekil 1 incelendiğinde Türkiye için 2019 ve 2020 yılları için sırasıyla GSYİH'nin %56.5 ve %56.4'nün yerleşik hanehalkı tüketim harcamasından oluştuğu ve yerleşik hanehalkı tüketim harcamasının 2019 ve 2020 yıllarına ait GSYİH'yi oluşturan en büyük kaleme olduğu açıkça görülmektedir. GSYİH içindeki payının yüksek olması nedeniyle, mali planlamada makroekonomik politikalarda tüketim harcamaları dikkate alınmaktadır. Politika yapıcılar, tüketicilerin gelir dalgalanmaları karşısında nasıl davranacağını tahmin etmeye çalışırlar. Tüketiciler açısından tüketim olgusu bir karar verme sürecini gerektirdiğinden dolayı tüketim fonksiyonu makroekonomide davranışsal bir ilişki ortaya koymaktadır (Tapsin ve Hespsag, 2014:2).

Devletin izlediği sosyo-ekonomik politikaların temel hedeflerinden biri olan bireylerin ve toplumun refahı, çeşitli göstergelerle ölçülebilmektedir. Çoğu refah analizi tüketime yerine gelire odaklanır. Bu durumun nedeni olarak hanehalkı kaynaklarının ölçüleri (ücretler, kazançlar, gelir vb.) hakkında bilgi içeren veri setlerinin, tüketim hakkında bilgi içeren veri setlerinden daha büyük örneklerle ve daha tutarlı değişken tanımlarına sahip olması ve daha sık elde edilebilir olması gösterilebilir. Gelir eşitsizliğinin analizi çok değerli olsa da, gelir ölçümleri, aileler arasında ve zaman içinde varlıkların birikiminde veya krediye erişimde farklılıkları yakalamakta başarısızdır. Ayrıca, tüketimin özel ve devlet transferlerini ve devlet programlarının sigorta değerini yansıtmaması daha olasıdır. Bireylerin faydası tipik olarak gelirin kendisinden ziyade malların tüketimi üzerinden tanımlandığından, tüketim eşitsizliği ölçümlerinin gelir eşitsizliğinden ziyade hanehalkı refahındaki ideal bir eşitsizlik ölçüsüne yaklaştığı iddia edilebilir. Dolayısıyla tüketici harcamalarının yapısının ve zaman içindeki değişikliklerinin hanelerin, sosyal grupların ve tüm toplumların ekonomik refahını daha iyi yansıtabileceği ifade edilebilir (OECD, 2013; Meyer ve Sullivan, 52-53; Attanasio vd., 2015:105).

Tüketim harcamalarının seviyeleri ve yapıları, birbiriyle ilişkili olan bir dizi faktör tarafından belirlenir. Bu faktörlere bağlı olarak hanehalkı tüketim harcamaları farklılıklar gösterebilirler. Bu farklılıklara sebep olan faktörler arasında, farklı talep düzeyleri ve türleri, gelir, fiyatlar, gelir dağılımı, bireylerin eğitim durumu, meslek, yaş ve sosyo-kültürel faktörler yer almaktadır. Bu faktörlerin yanı sıra, tüketici en çok yararlanan mal ve hizmetleri satın almaya öncelik vererek ve gelir nedeniyle en az ihtiyaç duyulan mal ve hizmetleri satın almayı erteleyerek faydayı en üst düzeye çıkardığında tüketicinin refahı artar. Tüketici davranışlarının analizi, kültürel yapıdaki değişimlerle birlikte hem sosyal hem de demografik faktörlerin takip edilmesini ve bu izlenimler sonucunda politikaların geliştirilmesini sağlar. Bu nedenle hanehalkı tüketim harcamaları literatürde önemli ve geniş bir yere sahiptir (Noll ve Weick, 2014:16; Çağlayan ve Astar: 2012:27). Hanehalkı tüketim harcamalarını etkileyen ekonomik, demografik ve sosyo-kültürel faktörlerin yanında psikolojik faktörlerin de dikkate alınması gerekir. Hanehalkı tüketim harcamalarını etkileyen psikolojik faktörlerin belirlenmesi ekonomik ve demografik faktörler kadar basit değildir. Bu tür faktörler, tüketicinin beklentilerine, planlarına, güdülerine, iyimserliğine veya karamsarlığına vb. ayrılabilir (Zeynalova ve Mammadli, 2020:242). Harcama istatistikleri ayrıca, farklı ürün ve hizmet kategorileri ve farklı sosyal gruplar için tüketim ve talebin analizleri için zengin bir veri kaynağı sağlar (OECD, 2013).

1. METODOLOJİ

Mevcut çalışmada ülkemizde hanehalklarının tüketim harcamaları dikkate alınarak TÜİK'in 26 bölgenin yer aldığı Düzey 2 olarak ifade edilen bölgelerin gruplandırılması amaçlanmıştır. Bu amaçla analiz yöntemi olarak hiyerarşik olmayan kümeleme yöntemi olan k-ortalamalar yöntemi, küme sayısının belirlenmesi için ise Silhouette indeksi kullanılmıştır. Bu yöntemler, RStudio'da NbClust (Charrad vd., 2014), factoextra (Kassambara ve Mundt, 2017) paketleri ve R'in kendi fonksiyonlarından biri olan kmeans() fonksiyonu ile uygulanmıştır. Analizde 12 harcama kaleminin verileri kullanılmıştır.

1.1. Veri

Çalışmada TÜİK'in Hanehalkı Tüketim Harcaması Bölgesel Sonuçları kullanılmıştır. TÜİK, 2002 yılından beri düzenli olarak Hanehalkı Tüketim Harcamasını ölçmektedir. Hanehalkı Tüketim Harcaması kapsamına giren harcamaların hanehalkı için ve tüketim amacıyla yapılmış olması gerekmektedir (TÜİK, 2020). Veride Türkiye İstatistik Bölge Birimleri Sınıflandırması Düzey 2 bölgeleri ve toplamda 12 harcama kalemi bulunmaktadır. Her bir bölge için 12 harcama kaleminin toplam tüketim harcaması içindeki oranı kullanılmıştır. Bölgelerin isimleri ve bölgelerde yer alan şehirler EK-1'de yer almaktadır. Türkiye'de hanehalklarının harcama dağılımlarının ortalaması Şekil 2'de görülmektedir. Grafiğe göre, harcamalar arasında en yüksek pay konut ve kiraya (%22,80), en düşük pay ise eğitim hizmetlerine (%1,70) aittir. Ayrıca Türkiye hanehalkları harcamalarının yaklaşık %63'ünü konut ve kira, gıda ve alkolsüz içecekler ile ulaştırma kalemleri oluşturmaktadır.



Şekil 2. Hanehalkları Tüketim Harcama Ortalamaları

Tablo 1’de, ilgili harcama kalemlerinde en yüksek ve en düşük değerlere sahip olan bölgeler gösterilmiştir. TRB2 bölgesinde yer alan iller, gıda ve alkolsüz içecekler, giyim ve ayakkabı harcamalarında en yüksek, ulaştırma, haberleşme, eğitim hizmetleri ile lokanta ve oteller harcamalarında en düşük değere sahiptir. TR10 bölgesi konut ve kira, eğlence ve kültür ile eğitim hizmetleri harcamalarında en yüksek, gıda ve alkolsüz içecekler harcamalarında en düşük değere sahiptir. TRA2 bölgesi alkollü içecekler, sigara ve tütün, TR21 bölgesi mobilya, ev aletleri ve bakım hizmetleri, TR22 bölgesi sağlık, TRB1 bölgesi haberleşme, TR41 bölgesi lokanta ve oteller harcamalarında en yüksek değere sahiptir.

Tablo 1. Harcamalara Göre En Yüksek ve En Düşük Bölgeler

Harcama Kalemi	En yüksek	%	En düşük	%
Gıda ve alkolsüz içecekler	TRB2	33,2	TR10	15,6
Alkollü içecekler, sigara ve tütün	TRA2	7,6	TR63	3,4
Giyim ve ayakkabı	TRB2	8,3	TR81, TR82	4,1
Konut ve kira	TR10	29,5	TRC3, TR32	20,7
Mobilya, ev aletleri ve bakım hizmetleri	TR21	7,9	TRC3	5,4
Sağlık	TR22	3,2	TRB2, TRC2	1
Ulaştırma	TR82	23,3	TRB2	11,2
Haberleşme	TRB1	4,2	TRB2	2,8
Eğlence ve kültür	TR10	3,3	TRC2	1,2
Eğitim hizmetleri	TR10	3,5	TRB2	0,6
Lokanta ve oteller	TR41	7,8	TRB2	3,6
Çeşitli mal ve hizmetler	TRC2	6,6	TRC3	2,8

1.2. K-Ortalamlar

Kümeleme, veri kümelerinde bulunabilecek gizli desenlerin aranmasıdır. Kümeleme teknikleri, veri analizleri, örüntü tanıma, görüntü işleme ve bilgi arama gibi birçok uygulama alanında kullanılmaktadır (Zalik, 2008:1385). K-ortalamlar algoritması, temelde verileri analiz etmek için uygulanan bir bölümlenmeli kümeleme yöntemidir ve gözlemleri, çeşitli giriş veri noktaları arasındaki konumlara ve mesafeye dayalı nesnelere olarak ele alır. Nesnelere birbirini dışlayan kümeler (k) bölümlenmesi, her bir küme içindeki nesnelere birbirine olabildiğince yakın, ancak diğer kümelerdeki nesnelere mümkün olduğunca uzakta kalacağı şekilde yapılır (Ghosh ve Dubey, 2013: 35).

Kümeleme algoritmaları arasında en sık kullanılan ve basit algoritmalarından biri k-ortalamlar algoritmasıdır (Kotu ve Deshpande, 2018: 226-227). K-ortalamlar, küme içi hatalarla, yerel optimum çözümü bulmayı amaçlamaktadır. K-ortalamlar algoritmasında her bir gözlemin, küme merkezlerine en yakın noktaya atanması hedeflenmektedir. Küme merkezleri, kümede yer alan tüm gözlemlerin ortalamasıdır (Madhulatha, 2012: 721). Jain ve Dubes (1988), k-ortalamların adımlarını şöyle özetlemiştir (Jain, 2010: 654):

1. Küme sayısı (k) belirlenerek ilk bölümlenme işlemi başlar ve küme üyeliği değişmeyene kadar 2. ve 3. adımlar tekrar edilir.
2. Her bir gözlem incelenerek, kendisine en yakın kümeye atanır.
3. Küme merkezi hesaplanır.

K-ortalama uygulanması kolay olmasına karşın bazı dezavantajlara sahiptir. Algoritma başlangıçta, daha önceden belirlenen küme sayısına göre küme merkezlerini rassal olarak belirlemekte, sonrasında ise küme içi hatayı minimize etmek için, küme merkezlerini değiştirerek ilerlemektedir. Algoritmanın en bilinen dezavantajlarından biri, küme sayısına duyarlılığının fazla olmasıdır. Bu dezavantajın etkilerinin azaltılması için, algoritma birkaç defa çalıştırılması önerilmektedir (Likas, Vlassis ve Verbeek, 2003: 452). Nihai kümeleme sonuçlarının kalitesi, başlangıç merkezinin keyfi seçimine bağlıdır. Dolayısıyla, başlangıç merkezi rastgele seçilirse, farklı başlangıç merkezleri için farklı sonuçlar olacaktır. Böylece istenilen segmentasyonun elde edilmesi için başlangıç merkezi dikkatlice seçilmelidir. Ayrıca veri sayısına, küme sayısına ve yineleme sayısına bağlı olan hesaplama karmaşıklığı, K-ortalama kümelemeyi tasarlarken dikkate alınması gereken bir başka önemli konudur (Dhanachandra vd., 2015: 767).

1.3. K Kırpılmış Ortalamalar

K-ortalama algoritması basit, sezgiseldir ve herhangi bir kısıtlama olmaksızın veriye doğrudan uygulanabilir. Sadeliği ve uygulanabilirliği yaygın kullanımına katkıda bulunmuştur. Ancak veri setinde aykırı değerlerin bulunması halinde k-ortalama kararlılığı, büyük k için bile aykırı değerlerden ciddi şekilde etkilenebilir (Dorabiala vd., 2022: 10). Kaufman ve Rousseeuw (1990) tarafından önerilen PAM yönteminin aykırı değerler açısından k-ortalamalardan daha dayanıklı (robust) olduğu ifade edilmektedir. Ancak Garcia-Escudero ve Gordaliza (1999) çalışmalarında PAM'ın dayanıklı olmadığını; tek bir aykırı değer bile yöntemi bozabileceğini belirtmişlerdir. Yazarlar aynı çalışmada k kırılmış ortalamalar tekniğinden bahsetmişler ve bu yaklaşımın aykırı gözlemler karşısında dayanıklı olduğunu ifade etmişlerdir yaklaşımını önermişlerdir (Yorulmaz, 2016:119).

K-kırılmış ortalamalar yöntemi

$$\operatorname{argmin}_Y \min_{m_1, \dots, m_k} \sum_{x_i \in Y} \min_{j=1, \dots, k} \|x_i - m_j\|^2 \quad (1)$$

Denklem 1'de verilen en küçükleme problemini çözmeye çalışır. Burada Y; örneklem (x_1, \dots, x_n) içindeki $n(1-\alpha)$ adet gözlem grubunu ifade etmektedir. k- kırılmış ortalamalar algoritmasının adımlarını şöyle özetlebilir:

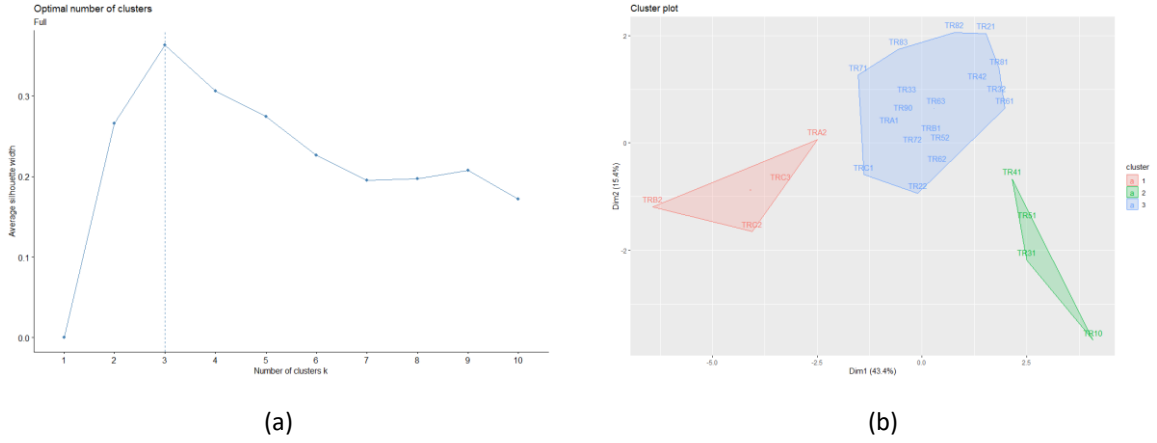
- K adet küme merkezi rastgele belirlenir m_1^0, \dots, m_k^0 .
- Yoğunlaşma Adımları
 - m_1^l, \dots, m_k^l merkezlerine en yakın $n(1-\alpha)$ gözlemden oluşan H seti belirlenir
 - H seti $\{H_1, \dots, H_k\}$ şeklinde k adet alt kümeye bölünür. Burada H_j , H seti içerisindeki m_j^l merkezine diğer merkezlere göre daha yakın olan gözlemleri içerir.
 - m_j^{l+1}, H_j 'de yer alan gözlemlerin örnek ortalaması olmak üzere tüm küme merkezleri $(m_1^{l+1}, \dots, m_k^{l+1})$ güncellenir.
- Adım 1 ve Adım 2'yi birkaç kez tekrarlanır ve Eşitlik 1'de verilen amaç fonksiyonunu minimize eden çözüm uygun çözümdür (García-Escudero vd., 2010:95,97).

2. BULGULAR

Mevcut çalışmada, Düzey 2 bölgelerinin tüketim harcamalarına göre kümelenmesi amaçlanmış ve yöntem olarak hiyerarşik olmayan kümeleme yöntemlerinden k-ortalama ve k-kırılmış ortalamalar kullanılmıştır.

2.1. K-Ortalama Sonuçları

K-ortalama yönteminde ilk olarak küme sayısı belirlenmelidir. Küme sayısının belirlenmesinde dirsek yöntemi, Silhouette indeksi, GAP istatistiği gibi farklı yöntemler mevcuttur. Dirsek yönteminde, küme içi hata kareleri kullanılarak çizilen grafikte birlikte, karar vericinin küme sayısına karar vermesi beklenmektedir. Silhouette indeksi ve GAP istatistiği yöntemlerinde ise karar vericiye, küme sayısının olması gereken değer belirtilmektedir. Bu bağlamda Silhouette indeksi kullanılmıştır. Şekil 3 (a)'da görüldüğü üzere küme sayısı 3 olarak belirlenmiştir. K-ortalama ile yapılan kümeleme analizi sonucunda oluşan kümeler Şekil 3 (b)'de yer almaktadır.



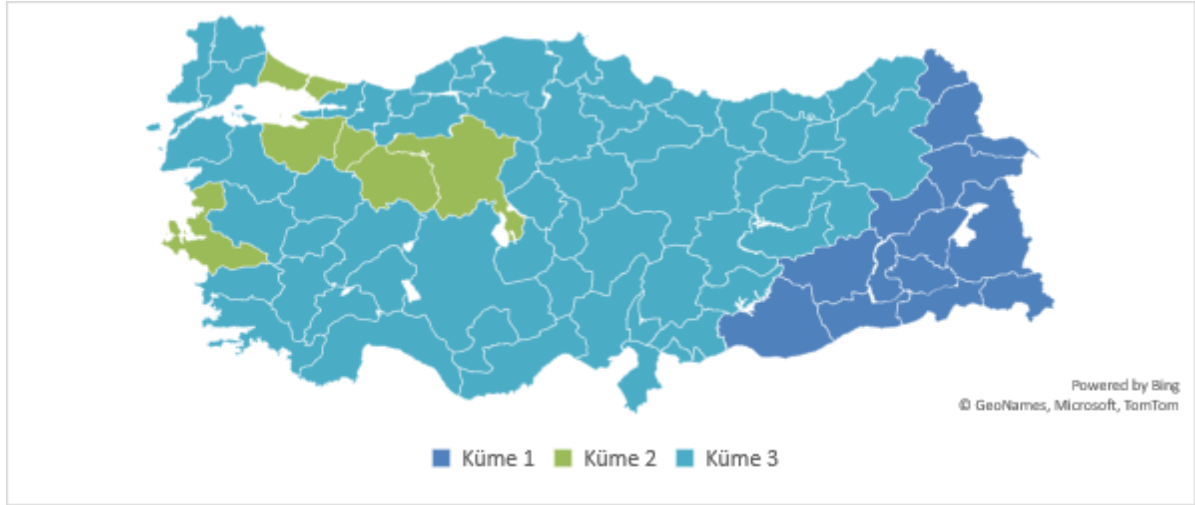
Şekil 3. Silhouette Sonucu Optimum Küme Sayısı Ve K-Ortalamlar Sonucu Oluşan Kümeler

Kümelerde yer alan iller Tablo 2’de gösterilmektedir. Kümelerde sırasıyla 4, 4 ve 18 İBBS-2 bölgesi yer almıştır.

Tablo 2. Analiz Sonucunda Kümelerde Yer Alan Bölgeler

Küme	Bölgeler
Küme 1	TRA2, TRB2, TRC2, TRC3
Küme 2	TR10, TR31, TR41, TR51
Küme 3	TRA1, TRB1, TRC1, TR21, TR22, TR32, TR33, TR42, TR52, TR61, TR62, TR63, TR71, TR72, TR81, TR82, TR83, TR90

Şekil 4’te oluşan kümeler haritada gösterilmektedir. Küme 1’de coğrafya etkisi açık bir şekilde görülmektedir. Bu kümede yer alan bölgeleri Doğu Anadolu ve Güneydoğu Anadolu’da yer alan iller oluşturmaktadır. Küme 2’de yer alan bölgelere bakıldığında ise nüfus etkisi göze çarpmaktadır. Türkiye’de nüfus anlamında en büyük ilk 4 il (İstanbul, Ankara, İzmir ve Bursa) bu kümede yer almaktadır. Bu kümede yer alan bölgelerin, Türkiye nüfusu içindeki payı %35,77’dir. TRB41 bölgesinde yer alan Eskişehir ve Bilecik illerinin, Bursa etkisiyle birlikte bu kümede yer aldığı ifade edilebilir.



Şekil 4. Oluşan Kümelerin Haritada Gösterimi

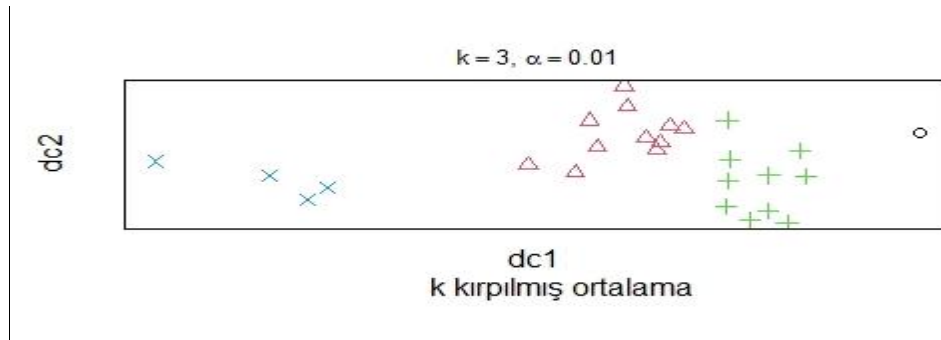
Analiz sonucunda elde edilen kümelerde yer alan illerin, harcamalarının ortalamaları ve standart sapmaları Tablo 3'te yer almaktadır. Tablo 3'ten elde edilen bilgilere göre, Küme 1 gıda ve alkolsüz içecekler, alkollü içecekler, sigara ve tütün ile giyim ve ayakkabı kalemlerinde daha fazla harcama yapmaktadır. Küme 2 konut ve kira, sağlık, haberleşme, eğlence ve kültür ile eğitim hizmetlerine daha fazla harcamışlardır. Küme 3'te mobilya, ev aletleri ve bakım hizmetleri, ulaştırma ile çeşitli mal hizmetler kalemlerine en fazla harcama yapmışlardır.

Tablo 3. Oluşan Kümelerde Yapılan Tüketim Harcamalarının Ortalamaları ve Standart Sapmaları

	Küme 1		Küme 2		Küme 3	
	Ort	Std	Ort	Std	Ort	Std
Gıda ve alkolsüz içecekler	29,60	2,17	17,43	1,07	22,04	2,07
Alkollü içecekler, sigara ve tütün	5,90	1,14	4,30	0,58	4,53	0,81
Giyim ve ayakkabı	7,23	1,15	4,68	0,45	5,11	0,56
Konut ve kira	21,45	0,77	26,30	2,25	22,32	1,05
Mobilya, ev aletleri ve bakım hizmetleri	5,70	0,27	6,08	0,40	6,84	0,58
Sağlık	1,30	0,31	2,13	0,33	2,01	0,46
Ulaştırma	13,65	1,53	17,98	1,37	19,52	1,79
Haberleşme	3,43	0,43	3,88	0,18	3,49	0,31
Eğlence ve kültür	1,68	0,34	3,13	0,13	2,51	0,33
Eğitim hizmetleri	1,03	0,41	2,83	0,51	1,59	0,40
Lokanta ve oteller	4,50	0,85	7,18	0,51	5,58	1,16
Çeşitli mal ve hizmetler	4,55	1,53	4,13	0,19	4,49	0,70

1.1. K Kırpılmış Ortalamalar Sonuçları

K-kırpılmış ortalamalar ile yapılan kümeleme analizi sonucunda elde edilen sonuçlar Tablo 4 ve Şekil 5'de özetlenmiştir.



Şekil 5. K Kırpılmış Ortalama Yöntemine Göre Elde Edilen Kümeler

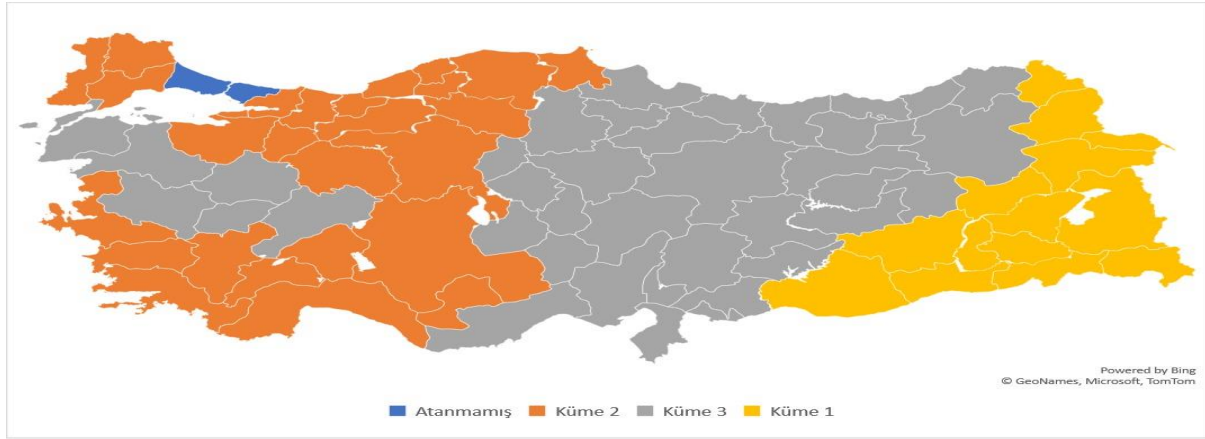
Şekil 5'de "x" küme 1'i, "+" küme 2'yi, "Δ" küme 3'ü ve "o" ise aykırı gözlem olan TR10 (İstanbul)'u ifade etmektedir.

K kırılmış ortalamalar yaklaşımı ile yapılan kümeleme analizi sonucunda oluşan kümeler Tablo 4’de yer almaktadır. Küme üyeliklerine ilişkin bilgiler değerlendirildiğinde kümelere sırasıyla 4, 10 ve 11 İBBS-2 bölgesi yer almıştır. TR10 (İstanbul)’un herhangi bir kümeye atanmadığı görülmektedir.

Tablo 4. Analiz Sonucunda Kümelerde Yer Alan Bölgeler

Küme	Bölgeler
Küme 1	TRA2, TRB2, TRC2, TRC3
Küme 2	TR21, TR31, TR32, TR41, TR42, TR51, TR52, TR61, TR81, TR82,
Küme 3	TRA1, TRB1, TRC1, TR22, TR33, TR62, TR63, TR71, TR72, TR83, TR90

Tablo 2 ve Tablo 4’de yer alan küme üyelikleri kıyaslandığında dayanıklı yaklaşımla ile aykırı gözlem olarak belirlenen TR10 (İstanbul)’un k-ortalamar kümeleme de TR31 (İzmir), TR41 (Bursa, Eskisehir, Bilecik) ve TR51 (Ankara) ‘inyer aldığı kümede yer aldığı görülmektedir. Her iki yaklaşımda da Küme 1 TRA2, TRB2, TRC2, TRC3 şehirlerinden oluşmaktadır. TR21, TR32, TR42, TR52, TR61, TR81 ve TR82’nin ait olduğu kümeler her iki yaklaşımda da farklıdır. Bu farklılık aykırı gözlem varlığının bir sonucu olarak değerlendirilebilir. Fakat bunun dışındaki sonuçlar uyumludur.



Şekil 6. Oluşan Kümelerin Haritada Gösterimi

Şekil 6’da oluşan kümeler haritada gösterilmektedir. Küme 1’de coğrafya etkisi açık bir şekilde görülmektedir. Bu kümede yer alan bölgeleri Doğu Anadolu ve Güneydoğu Anadolu’da yer alan iller oluşturmaktadır.

Tablo 5. Dayanıklı Kümeleme Analizi Sonucu Bulunan Küme Merkezleri

	Küme 1	Küme 2	Küme 3
Gıda ve alkolüz içecekler	29,60	19,38	23,36
Alkollü içecekler, sigara ve tütün	5,90	4,44	4,63
Giyim ve ayakkabı	7,23	4,76	5,35
Konut ve kira	21,45	22,93	22,55
Mobilya, ev aletleri ve bakım hizmetleri	5,70	6,85	6,66
Sağlık	1,30	2,11	1,92
Ulaştırma	13,65	20,19	18,57
Haberleşme	3,43	3,69	3,43
Eğlence ve kültür	1,68	2,85	2,35
Eğitim hizmetleri	1,03	1,91	1,58
Lokanta ve oteller	4,50	6,59	5,08
Çeşitli mal ve hizmetler	4,55	4,31	4,55

Tablo 5’ de yer alan bulgular değerlendirildiğinde; Küme 1 ve Küme 3’ün hanehalkı tüketim harcaması kalemleri içinde en fazla gıda ve alkolüz içecekler, en az ise eğitim hizmetlerine harcama yaptığı; Küme 2’nin ise en fazla konut ve kira harcaması yaptığı en az ise eğitim hizmetlerine harcama yaptığı görülmektedir.

3. SONUÇ

Bu çalışmada ülkemiz hanehalklarının tüketim harcamaları dikkate alınarak TÜİK’in 27 bölgenin yer aldığı Düzey 2 olarak ifade edilen bölgelerin gruplandırılması amaçlanmıştır. Bu amaçla analiz yöntemi olarak hiyerarşik olmayan kümeleme yöntemi olan k-ortalamar yöntemi kullanılmıştır. Analizde 12 harcama kalemi ele alınmıştır. 12 harcama kaleminin kümelere göre dağılımından elde edilen sonuçlar aşağıdaki gibi değerlendirilmiştir:

Herhangi bir sosyal güvenlik kurumunun güvencesi altında olmayan veya bir sosyal güvenlik kurumuna bağlı olduğu halde yönetmelikler gereği sağlık harcamaları karşılanmayan hanhalkları sağlık harcamalarını kendileri üstlenmektedir. Hanhalklarının yapmış oldukları bu tür sağlık harcaması “cepten yapılan sağlık harcaması” olarak adlandırılmaktadır. Türkiye’de cepten yapılan sağlık harcamalarına “Hanhalkı Bütçe Anketinden”nden elde edilen veriler yoluyla ulaşılmaktadır (Tokatlıoğlu ve Tokatlıoğlu, 2013:9). Türkiye OECD ülkeleri ile kıyaslandığında 2018 yılı için kişi başına 214 US doları ile en az cepten sağlık harcaması yapan sondan ikinci ülke konumundadır ayrıca Türkiye’yi kamunun yapmış olduğu zorunlu sağlık harcamaları açısından OECD ülkeleri ile kıyasladığımızda ise Türkiye GSYH’nın sadece 3,2’si kadar zorunlu sağlık harcaması yaparak 2018 yılı için OECD ülkeleri içinde sondan 3.sırada yer almaktadır. Türkiye için sağlık harcamaları açısından kümeleri değerlendirdiğimizde hem k- ortalamalar yöntemine göre hem de dayanıklı kümeleme yöntemine göre Doğu ve Güneydoğu Anadolu Bölge’lerindeki çoğu ilin (TRA2, TRB2, TRC2, TRC3) içinde bulunduğu küme 1’de yer alan hanhalkları diğer kümelerdeki hanhalklarına göre sağlık için daha az harcama yapmışlardır. Küme 1’deki hanhalklarının sağlık harcamalarının toplam tüketim harcamaları içindeki payı sadece %1,30’dur. Kümeler arasındaki sağlık harcamasına yönelik farklılıkların azaltılması için cepten yapılan sağlık harcamalarının artırılması çok önemlidir.

Beşeri sermayeye yapılan eğitim ve yatırım, herhangi bir ülkede, özellikle Türkiye gibi gelişmekte olan bir ülkede her zaman ekonomik kalkınmanın en önemli faktörlerinden biri olarak kabul edilmektedir. Eğitim makro ve düzeyde birçok etkiye sahiptir. Makro düzeyde, bireylerin daha iyi beceri ve bilgi edinmesine olanak tanıyarak ekonomik büyümeyi artırır. Mikro düzeyde ise, birçok aile için beşeri sermayeye yatırım, yoksulluktan kurtulmanın başlıca yolu olarak görülmektedir. Diğer bir neden de sosyal statüdür. İyi eğitilmiş insanlar genellikle toplumda daha saygı görürler (Huy, 2012:59). 2015 yılı için Türkiye’yi kamunun yapmış olduğu eğitim harcamaları açısından OECD ülkeleri ile kıyasladığımızda Türkiye GSYH’nın 3,8’i kadar ilköğretimden üniversiteye eğitim harcaması yapmış, üniversite için yapılan eğitim harcamalarına bakıldığında ise bu oran 1,24 olup OECD ülkeleri içinde 9. sırada yer almaktadır. Türkiye için eğitim harcamaları açısından kümeleri değerlendirdiğimizde hem k- ortalamalar yöntemine göre hem de dayanıklı kümeleme yöntemine göre Doğu ve Güneydoğu Anadolu Bölge’lerindeki çoğu ilin (TRA2, TRB2, TRC2, TRC3) içinde bulunduğu küme 1’de yer alan hanhalkları diğer kümelerdeki hanhalklarına göre eğitim için daha az harcama yapmışlardır. Küme 1’deki hanhalklarının eğitim harcamalarının toplam tüketim harcamaları içindeki payı sadece %1,03’tür. Kümeler arasındaki eğitim harcamasına yönelik farklılıkların azaltılması için eğitim harcamalarının artırılarak eğitimde fırsat eşitliği sağlanmalıdır.

Barınma temel bir ihtiyaçtır. Konutun başka bir nesneyle ikame edilememesi nedeni ile bireyler konuta diğer tüketim mallarından daha fazla bağımlıdır. Çoğu hanhalkı için konut, toplam aylık harcamaların en büyük payını oluşturmaktadır. Konuta ne kadar harcanacağına dair karar, kiralamak ve sahip olmak arasındaki seçimle büyük ölçüde ilgilidir. Kiracıların kira ödemesi gerekir. Ev sahipleri genellikle borçludur, yüksek bir ilk ödeme yapmak zorunda kalırlar ve ödenmemiş kredilerini ve tahakkuk eden faizi geri ödemeleri gerekir (Uğurlar ve Ecelal, 2014: 133; Beer ve Wagner, 2017: 48; Charlier vd., 2001: 72). Türkiye için konut ve kira harcamaları açısından k kırılmış ortalamalar sonucu elde edilen kümeleri değerlendirdiğimizde Ankara, İzmir, Bursa illerinin içinde bulunduğu küme 2’de yer alan hanhalkları diğer kümelerdeki hanhalklarına göre konut ve kira için daha fazla harcama yapmışlardır.

Gıda ve alkolsüz içecekler, alkollü içecekler, sigara ve tütün, giyim ve ayakkabı harcamaları açısından k kırılmış ortalamalar sonucu elde edilen kümeler kıyaslandığında söz konusu harcama kalemleri için küme 2 en az harcama yapan kümedir. Bu sektörlerden herhangi biri veya birinde faaliyet gösteren ya da gösterecek firmalar için küme 2 yeni bir pazar alanı oluşturmaktadır. Mobilya, ev aletleri ve bakım hizmetleri, ulaştırma, lokanta ve oteller, eğlence ve kültür harcamaları açısından k kırılmış ortalamalar sonucu elde edilen kümeler kıyaslandığında söz konusu harcama kalemleri için küme 1 en az harcama yapan kümedir. Bu sektörlerden herhangi biri veya birinde faaliyet gösteren ya da gösterecek firmalar için küme 1 yeni bir pazar alanı oluşturmaktadır. İlgili sektörlerden herhangi biri veya birinde faaliyet gösteren ya da gösterecek firmalar pazarlamanın 7P’si olan ürün, fiyat, dağıtım, tutundurma, insan, fiziki kanıt ve süreç değişkenleri ile pazarı etkileyerek müşteri tercihlerini kendilerine çevirebilirler.

YAZAR BEYANI

Araştırma ve Yayın Etiği Beyanı: Bu çalışma bilimsel araştırma ve yayın etiği kurallarına uygun olarak hazırlanmıştır.

Etik Kurul Onayı: Bu araştırma etik kurul izni gerektiren analizleri kapsamadığından etik kurul onayı gerektirmemektedir.

Yazar Katkıları: Yazarlar çalışmayı eşit katkı oranıyla gerçekleştirmiştir.

Çıkar Çatışması: Yazar açısından ya da üçüncü taraflar açısından çalışmadan kaynaklı çıkar çatışması bulunmamaktadır.

KAYNAKÇA (REFERENCES)

- Aslam, A. L. (2017). Does consumption expenditure induce the economic growth? An empirical evidence from Sri Lanka. *World Scientific News*, 81(2), 221-234.
- Attanasio, O., E. Hurst, and L. Pistaferri.(2015). The evolution of income, consumption, and leisure inequality in the us, 1980–2010.C.D. Carroll, T.F. Crossley, and J. Sabelhaus (Eds). *In Improving the Measurement of Consumer Expenditures* (100–140). University of Chicago Press
- Beer, C. and Wagner, K. (2017). Household's housing expenditure in Austria, Germany and Italy. *Monetary Policy and the Economy* (Q4/17). <https://www.oenb.at/en/Publications/Economics/Monetary-Policy-and-the-Economy.html>
- Berry, Francien, Brian Graf, Michael Stanger, and Mari Ylä-Jarkko (2019). Price Statistics Compilation in 196 Economies: The Relevance for Policy Analysis.” *International Monetary Fund Working Paper* 19/163
- Charlier, E., Melenberg, B., & van Soest, A. (2001).” An Analysis of Housing Expenditure Using Semiparametric Models and Panel Data”, *Journal of Econometrics*, 101, 71–107.
- Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., & Charrad, M. M. (2014). “Package ‘nbclust’”, *Journal of Statistical Software*”, 61, 1-36.
- Çağlayan, E. ve Astar, M. (2012). A Microeconomic Analysis of Household Consumption Expenditure Determinants for Both Rural and Urban Areas in Turkey. *American International Journal of Contemporary Research*, 2(2), 27-34.
- Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). “Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm”, *Procedia Computer Science*, 54, 764–771.
- Dorabiala, O., Kutz, J. N., & Aravkin, A. Y. (2022). Robust trimmed k-means. *Pattern Recognition Letters*, 161, 9-16.
- Eurostat (.2020). Household Budget Survey. (11.08.2020). <https://ec.europa.eu/eurostat/web/microdata/household-budget-survey>.
- García-Escudero, L. A., Gordaliza, A., Matrán, C., & Mayo-Iscar, A. (2010). A review of robust clustering methods. *Advances in Data Analysis and Classification*, 4(2), 89-109.
- Ghosh, S., & Dubey, S.K. (2013). “Comparative Analysis of K-Means and Fuzzy C Means Algorithms”, *International Journal of Advanced Computer Science and Applications*, 4 (4), 35-39.
- Huy, V.Q.(2012). “Determinants of Educational Expenditure in Vietnam”, *International Journal of Applied Economics*, 9(1), 59-72.
- Hronova, S., Hindls, R. (2013). Czech households in the years of crises. *Statistics and Economic Journal*, 93(4), 4-23.
- Jain, A. K. (2010).” Data clustering: 50 years beyond K-means”, *Pattern Recognition Letters*, 31(8), 651-666.
- Kassambara, A., & Mundt, F. (2017). “Package ‘factoextra’”, Extract and Visualize the Results of Multivariate Data Analyses, 76.
- Kotu, V., & Deshpande, B. (2018). *Data Science: Concepts and Practice*. Morgan Kaufmann.
- Likas, A., Vlassis, N., & Verbeek, J. J. (2003).” The Global K-Means Clustering Algorithm”, *Pattern recognition*, 36(2), 451-461.
- Madhulatha, T. S. (2012). An overview on clustering methods. arXiv preprint arXiv:1205.1117.
- McLean, I. W. (1999). Consumer Prices and Expenditure Patterns in Australia 1850-1914. *Australian Economic History Review*, 39(1), 1–28.
- Meyer, B. D., & Sullivan, J. X. (2011). Viewpoint: Further results on measuring the well-being of the poor using income and consumption. *Canadian Journal of Economics/Revue Canadienne D'économique*, 44(1), 52–87.
- Mont, O., & Power, K. (2010). “The Role of Formal and Informal Forces in Shaping Consumption and Implications for a Sustainable Society. Part I”, *Sustainability*, 2(7), 2232–2252.
- Noll, H.-H., & Weick, S. (2014). *Consumption expenditures and subjective well-being: empirical evidence from Germany. International Review of Economics*, 62(2), 101–119 .
- Obinna, o. (2020). Effect of Inflation on Household Final Consumption Expenditure in Nigeria. *Journal of Economics and Development Studies*, 8(1), 104-111
- OECD, (2013). *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*. Paris: OECD.
- Schultze, C. L. (2003). The Consumer Price Index: Conceptual Issues and Practical Suggestions. *Journal of Economic Perspectives*, 17(1), 3–22.
- Tapsin, G. ve Hepsag, A. (2014). An Analysis of Household Consumption Expenditures in EA-18. *European Scientific Journal* , 10(16), 1-12.
- Tokatlıoğlu, İ. ve Tokatlıoğlu, Y. (2013). “Türkiye’de 2002-2011 Yılları Arasında Katastrofik Sağlık Harcamalarının Yoksulluk Yaratma Kapasitesi”, *Ekonomik Yaklaşım*, 24(87), 1-36.
- TÜİK (2019). *Hanehalkı Tüketim Harcaması*. (10.08.2020). <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=30584>.
- TÜİK (2020). *Hanehalkı Tüketim Harcaması*. [https://data.tuik.gov.tr/Bulten/Index?p=Hanehalki-Tuketim-Harcamasi-\(Bolgesel\)-2019-33594](https://data.tuik.gov.tr/Bulten/Index?p=Hanehalki-Tuketim-Harcamasi-(Bolgesel)-2019-33594)
- Uğurlar, A. ve Eceral, T.Ö.(2014). “Ankara’da Mevcut Konut (Mülk ve Kiralık) Piyasasına İlişkin bir Değerlendirme”, *İdealkent*, (12), 132-159.
- Yorulmaz, Ö. (2016). *Dayanıklı İstatistiksel Yöntemler ve R Uygulamaları* (1 st ed.). İstanbul : Beta.
- Zalik, K.R. (2008). “An Efficient k’ -means Clustering Algorithm”, *Pattern Recognition Letters*, 29(9), 1385–1391.
- Zeynalova Z., Mammadli M., (2020). Analysis of the Economic Factors Affecting Household Consumption Expenditures in Azerbaijan. *Journal of Critical Reviews*, 7(7), 241-248.

EKLER (APPENDIX)**Ek 1 - Bölge kodları ve adları.**

Bölge Kodu	Bölge Adı
TRA1	Erzurum, Erzincan, Bayburt
TRA2	Agri, Kars, Iğdir, Ardahan
TRB1	Malatya, Elazığ, Bingöl, Tunceli
TRB2	Van, Mus, Bitlis, Hakkari
TRC1	Gaziantep, Adiyaman, Kilis
TRC2	Sanliurfa, Diyarbakir
TRC3	Mardin, Batman, Sırnak, Siirt
TR10	Istanbul
TR21	Tekirdag, Edirne, Kırklareli
TR22	Balıkesir, Çanakkale
TR31	Izmir
TR32	Aydin, Denizli, Mugla
TR33	Manisa, Afyon, Kütahya, Usak
TR41	Bursa, Eskisehir, Bilecik
TR42	Kocaeli, Sakarya, Düzce, Bolu, Yalova
TR51	Ankara
TR52	Konya, Karaman
TR61	Antalya, Isparta, Burdur
TR62	Adana, Mersin
TR63	Hatay, Kahramanmaraş, Osmaniye
TR71	Kirikkale, Aksaray, Nigde, Nevşehir, Kırşehir
TR72	Kayseri, Sivas, Yozgat
TR81	Zonguldak, Karabük, Bartın
TR82	Kastamonu, Çankırı, Sinop
TR83	Samsun, Tokat, Çorum, Amasya
TR90	Trabzon, Ordu, Giresun, Rize, Artvin, Gümüşhane

