

## K -MEANS KÜMELEME ALGORİTMASI İLE OECD ÜLKELERİNİN VERGİ YÜKÜ VE KAYITDIŞI EKONOMİ ANALİZİ<sup>1</sup>

Selcan ÜNAL<sup>2</sup>

### ÖZ

Kayıtdışı ekonomi ülkenin vergi kapasitesi ile toplanan vergi arasındaki farkı ifade etmektedir. Kayıtdışı ekonominin oranı, devletin kamu harcamalarını sürdürmek için ihtiyaç duyduğu gelir kaynağı olan vergi gelirini fazlasıyla etkilemektedir. Kayıtdışı ekonominin artması vergi yükünü arttırmakta bu durumda kayıtdışı ekonomi oranının artması ile sonuçlanmaktadır. Bu karşılıklı etkisi olan vergi yükü ve kayıtdışı ekonomi değişkenlerinin OECD ülkelerindeki durumunu görebilmek adına çalışmada keşif analizi olan k- means kümeleme yöntemi kullanılmıştır. İlk önce vergi yükü verisi dikkate alınarak analiz gerçekleştirilmiştir. Daha sonra kayıtdışı ekonomi verisi dahil edilerek analiz tekrarlanıp, ülkeler arası gerçekleşen küme ayrışmaları yorumlanmıştır. Meksika, Kolombiya, Kore, Kosta Rika ve Türkiye'nin her iki analizde de diğer ülkelerden ayrıştığı görülmektedir. Bu ülkelerin ortak özelliği vergi yükü oranının az ve kayıtdışı ekonomi oranının fazla olmasıdır. Analiz neticesinde bu ülkelerde subjektif vergi yükünün fazla olduğu yorumu yapılabilmektedir ve kayıtdışı ekonomi oranının düşürülmesi adına gerçekleştirilecek uygulamalar zaruridir.

**Anahtar Kelimeler:** Vergi Yükü, Kayıtdışı Ekonomi, Kümeleme, K-Means, Vergi Geliri

## ANALYSIS OF TAX BURDEN AND SHADOW ECONOMY OF OECD COUNTRIES WITH K-MEANS CLUSTERING ALGORITHM

### ABSTRACT

In order to see the situation of this mutually influential tax burden and shadow economy variable in OECD countries, the k-means clustering method, which is an exploratory analysis, was used in the study. First, the analysis was carried out by taking into account the tax burden data. Then, the analysis was repeated by including the shadow economy data, and the cluster divergences between countries were interpreted. It is seen that Mexico, Colombia, Korea, Costa Rica and Turkey differ from other countries in both analyzes. The common feature of these countries is that the tax burden rate is low and the shadow economy rate is high. As a result of the analysis, it can be interpreted that the subjective tax burden is high in these countries and practices to be carried out in order to reduce the rate of shadow economy are essential.

**Keywords:** Tax Burden, Shadow Economy, Clustering, K-Means, Tax Revenue

<sup>1</sup> Bu çalışma, 119B101 numaralı Tübitak 4005 projesi kapsamında düzenlenen eğitim ve etkinliklerle geliştirilmiştir.

<sup>2</sup>Araştırma Görevlisi, KTO Karatay Üniversitesi, selcan.unal@karatay.edu.tr, ORCID ID:0000-0001-7855-7085.

**Atıf/Citation:** Ünal, S. (2021). K -Means Kümeleme Algoritması İle OECD Ülkelerinin Vergi Yükü ve Kayıtdışı Ekonomi Analizi, Kırklareli Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi, 2 (2), 1-15.

## GİRİŞ

Kamu gelirleri içerisinde kuşkusuz en büyük payı vergiler almaktadır. Gelişmiş ve gelişmekte olan ülkelerin nerdeyse tamamında en önemli kamu gelir kalemi vergilerdir. Kamu hizmetlerinin devamlılığı esas iken, kamu harcamaları gün geçtikçe çeşitlenmekte ve yükü artmaktadır. Devletlerin kamu harcamalarını karşılama görevini idame ettirebilmesi için, devamlılığı olan gelir kaynaklarına ihtiyacı vardır. Doğal kaynak, maden, petrol vb. konularda zenginliği olmayan ülkeler için bu kaynak vergilerdir (Şenyüz, Yüce ve Gerçek, 2016).

Bir ülkede toplam vergi gelirinin miktarı, o ülkenin ekonomik, siyasi, coğrafi koşullarına bağlı değişkenlik gösterir. Ekonomik faktör olarak ülkelerin gelişmişlik seviyesi, ülke ekonomisinin dışa açıklık seviyesi, monetizasyon derecesi, kentleşme oranı gibi etkenler sayılabilir. Bu faktörlerin yanında ülkenin kayıtdışı ekonomi oranı da toplam vergi miktarını önemli oranda etkilemektedir.

Kayıtdışı ekonominin oluşmasında baskın neden ülkeden ülkeye değişmektedir fakat genel olarak ele aldığımızda kayıtdışı ekonomiyi var eden birçok ortak faktör mevcuttur. Ekonominin gelişmişlik düzeyinin yetersizliği, sosyolojik yapıdaki sorunlar, kronik yüksek enflasyon ve ekonomik krizler gibi yapısal sorunlar bu faktörlerden birkaçıdır. Bunun yanı sıra vergi mevzuatının karmaşıklığı, tekerrür eden vergi afları, vergi oranlarının yüksekliği, vergi tabanının yayılmaması, belge düzeninin yerleşik olmaması gibi mali nedenler ve vergi dairesi ve denetiminden kaynaklanan idari nedenlerde kayıtdışı ekonomiyi oluşturan nedenler arasında sayılabilmektedir.

Kayıtdışı ekonomi vergi kapasitesi ile toplanan vergi arasındaki farktır. Bir ülkede kayıtdışı ekonomi oranı ne kadar fazla olursa o ülkenin vergi kaybı da o kadar fazla olur. Vergi kaybı fazlalığı vergi ödeyen mükelleflerin vergi yükünü arttırır. Vergi yükü artışı kayıtdışılığı cazip hale getirmekte ve kayıtdışı ekonominin oranını attırmaktadır.

Ulusal ve uluslararası literatürde vergi yükü ile kayıtdışı ekonomi arasındaki ilişkinin analiz edildiği birçok çalışma bulunmaktadır. Bu ilişkinin pozitif mi, negatif mi olduğu farklı analiz teknikleri ile araştırılmıştır. Bunlar MIMIC yöntemi (Schneider, 2010) (Elgin, 2012), nedensellik testi (Ay, Sugözü, ve Erdoğan, 2014), modelleme tekniği (Kuehn, 2014), VAR tekniği (Kanca ve Akan , 2019) ve panel veri analizi (Kutbay, 2020) şeklindedir.

Bu araştırma, bir keşif analizidir. İki değişken arasındaki ilişkiyi bulmak amaçlanmamaktadır. Sarmal halde birbirini etkileyen vergi yükü ve kayıtdışı ekonomi oranının

OECD ülkelerinde görünümü ne şekildedir, hangi ülkeler birbirine benzemekte ve hangileri daha çok ayrılmaktadır sorularına cevap bulmak için gerçekleştirilmiştir. Bu amaçla çalışmanın ilk bölümünde OECD ülkelerinin kayıtdışı ekonomi ve vergi yükü oranları ile ilgili genel bilgiler sunulmuş, ardından araştırmanın yöntemi açıklanarak, analiz bulguları paylaşılmıştır.

## 1.OECD Ülkelerinde Vergi Yükü ve Kayıtdışı Ekonomi

Mükellefleri vergi kaçırmaya ve kayıtdışı kalmaya yönlendiren neden olarak denetim oranı, sürekli çıkan vergi afları, vergi mevzuatının karmaşıklığı gibi birçok sıralanacak neden arasında en önemlilerinden birisi vergi yüküdür. (Savaşan ve Çetintaş, 2008), (Akıncı ve Atalay, 2021).

Şekil 1. Kayıtdışı Ekonomi ve Vergi Yükü Etki Sarmalı



Mali sınırı aşan vergi yükü mükelleflerde vergi tazyiki veya baskısı olarak da ifade edilen subjektif vergi yükünü doğurmaktadır. Buna bağlı mükellefler ağır vergi yükü karşısında vergi kaçırma veya vergiden kaçınma yollarına gidebilmektedir. Yasal veya yasadışı olsun her iki tutumda kayıtdışı ekonominin artışına sebep olmaktadır. Kayıtdışı ekonominin artışı ülkelerin toplam vergi gelirini azaltır ve sürekli artış eğiliminde olan kamu harcamalarını karşılamak için ya yeni bir vergi konusu oluşturulur ya vergi oranları arttırılır yahut da borçlanma yoluna gidilir. İlk iki çözüm vergi yükünde artışa hemen etki ederken, borçlanma ile elde edilecek çözüm gelecekteki vergi yükünün artışına sebep olunur. Vergi yükünün artışı ise Şekil 1’de ki sarmalın oluşmasına sebep olur ve vergi yükü artışı da kayıtdışı ekonomiyi arttırır.

### 1.1.OECD Ülkelerinde Vergi Yükü

Vergi yükü objektif ve subjektif olmak üzere ikiye ayrılmaktadır. Ödenen vergi ile vergiyi ödeyenin gelir düzeyi arasında kurulan oransal ilişki objektif vergi yüküdür. Vergi nedeni ile mükellefin hissettiği psikolojik baskı subjektif vergi yüküdür. Analizde kullanılan

vergi yükü objektif vergi yüküdür; vergi gelirlerinin gayrisafi yurtiçi hasılaya bölünmesi ile elde edilen toplumsal vergi yükü oranları dikkate alınmaktadır (Şen ve Sağbaşı, 2017).

**Tablo 1. OECD Ülkelerinin Vergi Yükü Oranları (2009-2018)**

Ülkeler	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Avustralya	25,53	25,27	25,89	26,91	27,11	27,27	27,9	27,62	28,5	28,68
Avusturya	40,98	40,96	41,12	41,77	42,63	42,7	43,13	41,79	41,76	42,21
Belçika	42,7	42,88	43,51	44,33	45,04	44,76	44,13	43,32	43,8	43,95
Kanada	32,35	31,01	30,8	31,18	31,13	31,27	32,82	33,26	33,11	33,22
Şili	17,33	19,59	21,12	21,33	19,86	19,61	20,39	20,13	20,16	21,08
Kolombiya	18,82	18,1	18,93	19,72	20,02	19,55	19,9	19,08	18,98	19,3
Çekya	32,09	32,23	33,02	33,44	33,71	32,85	33,13	34,03	34,44	34,99
Danimarka	44,96	44,76	44,79	45,51	45,89	48,53	46,06	45,49	45,85	44,35
Estonya	34,76	32,93	31,21	31,46	31,46	31,93	33,08	33,24	32,52	32,89
Finlandiya	40,77	40,56	41,79	42,41	43,41	43,51	43,52	43,73	42,93	42,38
Fransa	41,53	42,15	43,33	44,36	45,37	45,45	45,28	45,37	46,07	45,93
Almanya	36,68	35,53	36,09	36,82	36,95	36,81	37,26	37,72	37,79	38,54
Yunanistan	30,75	32,04	33,64	35,81	35,74	36	36,39	38,36	38,62	38,92
Macaristan	38,9	37,19	36,64	39,16	38,59	38,55	38,9	39,41	38,28	37,45
İzlanda	31,18	32,3	33,34	34,01	34,52	37,34	35,4	50,81	37,56	37,23
İrlanda	27,97	27,66	28,38	28,63	28,91	29	23,37	23,61	22,76	22,67
İsrail	29,82	30,7	30,99	30,05	30,79	30,99	31,24	31,18	32,54	30,95
İtalya	41,97	41,7	41,64	43,62	43,83	43,33	42,96	42,24	41,91	41,87
Japonya	25,97	26,53	27,46	28,24	28,86	30,27	30,69	30,73	31,38	32,03
Kore	22,7	22,38	23,18	23,7	23,14	23,38	23,74	24,75	25,36	26,75
Letonya	28,24	28,73	28,34	29,11	29,36	29,72	30,02	31,16	31,39	31,24
Litvanya	30,17	28,36	27,24	26,98	26,75	27,5	28,7	29,65	29,47	30,17
Lüksemburg	38,29	37,51	37,06	38,35	38,16	37,46	36,98	36,83	37,61	39,74
Meksika	12,47	12,84	12,77	12,65	13,3	13,7	15,92	16,62	16,1	16,17
Hollanda	34,94	35,66	35,44	35,59	36,11	37,05	37,01	38,41	38,7	38,8
Yeni Zelanda	30,24	30,27	30,05	31,64	30,48	31,27	31,61	31,52	31,61	32,88
Norveç	41,11	41,8	41,92	41,41	39,82	38,75	38,42	38,88	38,78	39,57
Polonya	31,2	31,36	31,8	32,05	31,87	31,9	32,44	33,41	34,12	35,16
Portekiz	29,92	30,46	32,34	31,8	34,03	34,19	34,39	34,06	34,12	34,82
Slovakya	28,89	28,14	29,11	28,77	31,04	31,9	32,66	33,27	34,22	34,32
Slovenya	37,14	37,83	37,37	37,68	37,24	37,18	37,31	37,44	37,13	37,41
İspanya	29,68	31,28	31,19	32,37	33,12	33,89	33,84	33,6	33,87	34,6
İsveç	43,87	43,06	42,13	42,32	42,7	42,38	42,85	44,31	44,31	43,92
İsviçre	26,98	26,59	26,89	26,91	27	26,88	27,57	27,66	28,41	28,05
Türkiye	23,32	24,65	25,71	24,76	25,16	24,46	24,96	25,13	24,68	23,99
Birleşik Krallık	31,06	32,13	32,72	32,12	32,05	32,11	32,44	32,64	32,83	32,89
ABD	22,96	23,46	23,85	24	25,57	25,92	26,18	25,81	26,74	24,41

Kaynak: (OECD Data, 2021).

OECD ülkelerinin 2009-2018 yılları arası Tablo 1'deki vergi yükü oranları dikkate alındığında en düşük vergi yükü oranlarına Meksika, en yüksek vergi yükü oranlarına ise Danimarka ülkesinin sahip olduğu görülmektedir. En yüksek vergi yükü oranlarına sahip ülke olarak Danimarka'yı sırasıyla Fransa, Belçika, İtalya ve Finlandiya takip etmektedir. En düşük vergi yükü oranlarına sahip ülke olarak Meksika'yı sırasıyla Kolombiya, Şili, Kore ve Türkiye takip etmektedir. OECD ülkelerinin vergi yükü ortalaması 2008 yılında %32,57, 2012 yılında %32,73, 2014 yılında %33,23 ve 2018 yılında %33,88'dir.

## 1.2.OECD Ülkelerinde Kayıtdışı Ekonomi

'GSMH hesaplarını elde etmede kullanılan bilinen istatistik yöntemlerine göre tahmin edilemeyen, gelir yaratıcı ekonomik faaliyetlerin tümüdür' (Derdiyok, 1993). Tanzi'nin tanımıyla 'kayıtdışı ekonomi, bireylerin ve firmaların kendi ekonomik faaliyetlerini devletin müdahalesi olmaksızın yapma eğilimidir' (Tanzi, The Underground Economy in the United States: Annual Estimates 1930-1980, 1983).

Kayıtdışı ekonominin özellikleri; istatistiksel olarak ölçülemezlik, ahlaki normlara aykırılık, gelirin verilendirilmemiş olması şeklinde ifade edilebilir.

**Tablo 2. OECD Ülkelerinin Kayıtdışı Ekonomi Oranları (2008-2017)**

Ülkeler	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Avustralya	10.3	10.9	11.1	10.1	9.8	9.8	9.4	10.5	11.8	11.6
Avusturya	6.4	7.8	7.6	6.9	7.0	7.0	6.6	7.3	7.4	7.1
Belçika	15.5	17.4	16.9	16.0	16.7	16.6	15.9	17.2	16.9	16.5
Kanada	11.5	13.2	12.2	11.7	11.6	11.5	11.1	12.1	12.3	12.0
Şili	14.9	16.7	15.2	14.1	14.1	14.3	14.8	15.9	16.9	16.8
Kolombiya	28.9	31.4	29.9	25.7	25.2	25.8	26.0	28.6	30.0	29.9
Çekya	12.1	13.8	13.5	12.4	12.5	12.7	12.1	12.2	12.3	11.7
Danimarka	10.9	13.3	13.0	12.0	12.4	11.9	11.1	12.0	12.1	11.7
Estonya	21.1	24.1	22.7	20.1	20.0	19.6	19.3	21.0	20.9	20.1
Finlandiya	9.7	11.3	11.1	10.6	11.1	11.1	10.6	11.5	11.4	10.8
Fransa	10.1	12.1	11.8	11.1	11.7	11.6	11.4	12.2	12.2	11.7
Almanya	8.7	11.0	10.6	9.5	9.9	9.9	9.2	10.2	10.4	10.4
Yunanistan	19.7	22.8	23.1	23.0	24.3	23.7	23.6	25.3	25.4	24.8
Macaristan	18.9	21.2	20.7	19.7	20.4	19.9	19.8	20.8	20.5	19.8
İzlanda	12.0	13.9	13.5	12.9	12.9	12.7	12.1	12.2	11.6	10.8
İrlanda	11.8	12.6	12.3	12.0	12.0	11.7	11.0	9.5	9.7	9.6
İsrail	19.0	20.5	19.2	17.7	18.3	17.7	17.1	17.6	17.6	17.0
İtalya	18.1	21.2	20.8	18.9	20.0	20.0	19.7	20.9	20.6	19.8
Japonya	10.7	11.0	10.0	9.3	9.2	10.7	10.8	11.7	11.1	10.8
Kore	25.1	26.1	24.1	22.9	23.1	23.0	22.5	23.2	22.7	21.8

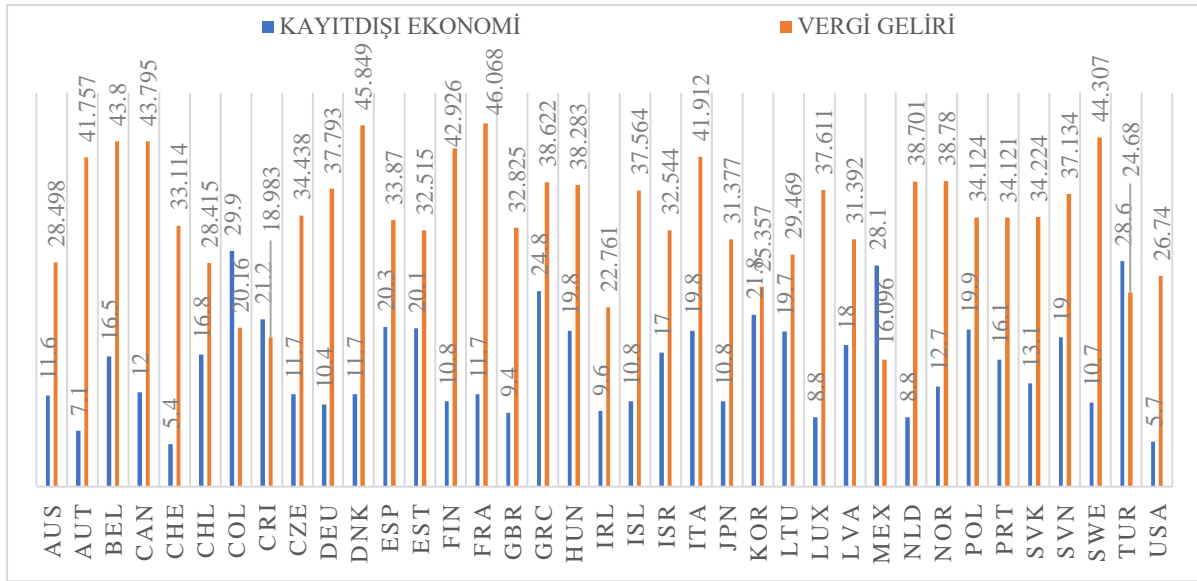
<b>Letonya</b>	19.5	21.3	20.8	19.3	18.7	18.4	17.9	19.1	18.8	18.0
<b>Litvanya</b>	22.5	25.2	24.1	22.0	21.5	20.5	20.0	21.3	21.0	19.7
<b>Lüksemburg</b>	7.9	9.2	8.8	8.4	9.0	8.7	7.9	8.5	8.7	8.8
<b>Meksika</b>	28.1	31.4	29.2	27.7	27.7	27.4	26.6	28.0	28.8	28.1
<b>Hollanda</b>	8.4	9.4	9.4	8.9	9.1	9.0	8.6	9.0	9.1	8.8
<b>Yeni Zelanda</b>	10.9	11.7	11.0	10.3	10.4	9.9	9.5	10.0	10.1	9.7
<b>Norveç</b>	9.4	12.4	11.7	10.9	10.8	10.9	11.1	12.5	13.1	12.7
<b>Polonya</b>	20.9	22.3	21.5	20.0	20.3	19.9	19.4	20.2	20.4	19.9
<b>Portekiz</b>	17.2	19.1	18.6	17.6	17.7	17.5	16.7	17.4	17.1	16.1
<b>Slovakya</b>	12.6	14.3	13.9	12.9	13.1	12.9	12.9	13.6	13.2	13.1
<b>Slovenya</b>	18.3	21.3	21.9	20.9	21.7	20.8	19.7	20.7	20.2	19.0
<b>İspanya</b>	18.6	21.3	21.3	20.6	21.3	21.1	20.9	21.9	21.3	20.3
<b>İsveç</b>	9.5	11.5	10.3	9.5	10.2	10.2	10.1	10.7	10.9	10.7
<b>İsviçre</b>	5.3	6.0	5.6	5.1	5.3	5.3	5.1	5.4	5.4	5.4
<b>Türkiye</b>	26.9	30.6	28.5	26.2	26.5	25.8	26.2	27.7	29.1	28.6
<b>Birleşik Krallık</b>	9.4	10.7	10.3	9.9	9.7	9.5	8.7	9.2	9.7	9.4
<b>ABD</b>	7.0	7.8	7.4	7.0	6.7	6.6	6.1	6.1	5.9	5.7

---

Kaynak: (Medina & Schneider, 2019).

Tablo 2’de ki MIMIC yöntemi kullanılarak hesaplanan kayıtdışı ekonomi verileri dikkate alındığında 2008-2017 yılları arasında OECD ülkeleri arasında en yüksek kayıtdışı ekonomi oranlarına sahip ülkeler aralarında çok az fark olmak üzere Meksika, Kolombiya ve Türkiye’dir. Bu ülkeleri Kore, Litvanya ve Estonya ülkeleri takip etmektedir. En düşük kayıtdışı ekonomi oranlarına sahip ülkeler ise İsviçre, ABD, Avusturya, Lüksemburg ve Hollanda’dır.

Şekil 2. OECD Ülkelerinde Kayıtdışı Ekonomi ve Vergi Yükü Oranları (2017)



<sup>3</sup>Kaynak: (OECD Data, 2021), (Medina ve Schneider, 2019).

Şekil 1’de 2017 yılı için OECD ülkelerinde kayıtdışı ekonomi ve vergi yükü verilerinin karşılaştırması yapılmıştır. Analizde kullanılacak veriler Şekil 1’de yer almaktadır.

### 3. Yöntem

Çalışmada kullanılacak değişkenler, 37 tane OECD ülkesinin 2017 yılına ait kayıtdışı ekonomi oranları ve vergi yükü oranlarıdır. Vergi yükü verisi için ülkelerin 2017 yılına ait vergi gelirlerinin gayrisafı yurtiçi hasılaya oranı dikkate alınmıştır. Vergi yükü verileri OECD Data’dan alınmıştır (OECD Data, 2021). Kayıtdışı ekonomi için Leandro Medina ve Friedrich Schneider’in MIMIC yöntemi ile ölçüldüğü veriler kullanılmıştır (Medina ve Schneider, 2019). Analiz için k-means kümeleme yöntemi uygulanmıştır. Önce veriler standartlaştırılmıştır. Daha sonra Ward Linkage metodu kullanılarak elde edilen dendrogram ile küme sayısına karar verilmiştir. Ardından belirlenen küme sayısı ile standartlaştırılmış veriler kullanılarak k-means kümeleme analizi gerçekleştirilmiştir.

Analiz için kullanılacak yöntem olan k-means kümeleme yöntemi; J.B. MacQueen tarafından 1967 yılında önerilen bir kümeleme algoritmasıdır (MacQUEEN, 1967). Bu algoritma n tane birimden oluşan veri setinin, k tane kümeye bölünmesine dayanmaktadır. En uygun çözüm bulunana kadar kümelerin yenilendiği döngüsel bir algoritmadır. Popüler kümeleme yöntemlerinden olan k-means algoritması öncelikle k sayıda nokta seçer. Sonraki aşamada seçilen her yeni nokta, en yakın merkezine taşınır. Üye noktaların ortalaması

<sup>3</sup>Şekil 2’de OECD ülke isimleri için 3 harfli ülke kodları kullanılmıştır. Şekil yazar tarafından oluşturulmuştur.

hesaplanarak merkez ve yakınsama kriterleri karşılanana kadar merkez değiştirme ve güncelleme işlemi devam eder (Jin ve Han, 2017).

K-means kümelemesinin başlangıcında, iki karşılaştırılabilir öge arasındaki fark nicel olarak nasıl hesaplanacağını bulunması gerekir. Burada farklılığın derecesi matematiksel olarak tanımlanır.

Diyelim ki;

$$(x_1, x_2, x_3, \dots, x_n), Y = (y_1, y_2, y_3, \dots, y_n) \text{ olsun.}$$

Burada  $X$  ve  $Y$ , her biri  $n$  ölçülebilir karakteristik niteliğe sahip iki unsurlu öğelerdir; daha sonra,  $X$  ve  $Y$  arasındaki farklılık derecesi şu şekilde tanımlanır:

$$d(X, Y) = f(X, Y) \rightarrow R \quad (1)$$

Bunların arasında  $R$ , gerçel sayılar kümesidir.

Başka bir deyişle, farklılık derecesi, iki ögenin bir gerçel sayıya eşlenmesidir ve bu gerçel sayı, iki ögenin farklı olma derecesini nicel olarak temsil eder. Farklılığın hesaplanmasında Öklid mesafesi, Manhattan mesafesi, Minkowski mesafesi vb. kullanılabilir. Genellikle kullanımı tercih edilen Öklid mesefasidir.

Öklid mesafesinin formülü şu şekildedir:

$$d(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (2)$$

K-means algoritması, verilerin doğal kategorisini bulmaya çalışır. Kullanıcı, iyi bir kategori merkezi bulmak için kategori sayısını belirler. Algoritma akışı aşağıdaki gibidir:

1-Veri kümelerinin ve  $k$  kategorilerinin sayısı girilir.

2- Kategorinin merkez noktası rastgele atanır.

3- Her nokta(veri), kendisine en yakın kategori merkez noktası kümesine konulur.

4- Veri, hangi kümenin ortalamasına en yakınsa o kümeye dâhil edilir.

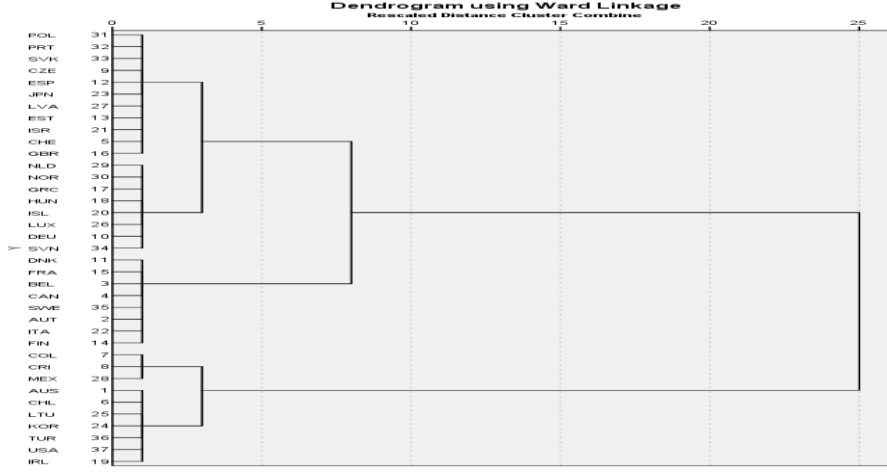
5-Yakınsamaya kadar 3. ve 4. adıma gidilir. (Yu, Zhu, Qin, Zhang, ve Huang, 2021), (Steinbach , Karypis, ve Kumar , 2000).



#### 4. Bulgular

Çalışmada iki kez k-means analizi uygulanmıştır. İlk olarak OECD ülkeleri için sadece vergi yükü verisi dikkate alınarak analiz gerçekleştirilmiştir. Daha sonra OECD ülkeleri için hem vergi yükü verileri hem de kayıtdışı ekonomi verileri dikkate alınarak k-means analizi uygulanmıştır. Sırası ile bulgular paylaşılmaktadır.

Şekil 3. Ward Linkage Metodu kullanılarak Elde Edilen Dendrogram



İlk olarak veriler standartlaştırılmıştır. Daha sonra Ward Linkage metodu kullanılarak elde edilen dendrogram ile (Şekil 3) küme sayısı 3 olarak belirlenmiştir.

Tablo 3. Son Küme Merkezleri Arasındaki Mesafe

Kümeler	1	2	3
1		1,322	2,463
2	1,322		1,142
3	2,463	1,142	

Yapılan analiz neticesinde üç kümenin birbirlerine olan mesafe bilgisi Tablo 1’de gösterilmektedir. Tablo 1 incelendiğinde en uzak mesafenin yani birbirinden en çok ayrılan ülkelerin 1. kümede bulunan ülkeler ile 3. kümede bulunan ülkeler olduğu ifade edilebilmektedir. 1. ve 2. kümede bulunan ülkelerin arasındaki ayrışma ile 2. ve 3. kümede bulunan ülkelerin arasındaki ayrışma nicelik olarak birbirlerine benzerdir.

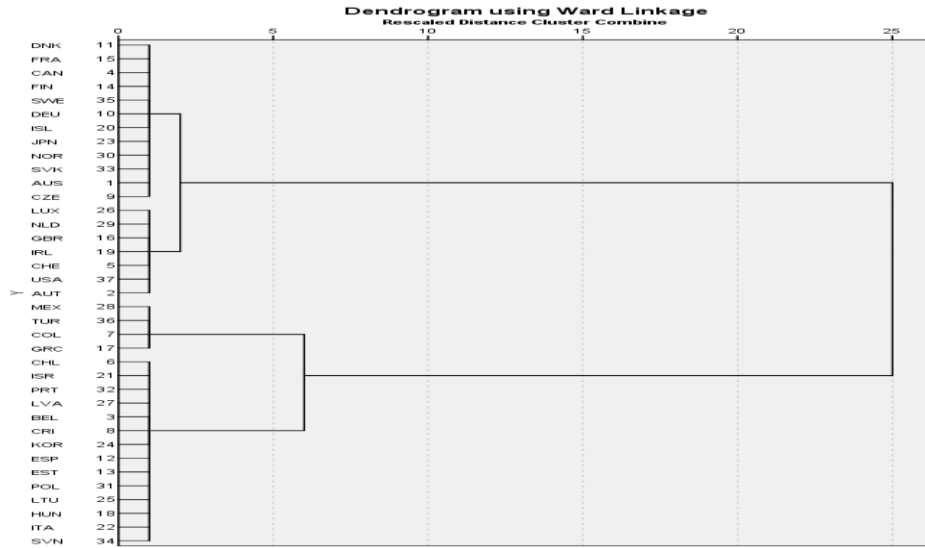
**Tablo 4. OECD Ülkelerine Ait Vergi Yükü Verilerinin K-Means Kümeleme Yöntemi ile Sınıflandırılması**

1. KÜME	2. KÜME	3. KÜME
1. Kolombiya	1. Avustralya	1. Avusturya
2. Kosta Rika	2. İsviçre	2. Belçika
3. İrlanda	3. Şili	3. Kanada
4. Güney Kore	4. Çekya	4. Almanya
5. Meksika	5. İspanya	5. Danimarka
6. Türkiye	6. Estonya	6. Finlandiya
7. Amerika Birleşik Devletleri	7. Birleşik Krallık	7. Fransa
	8. İsrail	8. Yunanistan
	9. Japonya	9. Macaristan
	10. Litvanya	10. İzlanda
	11. Letonya	11. İtalya
	12. Polonya	12. Lüksemburg
	13. Portekiz	13. Hollanda
	14. Slovakya	14. Norveç
		15. Slovenya
		16. İsveç

OECD ülkelerine ait vergi yükü verilerinin k-means kümeleme yöntemi ile sınıflandırılması neticesinde ülkelerin (Tablo 2) 1. kümede 7 ülke, 2. kümede 14 ülke ve 3. kümede 16 ülke olacak şekilde dağılımı gerçekleşmiştir.

Çalışmada gerçekleştirilen OECD ülkelerinin kayıtdışı ekonomi verilerinin de dahil edildiği ikinci k-means analizinin bulguları aşağıdaki gibidir.

**Şekil 4. Ward Linkage Metodu kullanılarak Elde Edilen Dendrogram**



İlk olarak veriler standartlaştırılmıştır. Daha sonra Ward Linkage metodu kullanılarak elde edilen dendrogram (Şekil 4) ile küme sayısı 4 olarak belirlenmiştir.

Tablo 5. Son Küme Merkezleri Arasındaki Mesafe

Kümeler	1	2	3	4
1		1,531	1,780	2,893
2	1,531		1,501	3,438
3	1,780	1,501		2,029
4	2,893	3,438	2,029	

Yapılan analiz neticesinde dört kümenin birbirlerine olan mesafe bilgisi Tablo 3'te gösterilmektedir. Tablo 3 incelendiğinde en uzak mesafenin yani birbirinden en çok ayrışan ülkelerin 4. kümede bulunan ülkeler ile 2. kümede bulunan ülkeler olduğu ifade edilebilmektedir. Bahsi geçen birbirinden en çok ayrışan ülkeler sıralamasını 4. kümedeki ülkeler ile 1. kümedeki ülkeler arasındaki ayrışma takip etmektedir. 1. ve 2. kümede bulunan ülkelerin arasındaki ayrışma ile 2. ve 3. kümede bulunan ülkelerin arasındaki ayrışma nicelik olarak birbirlerine benzerdir.

Tablo 6. OECD Ülkelerine Ait Kayıtdışı Ekonomi ve Vergi Yükü Verilerinin K-Means Kümeleme Yöntemi ile Sınıflandırılması

1.KÜME	2.KÜME	3.KÜME	4.KÜME
1. Avustralya	1. Avusturya	1. Şili	1. Kolombiya
2. İsviçre	2. Belçika	2. İspanya	2. Kosta Rika

---

3. Birleşik Krallık	3. Kanada	3. Estonya	3. Güney Kore
4. İrlanda	4. Çekya	4. Yunanistan	4. Meksika
5. Japonya	5. Almanya	5. Macaristan	5. Türkiye
6. Amerika Birleşik Devletleri	6. Danimarka	6. İsrail	
	7. Finlandiya	7. İtalya	
	8. Fransa	8. Litvanya	
	9. İzlanda	9. Letonya	
	10. Hollanda	10. Polonya	
	11. Lüksemburg	11. Portekiz	
	12. Norveç	12. Slovenya	
	13. Slovakya		
	14. İsveç		

---

OECD ülkelerine ait kayıtdışı ekonomi ve vergi yükü verilerinin k-means kümeleme yöntemi ile sınıflandırılması neticesinde ülkelerin 1.kümede 6 ülke, 2. kümede 14 ülke, 3. kümede 12 ülke ve 4. kümede 5 ülke olacak şekilde dağılımı gerçekleşmiştir (Tablo 4).

### SONUÇ ve DEĞERLENDİRME

Gerçekleştirilen iki analiz sonucu birlikte değerlendirildiğinde; ilk analizde 2. kümede bulunan Şili, İspanya, Estonya, İsrail, Litvanya, Letonya, Polonya, Portekiz, Slovenya ikinci analizde üçüncü kümede birlikte bulunmaktadır. Bu ülkeler hem vergi yükleri hem de kayıtdışı ekonomi çıktılarında birbirlerine benzeşen ülkeler olarak değerlendirilmektedir.

İlk analizde 3. kümeyi oluşturan Avusturya, Belçika, Kanada, Almanya, Danimarka, Finlandiya, Fransa, İzlanda, Lüksemburg, Hollanda, Norveç, Slovenya, İsveç ikinci analizde 2. kümede bulunmaktadır. Bu ülkelerde hem vergi yükü hem de kayıtdışı ekonomi verileri dikkate alındığında birbirlerine benzeşen ülkelerdir.

İlk analizde 3. kümede bulunan Yunanistan, Macaristan ve İtalya ilk analizde birlikte bulunduğu ülkelere ayrılarak ikinci analizde 3. kümeye geçmiştir. Yine ilk analizde 2. kümede bulunan Çekya kayıtdışı ekonomi verileri dâhil edildiğinde birlikte olduğu ülkelere ayrılarak 2. kümeye geçmiştir.

İlk analizde 2. kümede birlikte bulunan Avustralya, İsviçre, Birleşik Krallık ve Japonya ikinci analizde birlikteliklerini bozmayarak ve fakat kümedeki diğer ülkelere ayrılarak 1. kümeye geçmiştir. Geçtikleri birinci kümeyi oluşturan diğer ülkeler İrlanda ve Amerika Birleşik Devletleri'dir. Bu iki ülke birinci analizde Kolombiya, Kosta Rika, Güney

Kore, Meksika ve Türkiye'nin bulunduğu 1.kümeden ayrılarak ikinci analizde birinci kümeye geçmişlerdir. Birinci analizde 1. kümede bulunan Türkiye, Kolombiya, Güney Kore, Kosta Rika ve Meksika'nın birliktelikleri bozulmamış ve ikinci analizde 4. kümeyi oluşturmuşlardır.

Gelişmiş ülkeler genellikle, Laffer eğrisini dikkate aldığımızda, negatif eğimli bölgede bulunmaktadır ve vergi oranlarındaki indirimler vergi gelirlerine artış olarak yansiyabilmektedir. Pozitif eğimli bölgede ise vergi oranlarındaki düşüşler vergi gelirin'e negatif yansımaktadır. Az gelişmiş ülkelerde ise vergi oranları maksimum vergi geliri noktasının altında olduğundan ve kayıtdışı gelir elde etme potansiyeli yüksek olduğundan bu ülkelerde vergi oranlarındaki indirimlerde genellikle beklenen etki görülmemektedir. Az gelişmiş ülkelerde vergi oranında yapılan indirimin vergi gelirlerini artırma amacını gerçekleştirebilmesi için istisna ve muafiyetler mümkün olduğunca azaltılmalı, vergi tabanı genişletilmeli ve kayıtdışı ekonomi oranı minimuma çekilmelidir (Edizdoğan ve Çelikkaya, 2012).

İki verinin de dikkate alınarak gerçekleştirildiği son analizde 4. kümede bulunan ülkeler net bir şekilde diğer ülkelerden ayrılmaktadır. Bu ülkelerde vergi yükü oranı düşük olmakla birlikte kayıtdışı ekonomi oranları oldukça yüksektir. Vergi yüklerinde gerçekleştirecekleri bir birim azalış muhtemel vergi gelirlerinde artış gerçekleştirmeyecek bunun yanı sıra kayıtdışı ekonomi oranının artmasına neden olacaktır. Bunun sebebi ise mükelleflerin subjektif vergi yükü fazlalığıdır.

ABD ve İrlanda ülkelerinin vergi yükü oran azlığı ile son analizde 4. kümeyi oluşturan ülkeler ile benzeşse bile kayıtdışı oranı verisi ile onlardan ayrılmaktadır. Bu ülkelerin vergi uygulamaları ve gelişmişlik seviyeleri dikkate alınarak değerlendirildiğinde vergi yükünde gerçekleşen azalışın kamu gelirini arttırdığını ifade edebiliriz.

Uluslararası ekonomiye entegre olmak isteyen geliştirmekte olan ülkeler için etkin ve adil bir vergi sistemi oluşturmak hem mühim hem de basit olmaktan uzaktır. Bu ülkeler için ideal vergi sistemi; yüksek devlet borçlanmasından uzak, ekonomik faaliyetleri caydırmadan temel gelirleri arttırarak ve diğer ülkelerdeki vergi sistemlerinin yapısından çok fazla ayrılmadan gerçekleştirmelidir (Tanzi ve H.Zee, 2000).

## KAYNAKÇA

- AKINCI, A. ve ATALAY, M. (2021). “Türkiye’de Vergi Yüğü ile Kiři Bařına Düřen Gayri Safi Yurtiçi Hasıla Arasındaki İliři: Fourier Yaklařımı”, *Sayıřtay Dergisi*, 32(121), s.127-152.
- AY, A., SUGÖZÜ İ. ve ERDOĞAN, S. (2014). “Türkiye’de Vergi Yüğüünün, Enflasyonun ve Vergi Affı Beklentisinin Kayıt Dıřı Ekonomiye Etkisi Üzerine Ampirik Bir Uygulama (1985-2012)”, *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 31(1), s.23-32.
- DERDİYOK, T. (1993). “Türkiye’nin Kayıt Dıřı Ekonomisinin Tahmini”, *Türkiye İktisat Dergisi*, 14, s.45-63.
- EDİZDOĞAN, N. ve ÇELİKKAYA, A. (2012). *Vergilerin Ekonomik Analizi*. Bursa: Dora Yayınları.
- ELGİN, C. (2012). “Vergiler ve Kayıtdıřı Ekonomi: Bir Deęerlendirme ve Türkiye Örneęi”, *Middle East Technical University Studies in Development*, 39, s.237-258.
- JIN, X. and HAN, J. (2017). *Encyclopedia of Machine Learning* (Editörler) C. Sammut, & G. I. Webb. Expectation Maximization Clustering, New York: Springer Science+Business Media, p.563-564.
- KANCA, O. ve AKAN, Y. (2019). “Kayıtdıřı Ekonominin Dinamikleri: Ekonometrik Bir İnceleme”, *Vergi Raporu*, 232, s.53-73.
- KUEHN, Z. (2014). “Tax Rates, Governance, and the Informal Economy in High-Income Countries”. *Economic Inquiry*, 52(1), p.405-430.
- KUTBAY, H. (2020). “Vergi Yüğüünün Kayıtdıřı Ekonomiye Etkisi: Yükselen Piyasa Ekonomileri İçin Panel Veri Analizi”. *Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi*, 23(1), s.226-239.
- MacQUEEN, J. (1967). “Some Methods for Classification and Analysis of Multivariate Observations”, *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1, p.281-297.
- MEDİNA, L. ve SCHNEİDER, F. (2019). “Shedding Light on the Shadow Economy: A Global Database and the Interaction with the Official One”, *cesifo*, 7981(1), p.44-49.
- OECD DATA, <https://data.oecd.org>, Eriřim Tarihi: 10.11.2021.

- SAVAŞAN, F. ve ÇETİNTAŞ, H. (2008). “Vergi Yükü ve Vergi Kaçırma: Türkiye'de Uzun Dönem Karakteristikleri Üzerine Ampirik Bir Değerlendirme”, *Yönetim ve Ekonomi Bilimleri Konferansı, Kamu Ekonomisi ve Kamu Maliyesi*, 3, s.213-228.
- SCHNEIDER, F. (2010). “The Influence of Public Institutions on The Shadow Economy: An Empirical Investigation for OECD Countries”, *Review of Law & Economics*, 6(3), p.441-468.
- STEINBACH , M., KARYPIS, G. ve KUMAR , V. (2000). *A Comparison of Document Clustering Techniques*. Minnesota: Department of Computer Science and Engineering, University of Minnesota .
- ŞEN, H. ve SAĞBAŞ, İ. (2017). *Vergi Teorisi ve Politikası*. Ankara: Barış Arıkan Yayınları.
- ŞENYÜZ, D., YÜCE , M. ve GERÇEK, A. (2016). *Vergi Hukuku*. Bursa: Ekin Basım Yayın Dağıtım.
- TANZI, V. (1983). “The Underground Economy in the United States: Annual Estimates 1930-1980”, *Staff Papers (International Monetary Fund)*, 30, p.283-305.
- TANZI, V. ve ZEE, H. (2000). “Tax Policy for Emerging Markets: Developing Countries”, *National Tax Journal*, 53(2), p.5-31.
- YU, J., ZHU, L., QIN, R., ZHANG, Z. ve HUANG, T. (2021). “Combining K-Means Clustering and Random Forest to Evaluate the Gas Content of Coalbed Bed Methane Reservoirs”, *Hindawi Geofluids*, 2021, p.1-8.