

T.C.
DOKUZ EYLÜL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
EKONOMETRİ PROGRAMI
YÜKSEK LİSANS TEZİ

OECD ÜLKELERİNİN EKONOMİK ÖZGÜRLÜK
ENDEKSİ VE DÜNYA MUTLULUK RAPORU
AÇISINDAN KÜMELEME ANALİZİ

Hüseyin KAYA

Danışman
Prof. Dr. İpek DEVECİ KOCAKOÇ

İZMİR – 2022

TEZ ONAY SAYFASI



YEMİN METNİ

Yüksek Lisans Tezi olarak sunduğum “OECD Ülkelerinin Ekonomik Özgürlük Endeksi Ve Dünya Mutluluk Raporu Açısından Kümeleme Analizi ” adlı çalışmanın, tarafımdan, akademik kurallara ve etik değerlere uygun olarak yazıldığını ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.

26/08/2022

Hüseyin KAYA



ÖZET
Yüksek Lisans Tezi
OECD Ülkelerinin Ekonomik Özgürlük Endeksi Ve Dünya Mutluluk Raporu
Açısından Kümeleme Analizi
Hüseyin Kaya

Dokuz Eylül Üniversitesi
Sosyal Bilimler Enstitüsü
Ekonometri Anabilim Dalı
Ekonometri Programı

Toplumların ve ülkelerin refah düzeylerini arttırması, daha iyi hayat koşullarında hayatını sürdürmesi ülkelerin ekonomisiyle yakından ilişkilidir. Ekonomik özgürlük, devlet müdahalesinden olabildiğince uzak, özel sektörün olabildiğince serbestliğe sahip olduğu bir durumda en yüksek seviyelerde olacaktır. Ekonomik özgürlüğü yüksek toplumların daha iyi bir yaşam sürdüğü düşünülmektedir. Peki ekonomik özgürlüğü yüksek toplumlarda mutluluğun rolü nedir? Bu çalışmada “Heritage Foundation” kuruluşunun yayınladığı “Ekonomik Özgürlük Endeksi” ve “Dünya Mutluluk Raporu” veri setleri birleştirilip, ekonomik özgürlük ve mutluluk arasındaki ilişkiden yararlanarak bir kümeleme analizi gerçekleştirilmiştir. Araştırma 2017’den 2021’e 5 yılı kapsayacak şekilde Türkiye’nin de dahil olduğu 31 OECD ülkesi için gerçekleştirilmiştir. Araştırmada Ekonomik Özgürlük Endeksi’nden 4 ana faktör, Dünya Mutluluk Raporu’ndan 7 alt faktör toplamda 11 faktör kullanılmıştır. Analizde ilk önce küme sayısını belirlemek için Ward yöntemi ile bir dendrogram oluşturulmuş aynı zamanda dirsek yönteminden faydalanılmıştır ve optimal küme sayısı 4 olarak belirlenmiştir. Küme sayısı belirlendikten sonra kümeleme analizi k-ortalamalar yöntemi ile gerçekleştirilmiştir. Kümelerin oluşturulmasından sonra hangi faktörlerin kümelerin oluşmasında önemli rol oynadığını belirlemek amacıyla bir karar ağacı oluşturulmuştur. Bu bilgilerin ışığında ülkelerin bulunduğu kümeler yorumlanmıştır. Analizler sonucunda Meksika ve Slovakya 5 yıl, Macaristan ve

Polonya ise 4 yıl Türkiye ile aynı kümede yer almıştır. Kümelerin oluşmasında en etkili faktörler ise hukukun üstünlüğü, devlet büyüklüğü, doğumda sağlıklı yaşam beklentisi, sosyal destek, serbest piyasa, yolsuzluk algısı ve mevzuat verimliliği olarak gözlemlenmiştir.

Anahtar Kelimeler: OECD, Ward Yöntemi, k-ortalamlar Yöntemi, Karar Ağacı, Ekonomik Özgürlük Endeksi, Dünya Mutluluk Raporu.



ABSTRACT
Master's Thesis
Cluster Analysis of OECD Countries in terms of Economic Freedom Index and
World Happiness Report
Hüseyin KAYA

Dokuz Eylul University
Graduate School Of Social Sciences
Department of Econometrics
Econometrics Program

Increasing the welfare of societies and countries and maintaining their lives in better living conditions are closely related to the economy of the countries. Economic freedom will be at its highest in a situation where the private sector has as much freedom as possible, free from government interference. It is thought that societies with higher economic freedom maintain a better life. So, what is the role of happiness in societies with high economic freedom? In this study, a cluster analysis was carried out by combining the "Economic Freedom Index" and "World Happiness Report" datasets published by the "Heritage Foundation" and benefiting from the relationship between economic freedom and happiness. The research was carried out for 31 OECD countries, including Turkey, covering 5 years from 2017 to 2021. In the research, 4 main factors from the Economic Freedom Index and 7 sub-factors from the World Happiness Report, a total of 11 factors were used. In the analysis, firstly, a dendrogram was created with Ward's method to determine the number of clusters, and at the same time, the elbow method was used and the optimal number of clusters was determined as 4. After the number of clusters was determined, the cluster analysis was performed with the k-means method. A decision tree was created for this purpose. In the light of this information, the clusters formed by the countries were interpreted. As a result of the cluster analysis, Mexico and Slovakia have been in the same cluster as Turkey for 5 years, besides Hungary and Poland for 4 years. The most influential factors in the formation of clusters were observed to be the rule of law,

government size, healthy life expectancy at birth, social support, market openness, perception of corruption and regulatory efficiency.

Keywords: OECD, Ward Method, k-means method, Decision Tree, Economic Freedom Index, World Happiness Report.



**OECD ÜLKELERİNİN EKONOMİK ÖZGÜRLÜK ENDEKSİ VE DÜNYA
MUTLULUK RAPORU AÇISINDAN KÜMELEME ANALİZİ**

İÇİNDEKİLER

TEZ ONAY SAYFASI	ii
YEMİN METNİ	iii
ÖZET	iv
ABSTRACT	vi
İÇİNDEKİLER	vii
KISALTMALAR	xi
TABLolar LİSTESİ	xii
ŞEKİLLER LİSTESİ	xiii
EKLER LİSTESİ	xiv
 GİRİŞ	 1

**BİRİNCİ BÖLÜM
MAKİNE ÖĞRENMESİ**

1.1. DENETİMLİ ÖĞRENME	4
1.1.1. Regresyon	4
1.1.2. Sınıflandırma	5
1.2. DENETİMSİZ ÖĞRENME	10
1.3. YARI DENETİMLİ ÖĞRENME	10
1.4. PEKİŞTİRMELİ ÖĞRENME	10

İKİNCİ BÖLÜM

KÜMELEME ANALİZİ

2.1. UZAKLIK ÖLÇÜLERİ	14
2.1.1. Öklid Uzaklığı	14
2.1.2. Karesi Alınmış Öklid Uzaklığı	15
2.1.3. Manhattan (City-Block) Uzaklığı	15
2.1.4. Mahalanobis Uzaklığı	16
2.1.5. Minkowski Uzaklığı	16
2.1.6. Canberra Uzaklığı	17
2.1.7. Chebychev Uzaklığı	17
2.1.8. Pearson Korelasyon Ölçüsü	17
2.1.9. Hotelling T^2 Uzaklığı	18
2.2. HİYERARŞİK KÜMELEME ANALİZİ YÖNTEMLERİ	18
2.2.1. Tek Bağlantı Yöntemi	18
2.2.2. Tam Bağlantı Yöntemi	19
2.2.3. Ortalama Bağlantı Yöntemi	19
2.2.4. Centroid (Ağırlık Merkezi) Yöntemi	20
2.2.5. Ward Yöntemi	20
2.3. HİYERARŞİK OLMAYAN KÜMELEME ANALİZİ YÖNTEMLERİ	22
2.3.1. k-Ortalamalar Yöntemi	22
2.3.2. k-Medoid Yöntemi	23
2.4. KÜME SAYISININ BELİRLENMESİ	24

ÜÇÜNCÜ BÖLÜM

UYGULAMA

3.1. VERİ SETLERİ METODOLOJİSİ	27
3.1.1. Ekonomik Özgürlük Endeksi Metodoloji	27
3.1.2. Dünya Mutluluk Raporu Metodoloji	36
3.2. VERİ SETLERİ LİTERATÜR TARAMASI	40
3.2.1. Ekonomik Özgürlük Endeksi Literatür Taraması	40

3.2.2. Dünya Mutluluk Raporu Literatür Taraması	45
3.3. KÜME SAYISININ BELİRLENMESİ VE KÜMELEME ANALİZİ	
SONUÇLARI	49
3.3.1. 2017 Yılı İçin Analizler	51
3.3.2. 2018 Yılı İçin Analizler	56
3.3.3. 2019 Yılı İçin Analizler	61
3.3.4. 2020 Yılı İçin Analizler	66
3.3.5. 2021 Yılı İçin Analizler	71
3.4 UYGULAMA SONUÇ VE TARTIŞMASI	78
SONUÇ	79
KAYNAKÇA	80
EKLER	

KISALTMALAR

AID	Automatic Interaction Detector
BİST	Borsa İstanbul
CART	Classification and Regression Trees
CEMAUT	Critic Ve Entropy Tabanlı MAUT
CETOPSIS	Critic Ve Entropy Tabanlı TOPSIS
CHAID	Chi-Squared Automatic Interaction Detection
DMR	Dünya Mutluluk Raporu
E7	Emerging 7
HDI	İnsani Gelişme Endeksi
HPE	Henley Pasaport Endeksi
KBE	Küresel Barış Endeksi
MAUT	Multi Attribute Decision Making
OECD	Ekonomik İşbirliği ve Kalkınma Örgütü
SSE	Sum Squared Error
TOPSIS	Technique For Order Preference By Similarity To An Ideal Solution
WHR	Dünya Mutluluk Raporu
WVS	Dünya Değerler Araştırması
YAE	Yolsuzluk Algı Endeksi

TABLÖLAR LİSTESİ

Tablo 1: Ekonomik Özgürlük Endeksi Literatür Taraması Özeti	s. 41
Tablo 2: Dünya Mutluluk Raporu Literatür Taraması Özeti	s. 46
Tablo 3: Araştırmaya Dahil Edilen OECD Ülkeleri	s. 50
Tablo 4: Araştırmada Kullanılan Değişkenler ve Karşılıkları	s. 51
Tablo 5: 2017 Yılı İçin İlk Küme Merkezleri	s. 53
Tablo 6: 2017 Yılı İçin Nihai Küme Merkezleri	s. 54
Tablo 7: 2018 Yılı İçin İlk Küme Merkezleri	s. 58
Tablo 8: 2018 Yılı İçin Nihai Küme Merkezleri	s. 59
Tablo 9: 2019 Yılı İçin İlk Küme Merkezleri	s. 64
Tablo 10: 2019 Yılı İçin Nihai Küme Merkezleri	s. 64
Tablo 11: 2020 Yılı İçin İlk Küme Merkezleri	s. 68
Tablo 12: 2020 Yılı İçin Nihai Küme Merkezleri	s. 69
Tablo 13: 2021 Yılı İçin İlk Küme Merkezleri	s. 73
Tablo 14: 2021 Yılı İçin Nihai Küme Merkezleri	s. 74
Tablo 15: Türkiye İle Aynı Kümede Yer Alan Ülkeler	s. 76
Tablo 16: Kümeleme Analizine En Fazla Etki Eden Faktörler	s. 77

ŞEKİLLER LİSTESİ

Şekil 1: Doğrusal Sınıflandırma	s. 6
Şekil 2: Doğrusal Olmayan Sınıflandırma	s. 6
Şekil 3: Karar Ağacı Örneği	s. 8
Şekil 4: Kümeleme Analizi Yöntemleri	s. 13
Şekil 5: Dirsek Yönetimi	s. 25
Şekil 6: Dendrogram	s. 26
Şekil 7: Ekonomik Özgürlük Endeksi Değişkenleri	s. 36
Şekil 8: 2017 Yılı İçin Dendrogram	s. 52
Şekil 9: 2017 Yılı İçin Dirsek Grafiği	s. 53
Şekil 10: 2017 Yılı İçin Kümeler	s. 55
Şekil 11: 2017 Yılı İçin Karar Ağacı	s. 56
Şekil 12: 2018 Yılı İçin Dendrogram	s. 57
Şekil 13: 2018 Yılı İçin Dirsek Grafiği	s. 58
Şekil 14: 2018 Yılı İçin Kümeler	s. 60
Şekil 15: 2018 Yılı İçin Karar Ağacı	s. 61
Şekil 16: 2019 Yılı İçin Dendrogram	s. 62
Şekil 17: 2019 Yılı İçin Dirsek Grafiği	s. 63
Şekil 18: 2019 Yılı İçin Kümeler	s. 65
Şekil 19: 2019 Yılı İçin Karar Ağacı	s. 66
Şekil 20: 2020 Yılı İçin Dendrogram	s. 67
Şekil 21: 2020 Yılı İçin Dirsek Grafiği	s. 68
Şekil 22: 2020 Yılı İçin Kümeler	s. 70
Şekil 23: 2020 Yılı İçin Karar Ağacı	s. 71
Şekil 24: 2021 Yılı İçin Dendrogram	s. 72
Şekil 25: 2021 Yılı İçin Dirsek Grafiği	s. 73
Şekil 26: 2021 Yılı İçin Kümeler	s. 75
Şekil 27: 2021 Yılı İçin Karar Ağacı	s. 76

EKLER LİSTESİ

EK 1: Dirsek Metodu İçin Python Kodu	ek s. 1
EK 2: Kümeleme Analizi İçin Python Kodu	ek s. 2
EK 3: Karar Ağacı Oluşturulması İçin Python Kodları	ek s. 3
EK 4: Veri setleri İçin Github Linki	ek s. 5
EK 5: Değişkenler İçin Dağılım Matrisi	ek s. 6



GİRİŞ

Özgürlük, çok kısa bir şekilde ifade etmek gerekirse; “bireyin, başkasının keyfi iradesinin zorlamasına maruz bırakılmaması durumudur” (Taşar, 2007). Özgürlük, insanlığın varoluşundan bu zamana kadar kilit bir öneme sahip olan ancak anlamı ve sınırları tam olarak belirlenememiş bir kavramdır. Bu bağlamda Abraham Lincoln’ ün kullandığı bir ifadeye yer vermek oldukça ilişkili olacaktır: “Dünya hiçbir zaman özgürlük kelimesinin iyi bir tanımına sahip olamadı ve Amerikan halkı tam da şimdi böyle bir tanıma ihtiyaç duymaktadır. Biz hepimiz özgürlükten söz ederiz fakat aynı şeyi kastetmeyiz... Burada aynı kelimeyle adlandırılan fakat sadece farklı olmayıp aynı zamanda birbirleriyle bağdaştırılamayacak olan şeyler söz konusudur.” Bu tanımlama günümüz için de hala geçerliliğini korumaktadır. Bireylerin kişisel özgürlüğü kadar içinde bulundukları toplumun ve içinde bulundukları ülkenin ekonomik özgürlüğü de önem arz etmektedir.

Ülkelerin ekonomik açıdan özgür olması birçok unsura bağlıdır. Ülkedeki hukuk sistemi, vergi sistemi, ticaret özgürlükleri, iş ve emek özgürlüğü gibi birçok kavram ekonomik özgürlüğün sağlanmasına olanak sağlar. Ekonomik özgürlük genel olarak, kişinin refahını yükselten, ekonomik kalkınma dinamiklerini uyaran, ticaret sistemini hareketlendiren bir kavram olarak ifade edilebilir. Ekonomik özgürlük ve refah, Adam Smith’in 1776 yılında “Ulusların Zenginliği” isimli kitabında bahsettiği üzere, devlet müdahalesinden olabildiğince uzak, özel sektörün olabildiğince serbestliğe sahip olduğu bir durumda en yüksek seviyelerde olacaktır. Ülkelerin ekonomik özgürlükleri ile ilgili yorumlama ve karşılaştırma yapılabilmesi için üç farklı tarafından ekonomik özgürlük endeksi hesaplanmaktadır. Fraser Enstitüsü tarafından yayınlanan “Dünya Özgürlük Endeksi”, Freedom House tarafından yayınlanan “Özgürlük Endeksi” ve aynı zamanda bu araştırmada büyük yer sahibi olan 1995 yılında Adam Smith’in teorilerinden ilham alınarak The Wall Street Journal ve Heritage Foundation tarafından oluşturulan “Ekonomik Özgürlük Endeksi”. Ekonomik özgürlük endeksinde ülkeler, belli başlı unsurlara göre karşılaştırılıp, nihai bir skor oluşturulup bu skora göre sıralanmaktadır. Skor oluşturulurken 4 ana değişkenin altında birçok alt değişkenden yararlanılarak geniş bir hesaplama yapılmaktadır. Bütün değişkenlerin farklı hesaplama yöntemleri olmakla birlikte 0 ile

100 arasında en çok puan alan ülke ekonomik özgürlüğü en yüksek ülke olarak ifade edilmektedir.

Mutluluk, özgürlükte olduğu gibi tanımı ülkeden ülkeye, toplumdan topluma değişen belli bir sınırı ve tam olarak net bir tanımı olmayan bir kavram olarak karşımıza çıkmaktadır. Mutluluk terimi, genel olarak bireyin yaşamının iyi, anlamlı ve değerli olması durumu olarak açıklanabilir. Bu açıklama kişiden kişiye göre değişebildiği gibi birçok farklı faktöre göre de değişip anlam kazanabilmektedir. Mutluluk düzeylerini karşılaştırmak için yapılan ilk araştırma 1965 yılında Cantril tarafından gerçekleştirilmiş olup 14 ülkenin temsili örneklemelerini kapsamaktadır. O zamandan beri mutluluk Dünya Değerler Araştırması(WVS), the Euro-barometer, European Welfare Survey ve Dünya Mutluluk Raporu(WHR) gibi birçok uluslararası araştırma programlarında benimsenmiştir.

Dünya Mutluluk Raporu Cantril'in çalışmasından yola çıkarak dünyanın dört bir yanındaki ülkeleri, vatandaşların kendilerini ne kadar mutlu algıladıklarına göre sıralayan küresel mutluluğun durumuna ilişkin önemli bir ankettir. İlk rapor, 2012 yılında Birleşmiş Milletler toplantısına destek olarak halkın milli geliri kadar mutluluğunun da önemli olduğunu ortaya koyma motivasyonu ile ele alınmıştır. İlk rapordan beri, hükümetler, sivil toplum örgütleri vb. kuruluşlar tarafından küresel olarak tanınmaya ve karar verme mekanizmalarına yardımcı olması açısından her yıl yayınlamaya devam etmektedir. Dünya mutluluk raporu ülkeleri mutluluklarına göre sıralamak amacıyla "Mutluluk Endeksi" olarak adlandırılan bir skordan faydalanır. Mutluluk endeksi 7 alt faktörün toplanması ile elde edilir ve alt faktörler sırasıyla şunlardır; Kişi Başına Gayrisafi Yurtiçi Hasıla, Doğumda Sağlıklı Yaşam Beklentisi, Sosyal Destek, Yaşam Seçimleri Yapma Özgürlüğü, Cömertlik, Yolsuzluk Algısı, Distopya.

Mutluluk subjektif bakış açısının kuvvetli olduğu, aynı zamanda zaman, mekan ve gelir gibi faktörlerden de çokça etkilenen bir kavramdır. Bu sebeple sosyoloji, psikoloji ve iktisat gibi birçok alanda araştırmalara dahil edilen bir konudur. Bir ülkede refahın ve mutluluğun yükselmesi ekonomik gelişmenin önemli bir göstergesi olarak algılanmaktadır. Aynı şekilde gelir dağılımındaki eşitsizlikler ve yoksulluk mutluluğun düşüren faktörler olarak yorumlanmaktadır. Bu çerçevede sürdürülebilir kalkınmanın sağlanması ekonomik özgürlük ve mutluluk ile yakından ilişkilidir.

Bu alıřmada lkeler ekonomik zgrlk ve mutluluk aısından deęerlendirilmiřtir. Deęerlendirmede Heritage Foundation tarafından oluřturulan “Ekonomik zgrlk Endeksi” ve “Dnya Mutluluk Raporu” veri setleri birleřtirilmiřtir. Trkiye’nin de iinde bulunduęu 31 OECD lkesini kapsayan, 2017-2021 yılları iin toplamda 5 adet veri seti oluřturulmuř ve kmeleme analizi iin hazır hale getirilmiřtir. Kmeleme analizi iin oluřturulan veri seti iin ilk nce kme sayısı ward metodu ve dirsek ynteminden yararlanılarak belirlenmiř, daha sonra k-ortalamlar yntemi ile kmeleme analizi gerekleřtirilmiřtir. Kme sonuları her yıl iin ayrı ayrı elde edilip yorumlanmıřtır. Kme sonularını desteklemek ve hangi deęiřkenin kmelerin belirlenmesinde daha ok rol oynadıęını gzlemlemek amacıyla her yıl iin karar aęacı oluřturulmuřtur. Oluřan kmeler ve karar aęaları bu baęlamda yorumlanmıřtır.

BİRİNCİ BÖLÜM

MAKİNE ÖĞRENMESİ

Makine öğrenmesi, geçtiğimiz yıllarda bilgi teknolojilerinin temeli haline geldi. Farkında olmasak bile makine öğrenmesi hayatımızın büyük bir bölümünde kullanılıyor. Veri, günümüzün en etkili teknolojik gelişmelerinden biridir. Verinin gerekli bir bileşen olarak daha yaygın hale geleceği düşünülmektedir. Verilerden elde edilen makine öğrenmesi, anlamlı örüntülerin otomatik olarak elde edilmesi olarak ifade edilebilir. Makine öğrenmesini örneklem veya geçmiş deneyimlerinden yararlanarak herhangi bir problemi çözmeye yardımcı olan bir bilgisayar programı olarak düşünebiliriz. Makine öğrenmesinin önemi bir insan yardımı veya müdahalesi olmadan otomatik olarak öğrenebilme yeteneğidir. Bu durumda makineye sunulan veri girdi olarak adlandırılırken örnekleme veya geçmiş deneyimi temsil etmektedir. Makinenin nihai olarak elde ettiği sonuç ise bilgiye karşılık gelmektedir. Makine öğrenmesi; Denetimli Öğrenme, Denetimsiz Öğrenme, Yarı Denetimli Öğrenme ve Pekiştirmeli Öğrenme olarak 4 başlıkta incelenebilir.

1.1. DENETİMLİ ÖĞRENME

Veri setinde bulunan girdileri ve bu girdilerin sahip olduğu etiketleri kullanarak bir model üreten bir öğrenme biçimidir. Denetimli öğrenme ismini almasının sebebi, bir “öğretmen” gibi algoritmalara öğrendiği her örnek için istenen çıktılar doğrultusunda denetim sağlamasıdır. Girdi ve çıktılarından oluşan bir veri seti oluşturmak genellikle zahmetli bir süreç olsa da, denetimli öğrenme algoritmaları iyi anlaşılabilir ve performanslarının ölçülmesi diğer algoritmalara göre nispeten kolaydır. Denetimli makine öğrenmesi problemlerinin iki ana tipi bulunmaktadır, regresyon ve sınıflandırma.

1.1.1. Regresyon

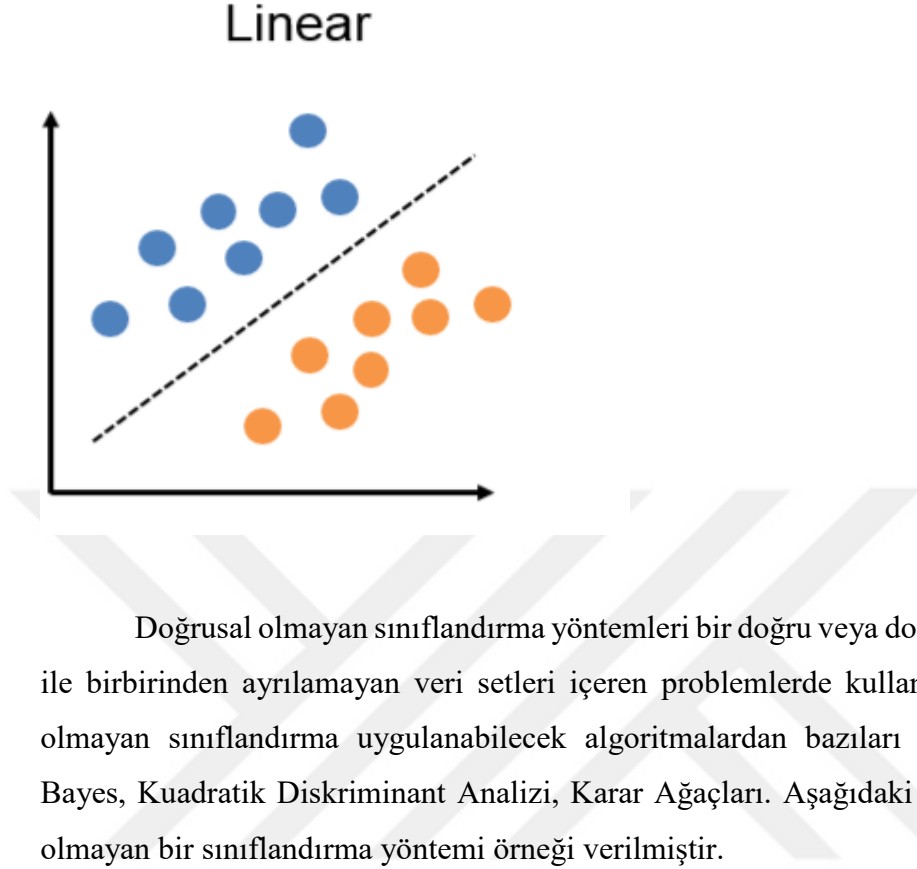
Regresyon için amaç, sürekli bir sayıyı veya gerçek bir sayıyı tahmin etmektir. Basit bir şekilde açıklamak gerekirse bir veya birden fazla değişkenin davranışının, bir

veya birden fazla deęişken üzerinden tahminlenmesidir. Bir kişinin eğitiminden, yaşından ve yaşadığı yerden yıllık gelirini tahmin etmek, bir regresyon görevine örnektir. Geliri tahmin ederken, tahmin edilen deęer bir miktardır ve belirli bir aralıktaki herhangi bir sayı olabilir. Bir başka regresyon görevi örneęi, önceki verimler, hava durumu ve çiftlikte çalışan çalışan sayısı gibi özellikler verilen bir mısır çiftliğinin verimini tahmin etmektir.

1.1.2. Sınıflandırma

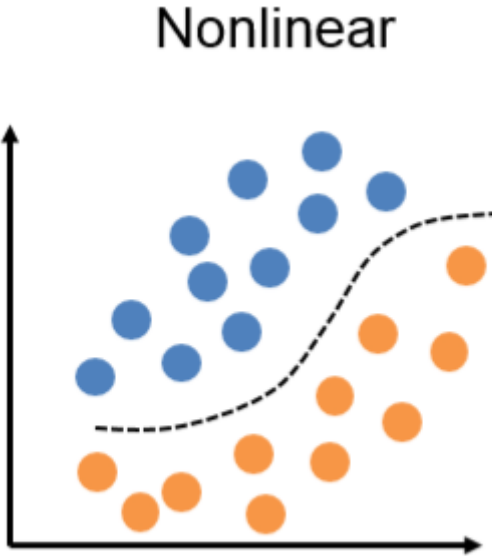
Sınıflandırmada amaç, önceden tanımlanmış bir olasılıklar listesinden bir seçim olan bir sınıf etiketini tahmin etmektir. Kısaca açıklamamak gerekirse verilerin hangi sınıfa ait olduęu bilgisi bulanana(etiketli) bir veri setinin eğitim verisi olarak kullanılıp, etiketi olmayan bir verinin hangi sınıfta yer alacağını tahmin etmeye sınıflandırma denir. Sınıflandırma bazen, tam olarak iki sınıfı ayırt etmenin özel durumu olan ikili sınıflandırmaya ve ikiden fazla sınıf durumunda ise çok sınıflı sınıflandırmaya ayrılır. İkili sınıflandırma bir soruya evet veya hayır olarak cevap vermeye benzetilebilir. Örneğin “e-posta kutusuna gelen mail bir spam mı?”. Sınıflandırma yöntemleri iki kategoride incelenebilir; “Doęrusal Sınıflandırma Yöntemleri” ve “Doęrusal Olmayan Sınıflandırma Yöntemleri”. Doęrusal sınıflandırma yöntemleri, örnekleri birbirine yakınlık/uzaklık deęerlerine göre sınıflandırma yaparken doęrusal bir yol izler. Yani benzer özellikteki verileri birbirinden ayırırken doęrusal bir çizgi kullanır. Doęrusal sınıflandırma uygulanabilecek algoritmalarından bazıları şunlardır; Lojistik Regresyon, Doęrusal Diskriminant Analizi, Kısmi En Küçük Kareler Diskriminant Analizi. Aşağıdaki şekilde doęrusal bir sınıflandırma yöntemi örneęi verilmiştir.

Şekil 1: Doğrusal Sınıflandırma



Doğrusal olmayan sınıflandırma yöntemleri bir doğru veya doğrusal bir düzlem ile birbirinden ayrılamayan veri setleri içeren problemlerde kullanılırlar. Doğrusal olmayan sınıflandırma uygulanabilecek algoritmalarından bazıları şunlardır; Naive Bayes, Kuadratik Diskriminant Analizi, Karar Ağaçları. Aşağıdaki şekilde doğrusal olmayan bir sınıflandırma yöntemi örneği verilmiştir.

Şekil 2: Doğrusal Olmayan Sınıflandırma



1.1.2.1. Karar Ağaçları Algoritmaları

Karar ağaçları, sınıflandırma ve regresyon görevleri için yaygın olarak kullanılan modellerdir. Kısaca bir ağaç yapısı şeklinde bir model oluşturan bir sınıflandırma yöntemi karar ağacı olarak adlandırılır. Karar ağaçlarının şekillendirilmesi, anlaşılması ve yorumlanması diğer yöntemlere göre daha kolaydır. Karar ağacı yönteminde, sınıflandırmak için elimizdeki verilerden bir ağaç oluşturulur. Veri kümesindeki kayıtlar bu ağaca uygulanır. Sonuçlara göre kayıtların sınıflandırılması yapılır. Diğer bir deyişle, veri tabanından elde edilen karar ağaçlarının hangi sınıfa ait olduğunu bilmediğimiz bir veriye sahip olduğumuzda, oluşturulan kural kümesine göre sınıf tahmin edilmektedir.

Karar ağaçları hiyerarşik bir yapıya sahiptir ve dallanma kök düğümden başlar. Kök Düğüm, heterojen bir yapıda aynı zamanda veri setindeki problemi en iyi temsil eden değişkenlerden biridir. Kök düğümler ikiye veya ikiden daha fazla dala ayrılabilir. Ayırma işlemindeki her düğüm basit bir şekilde Evet/Hayır ya da Doğru/Yanlış soruları kullanılarak yeni düğümler oluşturulur. Dalların ucundaki alt düğümler ise karar düğümleri olarak adlandırılır. Dallanmanın durduğu yani devam etmediği düğümler yaprak düğümler olarak adlandırılır. Yaprak düğümleri son sonuçları taşıyan düğümlerdir. Sınıflandırma problemlerinde yaprak düğüme ulaşıldığında, çoğunlukta olan kategorik değer o yaprak düğüme değerini verir. Diğer bir yandan regresyon problemlerinde çıktı değerlerinin ortalaması yaprak düğüme değerini verir.

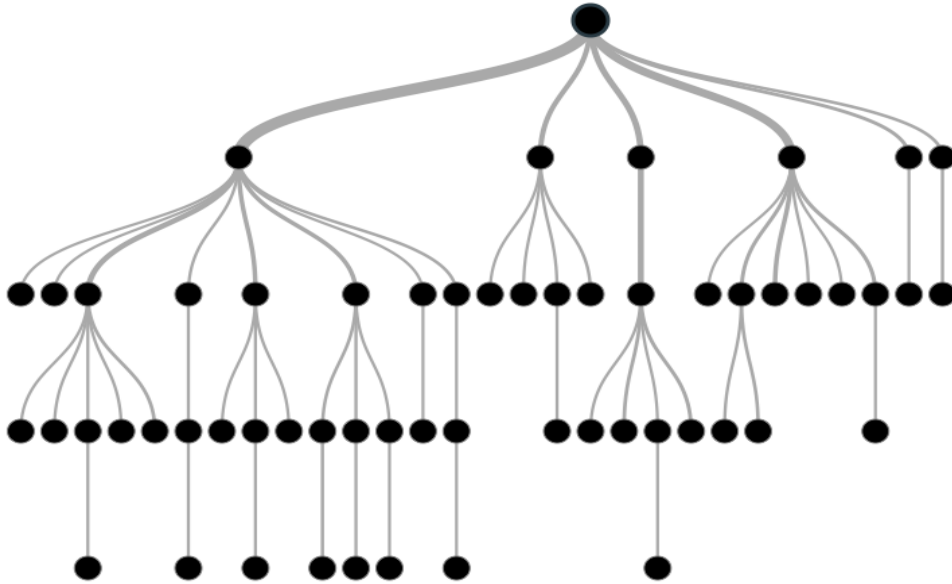
Güçlü bir karar ağacı oluşturmak için, hangi değişkenin kök, alt düğümler veya yaprak olacağına karar vermek çok önemlidir. Bu kararları vermek için kirlilik ve bilgi kazancı kavramları ortaya atılmıştır. Karar ağacı başlatılırken en düşük kirliliğe sahip değişken kök düğüm olarak seçilir, yapraklara ulaşana kadar her düğüm için kirlilik hesaplanır. Kirliliği hesaplamanın farklı yolları vardır ancak literatürde en çok kullanılan iki yöntem Gini İndeksi ve Entropi olarak belirlenmiştir.

Gini Endeksi(I_G) ve Entropi(I_H) kirlilik ölçülerinin formülleri yukarıda verilmiştir. p_n her c sınıfına ait gözlemlerin oranını vermektedir ve p 'nin 0'a eşit olması durumunda entropi değeri de 0 olarak kabul edilir.

Karar ağacı sınıflandırmak için kullanılan bir dizi kararı tanımladığı bir akış şemasına çok benzeyen bir model oluştursa da bu ağaçlar literatürde aynı zamanda sayısal tahminler yapmak için de kullanılmaktadır. Sayısal tahmin yapan karar ağaçları iki kategoriye ayrılır. İlki regresyon ağaçları olarak bilinen, 1980'lerde çığır açan Sınıflandırma ve Regresyon Ağacı (CART) algoritmasının bir parçası olarak tanıtılmıştır. Adına rağmen, regresyon ağaçları bu bölümde daha önce açıklandığı gibi doğrusal regresyon yöntemlerini kullanmazlar, bunun yerine bir yaprağa ulaşan örneklerin ortalama değerine dayanarak tahminlerde bulunurlar.

Sayısal tahmin için kullanılan ikinci tür ağaçlar, model ağaçlar olarak bilinir. Regresyon ağaçlarından birkaç yıl sonra literatüre kazandırılmışlardır ve daha az bilinirler. Model ağaçları, regresyon ağaçlarıyla aynı şekilde büyütülür, ancak her yaprakta, o düğüme ulaşan örneklerden çoklu doğrusal bir regresyon modeli oluşturulur. Yaprak düğümlerinin sayısına bağlı olarak, bir model ağacı bu türden onlarca, hatta yüzlerce model oluşturabilir. Bu, model ağaçlarının anlaşılmasını regresyon ağacından daha zor hale getirebilmektedir. AID, CHAID, CART gibi birçok karar ağacı algoritması bulunmaktadır.

Şekil 3: Karar Ağacı Örneği



1.1.2.1.1. AID Algoritması

Karar ağacı algoritmaları arasında matematiksel tabanlı olarak ortaya atılan ilk algoritma olan AID (Automatic Interaction Detector) algoritması, Morgan ve Sonquist tarafından 1963 yılında literatüre kazandırılmıştır. Algoritmanın varsayımları arasında bağımsız değişkenlerin nitel ve bağımlı değişkenin ise nicel olması şartı bulunmaktadır. Yalnızca ikili bölünmelere izin veren AID algoritması değişken sayısının fazla olduğu durumlarda değişkenlerin etkilerini aynı anda ölçebilme eksikliğinden doğmuştur.

1.1.2.1.2. CHAID Algoritması

CHAID (Chi-Squared Automatic Interaction Detection) algoritması Gordon V. Kass tarafından ilk olarak 1977’de literatüre kazandırılmış olsa da çeşitli revizelerden sonra 1980’de kullanılmaya başlanmıştır (Kass, 1980:119). CHAID algoritması; bağımlı değişkendeki varyasyonu gruplar içinde en düşük, gruplar arasında ise en yüksek olarak ayıran, bu amaçla değişkenler arasındaki ilişkiyi belirlemeye çalışan bir algoritmadır. CHAID algoritması varsayımlarında bağımlı değişken kategorik veya nice olarak ele alınabilirken, bağımsız değişkenin kategorik olması şartı aranmaktadır. AID algoritmasının eksik yönlerini tamamlayan bir yöntem olarak geliştirilmiştir. AID algoritmasına göre en önemli avantajı çoklu bölünmelere izin vermesidir aynı zamanda CHAID algoritmasının Pearson Ki-Kare analizi tabanlı bir ayırım yapması, bu algoritmayı daha güvenilir bir hale getirmektedir (Kayri ve Boysan, 2007:140).

1.1.2.1.3 CART Algoritması

1984 yılında Leo Breiman tarafından yayılan karar ağacı algoritması olarak literatüre kazandırılmıştır. CART algoritması (Classification and Regression Trees) CART algoritması, sınıflandırma için temel olarak Gini indeksini kullanır ve kesikli veya sürekli sınıflandırma niteliklerine dayalı olarak sınıflandırma ve regresyon ağaçları üretir. CART sınıflandırma algoritması temel olarak iki adıma bölünmüştür; birincisi, bir ağaç oluşturmak için numuneyi özyinelemeli olarak bölmek ve diğeri,

doğrulama verileriyle budama yapmaktır. Gini indeksi, modelin saflığını temsil eder. Gini indeksi ne kadar küçük olursa, saflık o kadar yüksek ve özellikler o kadar iyi olur(Zhu ve diğerleri, 2020).

1.2. DENETİMSİZ ÖĞRENME

Bir veri kümesinin denetimsiz dönüşümleri, verilerin orijinal temsiline kıyasla insanların veya diğer makine öğrenimi algoritmalarının anlamasını daha kolay hale getirecek şekilde yeniden düzenleyen algoritmalarlardır. Denetimsiz dönüşümlerin yaygın bir uygulaması olan boyut indirgeme, büyük boyutlardaki birçok değişkenden oluşan bir verinin temsilini alıp orijinal verinin niteliklerinden bir şey kaybettirmeden daha az değişkenle özetlemesinde kullanılır. Başka bir uygulama olan kümeleme algoritmaları ise verileri benzer öğelerden oluşan farklı gruplara ayırır. Bir sosyal medya sitesine fotoğraf yükleme örneğini düşünün. Site, resimlerinizi düzenlemenize izin vermek için aynı kişiyi gösteren resimleri bir araya getirmek isteyebilir. Ancak site, hangi resimlerin kimi gösterdiğini bilmiyor ve fotoğraf koleksiyonunuzda kaç farklı kişinin görüldüğünü bilmiyor. Mantıklı bir yaklaşım, tüm yüzleri çıkarmak ve onları benzer görünen yüz gruplarına bölmek olacaktır.

1.3. YARI DENETİMLİ ÖĞRENME

Yarı denetimli öğrenme, denetimli ve denetimsiz öğrenme tekniklerinin birleşimi olan bir makine öğrenmesi tekniğidir. Bu öğrenme stilinden eğitim verileri etiketli ve etiketsiz olarak bulunabilirler. Bu yöntem etiketli verilerden elde ettiği bilgileri, etiketsiz verileri etiketlemek için kullanır, yöntem genellikle etiketli verilerin etiketsiz verilerden sayıca az olduğu durumlarda kullanılır. Yöntemin amacı, eğitim verisine etiketli gözlemleri etiketsiz gözlemlerin yanına tekrarlı olarak dahil ederek algoritmanın tahmin performansını iyileştirmektir.

1.4. PEKİŞTİRMELİ ÖĞRENME

Mevcut durumunu algılayarak otonom kararlar alabilen bir sistemimiz olduğunu düşünelim. Sistem bir amaç doğrultusunda belli kararlar almakta ve çeşitli

aksiyonlarda bulunarak durumunu yenilemektedir. Pekiştirmeli öğrenme sistemin aldığı kararların arasından doğru olanı nasıl bulacağını veya bulmayı öğreneceğini gösteren bir öğrenme algoritmasıdır. Sistem doğru alınan kararlar için ödüllendirilirken, yanlış alınan kararlar için ya daha az ödül almaktadır ya da cezalandırılmaktadır. Sistemin amacı en yüksek ödüle ulaşacak kararları almak için gereken aksiyonların sırasını öğrenmektir.



İKİNCİ BÖLÜM

KÜMELEME ANALİZİ

Kümeleme analizi, varlıkları benzerliklerine ve farklılıklarına göre nesnel olarak gruplamak için kullanılan çeşitli yöntemler içeren genel bir terimdir. Bu yöntemleri uygularken amaç, varlıkları (öğeler, nesneler, vb.) birbirini dışlayan kümeler halinde gruplandırmaktır, böylece kümeler birbirinden farklı iken kümelerin içindeki öğeler homojen bir şekilde gruplanmış olur. Kümeleme analizinin temel amaçları, verilerin azaltılması, veri keşfi, doğal grupların belirlenmesi, gruplara dayalı tahmin, sınıflandırma, model kurma, hipotezlerin üretilmesi ve test edilmesidir(Iyer ve Aronson, 1999:91).

Kümeleme analizi, birincil amacı nesneleri sahip oldukları özelliklere göre bir araya getirmek olan çok değişkenli bir teknikler grubudur. Küme analizi, nesneleri (örneğin, yanıtlayanlar, ürünler veya diğer varlıklar) sınıflandırır, böylece her nesne önceden belirlenmiş bir seçim kriterine göre kümedeki diğerlerine benzer. Ortaya çıkan nesne kümeleri daha sonra küme içi homojenlik ve kümeler arası heterojenlik göstermelidir. Böylece, sınıflandırma başarılı olursa, kümeler içindeki nesneler geometrik olarak çizildiğinde birbirine yakın olacak ve farklı kümeler birbirinden daha uzak olacaktır.

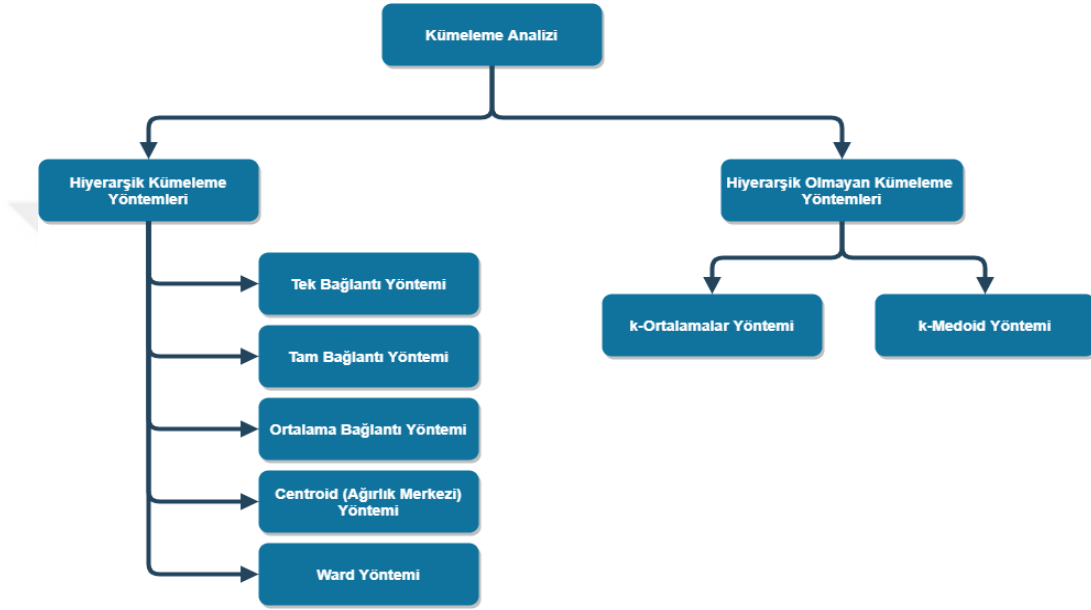
Küme analizi, "Q analizi", "tipoloji", "sınıflandırma analizi" ve "sayısal sınıflandırma" olarak adlandırılmıştır. Bu isim çeşitliliği kısmen psikoloji, biyoloji, sosyoloji, ekonomi, mühendislik ve işletme gibi çeşitli disiplinlerde kümeleme yöntemlerinin kullanılmasından kaynaklanmaktadır. Adları disiplinler arasında farklılık gösterse de, yöntemlerin hepsinin ortak bir boyutu vardır: doğal ilişkilere göre sınıflandırma. Bu ortak boyut, tüm kümeleme yaklaşımlarının özünü temsil eder. Bu nedenle, kümeleme analizinin en önemli özelliği, verilerin kendilerinin "doğal" gruplandırmaları tarafından önerildiği gibi, verilerin sınıflandırılmasında yatmaktadır.

Küme analizi birçok durumda yararlıdır. Örneğin, bir anketle veri toplayan bir araştırmacı, yönetilebilir gruplara ayrılmadıkça anlamsız olan çok sayıda gözlemle karşı karşıya kalabilir. Kümeleme analizini kullanarak araştırmacı, tüm bir popülasyondan veya örneklemden gelen bilgileri belirli, daha küçük alt gruplar

hakkındaki bilgilere indirgeyerek nesnel olarak veri indirgeme gerçekleştirebilir(Grimm ve Yarnold, 2000:147).

Kümeleme yöntemleri Hiyerarşik Kümeleme ve Hiyerarşik Olmayan Kümeleme şeklinde iki ana başlıkta incelenmektedir.

Şekil 4: Kümeleme Analizi Yöntemleri



Hiyerarşik kümeleme yöntemlerinde, öğelerin/değişkenlerin kümelenmesi için uygun bir uzaklık ölçüsü ile öğelerin/değişkenlerin birbirlerine olan uzaklıkları hesaplanmaktadır (Alpar 2011: 314). Değişkenlerin ölçüldüğü ölçekler, değişken sayısı arttıkça farklılık göstermektedir. Bu yüzden değişkenlerin ölçüldüğü ölçeklerin farklı olduğu durumlarda, verilerin analiz edilmeden önce standartlaştırılması veya belirli aralıklara dönüştürülmesi gerekmektedir, bunun için literatürde farklı yöntemler bulunmaktadır (Kalaycı 2010: 358).

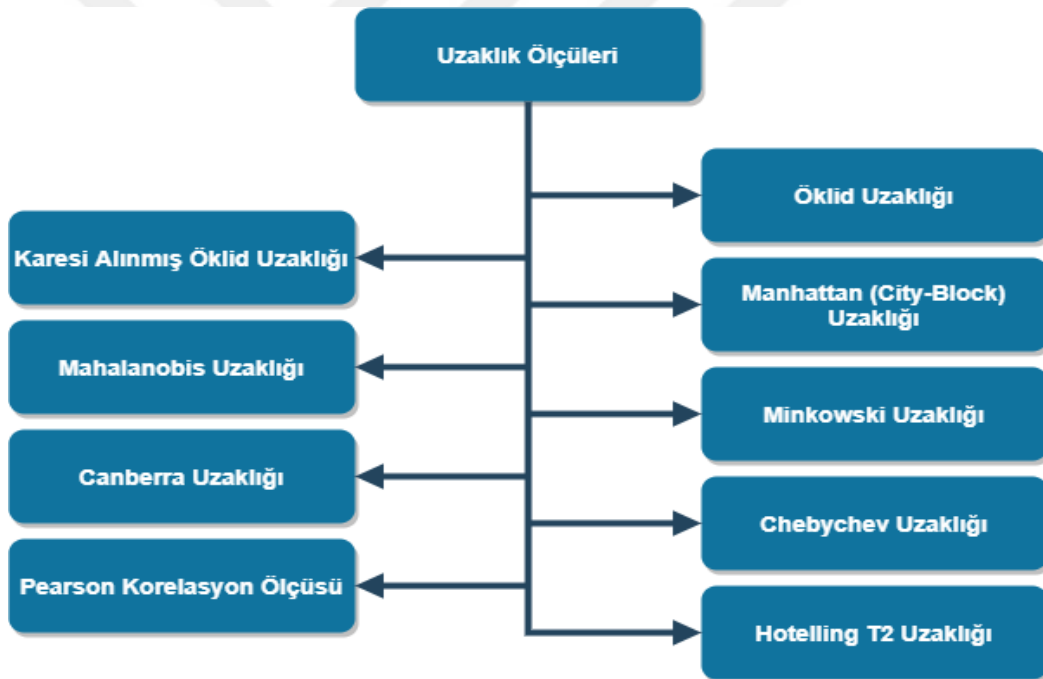
Kümeleme analizinin doğru bir sonuca ulaşması, araştırmacıların popülasyonu iyi temsil eden ve güvenilir bir örneklem seçimi yapmasına bağlıdır. Kümeleme analizinde, değişken sayısının artması gözlem sayısının artmasını gerektirmektedir.

Aynı zamanda gözlem sayısının değişken sayısının en az 3-4 katı olması gerektiği ile ilgili görüşler mevcuttur(Kalaycı 2010: 358).

2.1. UZAKLIK ÖLÇÜLERİ

Küme değişkenindeki değişkenler arasında gözlemlerin birbirine yakınlığı olarak benzerliği temsil eden uzaklık ölçüleri, en sık kullanılan benzerlik ölçüsüdür. Uzaklık ölçüleri aslında bir farklılık ölçüsüdür ve daha büyük değerler daha az benzerliği ifade eder. Mesafe, ters bir ilişki kullanılarak bir benzerlik ölçüsüne dönüştürülür. Aşağıda literatürde en çok kullanılan uzaklık ölçüleri incelenmiştir.

Şekil 5: Uzaklık Ölçüleri



2.1.1. Öklid Uzaklığı

En fazla bilinen ve çoğu zaman vektör uzaklığı olarak da adlandırılan bir uzaklık ölçüsüdür. Öklid uzaklığının özelliği genellikle ham veriler üzerinde kullanılmasıdır. Bu uzaklık aykırı değerlerden çok etkilenmezken verilerin farklı

ölçeklerden elde edilmesinden etkilenmektedir. Öklid(vektör) Uzaklığı kolaylıkla iki değişkenden fazla değişken bulunan durumlar için genelleştirilebilir. Genelleştirilmiş formül aşağıdaki gibi gösterilebilir:

$$d(x_i, x_k) = \sqrt{\sum_{j=0}^n w_j (x_{ij} - x_{kj})^2} \quad (1)$$

Literatürdeki çalışmaların bazılarında standart ölçü olarak kullanılan ve çok değişkenli analiz paket programlarının büyük çoğunluğunda hali hazırda bulunan bir ölçüdür. Hesaplama aşamasında değerlerin birbirinden bağımsız olduğu varsayımı vardır ve en yakın değer 0 olarak kabul edilir (Hair ve diğerleri, 2018: 253).

2.1.2. Karesi Alınmış Öklid Uzaklığı

Karesi alınmış Öklid uzaklığı kareköksüz hesaplama yapmasından kaynaklı hesaplamayı hızlandıran bir avantaja sahiptir aynı zamanda Ward ve ağırlık merkezi kümeleme yöntemleri için uygun bir uzaklık ölçüsüdür. Kareköksüz hesaplama yapılması bir avantaj olduğu kadar yöntemi aykırı gözlemlere karşı hassaslaştırdığı için aynı zamanda bir dezavantajdır (Hair ve diğerleri, 2018: 431).

$$d(x_i, x_k) = \sum_{j=0}^n (x_{ij} - x_{kj})^2 \quad (2)$$

2.1.3. Manhattan (City-Block) Uzaklığı

Manhattan Uzaklığı, Öklid Uzaklığı temelli bir uzaklık değildir. Hesaplanırken değişkenlerin arasından mutlak farkların toplamını esas alır. Hesaplaması kolay bir yöntem olsa da eğer değişkenler arasındaki korelasyon fazla ise geçersiz kümelerle sonuçlanabilir. Daha çok kesikli veriler içeren değişkenler ile kullanılması tavsiye edilmektedir ve formül aşağıdaki gibidir (Alpar, 2017: 172):

$$d(x_i, x_k) = \sum_{j=0}^n |x_{ij} - x_{kj}| \quad (3)$$

2.1.4. Mahalanobis Uzaklığı

Değişkenler arasındaki korelasyonu dikkate alarak hesaplanmak için ortaya konulmuştur aynı zamanda Mahalanobis Uzaklığı'nda ölçek sabittir. Standartlaştırılmamış yani ham verilerde değişkenlerin birbirinden bağımsız olduğu durumda Mahalanobis Uzaklığı, Öklid uzaklığının aynısıdır. Yani standartlaştırılmış veriler söz konusu olduğunda Öklid Uzaklığı, Mahalanobis Uzaklığı'nın özel bir durumu olarak karşımıza çıkar. Değişkenlerin standartlaştırılmış ve birbiri ile bağımlı olması durumunda Mahalanobis Uzaklığı en uygun uzaklıktır. Formülü aşağıdaki gibidir (Sharma, 1996:220):

$$d(x_i, x_k) = (x_i - x_k)'S^{-1}(x_i - x_k) \quad (4)$$

2.1.5. Minkowski Uzaklığı

Minkowski Uzaklığı $m = 1$ olması durumunda iki nokta arasındaki “Manhattan” uzaklığını ölçerken, $m = 2$ olması durumunda ise Öklid uzaklığını ölçer. Formüldeki m değerinin değişimine göre farklar değişiklik göstermektedir. Formül aşağıdaki gibidir:

$$d(x_i, x_k) = \left[\sum_{j=1}^p |x_{ij} - x_{kj}|^m \right]^{\frac{1}{m}} m \geq 1 \quad (5)$$

2.1.6. Canberra Uzaklığı

Canberra Uzaklığı sadece pozitif değerlere sahip değişkenler için kullanılmaktadır. Canberra denklemindeki pay farkı belirtirken, payda farkı normalleştirmektedir. Canberra Uzaklığı karmaşık ve gürültülü veri setleri için uygun değildir. Formül aşağıdaki gibidir:

$$d(x_i, x_k) = \sum_{i=1}^p \frac{|x_i - x_k|}{|x_i + x_k|} \quad (6)$$

Burada p boyut sayısını x_i ve x_k farklı gözlem değerlerini ifade etmektedir.

2.1.7. Chebychev Uzaklığı

Başka bir ifadeyle Maksimum Değer Uzaklığı, aşağıda verilmiş olan denklemden de anlaşılabileceği gibi iki gözlem arasındaki farkın mutlak değerinin maksimumunu bulan bir uzaklık ölçüsüdür. Formül aşağıdaki gibidir (Irani ve diğerleri, 2016: 10):

$$d(x_i, x_k) = \max |x_{ij} - x_{kj}| \quad (7)$$

2.1.8. Pearson Korelasyon Ölçüsü

Bilinen en eski uzaklık ölçülerinden biridir. İki sürekli değişken arasındaki doğrusal ilişkiyi göstermektedir. Formül aşağıdaki gibidir:

$$d(i, j) = \sqrt{\frac{\sum_{k=1}^p (x_{ik} - x_{jk})^2}{S_k^2}} \quad , \quad i, j = 1, 2, \dots, n \quad (8)$$

2.1.9. Hotelling T² Uzaklığı

Genellikle normal dağılıma sahip iki kümenin ya da grubun ortalama vektörlerinin karşılaştırılması için kullanılan Hotelling T² değeri bir uzaklık ölçüsü olarak da kullanılabilir. Formülü aşağıdaki gibidir (Johnson, 1998: 329):

$$d(x_i, x_k) = \frac{n_1 - n_2}{n_1 + n_2} (x_i - x_k)' S^{-1} (x_i - x_k) \quad (9)$$

2.2. HİYERARŞİK KÜMELEME ANALİZİ YÖNTEMLERİ

Aşama aşama kümeleme yapmasından kaynaklı bu ismi almıştır. Gruplayıcı ve bölücü olarak iki tür hiyerarşik kümeleme yöntemi mevcuttur. Gruplayıcı yöntemler küçük kümelerden başlayarak en yakın kümeleri birleştirip büyük kümeler oluştururlar, benzer kümeler birleştikçe küme sayısı azalır. Bu yöntemler dendrogram veya ağaç grafikleri ile görselleştirilebilir. Bölücü yöntemler gruplayıcı yöntemlerin tam ters mantığı ile çalışırlar. Başta bütün gözlemleri içeren tek bir küme varken kümeler kendi aralarında bölünmeye başlarlar bu şekilde büyük kümeler birbirine benzeyen birimlerden oluşan küçük kümelere dönüşürler. Hiyerarşik kümeleme analizi yöntemlerinden bazıları bir sonraki bölümde anlatılmıştır.

2.2.1. Tek Bağlantı Yöntemi

İlk olarak 1951 yılında Florek ve arkadaşları tanımlanıp literatüre kazandırılmıştır (Hartigan, 1981: 388). Tek bağlantı yöntemi literatürde en yakın komşuluk ismiyle de birçok çalışmada kullanılmaktadır. Bu yöntemde, birbirine en benzer birimler veya nesneler birleştirilerek kümeler oluşturulur. Her bir birim veya nesne bir kümeye ait olana kadar bu işlem devam eder. Birimlerin birbirine benzeme dereceleri uzaklık matrisi kullanarak ölçülür. Kümeleme işlemi ilk olarak birbirine en benzer iki birimin bir küme oluşturması ile başlar. Daha sonra bu kümeye en yakın diğer birimin eklenmesi ile işlem bütün kümeler oluşturuluncaya kadar devam eder.

Bu yöntem bazı çalışmalarda eleştirilmektedir bunun sebebi küme yapısını hesaba katmaması ve oluşan kümelerin seyrek, uzun olmasından kaynaklanmaktadır. Aynı zamanda bu yöntem sapan değerlere karşı hassastır (Everitt ve diğerleri, 2011: 102). Bu yöntemde i . ve j . birimler birleştirilmiş ise birleştirilen kümenin k . küme ile ilişkisi uzaklık ölçüsü olarak,

$$d_{k(i,j)} = \text{Min}(d_{ki}, d_{kj}) \quad (10)$$

biçiminde ifade edilir. Denklemden;

$d_{k(i,j)}$: k . kümenin daha önce oluşmuş olan i . ve j . kümeye olan uzaklığını,

d_{ki} : k . kümenin i . kümeye olan uzaklığını,

d_{kj} : k . kümenin j . kümeye olan uzaklığını göstermektedir.

2.2.2. Tam Bağlantı Yöntemi

İlk olarak 1948 yılında Sorensen tarafından önerilmiş ve literatüre kazandırılmıştır(Hartigan,1981:388). Tek bağlantı yöntemine benzer bir hesaplama kullanır ancak mantık açısından tam tersi olarak düşünülebilir. Tek bağlantı yöntemi için yapılan eleştiriler genellikle bu yöntem içinde geçerlidir. Tam bağlantı yöntemi iki birim arasındaki en büyük uzaklığı baz alır ve kümeleri oluşturur. Bu yöntemde uzaklıklar,

$$d_{k(i,j)} = \text{Max}(d_{ki}, d_{kj}) \quad (11)$$

şeklinde gösterilmektedir.

2.2.3. Ortalama Bağlantı Yöntemi

İlk olarak 1958 yılında Sokal ve Michener tarafından literatüre kazandırılmıştır. Bu yöntem tek ve tam bağlantı yöntemlerindeki gibi kümelemeye başlar ancak kümeleme kriteri olarak bir kümedeki bütün bireylerin başka bir

kümedeki bütün bireylerle olan ortalama benzerliğini kullanır. Bu yöntemin tek ve tam bağlantı yöntemlerine göre avantajı en yakın veya en uzak çiftlere bağlı olmamasıdır. Yani kümeleme kriterindeki benzerlik kavramı kümeyi oluşturan bütün birimlere dayandırılır. Yöntem oluşan kümeleri homojen tutma eğilimindedir (Hair ve diğerleri, 2014: 441).

$$d_{(uv)w} = \frac{\sum_i \sum_k d_{ik}}{N_{(uv)}N_w} \quad (12)$$

$N_{(uv)}$: (uv) kümesindeki gözlem sayısı,

N_w : w kümesindeki gözlem sayısı,

d_{ik} : w kümesindeki k değeri ile (uv) kümesindeki i değeri arasındaki uzaklıktır.

2.2.4. Centroid (Ağırlık Merkezi) Yöntemi

Ortalama bağlantı yöntemi gibi Sokal ve Michener tarafından 1958 yılında literatüre kazandırılmıştır. Ortalama bağlantı yönteminin özel bir durumudur. Bu yöntemde kümelerin ağırlık merkezleri kullanılır kümeler arasındaki benzerlik ağırlık merkezleri arasındaki uzaklık ile ölçümlenir. Her bir birim kümeye eklendiğinde aşağıdaki formülle yeni küme ağırlık merkezleri hesaplanır:

$$\bar{x}_{AB} = \frac{n_A \bar{x}_A + n_B \bar{x}_B}{n_A + n_B} \quad (13)$$

$\bar{x}_A$: A kümesinin değerlerinin ortalamasını,

$\bar{x}_B$: B kümesinin değerlerinin ortalamasını göstermektedir.

2.2.5. Ward Yöntemi

İlk olarak 1963 yılında Ward tarafından ortaya atılan bu yöntem hiyerarşik kümeleme yöntemlerinde en çok kullanılan yöntemlerden biridir. Kümeleme kriteri olarak uzaklık ya da benzerlik ölçütleri kullanmak yerine varyans analizi

benimsenmiştir. Yöntemin başlangıcında her bir küme içerisinde tek birim bulunacak şekilde n tane küme oluşturulur. Başlangıçta her küme sadece bir birimden oluştuğu için hata kareler toplamı sıfırdır ve amaç hata kareler toplamı en az olacak kümeleri oluşturmaktır. Yöntemin her adımında iki küme birleştirilir ve bu iki küme, tüm kümelerdeki değişkenlerin toplam kareler toplamındaki artışı en aza indiren kümelerdir. Yöntemde toplam kareler kullanılması aykırı değerlere hassas olmasına yol açmaktadır ancak homejen kümeler elde etmek için uygundur.

Ward yöntemi küme içerişindeki uzaklıkları en küçük, kümeler arasındaki uzaklıkları en büyük yapmayı amaçlamaktadır. Yöntem aşağıdaki formüllerle ifade edilebilir:

$$KT_A = \sum_{j=1}^{n_A} (x_j - \bar{x}_A)' (x_j - \bar{x}_A) \quad (14)$$

$$KT_B = \sum_{j=1}^{n_B} (x_j - \bar{x}_B)' (x_j - \bar{x}_B) \quad (15)$$

A ve B kümeleri birleştirildikten sonra elde edilen denklem:

$$KT_{AB} = \sum_{k=1}^{n_{AB}} (x_k - \bar{x}_{AB})' (x_k - \bar{x}_{AB}) \quad (16)$$

Denklemde geçen $\bar{x}_{AB}$ için formül aşağıdaki gibidir:

$$\bar{x}_{AB} = \frac{n_A \bar{x}_A + n_B \bar{x}_B}{n_A + n_B} \quad (17)$$

Kareler toplamının küçüklüğü aşağıdaki formüller belirlenir, formülde en küçük değer alan kümeler bir araya getirilir:

$$I_{AB} = KT_{AB} - (KT_A + KT_B) \quad (18)$$

Ward yöntemi sonucunda kümelerin birleştirilme durumları “dendrogram” adlı şemada gösterilir.

2.3. HİYERARŞİK OLMAYAN KÜMELEME ANALİZİ YÖNTEMLERİ

Hiyerarşik olmayan kümeleme yöntemlerinin hiyerarşik kümeleme yöntemlerinden en büyük farkı küme sayısı hakkında bilgi sahibi olup olunmamasıdır. Araştırmacı küme sayısı hakkında bilgi sahibi veya oluşturulacak küme sayısına kendi karar vermişse hiyerarşik olmayan kümeleme yöntemlerini kullanabilir. Bu yöntemlerde ana mantık; n tane birimin k tane kümeye ayrıştırılmasıdır. Bu şekilde ilk bölünme yapıldıktan sonra birimlerin kümeler arası yer değiştirmesi yapılarak en uygun kümeleme elde edilebilir. Hiyerarşik olmayan kümeleme yöntemlerinde dendrogram gibi karmaşık tekniklere başvurulmaz. Kümeleme sonucunda elde edilen kümeler artık alt kümelere ayrılmaz ve kümeler arasında kesin bir bölünme oluşmuştur. Hiyerarşik olmayan kümeleme yöntemleri arasında en çok kullanılanları k -ortalamalar kümeleme ve k -medoid kümeleme yöntemleri yer alır. Bu yöntemlere devam eden bölümlerde yer verilmiştir.

2.3.1. k -Ortalamalar Yöntemi

k -ortalamalar yöntemi İlk olarak 1967 yılında Macquenn tarafından literatüre kazandırılmıştır. Kümeleme algoritmaları arasında en popülerlerinden birisidir. k -ortalamalar yöntemi birimleri önceden belirlenmiş küme sayısına göre gruplayarak analize başlar. Yani gözlemler k adet kümeye her birinin içerisinde bir gözlem bulunacak şekilde ayrılmış olur. Daha sonra gözlemler en yakın ortalama grubu eklenmeye başlarlar. Her bir gözlemin eklenmesinden sonra ortalama tekrar hesaplanır ve bütün gözlemler kümeler yerleşene kadar bu adım tekrarlanır. Bütün gözlemler kümeleri oluşturduktan sonra herhangi bir gözlemin ortalaması, başka bir küme ortalamasına kendi küme ortalamasından daha yakınsa yakın olduğu kümeye geçiş yapar. Bu şekilde kümeler içi homojenlik ve kümeler arası heterojenlik sağlanmış olur.

Yöntemin değerlendirilmesinde en çok hata kareler toplamı(Sum of Squared Error-SSE) kullanılır. SSE'nin küçük olması, kümelemede ortalamaların kümeleri iyi temsil ettiği anlamına gelir. Hata kareler toplamı aşağıdaki gibi hesaplanır:

$$SSE = \sum_{j=1}^k \sum_{x \in c_j} \|x_i - c_j\|^2 \quad (19)$$

Burada x_i : veri noktaları, c_j : merkezleri, k : küme sayısını göstermektedir.

2.3.2. k-Medoid Yöntemi

k-ortalamalar yöntemine çok benzer bir yöntemdir. İlk olarak 1987 yılında Rousseeuw ve Kaufman tarafından literatüre kazandırılmıştır. k-ortalamalar yönteminde kullanılan merkezler yerine burada merkeze konum olarak en yakın (medoid) birim küme merkezi olarak seçilir. Bu yöntem merkezin ilk başta rastgele seçilmesi ile başlar kümeye yeni elemanlar eklendikçe kümelerin yeni medoidleri tekrar seçilir ve bu işlem sürekli tekrarlanır. Yöntem her bir birim ile ilgili medoid arasındaki farkların toplamını minimuma indirmeye çalışır. Yöntem sonucunda oluşan kümeler konum olarak medoidlerin çevresinde toplanmış ve birimler kendisine en yakın medoidin oluşturduğu kümeye dahil olmuştur. k-ortalamalar yöntemine göre avantajı aykırı değerlere karşı çok hassas olmaması dezavantajı ise hesaplama karmaşıklığından kaynaklı büyük veri setlerine uygun olmamasıdır. k-medoid yönteminin fonksiyonu aşağıdaki gibidir:

$$S = \sum_{k=1}^K \sum_{i \in C_k: j \in C_k} d(x_i, x_j) \quad (20)$$

Burada S küme içindeki gözlemlerin arasındaki toplam uzaklıktır, S değerinin minimum olması en iyi kümeleme sonucuna ulaşıldığı anlamına gelir.

2.4. KÜME SAYISININ BELİRLENMESİ

Kümeleme analizinin en önemli kısmı küme sayısına karar vermektir. Literatüre bakıldığında küme sayısına karar vermek için çoğunlukla kullanılan bir yöntem olmamasına rağmen kullanılan birkaç yöntem bulunmaktadır. Bu yöntemlerden ilki ve en çok bilineni;

$$k = \sqrt{n/2} \quad (21)$$

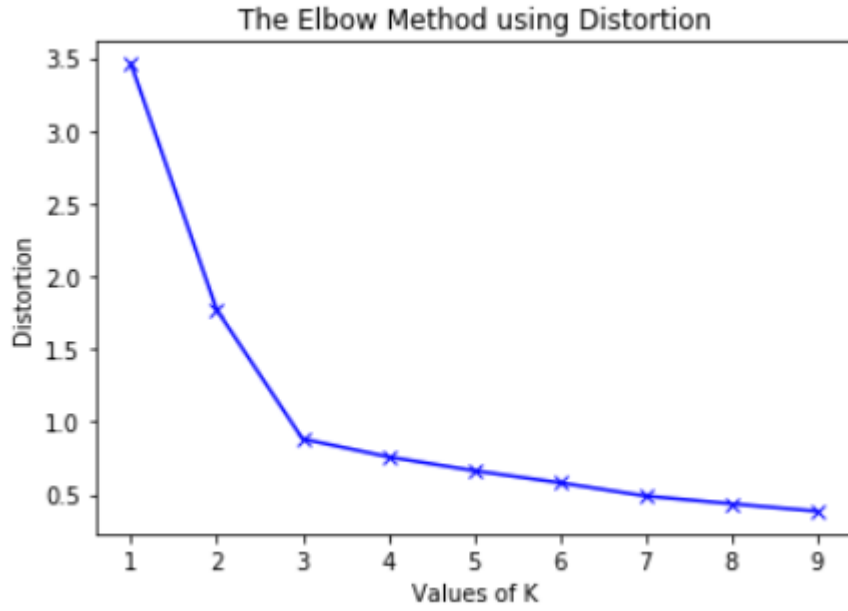
şeklinde hesaplanmaktadır. Bu yöntemde k küme sayısını n ise birim sayısını göstermektedir. Bu yöntemin küçük örneklerde kullanılması tavsiye edilmektedir, büyük örneklerde kullanılması durumunda sonuçların sağlıklı olmayacağı öne sürülmüştür (Everitt, 1974).

Bu yöntemlerden ikincisi literatürde dirsek yöntemi olarak adlandırılan, genellik k-ortalamalar yöntemi ile birlikte kullanılan bir yöntemdir. Dirsek yöntemi, küme sayısının bir fonksiyonu olarak açıklanan varyans yüzdesine bakan bir yöntemdir. Bu yöntem, başka bir küme eklenmesinin verilerin daha iyi modellenmesini sağlamaması fikrine dayanmaktadır. Bu yöntem bir grafik yöntemidir, kümeler tarafından açıklanan varyans yüzdesi, küme sayısına karşı çizilir. İlk kümeler çok fazla bilgi içerecektir, ancak bir noktada marjinal kazanç önemli ölçüde düşecek ve grafikte bir açığı oluşturacaktır. Doğru "k" yani küme sayısı bu noktada seçilir, literatürde bu durum "dirsek kriteri" olarak adlandırılır. Buradaki fikir, $k = 2$ ile başlamak ve küme sayısını her adımda 1'er artırmaya devam etmektir. k için bir değerde kazanç çarpıcı biçimde düşer ve bundan sonra küme sayısını daha da artırdığımızda bir düzlüğe ulaşır. Bu, istediğimiz k değeridir. Dirsek yönteminin denklemi aşağıda verilmiştir:

$$SSE = \sum_{i=1}^k \sum_{j \in C_i} \|X_j - C_i\|^2 \quad (22)$$

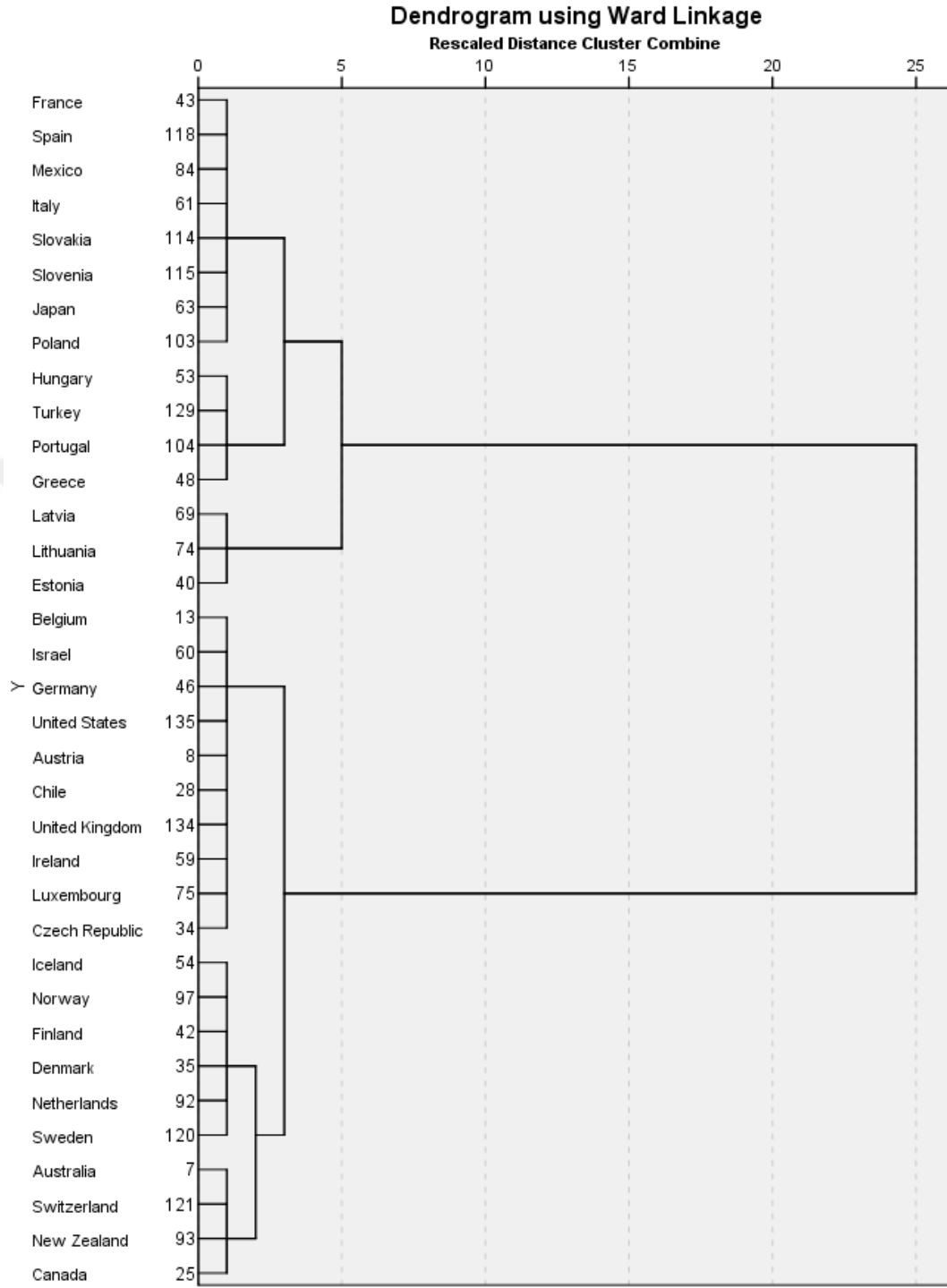
bu denklemde, k küme sayısını, i i. kümeyi, X_j i kümesindeki j verisini, C_i ise rastgele başlatılan merkezleri ifade etmektedir. Aşağıda dirsek yönteminin örnek bir grafiği gösterilmiştir(Kodinariya ve Makwana, 2013).

Şekil 5: Dirsek Yönetimi



Küme sayısının belirlenmesinde kullanılan üçüncü bir yöntemde “Dendrogram” yöntemidir. Dendrogram, çeşitli veri kümeleri arasındaki yapısal ilişkileri gösteren bir ağaç diyagramıdır. Bir Dendrogram, bir hiyerarşik kümeleme algoritması belleğini içerir, böylece kümenin nasıl şekillendiğini yalnızca Dendrogram'a bakarak görebilirsiniz. Dendrogram, kümelere birimleri atamanın en iyi yolunu bulmak için kullanılır. Yöntem kullanılırken y eksenine paralel bir doğru çizilir ve bu çizginin grafikteki yer alan çizgilerle kesişme sayısı küme sayısı olarak kabul edilir. Aşağıda Dendrogram yönteminin örnek bir grafiği gösterilmiştir.

Şekil 6: Dendrogram



Yukarıda SPSS programından alınan bir dendrogram örneği verilmiştir.

ÜÇÜNCÜ BÖLÜM

UYGULAMA

Bu bölümde çalışmada kullanılan veri setlerinin metodolojisi, uygulamanın içeriği, yıllara göre analiz kümeleme sonuçları ve yıllara göre kümelerin belirlenmesinde rol oynayan faktörler incelenmiştir.

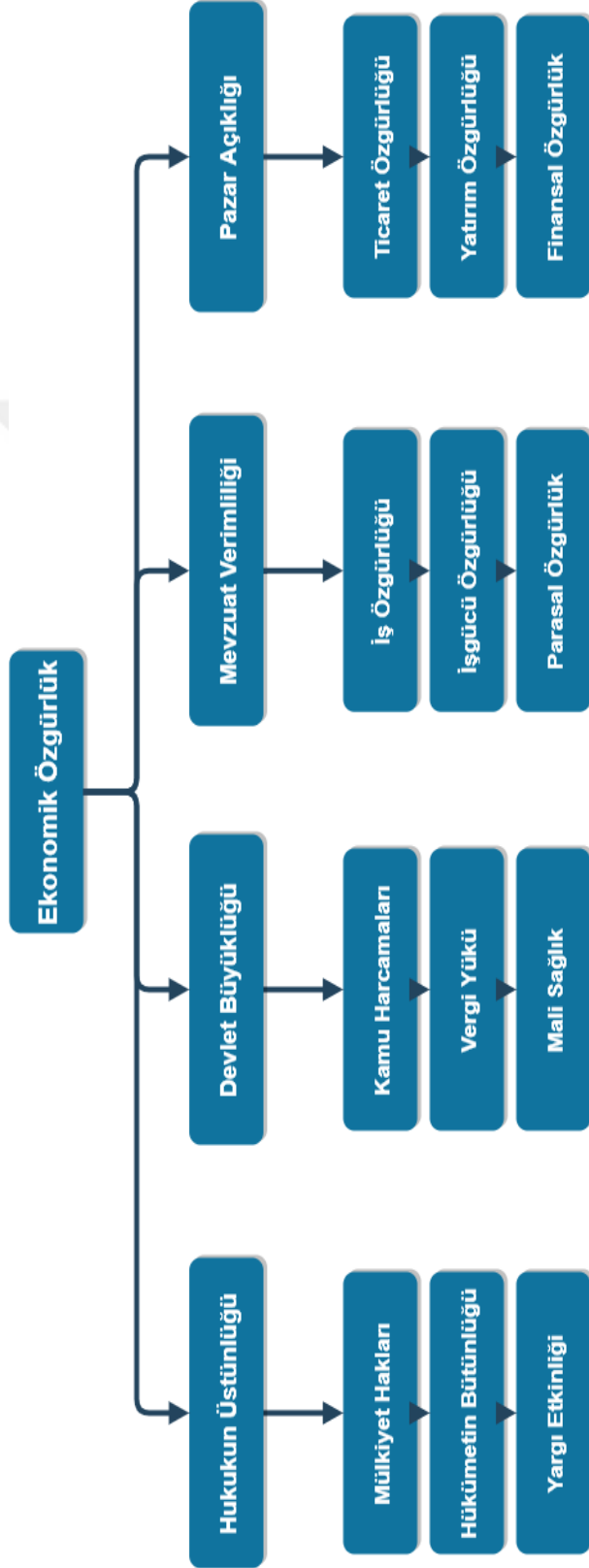
3.1. VERİ SETLERİ METODOLOJİSİ

Bu bölümde ekonomik özgürlük endeksi ve dünya mutluluk raporu verisetlerinde yer alan faktörlerin detaylı açıklamaları ve hesaplama yöntemleri incelenmiştir.

3.1.1. Ekonomik Özgürlük Endeksi Metodoloji

Ekonomik özgürlük endeksi dört ana faktör ve bu dört faktörü oluşturan on iki faktörden oluşmaktadır. Bu alt faktörlerin her biri 0 ile 100 arasında değer almaktadır. Bu on iki değişkenin aritmetik ortalaması o ülkenin o sene elde ettiği ekonomik özgürlük skorunu oluşturur. Aritmetik ortalamaya katılan faktörlerin hepsi eşit ağırlığa sahiptir. Faktörlerle ilgili detaylı bilgi Ekonomik Özgürlük Endeksi resmi sitesinden elde edilmiştir. Bütün değişkenlerin detaylı açıklamalarına bir sonraki bölümde değinilecektir.

Şekil 7: Ekonomik Özgürlük Endeksi Değişkenleri



3.1.1.1. Hukukun Üstünlüğü

Hukukun üstünlüğü üç faktörün ortalamasının alınması ile elde edilmektedir, bu faktörler Mülkiyet Hakları, Hükümetin Bütünlüğü ve Yargı Etkinliği olarak adlandırılmıştır.

3.1.1.1.1. Mülkiyet Hakları

Mülkiyet hakları faktörü bireylerin bir mülkü elde etme, elinde tutma ve kullanmasına hükümet tarafından yasalarca ne kadar izin verildiğini değerlendiren bir faktördür. Faktör anket verilerinin ve bağımsız değerlendirmelerin bir karışımına dayanarak, ülkenin yasalarının özel mülkiyet haklarını ne ölçüde koruduğuna ve bu yasalara ne kadar saygı duyulduğuna ilişkin bir ölçüm sağlar. Mülkiyetin yasalar tarafından korunması ne kadar etkili ise ülkenin skoru da o kadar yüksek olacaktır. Aynı şekilde hükümet tarafından yapılan kamulaştırma ne kadar fazla ise ülkenin skoru o kadar düşük olacaktır. Mülkiyet hakları faktörü aynı ağırlıktaki 5 alt faktörün ortalaması ile elde edilmektedir. Bu alt faktörler; fiziksel mülkiyet hakları, zihinsel mülkiyet hakları, yatırımcı korunmasının gücü, kamulaştırma riski ve arazi idaresinin kalitesi olarak isimlendirilmiştir. Bu alt faktörlerin her biri karşılaştırma amacıyla aşağıdaki denklem kullanılarak normalize edilmiş veri setlerinden türetilmektedir.

$$Alt Faktör Skoru(i) = 100 * \frac{Alt Faktör(Max) - Alt Faktör(i)}{Alt Faktör(Max) - Alt Faktör(Min)} \quad (23)$$

Alt Faktör(i) : (i) ülkesi için orijinal skoru,

Alt Faktör(Max) : Veri setinin en yüksek değerini,

Alt Faktör(Min) : Veri setinin en düşük değerini,

Alt Faktör Skoru(i) : (i) ülkesi için alt faktör skorunun temsil etmektedir.

3.1.1.1.2. Hükümetin Bütünlüğü

Yolsuzluk, güvensizlik ve ekonomik ilişkilerde baskı ortamı oluşturmamasından kaynaklı ekonomik özgürlüğü sarsmaktadır. Rüşvet, haraç, adam kayırma, zimmetine para geçirme gibi uygulamalarla hükümet kurumlarının yozlaşması en büyük endişe

kaynağıdır. Bu tarz uygulamalar hükümet bütünlüğünü bozar, hükümet bütünlüğünün bozulması ekonomik aktivitelerin maliyetini arttırdığı için toplumdaki güven ve ekonomik güç düşer. Bu faktör aynı ağırlıktaki 6 alt faktörün ortalaması ile elde edilmektedir. Bu faktörler, halkın siyasetçilere güveni, çarpık ödemeler ve rüşvetler, hükümet politikalarında şeffaflık, yolsuzluğun yokluğu, yolsuzluk algısı, hükümet ve kamu hizmeti şeffaflığı olarak isimlendirilmiştir. Alt faktörlerin hesaplama biçimi bir üstteki faktör ile aynı yapılmaktadır.

3.1.1.1.3. Yargı Etkinliği

Bütün vatandaşların haklarını yasadışı insanların, hükümetlerin ve özel grupların eylemlerinden korumak için iyi çalışan bir yasal çerçeve şarttır. Yargı etkinliği, yasalara tam olarak uyulmasını ve ihlallere karşı uygun yasal işlemlerin yapılmasını sağlamak için verimli ve adil yargı sistemleri gerektirir. Yargı etkinliği faktörü aynı ağırlıktaki 3 alt faktörün ortalaması ile elde edilmektedir. Bu alt faktörler, yargı bağımsızlığı, yargı sürecinin kalitesi ve hükümet yetkililerinin kararlarında tarafsızlık olarak isimlendirilmiştir. Alt faktörlerin hesaplama biçimi bir üstteki faktör ile aynı yapılmaktadır.

3.1.1.2. Devlet Büyüklüğü

Devlet Büyüklüğü üç faktörün ortalamasının alınması ile elde edilmektedir, bu faktörler Kamu Harcamaları, Vergi Yüğü, ve Mali Sağlık olarak adlandırılmıştır.

3.1.1.2.1. Kamu Harcamaları

Kamu harcamaları faktörü, devlet tarafından yapılan tüketimi ve çeşitli yetkilendirme programlarıyla ilgili tüm transfer ödemelerini içeren devlet harcamalarının yüklediği yükü ölçmeyi amaçlar. Bu faktör için optimal bir kamu harcaması belirlenmemiştir çünkü ideal bir kamu harcaması miktarı coğrafi, kültürel ve ekonomik gelişmişlik gibi birçok değişken açısından ülkeden ülkeye değişiklik göstermektedir. Yapılan araştırmalar sonucunda aşırı kamu harcamalarının kronik

bütçe eksikliğine ve kamu borçlarının birikmesine neden olduğu ortaya konmuştur. Hesaplama yöntemi kamu harcamalarının varsayılan değerini 0 olarak alır bunun sonucunda gelişmemiş ülkeler özellikle hükümet kapasitesi küçük olan ülkeler yapay olarak yüksek puan almış olurlar. Yine de gelişmemiş ülkeler bu değişkenden yüksek puan alacak olsalar da diğer değişkenlerden düşük puan almaları muhtemeldir.

Kamu harcamaları ekonomik özgürlük üzerinde büyük bir etkiye sahiptir ancak aynı şekilde etki eden bir birçok değişken bulunmaktadır. Kamu harcamaları skorunu belirlenen ölçek doğrusal değildir yani kamu harcamaları 0'a yakın olan ülkeler çok az cezalandırılırken, kamu harcamaları GSYİH' nın 30%'unu geçen ülkeler çok daha fazla ceza alırlar. Kamu harcamalarını hesaplamak için kullanılan denklem aşağıda verilmiştir.

$$KH_i = 100 - a * (Harcamalar_i)^2 \quad (24)$$

KH_i : i ülkesindeki kamu harcamaları skorunu,

$Harcamalar_i$: son üç yıl için GSYİH' nın yüzdesi olarak ortalama toplam kamu harcamasını temsil eder.

3.1.1.2.2. Vergi Yükü

Vergi yükü, hem kişisel hem de kurumsal gelir üzerindeki marjinal vergi oranlarını ve gayri safi yurtiçi hasılanın (GSYİH) yüzdesi olarak genel vergilendirme düzeyini yansıtan bileşik bir ölçüdür. Bileşik puan üç alt faktörden elde edilir, bireysel gelir üzerindeki en yüksek marjinal vergi oranı, kurumların geliri üzerindeki en yüksek marjinal vergi oranı, toplam vergi yükünün GSYİH üzerindeki yüzdesi. Bu alt faktörlerin hepsi eşit olarak ağırlıklandırılmıştır ve üç faktörün toplamı vergi yükü faktörünün skorunu vermektedir yani bir alt faktörün 0 olması sonucunda diğer iki faktörden toplam en yüksek 67 puan alınabilmektedir. Üç faktörden ayrı ayrı alınan sonuçlar aşağıdaki denklem ile 100 puanlık ölçeğe uygun hale getirilmiştir.

$$Vergi Y\ddot{u}k\ddot{u}_{ij} = 100 - a * (Fakt\ddot{o}r)_{ij}^2 \quad (25)$$

Vergi Y\ddot{u}k\ddot{u}_{ij} : i \ddot{u}lkesindeki j fakt\ddot{o}r\ddot{u} i\ddot{c}in vergi y\ddot{u}k\ddot{u},

Fakt\ddot{o}r_{ij} : i \ddot{u}lkesindeki j fakt\ddot{o}r\ddot{u} i\ddot{c}in deęer,

a : 0.03 deęerini alan bir sabit olarak tanımlanmıřtır.

3.1.1.2.3. Mali Saęlık

H\ddot{u}k\ddot{u}metin zayıf b\ddot{u}t\ddot{c}e y\ddot{o}netiminden kaynaklanan artan bor\ddot{c} y\ddot{u}k\ddot{u} ve a\ddot{c}ıklar \ddot{u}lkenin mali saęlıęının ařınmasına yol a\ddot{c}ar. K\ddot{o}t\ddot{u}leřen mali saęlık, makroekonomik istikrarsızlık ve ekonomik belirsizlięe yol a\ddot{c}ar. Kamu borcu, b\ddot{u}t\ddot{c}e a\ddot{c}ıklarının zamanla birikmesi ile olur. Teoride, kamu harcamalarının bor\ddot{c} finansmanı, \ddot{u}retken yatırımlara ve nihayetinde ekonomik b\ddot{u}y\ddot{u}meye olumlu katkı saęlayabilir ancak sadece kamu t\ddot{u}ketimini veya transfer \ddot{o}demelerini artıran harcamalar kalıcı kamu borcuna sebep olup \ddot{u}retken yatırımların b\ddot{u}y\ddot{u}mesini baltalar ve ekonomik durgunluęa yol a\ddot{c}ar. Mali saęlık fakt\ddot{o}r\ddot{u} iki alt fakt\ddot{o}r\ddot{u}n hesaplanması ile elde edilir ancak aęırlıkları eřit daęılmamaktadır, 80% aęırlıkla son \ddot{u}\ddot{c} yılın ortalama a\ddot{c}ıęının GSY\ddot{I}H' ya oranı hesaplanırken, 20% aęırlıkla borcun GSY\ddot{I}H' ya oranı hesaplamaya dahil edilmektedir.

$$Alt Fakt\ddot{o}r Skoru_i = 100 - a * (Alt Fakt\ddot{o}r)_i^2 \quad (26)$$

Alt Fakt\ddot{o}r Skoru_i : i \ddot{u}lkesindeki a\ddot{c}ık veya bor\ddot{c} skorunu,

Alt Fakt\ddot{o}r_i : i \ddot{u}lkesindeki GSY\ddot{I}H' nın deęerini,

a : a\ddot{c}ık i\ddot{c}in 2, bor\ddot{c} i\ddot{c}in 0.01 deęerini alan bir sabit olarak tanımlanmıřtır.

Genel kamu a\ddot{c}ıęı ve bor\ddot{c}ları \ddot{c}oęu \ddot{u}lke i\ddot{c}in her seviyede ulařılabilir olsa da bazı \ddot{u}lkelerdeki veri eksikliklerinden kaynaklı merkezi kamu harcamaları bu hesaplama i\ddot{c}in kullanılmıřtır.

3.1.1.3. Mevzuat Verimliliği

Mevzuat verimliliği üç faktörün ortalamasının alınması ile elde edilmektedir, bu faktörler İş Özgürlüğü, İşgücü Özgürlüğü ve Parasal Özgürlük olarak adlandırılmıştır.

3.1.1.3.1. İş Özgürlüğü

İş özgürlüğü faktörü, düzenleyici ve altyapı ortamlarının işletmelerin verimli çalışmasını ne ölçüde kısıtladığını ölçer. Bu faktördeki nicel puan, bir işletmeyi başlatma, işletme ve kapatma kolaylığını etkileyen bir dizi faktörden elde edilir. İş özgürlüğü skoru 0 ile 100 arasında puan alır 100 puan en özgür iş çevresini temsil etmektedir. Skor eşit ağırlıklı 13 alt faktörden oluşmaktadır ve veriler Dünya Bankası'ndan elde edilmiştir. Alt faktörler; işletmeyi kurma(prosedür sayısı), işletmeyi kurma(gün sayısı), işletmeyi kurma(maliyet), işletmeyi kurma(asgari sermaye), lisans alma(prosedür sayısı), lisans alma(gün sayısı)², lisans alma(maliyet), işletme kapatma(yıl sayısı), işletme kapatma(maliyet), işletme kapatma(geri dönüşüm oranı), elektrik alma(prosedür sayısı), elektrik alma(gün sayısı) ve elektrik alma(maliyet)³ olarak belirlenmiştir. Her bir alt faktör aşağıdaki denklemi kullanarak 0 ile 100 arası bir skor elde edilmiştir.

$$Alt Faktör Skoru = 50 * \left(\frac{Alt Faktör_{ortalama}}{Alt Faktör_i} \right) \quad (27)$$

3.1.1.3.2. İşgücü Özgürlüğü

Çalışma özgürlüğü faktörü, bir ülkenin işgücü piyasasının yasal ve düzenleyici çerçevesinin çeşitli yönlerini, asgari ücretlerle ilgili düzenlemeleri, işten çıkarmaları önleyen yasaları, kıdem tazminatı gerekliliklerini, işe alma ve çalışılan saatlerle ilgili ölçülebilir düzenleyici kısıtlamaları ve ayrıca işgücünü dikkate alan nicel bir ölçüdür. 7 alt faktörden oluşmaktadır ve alt faktörler eşit ağırlıklandırılmıştır. Alt faktörler, asgari ücretin işçi başına ortalama katma değere oranı, ek işçi çalıştırmanın engellenmesi, çalışma saati esnekliği, gereksiz çalışanları işten çıkarmanın zorluğu, yasal olarak zorunlu ihbar süresi, zorunlu kıdem tazminatı ve iş gücüne katılım oranı

olarak adlandırılmıştır. Faktörler aşağıdaki denklem ile 0 ile 100 arası bir skora dönüştürülmektedir.

$$Alt Faktör Skoru = 50 * \left(\frac{Alt Faktör_{ortalama}}{Alt Faktör_i} \right) \quad (28)$$

3.1.1.3.3. Parasal Özgürlük

Parasal Özgürlük, fiyatlardaki istikrarı ve fiyat kontrolünü birleştirerek değerlendirir. Hem enflasyon hem de fiyat kontrolleri piyasa faaliyetlerini bozar. Mikro ekonomik müdahale olmaksızın fiyat istikrarı, serbest piyasa için ideal durumdur. Parasal Özgürlük iki alt faktör ile ölçümlenir bu alt faktörler, son üç yılın enflasyonunun ağırlıklı ortalaması ve fiyat kontrolü olarak belirlenmiştir.

$$Enflasyon A.O. = \theta_1 Enflasyon_{it} + \theta_2 Enflasyon_{it-1} + \theta_3 Enflasyon_{it-2} \quad (29)$$

$$Parasal Özgürlük = 100 - a * \sqrt{Enflasyon A.O._i} - Fiyat Kontrolü_i \quad (30)$$

Üstte yer alan denklem son üç yılın enflasyonunun ağırlıklı ortalamasını verirken, bir altında yer alan denklem nihai parasal özgürlük skorunu vermektedir. Fiyat kontrolü değişkeni 0 ile 20 arasında değer alan bir cezayı temsil etmektedir.

3.1.1.4. Serbest Piyasa

Serbest piyasa üç faktörün ortalamasının alınması ile elde edilmektedir, bu faktörler Ticaret Özgürlüğü, Yatırım Özgürlüğü ve Finansal Özgürlük olarak adlandırılmıştır.

3.1.1.4.1. Ticaret Özgürlüğü

Ticaret özgürlüğü, mal ve hizmetlerin ithalat ve ihracatını etkileyen gümrüğe vergisi içeren ve gümrük vergisi içermeyen engellerin kapsamının bileşik bir ölçüsüdür. Ticaret özgürlüğü skoru iki girdiyi esas alır, ticaret ağırlıklı ortalama gümrük vergisi oranı ve gümrük vergisi içermeyen mallardaki engeller. Bir ülkeye giren farklı ithalatlar, farklı gümrük vergileri ile karşı karşıya kalabilir. Ağırlıklı

ortalama gümrük vergisi, her bir mal için ithalat payına dayalı olarak her gümrük vergisi için ağırlıklar kullanır. Ağırlıklı ortalama gümrük vergisi, tamamen nicel bir ölçüdür ve aşağıdaki denklem kullanılarak ticaret özgürlüğü skoru hesaplanır.

$$Ticaret Özgürlüğü_i = 100 * \frac{GV_{max} - GV_i}{GV_{max} - GV_{min}} - GVIME_i \quad (31)$$

GV_i : i ülkesi için gümrük vergisi,

GV_{max} :Maksimum gümrük vergisi,

GV_{min} :Minimum gümrük vergisi,

$GVIME_i$:i ülkesi için gümrük vergisi içermeyen mallardaki engeller olarak tanımlanmıştır.

3.1.1.4.2. Yatırım Özgürlüğü

Ekonomik olarak özgür bir ülkede, yatırımların akışı üzerinde hiçbir kısıtlama olmayacaktır. Bireylerin ve firmaların kaynaklarını kısıtlama olmaksızın hem ülke içinde hem de ülke sınırlarının ötesinde belirli faaliyetlerden geçirmelerine izin verilecektir. Böyle ideal bir ülke, yatırım özgürlüğü faktöründen 100 puan alacaktır. Ancak birçok ülkenin yatırımlar üzerinde çeşitli kısıtlamaları bulunmaktadır. Bazı ülkelerin yurtiçi ve yurtdışı yatırımlarının üzerindeki farklı kuralları varken bazılarında dövizle ulaşmak bile kısıtlıdır. Bazıları ödemeler, transferler ve sermaye piyasası işlemleri üzerinde kısıtlamalar getirirken bazılarında ise belirli endüstriler yabancı yatırımlara kapalıdır. Endeks, kısaca yatırıma uygulanan çeşitli düzenleyici kısıtlamaları değerlendirir. Bu faktörde varsayılan skor 100'dür, bir ülkenin yatırım rejiminde bulunan kısıtlamaların her biri için ideal 100 puandan düşülür.

3.1.1.4.3. Finansal Özgürlük

Finansal özgürlük, bankacılık verimliliğinin bir göstergesi olduğu kadar, hükümet kontrolünden ve finans sektörüne müdahaleden bağımsızlığın bir ölçüsüdür. Sigorta şirketleri ve sermaye piyasaları finansal hizmetlerin ve bankaların kamu mülkiyetinde olması rekabeti azaltır ve genel olarak krediye ulaşım seviyesini azaltır.

İdeal bir bankacılık ve finansal ortam minimum devlet müdahalesi, bağımsız bir merkez bankası, finansal kurumlarda regülasyonlar ve dolandırıcılığı engelleme ile elde edilir. Faktör bir ekonominin finansal özgürlüğünü beş geniş alanı değerlendirerek belirler. Alanlar, mali hizmetlere ilişkin hükümet düzenlemelerinin kapsamı, doğrudan ve dolaylı mülkiyet yoluyla bankalara ve diğer finansal firmalara devlet müdahalesinin derecesi, kredi tahsisinde hükümetin etkisi, finansal ve sermaye piyasası gelişiminin kapsamı ve dış rekabete açıklık olarak belirlenmiştir. Bu beş alanın, ekonomideki insanlar ve işletmeler için finansman fırsatlarına kolay ve etkili erişim sağlayan bir ekonominin genel finansal özgürlük düzeyini değerlendirdiği düşünülmektedir. Metodolojide belirlenen birçok kritere göre ideal 100 puandan kesintiler yapılarak bir ekonominin finansal özgürlüğüne 0 ila 100 arasında bir genel puan verilir.

3.1.2. Dünya Mutluluk Raporu Metodoloji

Dünya Mutluluk Raporu, 2012 yılının nisan ayından itibaren her yıl yayınlanan ülkelerdeki mutluluğu ölçmeyi amaçlayan bir rapordur. İlk rapor “Refah ve Mutluluk: Yeni Bir Ekonomik Paradigmanın Tanımlanması” konulu Birleşmiş Milletler yüksek düzeyli toplantısını desteklemek amacı ile hazırlanmıştır. Raporda toplumların mutluluğu ile ilgili küresel verileri ortaya koymuş, yaşam kalitesinin güvenilir ve tutarlı bir biçimde bazı faktörlerle değerlendirilebileceğini göstermiştir. Raporun içeriğinde, belirli bölge ve ülkelerin mutluluk çerçevesinden değerlendirmeleri aynı zamanda refah biliminin detaylarına inen bazı bölümler yer almaktadır. Raporları yıllara göre değerlendirdiğimizde çoğunlukla benzer bir çerçevede yazılmış olsalar da bazı raporların odaklandığı noktalar değişiklik göstermektedir.

Rapordaki veriler “Gallup World Poll” isimli bir şirket tarafından anket yoluyla toplanmaktadır. Gallup 150’den fazla ülkeye, her yıl 15 yaşından büyük yaklaşık 1000 kişiye bu anketi uygulamaktadır. Gallup anketleri yapılırken katılımcılardan mümkün olan en iyi yaşam 10, mümkün olan en kötü yaşam ise 0 olacak şekilde 10 basamaklı bir merdiven düşünmelerini ve kendilerini bu merdivenin kaçınıcı basamağında gördüklerini sorar. Bu merdiven “Cantril Merdiveni” olarak adlandırılır. Bu anket verileri Dünya Mutluluk Konseyi tarafından başka veri kaynakları ile birleştirilerek

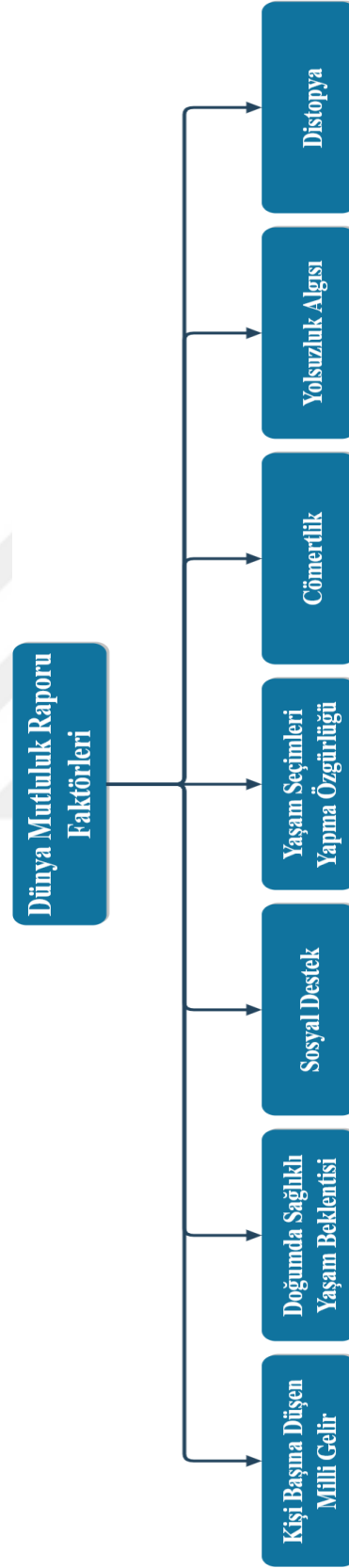
Dünya Mutluluk Raporu oluşturulur. Dünya Mutluluk Raporu birçok değişken ve veriden oluşmasına rağmen tek bir endekse indirgenir ve bu endeks “Mutluluk Endeksi” olarak adlandırılır.

Mutluluk endeksi hesaplanırken kişi başına gayrisafi yurtiçi hasıla, sosyal destek, doğumda sağlıklı yaşam beklentisi, yaşam seçimleri yapma özgürlüğü, yolsuzluk algısı ve cömertlik isimli mutluluk ile en çok ilişkili 6 faktör ve ek olarak “Distopya” adı verilen bir değişken hesaplamaya katılmaktadır. Mutluluk skoru bahsi geçen 6 faktör ve Distopya değişkeninin toplanması ile elde edilmektedir. Raporda sadece bu 6 faktöre yer verilmesinin birkaç nedeni vardır. Bu sebeplerden biri literatürde ulusları karşılaştırırken en çok kullanılan faktörler olmaları, bir diğeri ise bu faktörlerin hemen hemen her ülke için elde edilebiliyor olması. Örneğin işsizlik ve eşitsizlik gibi bazı önemli faktörlerin hesaplamada yer almamasının sebebi verilerin elde edilememesinden kaynaklanmaktadır. Kullanılan faktörler seçilirken belli bir oranda korelasyon göstermeleri amaçlanmıştır, bu korelasyon bazı durumlarda iki yönlüde olabilmektedir. Örneğin 2013 yılı raporunda da belirtildiği gibi daha sağlıklı insanlar genel olarak daha mutludur ancak aynı zamanda daha mutlu insanlar da genel olarak daha sağlıklıdır.

Hesaplamalar yapılırken güncel bir ölçüt elde etmek ve zamanla endekste değişen skorları görmek amacıyla önceki yılların skorları kullanılmaktadır. Örneğin 2021 yılı için yapılan hesaplamada rastgele örnekleme hatasını azaltmak için 2018, 2019 ve 2020 yılı verileri de kullanılmıştır. Eğer ülkenin her yıl anket yaptığını düşünürsek örneklem boyutu yaklaşık 3000 olacaktır ancak her yıl anket yapmayan ülkeler olmasından kaynaklı o ülkelerin örneklem boyutu 3000’in altında kalacaktır. Yapılan araştırmalara göre 2000-3000 verilik bir örneklemin bir ülkenin mutluluk skorunun belirlemede yeterli olduğu raporda belirtilmiştir.

Mutluluk skoru ya da diğer ismiyle sübjektif mutluluk 6 ana faktör ve 1 distopya değişkeninin toplanması ile elde edilmektedir. Faktörlerle ilgili detaylı bilgi Dünya Mutluluk Raporu resmi sitedinden elde edilmiştir. Verilerin kaynağı genellikle Gallup World Poll isimli şirketin anketlerden elde ettiği veriler olsa da başka organizasyonların elde ettiği verilerde yıllık raporda kullanılmaktadır.

Şekil 9: Dünya Mutluluk Raporu Değişkenleri



3.1.2.1. Kişi Başına Gayrisafi Yurtiçi Hasıla

Satın alma gücü paritesinde ismi geçen kişi başına gayrisafi yurtiçi hasıla Dünya Kalkınma Göstergeleri verilerinden elde edilmektedir. Verilerin eksik olması durumunda bir önceki yılın verilerinden yararlanılıp ileriye dönük tahminler kullanılmaktadır. Tahminlerde Dünya Bankası Küresel Ekonomik Beklentiler verilerinden yararlanılmaktadır. Verisetinde bu faktör logaritması alınmış biçimde bulunmaktadır.

3.1.2.2. Doğumda Sağlıklı Yaşam Beklentisi

Doğumda sağlıklı yaşam beklentisi, Dünya Sağlık Örgütü verilerine dayanmaktadır. Veri kaynağında veriler 2016 yılına kadar bulunduğu için sonuçlarda ekleme ve genişletme yapılmıştır.

3.1.2.3. Sosyal Destek

Sosyal destek metodolojide zor zamanlarda güvenilecek birileri olarak açıklanmaktadır. Bu faktörün değerleri anket verilerinden oluşmaktadır. Katılımcılara cevabı evet(1) ya da hayır(0) olacak şekilde “Zor bir durumda olsanız, size her durumda yardım edecek bir yakınınız veya arkadaşınız var mı?” sorusu sorularak elde edilir.

3.1.2.4. Yaşam Seçimleri Yapma Özgürlüğü

Yaşam seçimleri yapma özgürlüğü Gallup World Poll tarafından sorulan “Hayatınla ilgili yaptığın seçimlerde kendini yeterince özgür hissediyor musun?” sorusunun ulusal ortalaması alınarak hesaplanır.

3.1.2.5. Cömertlik

Cömertlik, Gallup World Poll’ un sorduğu “Geçen ay bir hayır kurumuna bağış yaptınız mı?” sorusunun cevabından gelen sonuçlardan elde edilir.

3.1.2.6. Yolsuzluk Algısı

Yolsuzluk algısı iki soru ile ölçülmektedir. Birincisi “Hükümetinizde yüksek düzeyde bir yolsuzluk bulunmakta mıdır?” , ikincisi ise “İşletmelerde yüksek düzeyde bir yolsuzluk bulunmakta mıdır?”. İlk sorunun cevabı alınamadığı durumlarda ikinci sorunun cevapları kullanılmaktadır.

3.1.2.7. Distopya

Dünya mutluluk raporuna göre Distopya dünyanın en mutsuz insanlarını barındıran hayali bir ülkedir. Ülkeler distopya ülkesinden ne kadar uzaksa o kadar yüksek puan alırlar.

3.2. VERİ SETLERİ LİTERATÜR TARAMASI

Aşağıda araştırma için kullanılan Ekonomik Özgürlük Endeksi ve Dünya Mutluluk Raporu için yapılan literatür taramasına yer verilmiştir.

3.2.1. Ekonomik Özgürlük Endeksi Literatür Taraması

Ekonomik özgürlük endeksi literatür taraması özet olarak aşağıdaki tabloda daha detaylı olarak tablodan sonra paylaşılmıştır.

Tablo 1: Ekonomik Özgürlük Endeksi Literatür Taraması Özeti

Kim	Ne Zaman	Hangi Değişkenler Kullanılmış	Yöntem
Kangallı ve diğerleri	2014	İş Özgürlüğü, Ticaret Özgürlüğü, Mali Özgürlük, Kamu Harcaması, Parasal Özgürlük, Yatırım Özgürlüğü, Finansal Özgürlük, Mülkiyet Hakları, Yolsuzlukla Mücadele, İşgücü Özgürlüğü	Ward, K-Ortalamalar
Şahin	2017	Mülkiyet Hakları, Yolsuzlukla Mücadele, Mali Özgürlük, Kamu Harcaması, İş Özgürlüğü, İşgücü Özgürlüğü, Parasal Özgürlük, Ticaret Özgürlüğü, Yatırım Özgürlüğü Ve Finansal Özgürlük Endeksi	Ward
Kılcı	2019	Ekonomik Özgürlük Endeksi	Fourier KPSS Testi, Fourier Granger Nedensellik Testi
Güliden	2019	Ekonomik Özgürlük Endeksi'ni oluşturan 12 faktör	Ward, K-Ortalamalar
Taşcı	2019	İş Özgürlüğü, Ticaret özgürlüğü, İşgücü Özgürlüğü, Yatırım Özgürlüğü, Parasal Özgürlük, Finansal özgürlük	Gri İlişkiler Analiz, K-Ortalamalar
Kutbay	2020	Hukukun Üstünlüğü, Mevzuat Verimliliği, Serbest Piyasa, Devlet Büyüklüğü	Panel Veri Analizi
Özcan ve Akar	2020	Ekonomik Özgürlük Endeksi	Granger Nedensellik Testi
Deveci	2020	Ekonomik Özgürlük Endeksi	K-En Yakın Komşu
Altın	2020	Ekonomik Özgürlük Endeksi'ni oluşturan 12 faktör	TOPSIS, MAUT, CETOPSIS, CEMAUT
Coşkuner ve diğerleri	2021	Ekonomik Özgürlük Endeksi'ni oluşturan 12 faktör	Bulanık K-Ortalamalar

“Heritage Foundation” tarafından 1995 yılından itibaren “Ekonomik Özgürlük Endeksi” isimli bir endeks yayınlanmaktadır. Yayınlanmaya başlanmasından bu zamana kadar birçok değişken barındırmasına rağmen hukukun üstünlüğü, devlet büyüklüğü, mevzuat verimliliği ve serbest piyasa gibi dört ana faktörle de açıklanabileceğini göstermiştir. Ekonomik özgürlüğü yüksek olan ülkeler daha fazla refah içinde olan ülkelerdir varsayımı ile yola çıkılmış bir endekstir. Endeks çeşitli ekonomik ve sosyal hedefler ile ekonomik özgürlük arasındaki pozitif yönlü ilişkiyi belgelemektedir. Ekonomik özgürlük daha sağlıklı ve temiz toplumlar, kişi başına daha fazla refah ve yoksulluğun ortadan kaldırılması ile kuvvetli bir biçimde ilişkilidir. Özgürlük kendi başına çok önemli bir değerdir, ekonomik özgürlük ise toplumun ilerlemesinin ölçülebildiği çeşitli ekonomik ve sosyal başarıları mümkün kılan zenginliği üretmek için gereken olmazsa olmaz bir mekanizmadır. Ekonomik özgürlük endeksi on iki nicel ve nitel faktörün dört ana faktörü oluşturduğu bir endekstir.

Liberal ekonomi politikalarının 1980’li yıllardan itibaren gündeme gelmesiyle ekonomik özgürlük kavramı ekonomi ve hukuk arasındaki ilişkini vazgeçilmezi haline gelmiştir. Klasik liberal görüşe açısından ekonomik özgürlük kavramı ekonomik faaliyetlerin baskı ve zorlamadan arındırılmış bir şekilde uygulanmasını ifade etmektedir(Hickerson, 1984: 91).

Ekonomik özgürlük kavramını daha detaylı incelemek ve ülkeler arası karşılaştırma yapmak amacıyla bir ölçüte ihtiyaç duyulmuştur. Temel faktörlerden yola çıkarak geliştirilen ekonomik özgürlük endeksi birbiriyle yakın sonuçlar üreten üç kurum tarafından yayınlanmaktadır(Acar,2010: 10-11). Bu endeksler; Fraser Enstitüsü tarafından yayınlanan Dünya Ekonomik Özgürlüğü Endeksi, Heritage Foundation tarafından hazırlanan Ekonomik Özgürlük Endeksi ve Freedom House tarafından hazırlanan Ekonomik Özgürlük Endeksi’dir. Bu çalışmada Heritage Foundation tarafından hazırlanan Ekonomik Özgürlük Endeksi verilerinden yararlanılmıştır.

Kangallı ve diğerleri (2014) “OECD ülkelerinde ekonomik özgürlük: bir kümeleme analizi” isimli çalışmalarında, OECD üyesi ülkelerin 2011 yılına ait Ekonomik Özgürlük Endeksi verilerinden yararlanmışlardır. Çalışmalarında kümeleme analizi uygulamış, Ward ve K-Ortalamlar yöntemlerinden faydalanmışlardır. Uzaklık ölçüsü olarak karesi alınmış öklid uzaklığı kullanılmıştır.

2011 yılında endeksi oluşturan 10 adet değişken olmasından kaynaklı çalışmada 10 adet değişken kullanılmıştır. Yapılan analizler sonucunda üç kümeye ulaşmış ve bu kümeleri yorumlamışlardır. Üç kümeye karar vermelerindeki amaç ülkelerin gelişmiş, gelişmekte olan ve geri kalmış ülkeler olarak nitelendirilebileceğini düşünmeleridir.

Şahin (2017) “Kümeleme Analizi İle Doğu Avrupa Ülkelerinin Ekonomik Özgürlükler Açısından Değerlendirilmesi” isimli çalışmasında, Ekonomik Özgürlük Endeksi 2015 yılı verilerini kullanarak Doğu Avrupa ülkelerini sınıflandırmaktadır. Çalışmada kümeleme analizinden faydalanılmıştır. Kümeleme analizi olarak hiyerarşik kümeleme analizi yöntemlerinden Ward yöntemi kullanılmıştır. Çalışmada 23 ülke 10 adet değişken kullanılarak analiz edilmiştir. Hiyerarşik kümeleme analizi sonucunda ülkeler 4 kümeye ayrılmışlardır ve Türkiye üçüncü kümede yer almıştır.

Kılıcı (2019) “Analysis Of The Relationship Between Economic Freedom Index And Stock Market Indices; Evidence From Turkey” isimli çalışmasında, 1995-2018 yılları arasındaki ekonomik özgürlük endeksi verileri ile BİST 100 arasındaki ilişkiyi araştırmıştır. Araştırmasında Fourier KPSS Testi ve Fourier Granger Nedensellik Testi’ni kullanmıştır. Araştırma sonucunda Türkiye için ekonomik özgürlük endeksi ve BİST 100 arasında kısa dönemde bir ilişki bulamamıştır.

Güliden (2019) “OECD Ülkelerinin Ekonomik Özgürlüklerine Göre Kümeleme Analizi İle Sınıflandırılması” isimli çalışmasında, Ekonomik Özgürlük Endeksi 2015-2019 yılları arasındaki verilerini kullanarak bir kümeleme analizi gerçekleştirmiştir. Çalışmada küme sayısını belirlemek için Ward yöntemi daha sonra hiyerarşik olmayan kümeleme yöntemlerinden k-ortalamlar yöntemi kullanılmıştır. Çalışmada küme sayısı beş olarak belirlenmiştir. Her yıl için kümeler oluşturulup yorumlanmıştır.

Taşcı (2019) “OECD Ülkelerinin Ekonomik Özgürlük Göstergelerinin K-Ortalamlar Kümeleme Yöntemi Ve Gri İlişkisel Yöntem İle Analizi” isimli çalışmasında, Ekonomik Özgürlük Endeksi faktörlerinden ekonomik özgürlük ile en çok ilgili olanlar alınarak 2015-2019 yılları arası veriler çalışma için kullanılmıştır. Çalışmada Çok Kriterli Karar Verme yöntemlerinden Gri İlişkiler Analiz yöntemi ve kümeleme yöntemlerinden K-Ortalamlar yöntemi kullanılmıştır. Çalışmanın amacı iki yöntemin sonuçlarının birbiri ile uyuşup uyuşmadığını belirlemek ve bu yöntemlerin bu tür çalışmalarda uygulanabileceğini göstermektir. Gri İlişkisel Analiz yöntemi ile seçilen kriterleri kullanarak ülkeler en iyiden en kötüye sıralanıp, dönemsel

farklılıklar incelenmiştir. K-Ortalamalar yöntemi ile de aynı kriterler kullanılarak ülkeler iki kümeye ayrılmıştır. Çalışma sonunda sıralamada birbiri ile yakın olan ülkelerin aynı kümelerde, sıralamada birbirine uzak olan ülkelerin farklı kümelerde yer alıp almadığı incelenmiştir

Kutbay (2020) “Ekonomik Özgürlük Ve Vergi Gelir Performansı Arasındaki İlişki: Yükselen Piyasa Ekonomileri İçin Panel Veri Analizi” isimli çalışmasında bazı ülkelerde ekonomik özgürlük endeksinin alt faktörlerinin, vergi gelirleri üzerindeki etkisini incelemiştir. Veri setleri arasındaki ilişki panel veri analizi ile incelenmiştir. Tahminlemeler Driscoll-Kraay model tahmincisi ile yapılmıştır. Analizler sonucunda Ekonomik özgürlük endeksi alt bileşenlerinden hukukun üstünlüğü, mevzuat verimliliği ve serbest piyasadaki artışların vergi gelirlerini artırdığı, devlet büyüklüğündeki artışların ise vergi gelirlerini azalttığı tespit edilmiştir.

Özcan ve Akar (2020) “E7 Ülkelerinde Ekonomik Özgürlük Ve Doğrudan Yabancı Sermaye Yatırımları: Panel Veri Analizi” isimli çalışmalarında, E7 ülkeleri olarak adlandırılan Çin, Hindistan, Brezilya, Rusya, Endonezya, Meksika ve Türkiye için doğrudan yabancı yatırımlar ve ekonomik özgürlük arasındaki ilişkiyi incelemektedir. Çalışmada panel veri analizi için 1996-2018 yılı arasındaki verilerden yararlanılmıştır. Çalışmada sonuçlar Dumitrescu ve Hurlin (2012) Granger nedensellik testi sonuçlarından elde edilmiştir ancak doğrudan yabancı yatırımlar ve ekonomik özgürlük arasında bir ilişki tespit edilememiştir.

Deveci (2020) “K-En Yakın Komşu Algoritması ile ekonomik özgürlük endeksinin tahmin edilmesi: OECD ülkelerinde örnek uygulama” isimli çalışmasında, 36 OECD ülkesinin 2017 yılı ekonomik özgürlük endeksi verilerinden yararlanmıştır. Çalışmasında ekonomik özgürlük endeksi verilerine ek olarak 2016 yılı ithalat, ihracat, işsizlik ve enflasyon makroekonomik değişkenlerini dahil etmiştir. Çalışmada K-En Yakın Komşu algoritması kullanılmıştır. Çalışmanın sonucunda 7 ülke için tüm k değerlerinin doğru tahmin edildiği, iki ülke için tüm k değerlerinin yanlış tahmin edildiği saptanmıştır.

Altın (2020) “Ekonomik Özgürlük Endeksinin Çok Kriterli Karar Verme Yöntemleriyle Analizi” isimli çalışmasında, ekonomik özgürlük endeksinin oluşturan 12 faktör için çok kriterli karar verme yöntemleri kullanmıştır. Ekonomik özgürlük sıralama sonuçları hesaplanmasında TOPSIS VE MAUT yöntemleri kullanılmıştır.

Sıralama sonuçları hesaplandıktan sonra CETOPSIS VE CEMAUT yöntemlerinden elde edilen sonuçlar ile karşılaştırılmıştır. Karşılaştırma için Spearman Korelasyon yaklaşımı kullanılmıştır. Çalışmanın sonucunda birkaç bulgu elde edilmiştir, CRITIC VE ENTROPY yöntemleri ile yapılan TOPSIS sıralaması arasında tam korelasyon, CETOPSIS ve CEMAUT sıralaması ve ekonomik özgürlük endeksi sıralama sonuçları arasında pozitif yönlü kuvvetli bir ilişki vardır.

Coşkuner ve diğerleri (2021) “Ülkelerin Girişimcilik Ve Ekonomik Özgürlük Özelliklerinin Karşılaştırılması: Bulanık K-Ortalamlar Yöntemi İle Örnek Bir Uygulama” çalışmalarında, Birleşmiş Milletler’ in ülkeleri karşılaştırmasında kullandığı faktörlerden yola çıkarak iş yapma kolaylığı endeksinin, ekonomik özgürlük endeksi başlığı altında yer aldığını bahsetmişlerdir. Aynı zamanda bu iki endeksi etkileyen çok fazla faktör olduğunu ve endekslerin faktörlerden hangi yönde ve derecede etkilendiğinin örneklem grubuna göre değiştiği tespit edilmiştir. Çalışmada ülkelerin finansal serbestlik ve girişimcilik faktörleri baz alınarak kümeleme analizi yapılmıştır. Yöntem olarak Bulanık K-Ortalamlar analizi 159 farklı ülke için yapılmıştır. Ülkeler önce beş kümeye ayrılmış daha sonra bu kümeler nitelendirilmiştir. Çalışmanın sonucunda, iş yapma kolaylığı ile ekonomik özgürlükler arasında aynı zamanda bu iki endeksin gelir grupları arasında pozitif yönlü anlamlı bir ilişki olduğuna ve finansal serbestlik arttıkça iş yapma kolaylığının da arttığı bulgularına ulaşılmıştır.

3.2.2. Dünya Mutluluk Raporu Literatür Taraması

Dünya mutluluk raporu literatür taraması özet olarak aşağıdaki tabloda daha detaylı olarak tablodan sonra paylaşılmıştır.

Tablo 2: Dünya Mutluluk Raporu Literatür Taraması Özeti

Kim	Ne Zaman	Hangi Değişkenler Kullanılmış	Yöntem
Bradshaw ve Rees	2017	Log(Kişi Başına GSYİH), Sosyal Destek, Doğumda Sağlıklı Yaşam Beklentisi, Yaşam Seçimleri Yapma Özgürlüğü, Cömertlik, Yolsuzluk Algısı	Regresyon Analizi
De Neve ve Ward	2017	Mutluluk Skoru	Korelasyon Analizi
Okagbue ve diğerleri	2019	Mutluluk Skoru	Regresyon Analizi
Teague ve Locker	2019	Mutluluk Skoru	Regresyon Analizi
Georgescu ve diğerleri	2020	Mutluluk Skoru	Kümeleme Analizi, Lojistik Regresyon
Ulkhaq M. M.	2020	Log(Kişi Başına GSYİH), Sosyal Destek, Doğumda Sağlıklı Yaşam Beklentisi, Yaşam Seçimleri Yapma Özgürlüğü, Yolsuzluk Algısı	Dokuz adet kümeleme analizi algoritması
Kandemir O.	2021	Mutluluk Skoru	Lojistik Regresyon Analizi

Bradshaw ve Rees (2017) “Exploring national variations in child subjective well-being” isimli çalışmalarında, Children's Worlds anketine katılan 10 ve 12 yaş aralığındaki çocuklar için ortalama SLSS(Student's life satisfaction scale) puanlarını kullanarak ulusal mutluluk düzeylerindeki farklılıkları araştırmışlardır. Çalışmanın başlarında ülkenin ekonomik, sosyal, politik veya kültürel özelliklerinin göstergelerini kullanarak mutluluktaki çeşitliliğin çoğunu açıklamakta zorlanmışlardır. Bunun nedenini ülke sayısının sınırlı olması ve Romanya'nın yüksek aykırı değer, G. Kore'nin ise düşük aykırı olmasına bağlamışlardır. Ancak Dünya Mutluluk Raporu'nda bulunan sosyal destek değişkeni ile çocukların SLSS arasında bir ilişki bulmuşlardır ve modellerini Dünya Mutluluk Raporu'ndan değişkenler ekleyerek oluşturmaya karar vermişlerdir. Dünya Mutluluk Raporu'nda olduğu gibi, çocukların SLSS puanları ile

“Children's Worlds Survey”deki diğer göstergelerle, özellikle de arkadaşlık ve zaman kullanımını seçimi ile ilgili bazı güçlü ilişkiler bulmuşlardır. Tüm örneklemin SLSS puanlarındaki varyasyonun %57'sini açıklayan açıklayıcı bir SLSS modeli geliştirmişlerdir. Sonuç olarak özgürlüğün en önemli, ev ve arkadaşlık ortamının da en önemsiz değişkenler olduğuna karar vermişlerdir.

De Neve ve Ward (2017) “Does Work Make You Happy? Evidence from the World Happiness Report” isimli bu çalışmada Gallup Dünya Anketi verilerinden bireylerin çalıştıkları işlerden memnuniyetleri ile mutluluğun ilişkisi incelenmiştir. “Memnun” olduğunu bildiren katılımcıların yüzdesinin Kuzey ve Güney Amerika, Avrupa, Avustralya ve Yeni Zelanda'daki ülkelerde daha yüksek olduğu gözlemlenmiştir. Özellikle, Avusturya'da, katılımcıların% 95'inin işlerinden memnun olduğunu bildirerek en üst sırada yer almıştır. Avusturya'yı Norveç ve İzlanda izlemektedir. Gallup Dünya Anketi'ndeki bireylerin iş doyumu yanıtları ile yaşam değerlendirmesi arasında orta düzeyde bir ilişki (0.28) olduğunu görmüşlerdir. Bazı toplumların neden diğerlerinden daha fazla iş memnuniyeti sağladığını anlamak için Avrupa Sosyal Anketinden daha ayrıntılı verilere yönelmiş ve beklenebileceği gibi, iyi ücretli işlerde yaşayan insanların yaşamlarından ve işlerinden daha mutlu ve daha memnun olduklarını görmüşlerdir. Ancak insanların işlerinin diğer bazı yönleri de çeşitli mutluluk ölçülerini güçlü bir şekilde öngörmektedir. İş-yaşam dengesi insanların mutluluğunun güçlü bir belirleyicisi olarak ortaya çıkmaktadır. Diğer faktörler arasında iş çeşitliliği ve yeni şeyler öğrenme ihtiyacı ile çalışanın sahip olduğu bireysel özerklik düzeyi yer almaktadır. Dahası, iş güvenliği ve sosyal sermaye mutluluk ile pozitif olarak ilişkiliyken, sağlık ve güvenlik için risk içeren işler genellikle daha düşük refah seviyeleri ile ilişkilidir. İş tatmini açısından üst sıralarda yer alan ülkelerin, bu manevi iş özelliklerine hitap ederek daha kaliteli işler sunduğundan şüphelenmişlerdir.

Okagbue ve diğerleri (2019) “Significant Predictors of Henley Passport Index” adlı çalışmalarında, Henley Pasaport Endeksi (HPE) ile Yolsuzluk Algı Endeksi (YAE), Küresel Barış Endeksi (KBE) ve Dünya Mutluluk Raporu (DMR) üçlüsü arasında regresyon analizi yardımıyla anlamlı ilişki kuran bir sonuç sunmuştur. HPE'da üst sıralarda yer alan ülkelerin, YAE, KBE ve DMR' de eşit derecede üst sıralarda yer aldığını gözlemlemişlerdir. Bu araştırma, yüksek HPE'ye sahip ülkeler

için kamu sektörü yolsuzluğunu azaltmak, iç ve dış çatışmaları azaltmak ve vatandaşlarının yaşam kalitesini iyileştirmek için çalışmaları gerektiğini savunmaktadır.

Teague ve Locker (2019) “Analyzing the Relationship Between Happiness and European State-Mandated Parental Leave: Lessons For Human Resource Policymakers” isimli bu çalışmada, işyerinde sağlanan ebeveyn izninin miktarı ile doğum yapması ya da yakın zamanda bir çocuğu evlat edinmesi söz konusu olan bireylerin genel olarak toplum ve ülkelerin mutluluğuna etkisi arasında istatistiksel olarak anlamlı bir ilişki olup olmadığını araştırmıştır. Bu araştırmada Dünya Mutluluk Raporu verilerinden ve regresyon analizi yönteminden faydalanmışlardır. Avrupa'daki bazı bölgeler için ebeveyn izni ve mutluluk puanı arasındaki çelişkili sonuçlar, daha fazla araştırmaya ihtiyaç olduğunu göstermiştir. Ebeveynlik izni kişinin mutluluğuna katkıda bulunan bir faktör olsa da, tek katılımcı olarak görülmesinin yeterli olmayacağı sonucuna ulaşmışlardır.

Georgescu ve diğerleri (2020) “A Computational Approach to Economic Inequality, Happiness and Human Development” isimli çalışmalarında, 98 ülke için İnsani Gelişme Endeksi (HDI), kişi başına GSYİH, Dünya Mutluluk Endeksi, Gini Gelir ve Zenginlik eşitsizlikleri endekslerinin kategorik düzeylerinin ve yoksulluk oranının bağlantılarını incelemişlerdir. Çalışmalarında kümeleme analizi ile benzer özelliklere sahip dört ülke grubu belirlemişlerdir. Verilerdeki toplam değişimin %68,3'ünü açıklayarak 21-26 ülke büyüklüğündeki dört kümeyi elde etmek için K-ortalama kümeleme algoritması uygulamışlardır. Analiz, kümeler arasındaki önemli farklılıkları ortaya koyarken, kümeler içinde en büyük farklılıklara sahip faktörleri de ortaya çıkarmıştır. İkinci olarak, 2018 yılı için 98 ülkede insani gelişme endeksinin kategorilerinin tahmininde çok değişkenli lojistik regresyon (MLR) uygulanmaktadır. Lojistik regresyon modeli sonucunda 91.8% genel doğruluk elde etmişlerdir.

Ulkhaq (2020) “Clustering Countries According To The World Happiness Report” isimli çalışmasında, ülkeleri 2020 Dünya Mutluluk Raporu'na göre kümelemeyi amaçlamaktadır. Çalışmada dokuz kümeleme algoritması kullanılmıştır, kullanılan algoritmalar; k-ortalamlar, k-ortalamlar++, k-medoid, büyük uygulamalarda kümeleme, yakınlık yayılımı, spektral kümeleme, gürültülü kümeleme uygulamalarında yoğunluk tabanlı uzamsal kümeleme, kümelenmiş gruplama ve

ayırıcı analizdir(sırasıyla; k-means, k-means++, k-medoids, clustering large applications, affinity propagation, spectral clustering, density-based spatial clustering of applications with noise, agglomerative nesting, and divisive analysis). Aynı zamanda algoritmaları karşılaştırmak için Siluet indeksi, Dunn indeksi ve Calinski-Harabasz indeksi kullanılmıştır.

Kandemir (2021) “Dünya Ülkelerinde Sosyal Eşitsizliklerin Mutluluk Düzeyine Etkisi: Lojistik Regresyon Analizi” isimli çalışmasında, ülkelerdeki sosyal eşitsizliklerin mutluluk üzerindeki etkisini araştırmıştır. Çalışmada lojistik regresyon analizinden yararlanmıştır. Çalışmada 2019 Dünya Mutluluk Raporu veri setinden Mutluluk Skoru bağımlı, 2019 İnsani Gelişme Raporu’ndan eğitim, sağlık ve gelir bağımsız değişkenlerden olarak kullanılmıştır. Çalışmada 104 ülke için mutluluk skoru ortalaması alınmış ve ortalamanın üzerinde olan ülkeler mutlu, ortalamanın altında kalan ülkeler ise mutsuz olarak gruplandırılmıştır. Çalışmanın sonucunda eğitim ve sağlık alanındaki eşitsizliklerin mutluluk düzeyini düşürdüğü belirlenmiş özellikle eğitim alanındaki eşitsizliklerin mutluluk düzeyi ile doğrudan ilişkili olduğu tespit edilmiştir.

3.3. KÜME SAYISININ BELİRLENMESİ VE KÜMELEME ANALİZİ SONUÇLARI

Bu çalışmada Dünya Mutluluk Raporu ve Ekonomik Özgürlük Endeksi verilerinden yararlanılarak bir kümeleme analizi gerçekleştirilmiştir. Kümeleme analizi için kullanılan veri setinde bu iki veri seti 2017, 2018, 2019, 2020 ve 2021 yılları için birleştirilmiştir. Veri setleri sadece OECD ülkelerini kapsayacak şekilde düzenlenmiştir ve eksik verisi olan ülkeler değerlendirmeye katılmamıştır dolayısıyla 31 ülke üzerinden bir değerlendirme yapılmıştır. Veri setleri birleştirilirken ölçeklerin birbirine benzemesi amacıyla Dünya Mutluluk Raporu verileri 0-10 ölçeğinden 10 ile çarpılarak 0-100 ölçeğine getirilmiştir. Analiz için Ekonomik Özgürlük Endeksi veri setinin; Hukukun Üstünlüğü, Devlet Büyüklüğü, Mevzuat Verimliliği ve Serbest Piyasa değişkenleri, Dünya Mutluluk Raporu veri setinin; log(Kişi Başına Gayrisafi Yurtiçi Hasıla), Doğumda Sağlıklı Yaşam Beklentisi, Sosyal Destek, Yaşam Seçimleri Yapma Özgürlüğü, Cömertlik, Yolsuzluk Algısı ve Distopya değişkenleri

olmak üzere toplamda 11 değişken kullanılmıştır. Küme sayısını belirlemek için birden fazla yöntemden yararlanılmıştır ancak en başta Ward yöntemi sonucunda elde edilen Dendrogram'a bağlı kalınmıştır. Ward yöntemi uygulanırken uzaklık ölçütü olarak Karesi Alınmış Öklid Uzaklığı'ndan yararlanılmıştır. Bu yöntemler neticesinde optimum küme küme sayısı 4 olarak belirlenmiştir. Kümeler K-Ortalamlar kümeleme analizi sonuçlarından elde edilmiştir. Kümeleme analizi sonuçlarını değerlendirmek amaçlı karar ağacı yöntemi kullanılarak kümelerin oluşmasında hangi faktörlerin etkili olduğu değerlendirilmiştir. Karar ağacı yaklaşımında CART algoritması ile gini endeksi kullanılmıştır. Bu yöntemler uygulanırken IBM SPSS Statistics paket programı ve Python programlama dilinden yararlanılmıştır.

Tablo 3: Araştırmaya Dahil Edilen OECD Ülkeleri

Araştırmaya Dahil Edilen OECD Ülkeleri	
Çekya	Meksika
Danimarka	Hollanda
Estonya	Yeni Zelanda
Finlandiya	Norveç
Fransa	Polonya
Almanya	Portekiz
Yunanistan	Slovakya
Macaristan	Slovenya
İzlanda	Güney Kore
İrlanda	İspanya
İsrail	İsveç
İtalya	İsviçre
Japonya	Türkiye
Letonya	Birleşik Krallık
Litvanya	Birleşik Devletler
Lüksemburg	

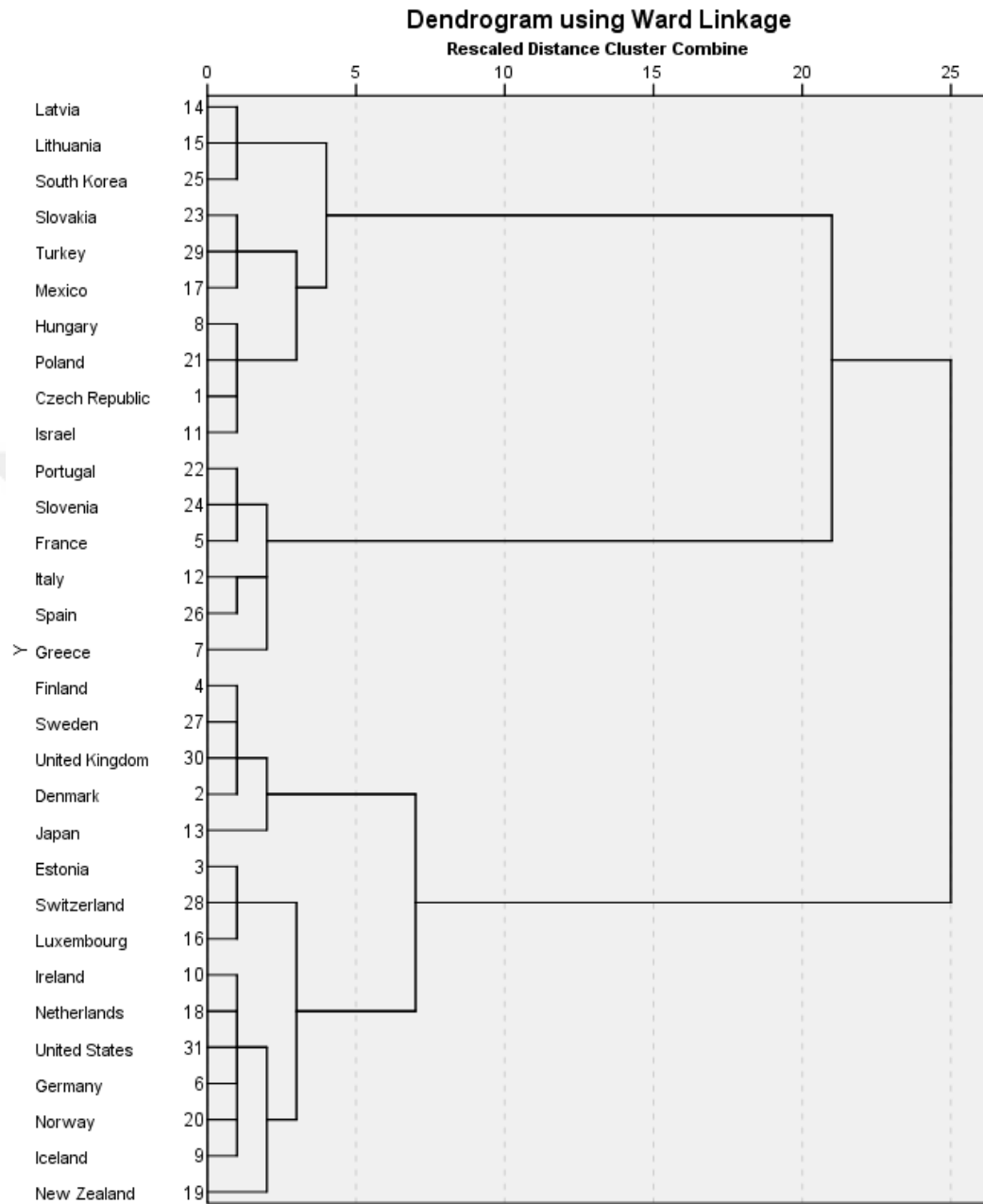
Tablo 4: Araştırmada Kullanılan Değişkenler ve Karşılıkları

Değişken İsimleri	Değişkeni İçeren Veri Seti	Veri Setinde Karşılığı
Hukukun Üstünlüğü	Ekonomik Özgürlük Endeksi	EF1_Rule of law
Devlet Büyüklüğü	Ekonomik Özgürlük Endeksi	EF2_Government size
Mevzuat Verimliliği	Ekonomik Özgürlük Endeksi	EF3_Regulatory efficiency
Serbest Piyasa	Ekonomik Özgürlük Endeksi	EF4_Market openness
Log(Kişi Başına GSYİH)	Dünya Mutluluk Raporu	WHI_11_Log GDP per capita
Sosyal Destek	Dünya Mutluluk Raporu	WHI_12_Social support
Doğumda Sağlıklı Yaşam Beklentisi	Dünya Mutluluk Raporu	WHI_13_Healthy life expectancy
Yaşam Seçimleri Yapma Özgürlüğü	Dünya Mutluluk Raporu	WHI_14_Freedom to make life choices
Cömertlik	Dünya Mutluluk Raporu	WHI_15_Generosity
Yolsuzluk Algısı	Dünya Mutluluk Raporu	WHI_16_Perceptions of corruption
Distopya + Residual	Dünya Mutluluk Raporu	WHI_17_Dystopia + residual

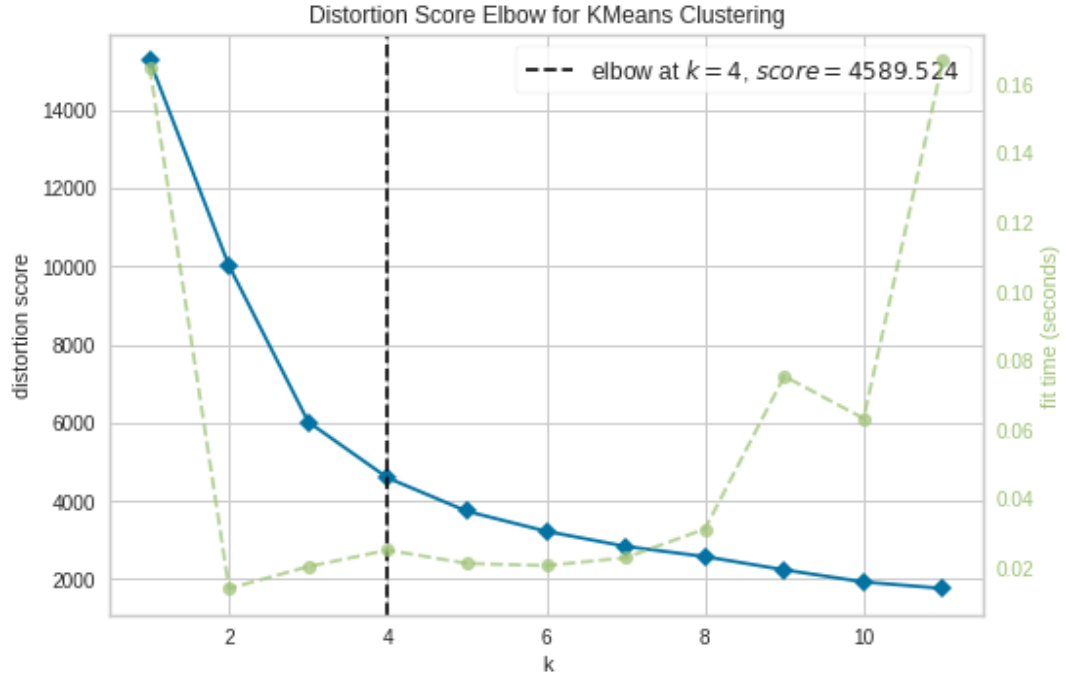
3.3.1. 2017 Yılı İçin Analizler

Veri seti ile 2017 yılı için analizler 31 OECD ülkesi içermektedir. Veri setine ilk önce küme sayısının belirlenmesi için dendrogram ve dirsek grafiği yöntemleri uygulanmıştır. Küme sayısı belirlendikten sonra k-ortalamalar yöntemi ile ilk ve nihai küme merkezleri aynı zamanda küme sonuçları elde edilmiştir. Kümelerin belirlenmesinde hangi değişkenin daha çok rol oynadığını belirlemek için bir karar ağacı oluşturulmuştur.

Şekil 8: 2017 Yılı İçin Dendrogram



Şekil 9:2017 Yılı İçin Dirsek Grafiği



Yukarıda bulunan dendrogram ve dirsek grafiği doğrultusunda 2020 yılı için optimum küme sayısının 4 olduğuna karar verilmiştir ve K-Ortalamalar yöntemi 4 küme için gerçekleştirilmiştir. Aşağıda K-Ortalamalar yönteminden elde edilen ilk ve nihai küme merkezleri tablo olarak gösterilmiştir.

Tablo 5: 2017 Yılı İçin İlk Küme Merkezleri

İlk Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	49,977	91,486	51,483	75,785
Devlet Büyüklüğü	41,518	71,685	76,300	35,541
Mevzuat Verimliliği	67,850	89,358	61,661	67,881
Serbest Piyasa	60,660	82,453	71,480	73,993
Log(Kişi Başına GSYİH)	12,895	14,057	11,983	14,309
Sosyal Destek	12,394	15,482	13,378	13,878
Doğumda Sağlıklı Yaşam Beklentisi	8,102	8,168	6,376	8,445

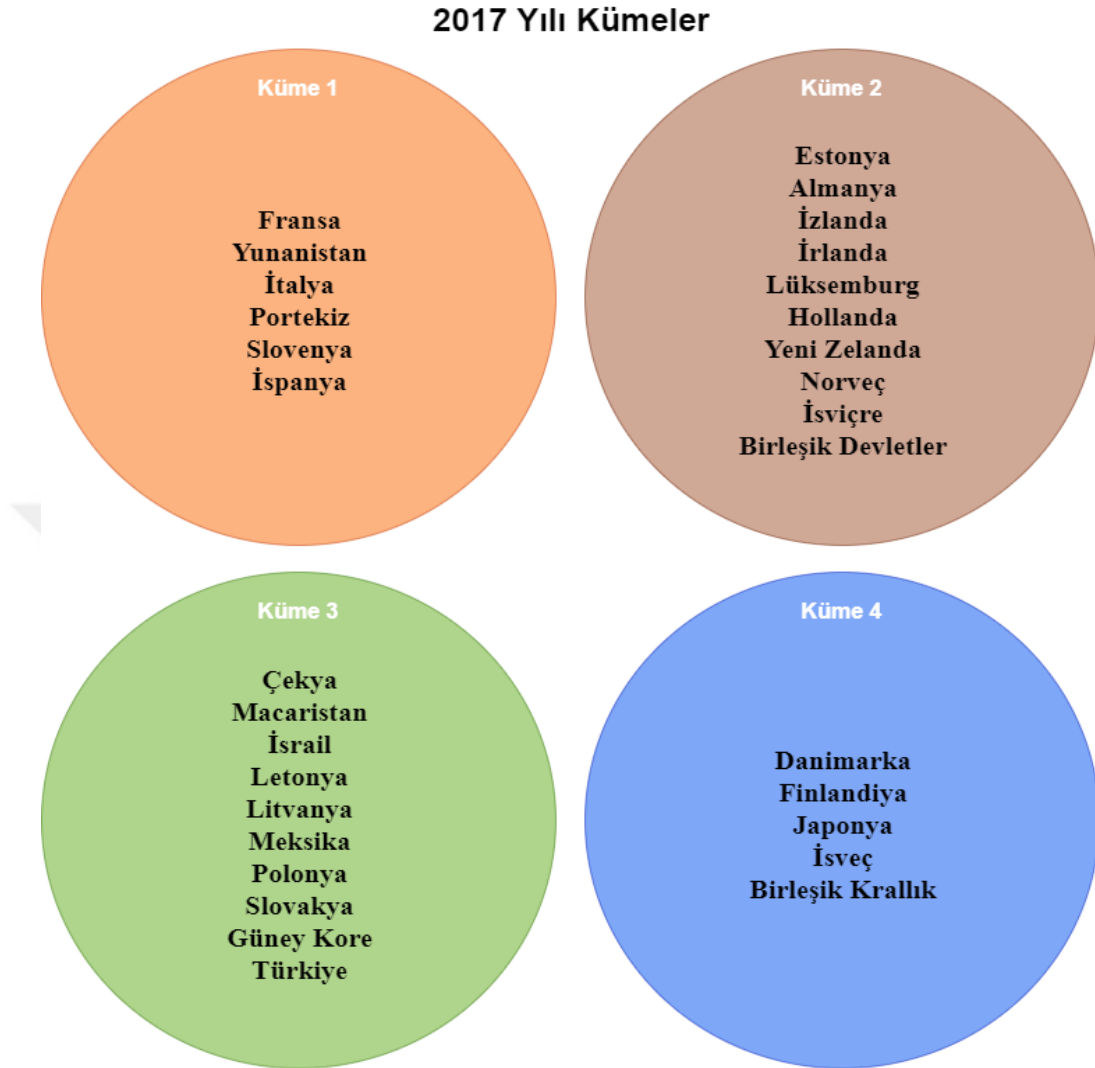
Yaşam Seçimleri Yapma Özgürlüğü	0,957	6,141	3,007	4,702
Cömertlik	0,000	5,000	0,467	1,298
Yolsuzluk Algısı	0,433	3,828	0,997	1,725
Distopya + Residual	17,492	20,465	18,793	20,060

İlk küme merkezleri 7 iterasyon sonucunda optimum hale getirilmiş ve nihai küme merkezleri elde edilmiştir.

Tablo 6: 2017 Yılı İçin Nihai Küme Merkezleri

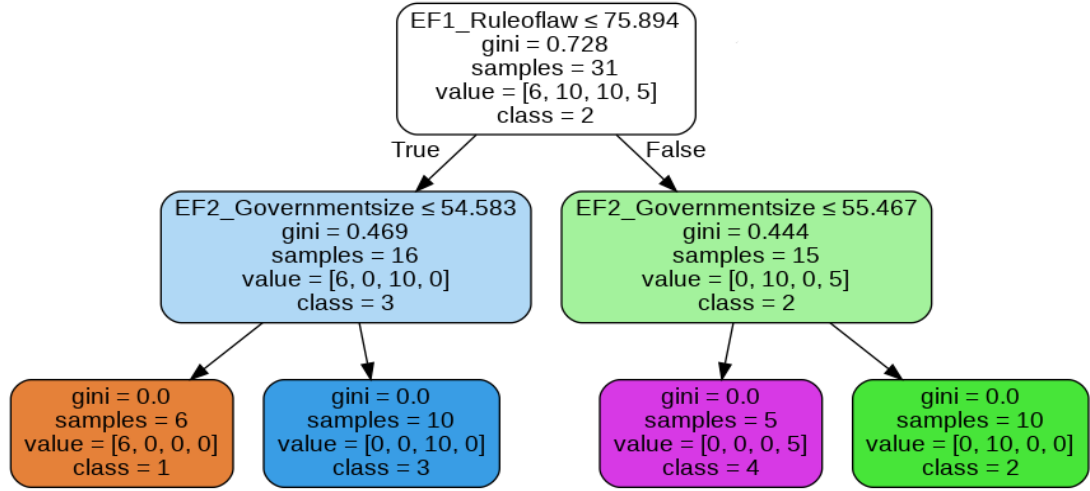
Nihai Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	62,185	81,406	58,297	85,059
Devlet Büyüklüğü	39,816	67,392	72,178	47,967
Mevzuat Verimliliği	70,354	77,215	72,641	80,908
Serbest Piyasa	71,771	82,001	75,463	82,036
Log(Kişi Başına GSYİH)	13,594	15,204	12,960	14,558
Sosyal Destek	14,040	15,023	13,659	15,005
Doğumda Sağlıklı Yaşam Beklentisi	8,306	8,039	7,208	8,303
Yaşam Seçimleri Yapma Özgürlüğü	3,837	5,799	3,420	5,741
Cömertlik	1,384	3,637	1,329	3,199
Yolsuzluk Algısı	0,626	2,715	0,624	3,194
Distopya + Residual	16,529	20,202	21,684	19,818

Şekil 10: 2017 Yılı İçin Kümeler



2017 yılı için kümeler k-ortalamalar yöntemi sonucunda yukarıdaki gibi oluşmuştur. Türkiye yer aldığı kümede Çekya, Macaristan, İsrail, Letonya, Litvanya, Meksika, Polonya, Slovakya ve Güney Kore ülkeleri ile birlikte yer almaktadır.

Şekil 11: 2017 Yılı İçin Karar Ağacı

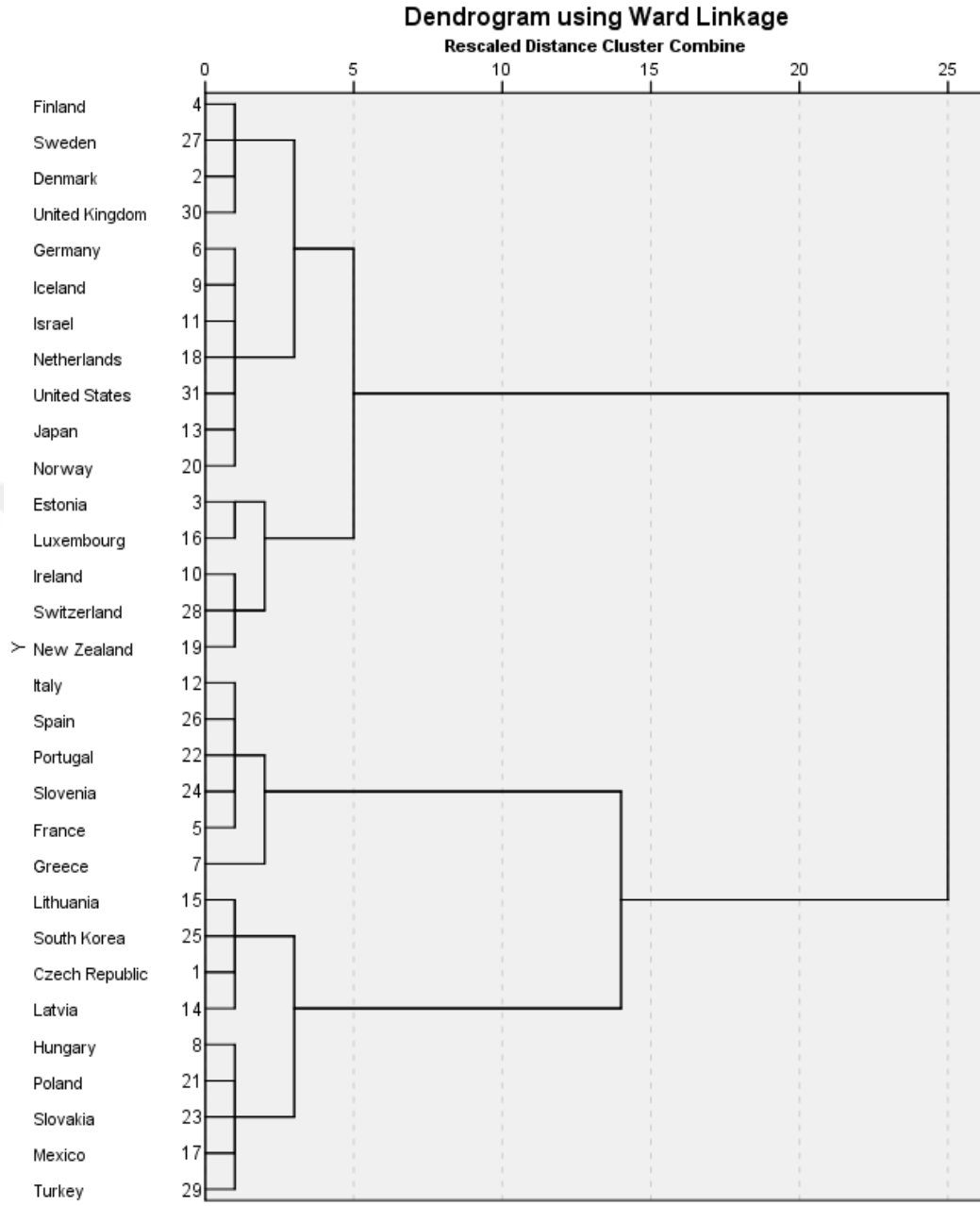


Karar ağacı oluşturulurken gini endeksinden faydalanılmıştır. Karar ağacı sonuçları incelendiğinde 2017 yılı için kümelerin belirlenmesinde en çok rol oynayan faktörler sırasıyla, hukukun üstünlüğü ve devlet büyüklüğü olarak belirlenmiştir.

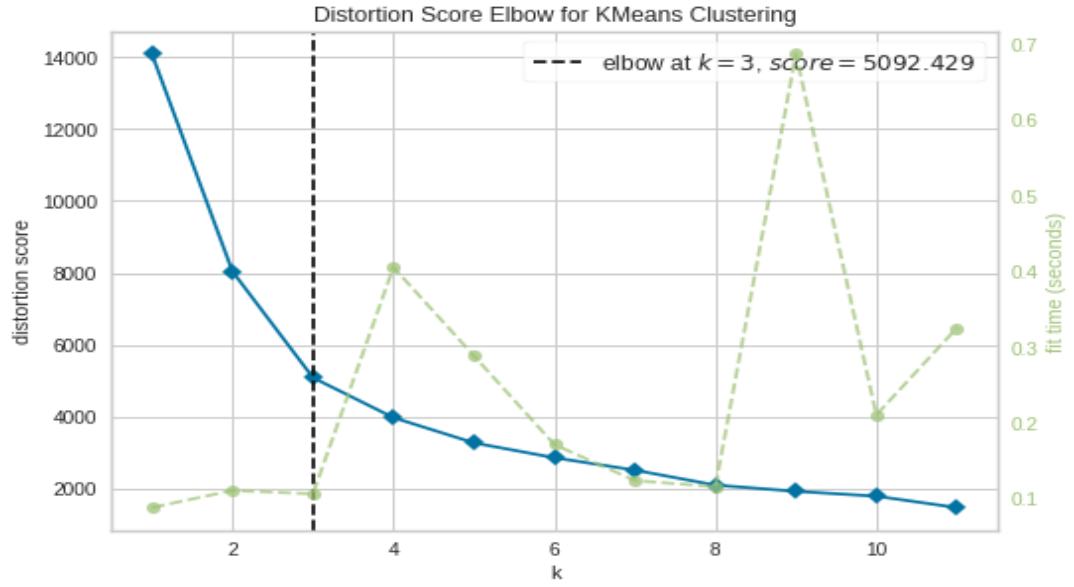
3.3.2. 2018 Yılı İçin Analizler

Veri seti ile 2018 yılı için analizler 31 OECD ülkesi içermektedir. Veri setine ilk önce küme sayısının belirlenmesi için dendrogram ve dirsek grafiği yöntemleri uygulanmıştır. Küme sayısı belirlendikten sonra k-ortalamalar yöntemi ile ilk ve nihai küme merkezleri aynı zamanda küme sonuçları elde edilmiştir. Kümelerin belirlenmesinde hangi değişkenin daha çok rol oynadığını belirlemek için bir karar ağacı oluşturulmuştur.

Şekil 12: 2018 Yılı İçin Dendrogram



Şekil 13: 2018 Yılı İçin Dirsek Grafiği



Yukarıda bulunan dendrogram ve dirsek grafiği doğrultusunda 2018 yılı için optimum küme sayısının 3 ya da 4 olması gerektiği görülmektedir. Ülkeler için küme sayısı 3 ve 4 olarak belirlenip analiz yapıldığında; 4 kümeli durumdaki “Küme 2” ve “Küme 3” kümelerinin, 3 kümeli durumda birleşip “Küme 2” olarak karşımıza çıktığı ve diğer kümelerde bir değişim gözlemlenmediği belirlenmiştir. Çalışmanın yıllar arasında sağlıklı karşılaştırılması açısından 2018 yılı için optimum küme sayısı 4 olarak belirlenmiştir ve K-Ortalamlar yöntemi 4 küme için gerçekleştirilmiştir. Aşağıda K-Ortalamlar yönteminden elde edilen ilk ve nihai küme merkezleri tablo olarak gösterilmiştir.

Tablo 7: 2018 Yılı İçin İlk Küme Merkezleri

İlk Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	41,500	84,167	80,000	62,133
Devlet Büyüklüğü	74,533	49,567	77,700	52,067
Mevzuat Verimliliği	68,833	87,233	71,833	76,033
Serbest Piyasa	74,333	85,633	85,633	68,967
Log(Kişi Başına GSYİH)	10,382	13,512	12,000	12,190
Sosyal Destek	12,522	15,905	15,324	15,064
Doğumda Sağlıklı Yaşam Beklentisi	7,607	8,679	7,373	8,562

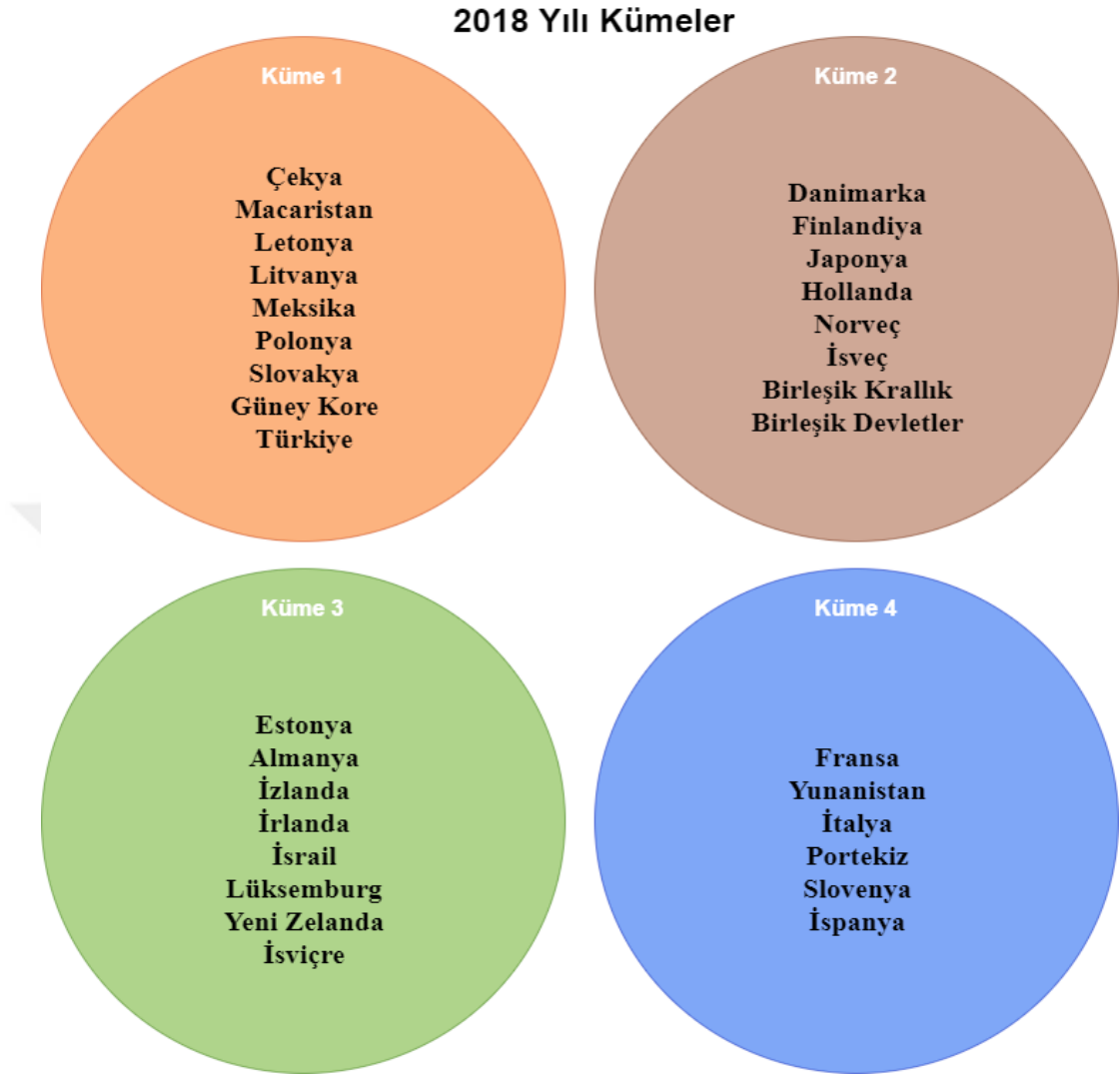
Yaşam Seçimleri Yapma Özgürlüğü	4,792	6,833	5,528	6,329
Cömertlik	0,687	2,844	0,857	1,598
Yolsuzluk Algısı	0,947	4,079	1,739	0,510
Distopya + Residual	27,945	23,702	14,571	15,223

İlk küme merkezleri 4 iterasyon sonucunda optimum hale getirilmiş ve nihai küme merkezleri elde edilmiştir.

Tablo 3: 2018 Yılı İçin Nihai Küme Merkezleri

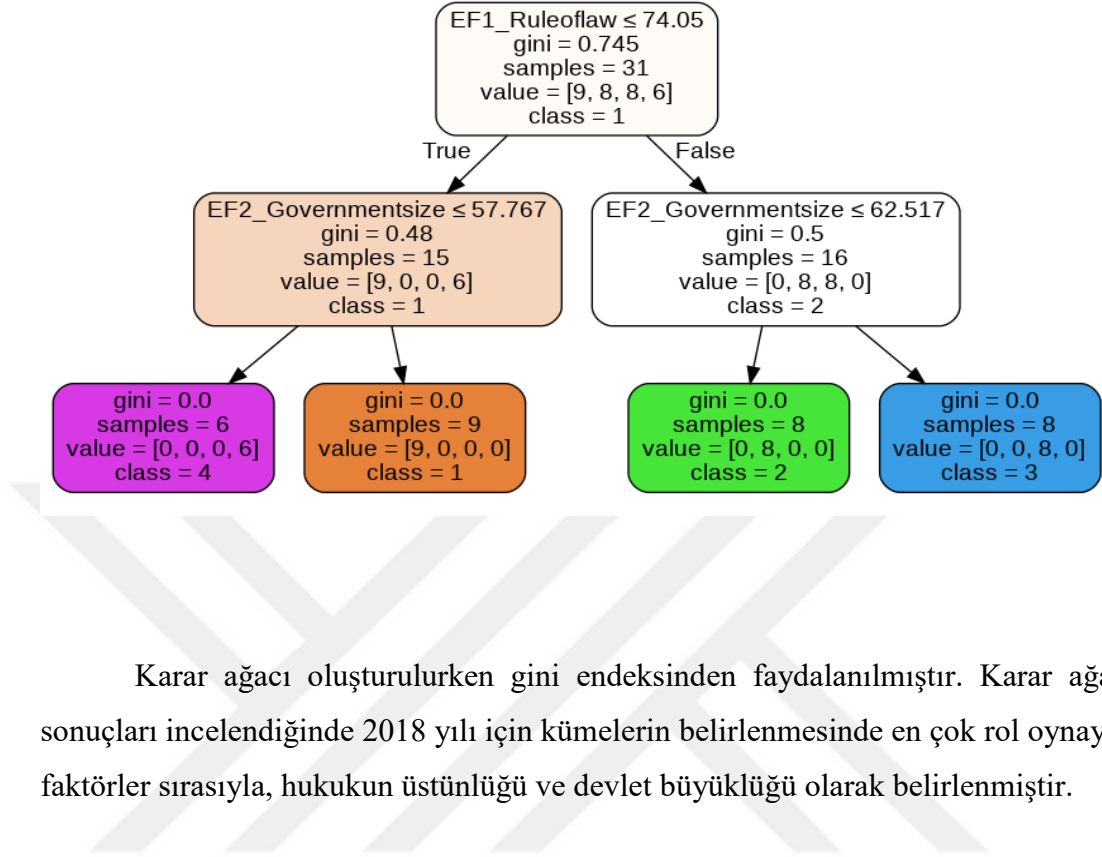
Nihai Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	54,826	84,717	81,125	61,794
Devlet Büyüklüğü	74,659	55,638	71,583	46,822
Mevzuat Verimliliği	72,678	79,708	76,900	71,006
Serbest Piyasa	76,793	81,725	82,875	71,744
Log(Kişi Başına GSYİH)	11,764	13,527	13,548	12,280
Sosyal Destek	14,104	15,307	15,420	14,404
Doğumda Sağlıklı Yaşam Beklentisi	7,700	8,855	8,719	9,064
Yaşam Seçimleri Yapma Özgürlüğü	3,846	6,226	6,068	4,293
Cömertlik	0,907	2,629	2,623	0,984
Yolsuzluk Algısı	0,512	2,967	2,557	0,650
Distopya + Residual	21,564	21,929	21,174	17,516

Şekil 14: 2018 Yılı İçin Kümeler



2018 yılı için kümeler k-ortalamalar yöntemi sonucunda yukarıdaki gibi oluşmuştur. Türkiye yer aldığı kümede Çekya, Macaristan, Letonya, Litvanya, Meksika, Polonya, Slovakya ve Güney Kore ülkeleri ile birlikte yer almaktadır.

Şekil 15: 2018 Yılı İçin Karar Ağacı

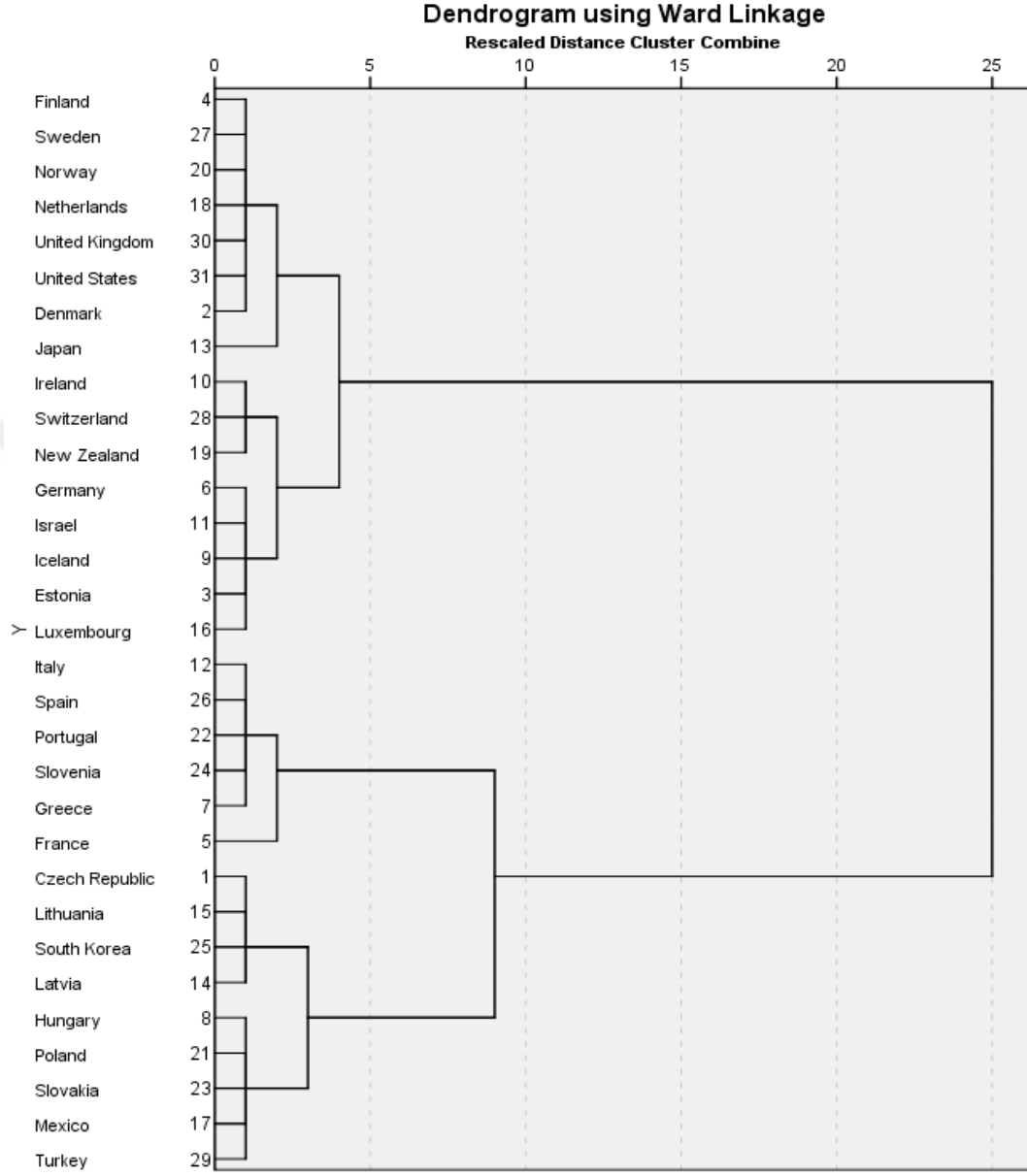


Karar ağacı oluşturulurken gini endeksinden faydalanılmıştır. Karar ağacı sonuçları incelendiğinde 2018 yılı için kümelerin belirlenmesinde en çok rol oynayan faktörler sırasıyla, hukukun üstünlüğü ve devlet büyüklüğü olarak belirlenmiştir.

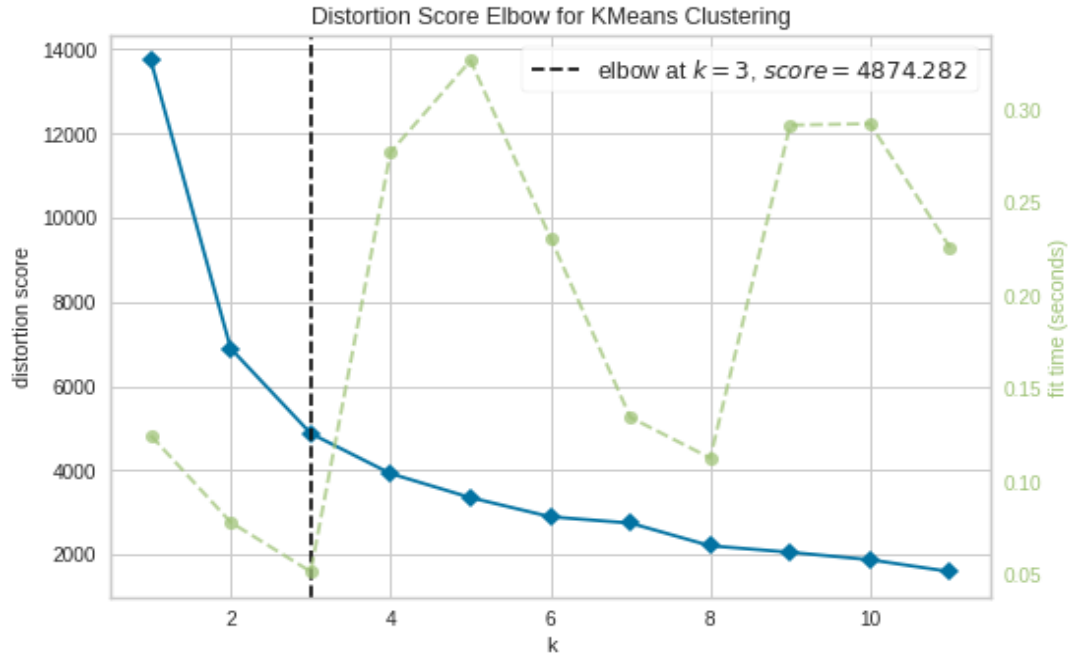
3.3.3. 2019 Yılı İçin Analizler

Veri seti ile 2019 yılı için analizler 31 OECD ülkesi içermektedir. Veri setine ilk önce küme sayısının belirlenmesi için dendrogram ve dirsek grafiği yöntemleri uygulanmıştır. Küme sayısı belirlendikten sonra k-ortalamalar yöntemi ile ilk ve nihai küme merkezleri aynı zamanda küme sonuçları elde edilmiştir. Kümelerin belirlenmesinde hangi değişkenin daha çok rol oynadığını belirlemek için bir karar ağacı oluşturulmuştur.

Şekil 16: 2019 Yılı İçin Dendrogram



Şekil 17: 2019 Yılı İçin Dirsek Grafiği



Yukarıda bulunan dendrogram ve dirsek grafiği doğrultusunda 2019 yılı için optimum küme sayısının 3 ya da 4 olması gerektiği görülmektedir. Ülkeler için küme sayısı 3 ve 4 olarak belirlenip analiz yapıldığında; 4 kümeli durumdaki “Küme 3” ve “Küme 4” kümelerinin, 3 kümeli durumda birleşip “Küme 3” olarak karşımıza çıktığı ve diğer kümelerde bir değişim gözlemlenmediği belirlenmiştir. Çalışmanın yıllar arasında sağlıklı karşılaştırılması açısından 2019 yılı için optimum küme sayısı 4 olarak belirlenmiştir ve K-Ortalamlar yöntemi 4 küme için gerçekleştirilmiştir. Aşağıda K-Ortalamlar yönteminden elde edilen ilk ve nihai küme merkezleri tablo olarak gösterilmiştir.

Tablo 9: 2019 Yılı İçin İlk Küme Merkezleri

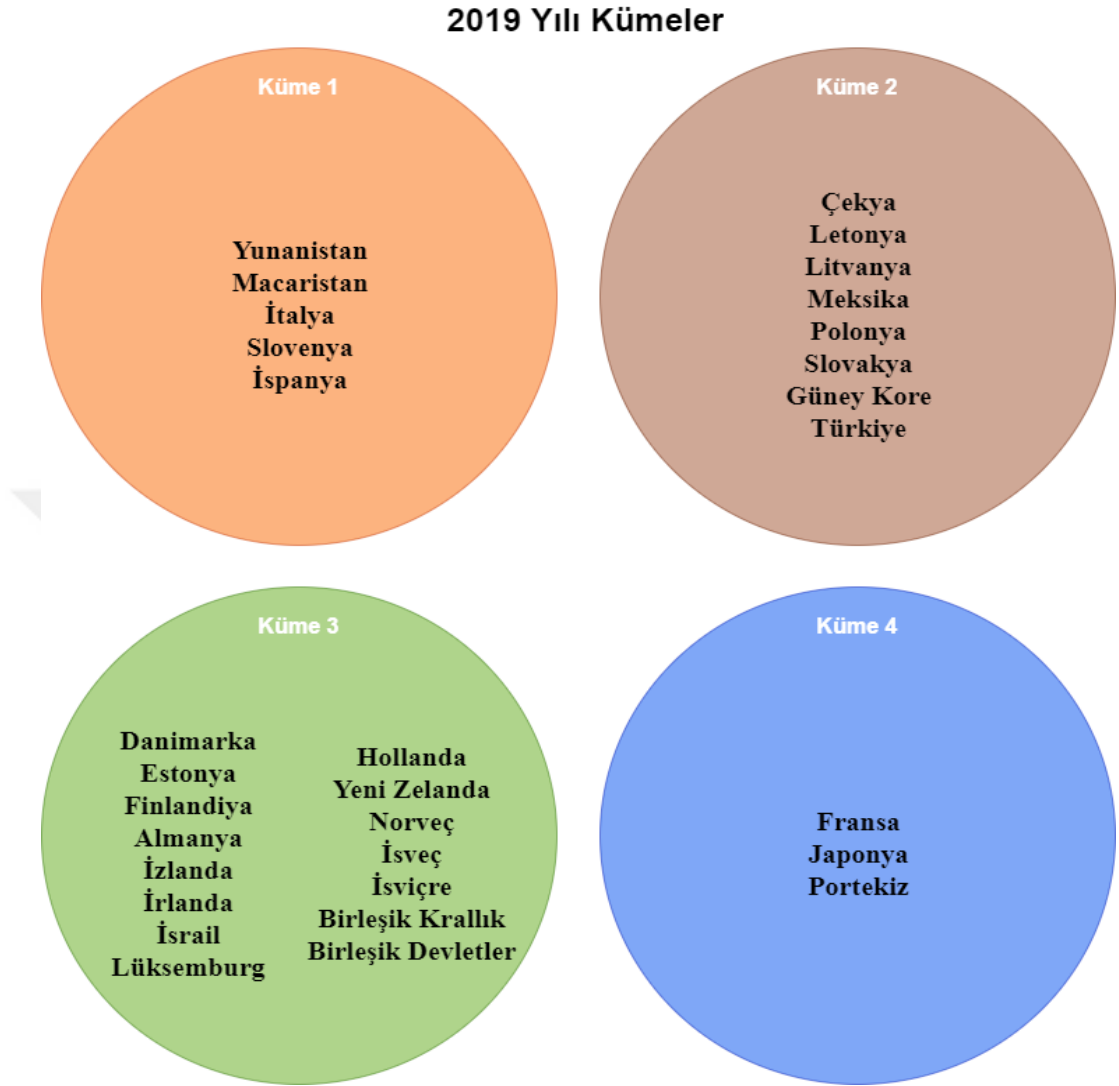
İlk Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	46,533	62,433	91,733	72,167
Devlet Büyüklüğü	53,800	76,533	73,333	39,067
Mevzuat Verimliliği	68,567	76,900	88,400	68,500
Serbest Piyasa	62,000	73,467	84,133	75,333
Log(Kişi Başına GSYİH)	11,813	13,010	13,026	13,239
Sosyal Destek	11,562	12,192	15,572	14,721
Doğumda Sağlıklı Yaşam Beklentisi	9,993	10,356	10,256	10,454
Yaşam Seçimleri Yapma Özgürlüğü	0,671	1,593	5,851	4,355
Cömertlik	0,000	1,747	3,298	1,111
Yolsuzluk Algısı	0,344	0,560	3,803	1,832
Distopya + Residual	18,484	19,489	21,268	20,209

İlk küme merkezleri 3 iterasyon sonucunda optimum hale getirilmiş ve nihai küme merkezleri elde edilmiştir.

Tablo 10: 2019 Yılı İçin Nihai Küme Merkezleri

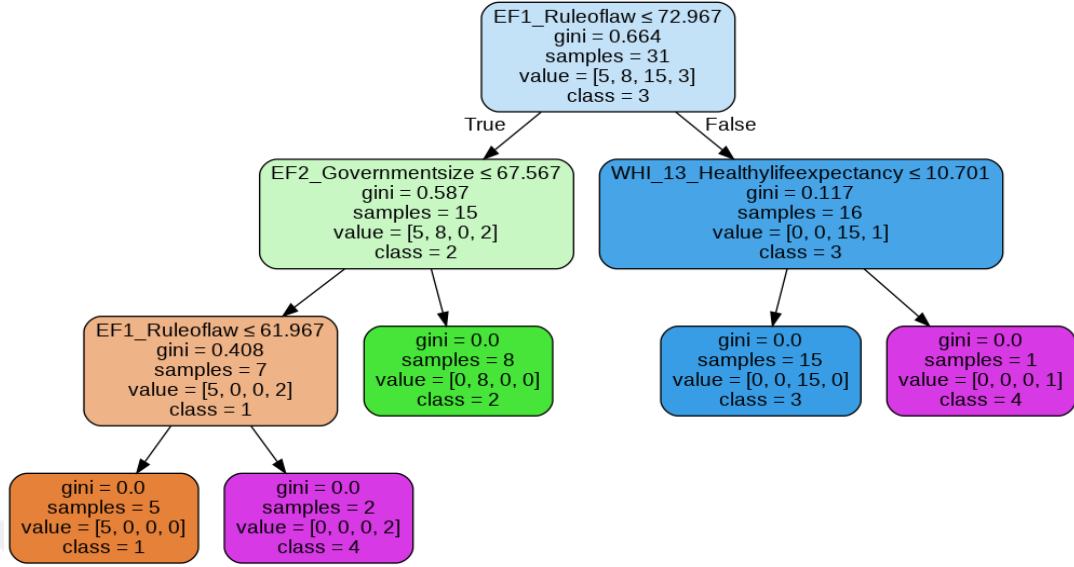
Nihai Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	53,260	52,592	82,460	71,378
Devlet Büyüklüğü	56,600	76,442	65,744	51,267
Mevzuat Verimliliği	70,420	71,258	77,227	73,100
Serbest Piyasa	72,667	76,100	82,556	72,444
Log(Kişi Başına GSYİH)	12,441	12,126	13,925	12,904
Sosyal Destek	14,123	14,141	15,281	14,405
Doğumda Sağlıklı Yaşam Beklentisi	9,763	8,775	9,929	10,443
Yaşam Seçimleri Yapma Özgürlüğü	2,846	3,271	5,318	4,629
Cömertlik	1,071	0,916	2,640	0,756
Yolsuzluk Algısı	0,440	0,552	2,797	1,161
Distopya + Residual	18,796	21,697	22,182	16,271

Şekil 18: 2019 Yılı İçin Kümeler



2019 yılı için kümeler k-ortalamalar yöntemi sonucunda yukarıdaki gibi oluşmuştur. Türkiye yer aldığı kümede Çekya, Letonya, Litvanya, Meksika, Polonya, Slovakya ve Güney Kore ülkeleri ile birlikte yer almaktadır.

Şekil 19: 2019 Yılı İçin Karar Ağacı

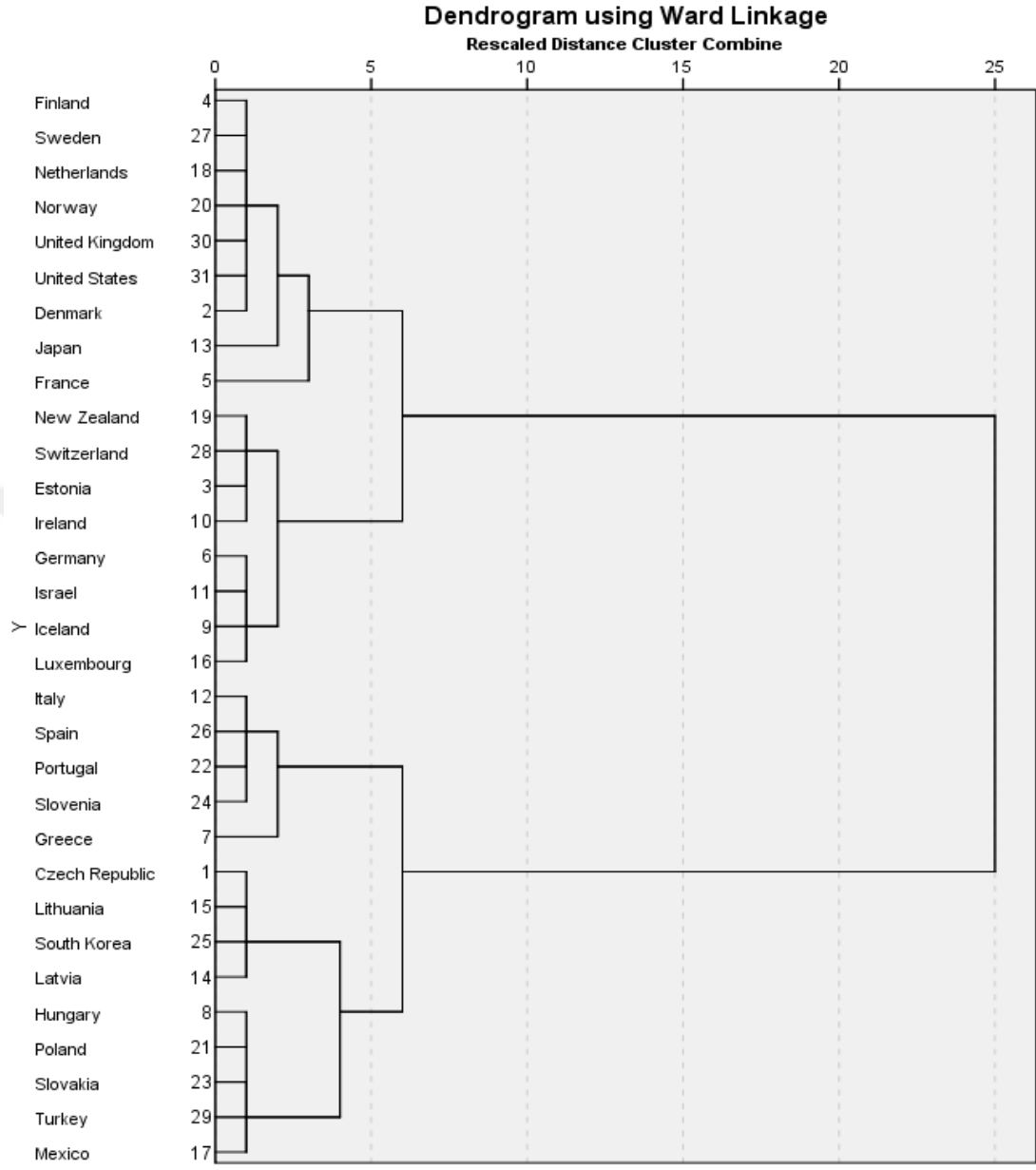


Karar ağacı oluşturulurken gini endeksinden faydalanılmıştır. Karar ağacı sonuçları incelendiğinde 2019 yılı için kümelerin belirlenmesinde en çok rol oynayan faktörler sırasıyla, hukukun üstünlüğü, devlet büyüklüğü ve doğumda sağlıklı yaşam beklentisi olarak belirlenmiştir.

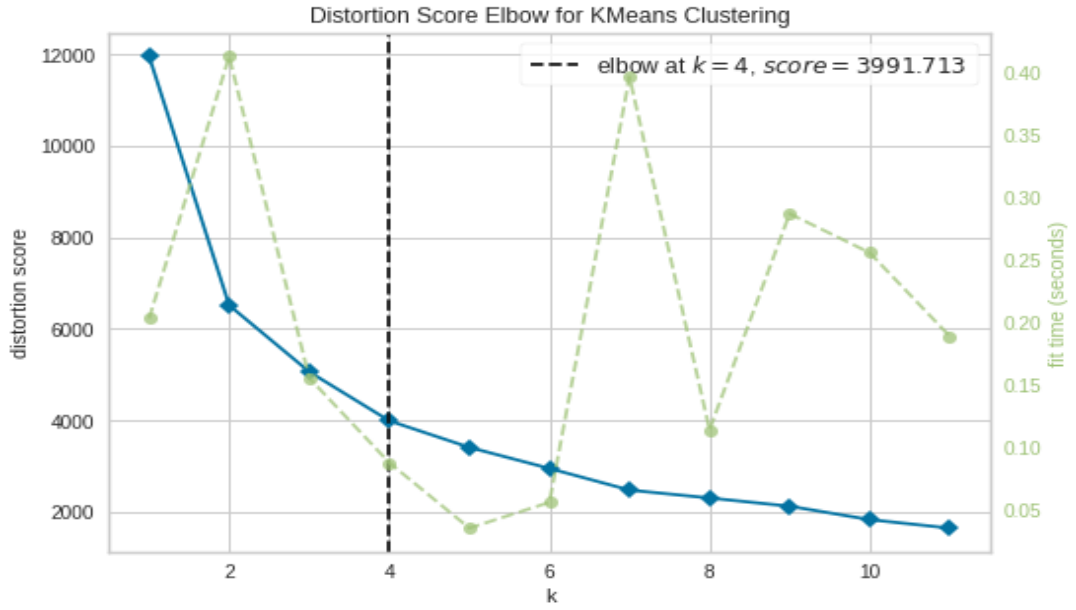
3.3.4. 2020 Yılı İçin Analizler

Veri seti ile 2020 yılı için analizler 31 OECD ülkesi içermektedir. Veri setine ilk önce küme sayısının belirlenmesi için dendrogram ve dirsek grafiği yöntemleri uygulanmıştır. Küme sayısı belirlendikten sonra k-ortalamalar yöntemi ile ilk ve nihai küme merkezleri aynı zamanda küme sonuçları elde edilmiştir. Kümelerin belirlenmesinde hangi değişkenin daha çok rol oynadığını belirlemek için bir karar ağacı oluşturulmuştur.

Şekil 20: 2020 Yılı İçin Dendrogram



Şekil 21: 2020 Yılı İçin Dirsek Grafiği



Yukarıda bulunan dendrogram ve dirsek grafiği doğrultusunda 2020 yılı için optimum küme sayısının 4 olduğuna karar verilmiştir ve K-Ortalamalar yöntemi 4 küme için gerçekleştirilmiştir. Aşağıda K-Ortalamalar yönteminden elde edilen ilk ve nihai küme merkezleri tablo olarak gösterilmiştir.

Tablo 11: 2020 Yılı İçin İlk Küme Merkezleri

İlk Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	64,167	43,233	88,767	80,133
Devlet Büyüklüğü	64,133	80,967	75,700	40,133
Mevzuat Verimliliği	74,000	65,433	88,033	68,433
Serbest Piyasa	68,800	74,200	84,067	75,467
Log(Kişi Başına GSYİH)	12,087	10,244	12,423	12,681
Sosyal Destek	14,647	12,263	14,872	14,588

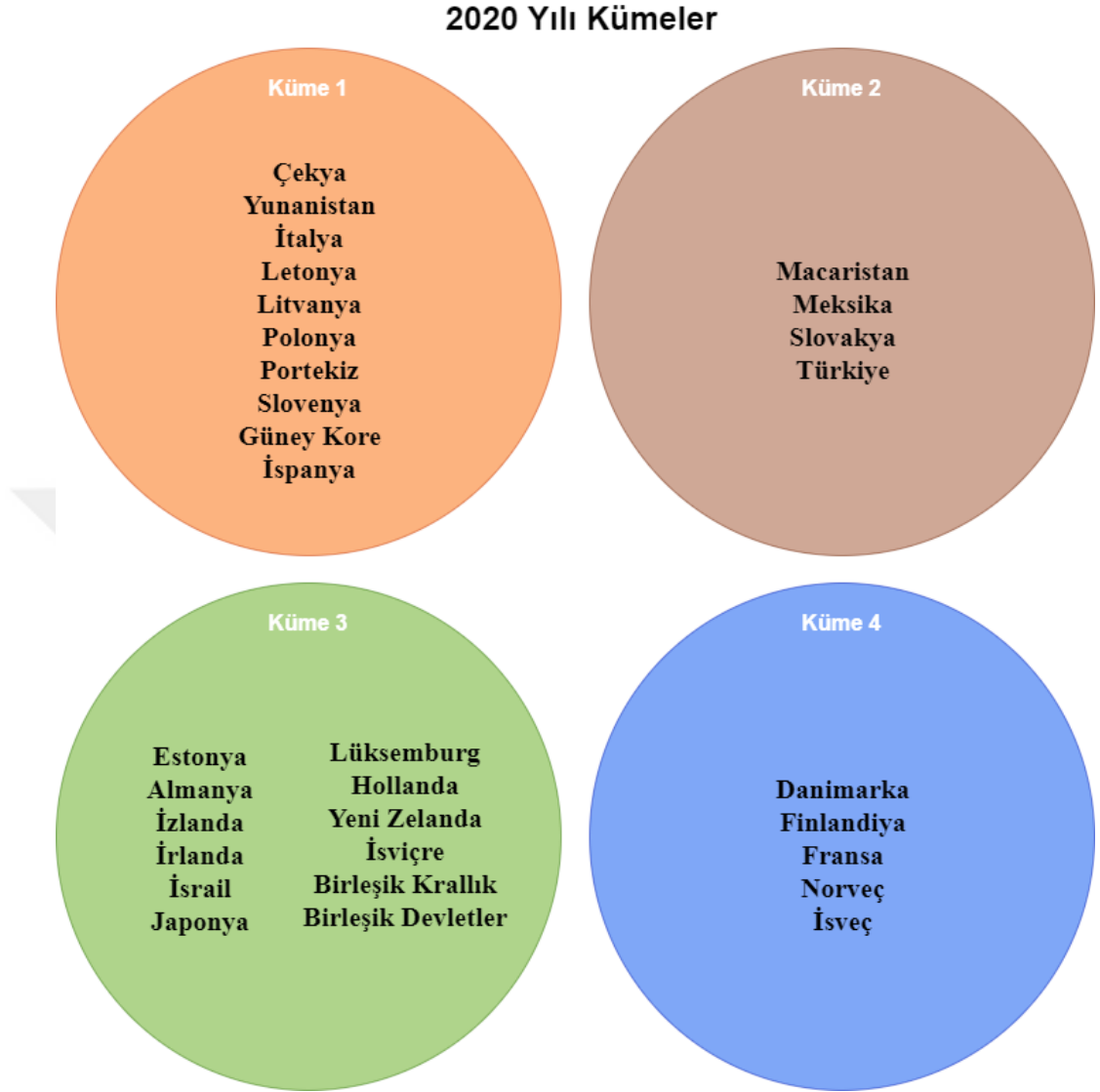
Doğumda Sağlıklı Yaşam Beklentisi	9,325	8,316	10,081	10,297
Yaşam Seçimleri Yapma Özgürlüğü	6,467	5,539	6,468	5,141
Cömertlik	1,457	0,831	3,257	1,126
Yolsuzluk Algısı	0,765	0,831	4,613	2,273
Distopya + Residual	18,886	26,625	21,281	20,532

İlk küme merkezleri 4 iterasyon sonucunda optimum hale getirilmiş ve nihai küme merkezleri elde edilmiştir.

Tablo 12: 2020 Yılı İçin Nihai Küme Merkezleri

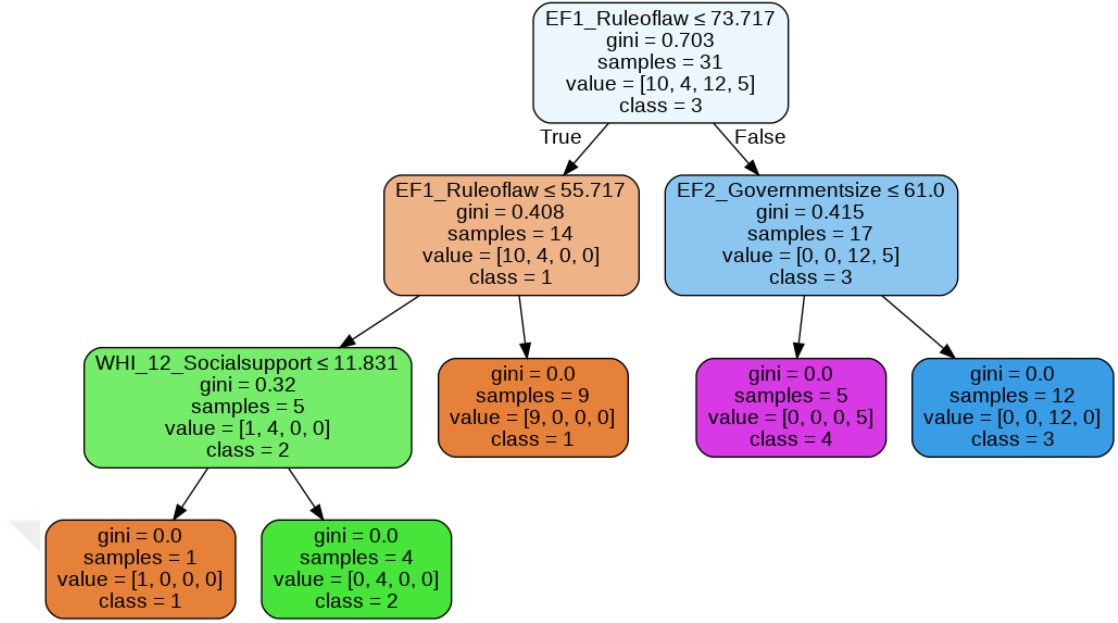
Nihai Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	62,650	50,658	82,306	86,400
Devlet Büyüklüğü	67,393	74,092	69,900	53,080
Mevzuat Verimliliği	72,097	63,925	76,717	74,807
Serbest Piyasa	74,587	74,867	81,528	80,293
Log(Kişi Başına GSYİH)	11,934	11,277	13,266	13,253
Sosyal Destek	13,438	13,177	14,374	14,781
Doğumda Sağlıklı Yaşam Beklentisi	9,324	8,183	9,742	9,930
Yaşam Seçimleri Yapma Özgürlüğü	4,230	4,046	5,722	6,324
Cömertlik	0,943	0,891	2,613	2,152
Yolsuzluk Algısı	0,780	0,607	2,825	4,153
Distopya + Residual	21,064	21,514	21,745	23,327

Şekil 22: 2020 Yılı İçin Kümeler



2020 yılı için kümeler k-ortalamlar yöntemi sonucunda yukarıdaki gibi oluşmuştur. Türkiye yer aldığı kümede Macaristan, Meksika ve Slovakya ülkeleri ile birlikte yer almaktadır.

Şekil 23: 2020 Yılı İçin Karar Ağacı

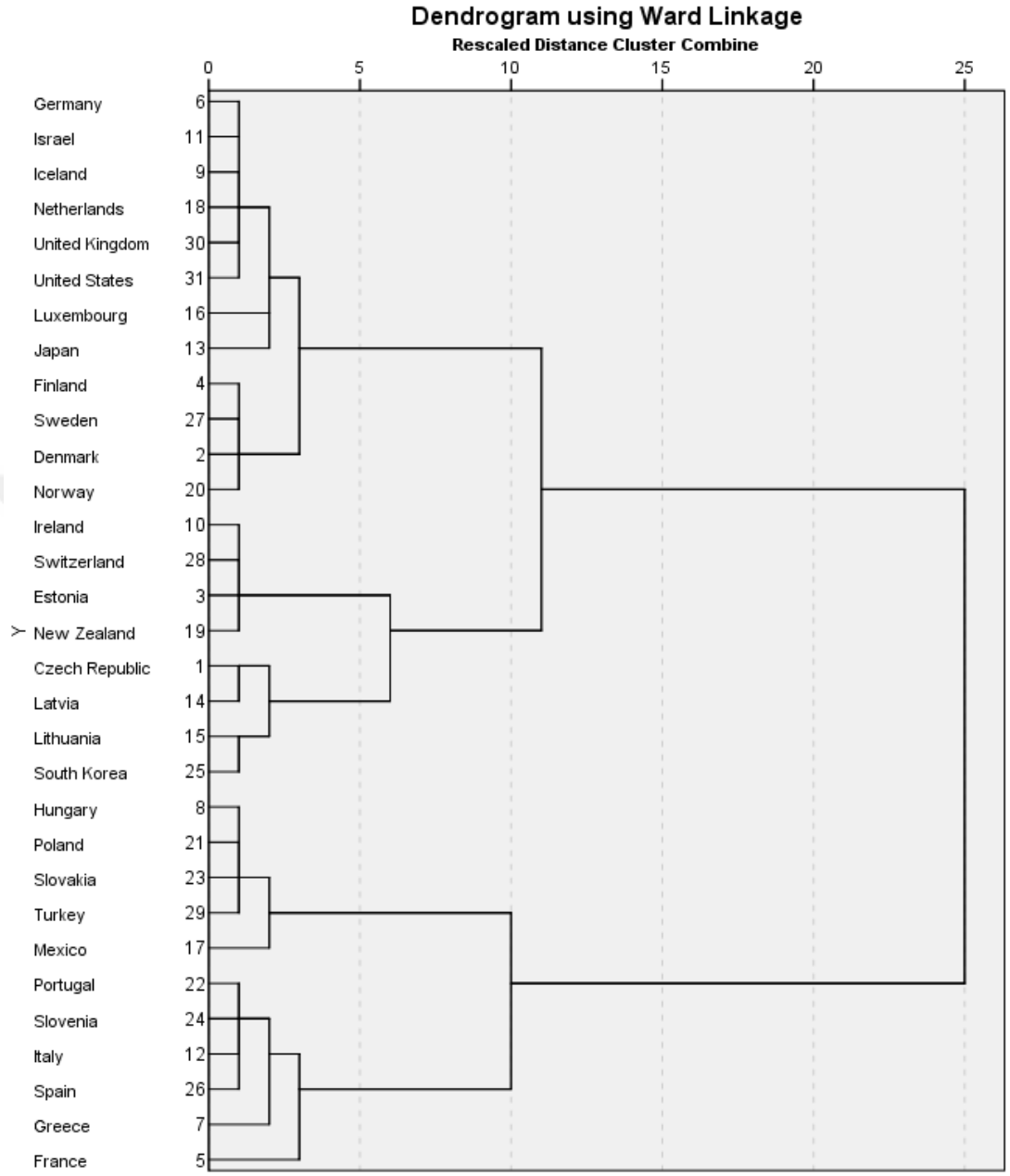


Karar ağacı oluşturulurken gini endeksinden faydalanılmıştır. Karar ağacı sonuçları incelendiğinde 2020 yılı için kümelerin belirlenmesinde en çok rol oynayan faktörler sırasıyla, hukukun üstünlüğü, devlet büyüklüğü ve sosyal destek olarak belirlenmiştir.

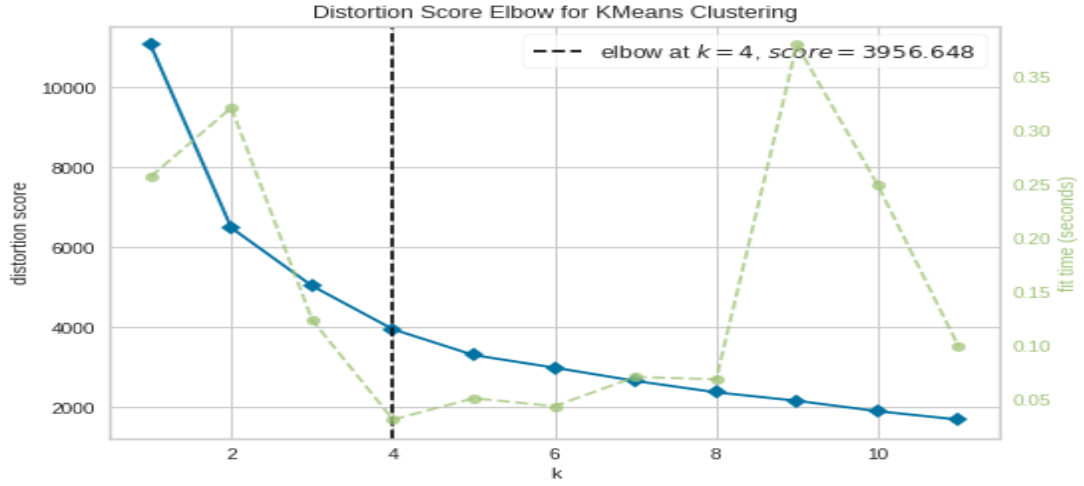
3.3.5. 2021 Yılı İçin Analizler

Veri seti ile 2021 yılı için analizler 31 OECD ülkesi içermektedir. Veri setine ilk önce küme sayısının belirlenmesi için dendrogram ve dirsek grafiği yöntemleri uygulanmıştır. Küme sayısı belirlendikten sonra k-ortalamalar yöntemi ile ilk ve nihai küme merkezleri aynı zamanda küme sonuçları elde edilmiştir. Kümelerin belirlenmesinde hangi değişkenin daha çok rol oynadığını belirlemek için bir karar ağacı oluşturulmuştur.

Şekil 24: 2021 Yılı İçin Dendrogram



Şekil 25: 2021 Yılı İçin Dirsek Grafiği



Yukarıda bulunan dendrogram ve dirsek grafiği doğrultusunda 2021 yılı için optimum küme sayısının 4 olduğuna karar verilmiştir ve K-Ortalamlar yöntemi 4 küme için gerçekleştirilmiştir. Aşağıda K-Ortalamlar yönteminden elde edilen ilk ve nihai küme merkezleri tablo olarak gösterilmiştir.

Tablo 13: 2021 Yılı İçin İlk Küme Merkezleri

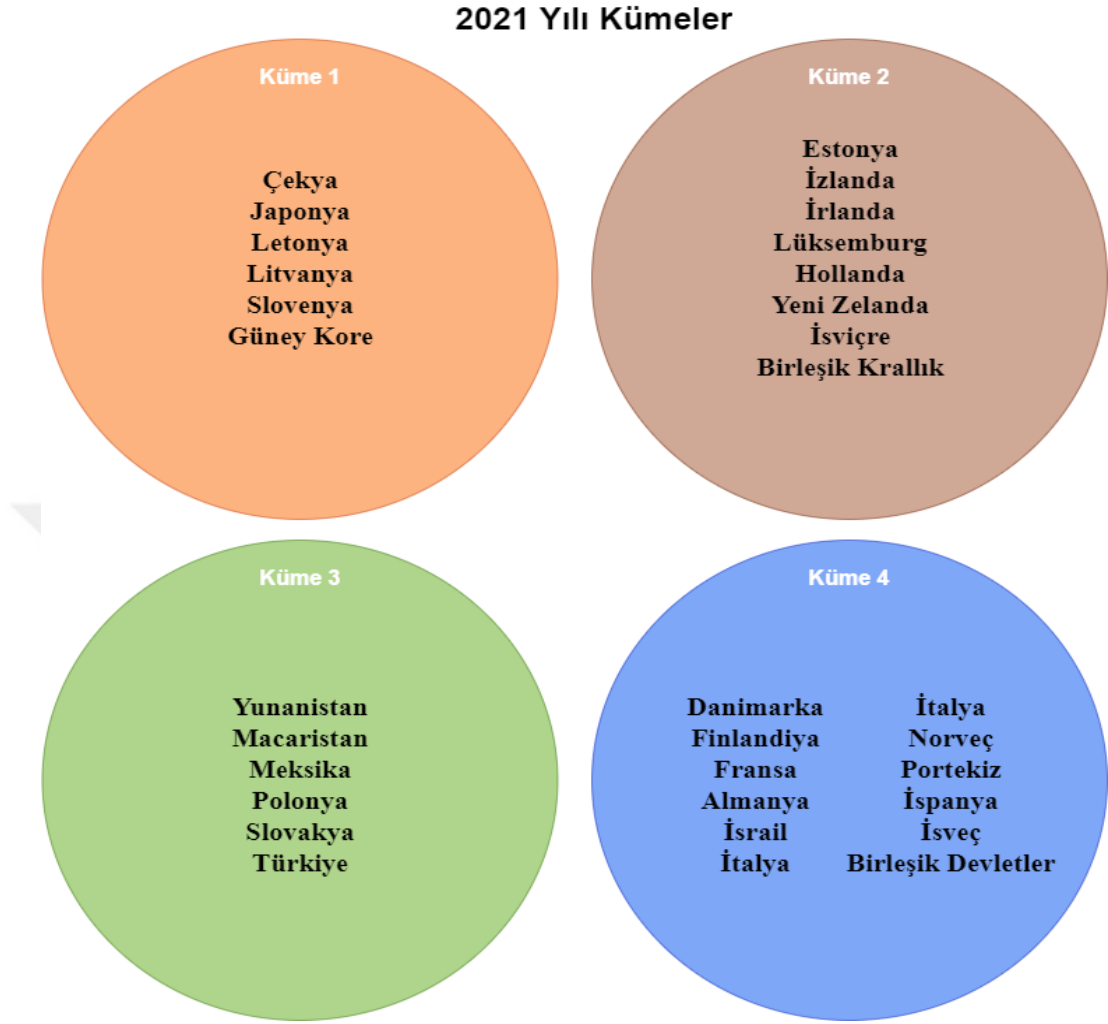
İlk Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	71,000	89,033	43,033	77,167
Devlet Büyüklüğü	82,200	75,633	81,533	42,000
Mevzuat Verimliliği	76,567	87,700	65,300	67,433
Serbest Piyasa	66,333	83,400	72,200	76,333
Log(Kişi Başına GSYİH)	14,029	14,001	11,262	14,214
Sosyal Destek	7,580	10,939	8,297	10,808
Doğumda Sağlıklı Yaşam Beklentisi	8,007	7,849	6,337	8,038
Yaşam Seçimleri Yapma Özgürlüğü	3,532	6,654	5,850	5,355
Cömertlik	1,337	2,756	0,922	0,924
Yolsuzluk Algısı	1,354	4,449	0,894	2,351
Distopya + Residual	22,617	26,118	29,606	25,208

İlk küme merkezleri 4 iterasyon sonucunda optimum hale getirilmiş ve nihai küme merkezleri elde edilmiştir.

Tablo 14: 2021 Yılı İçin Nihai Küme Merkezleri

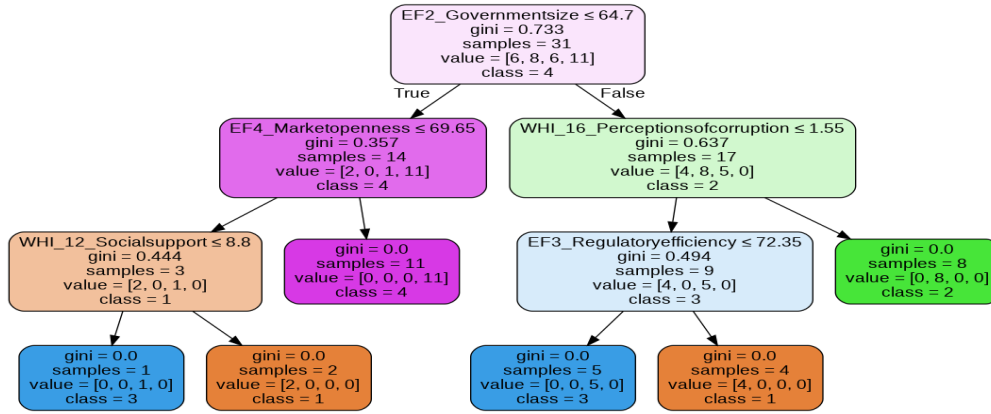
Nihai Küme Merkezleri				
	Kümeler			
	1	2	3	4
Hukukun Üstünlüğü	69,883	84,213	53,656	79,282
Devlet Büyüklüğü	74,422	73,504	70,528	57,400
Mevzuat Verimliliği	76,944	76,246	65,489	72,985
Serbest Piyasa	71,689	83,042	72,700	78,048
Log(Kişi Başına GSYİH)	13,596	15,138	12,621	14,429
Sosyal Destek	10,006	10,826	9,303	10,452
Doğumda Sağlıklı Yaşam Beklentisi	7,104	7,546	6,532	7,637
Yaşam Seçimleri Yapma Özgürlüğü	5,013	6,412	4,242	5,841
Cömertlik	0,776	2,422	0,739	1,699
Yolsuzluk Algısı	1,032	3,376	0,760	2,597
Distopya + Residual	24,970	26,190	24,930	27,411

Şekil 26: 2021 Yılı İçin Kümeler



2021 yılı için kümeler k-ortalamlar yöntemi sonucunda yukarıdaki gibi oluşmuştur. Türkiye yer aldığı kümede Yunanistan, Macaristan, Meksika, Polonya ve Slovakya ülkeleri ile birlikte yer almaktadır.

Şekil 27: 2021 Yılı İçin Karar Ağacı



Karar ağacı oluşturulurken gini endeksinden faydalanılmıştır. Karar ağacı sonuçları incelendiğinde 2021 yılı için kümelerin belirlenmesinde en çok rol oynayan faktörler sırasıyla, devlet büyüklüğü, serbest piyasa, yolsuzluk algısı, mevzuat verimliliği ve sosyal destek olarak belirlenmiştir.

2017-2021 yılları arasında bütün analizler genel olarak değerlendirildiğinde hangi ülkenin kaç kez Türkiye ile aynı kümede yer aldığını gösteren tablo aşağıda verilmiştir.

Tablo 15: Türkiye İle Aynı Kümede Yer Alan Ülkeler

Türkiye İle Aynı Kümede Yer Alan Ülkeler	Kaç Kez Aynı Kümede Yer Aldı	Hangi Yıllarda Aynı Kümede Yer aldı
Meksika	5	2017, 2018, 2019, 2020, 2021
Slovakya	5	2017, 2018, 2019, 2020, 2021
Macaristan	4	2017, 2018, 2020, 2021
Polonya	4	2017, 2018, 2019, 2021
Çekya	3	2017, 2018, 2019
Letonya	3	2017, 2018, 2019
Litvanya	3	2017, 2018, 2019
Güney Kore	3	2017, 2018, 2019
İsrail	1	2017
Yunanistan	1	2021

2017-2021 yılları arasında bütün analizler genel olarak değerlendirildiğinde kümeleme analizine yıllara göre en çok etki eden faktörler sırasıyla aşağıdaki tabloda verilmiştir.

Tablo 16: Kümeleme Analizine En Fazla Etki Eden Faktörler

Analiz Yılı	Kümelemeye En Fazla Etki Eden Faktörler
2017	Hukukun Üstünlüğü, Devlet büyüklüğü
2018	Hukukun Üstünlüğü, Devlet büyüklüğü
2019	Hukukun Üstünlüğü, Devlet büyüklüğü, Doğumda Sağlıklı Yaşam Beklentisi
2020	Hukukun Üstünlüğü, Devlet büyüklüğü, Sosyal Destek
2021	Devlet Büyüklüğü, Serbest Piyasa, Yolsuzluk Algısı, Mevzuat Verimliliği, Sosyal Destek

3.4 UYGULAMA SONUÇ VE TARTIŞMASI

Yapılan analizler sonucunda, 2017 yılı için Türkiye'nin kümesinde yer alan ülkeler Çekya, Macaristan, İsrail, Letonya, Litvanya, Meksika, Polonya, Slovakya ve Güney Kore olmakla birlikte 2017 yılı için kümelerin belirlenmesinde en büyük rolü oynayan faktörler sırasıyla hukukun üstünlüğü ve devlet büyüklüğü olarak belirlenmiştir. 2018 yılı için Türkiye'nin kümesinde yer alan ülkeler Çekya, Macaristan, Letonya, Litvanya, Meksika, Polonya, Slovakya ve Güney Kore olmakla birlikte 2018 yılı için kümelerin belirlenmesinde en büyük rolü oynayan faktörler sırasıyla hukukun üstünlüğü ve devlet büyüklüğü olarak belirlenmiştir. 2019 yılı için Türkiye'nin kümesinde yer alan ülkeler Çekya, Letonya, Litvanya, Meksika, Polonya, Slovakya ve Güney Kore olmakla birlikte 2019 yılı için kümelerin belirlenmesinde en büyük rolü oynayan faktörler sırasıyla hukukun üstünlüğü, devlet büyüklüğü ve doğumda sağlıklı yaşam beklentisi olarak belirlenmiştir. 2020 yılı için Türkiye'nin kümesinde yer alan ülkeler Macaristan, Meksika ve Slovakya olmakla birlikte 2020 yılı için kümelerin belirlenmesinde en büyük rolü oynayan faktörler sırasıyla hukukun üstünlüğü, devlet büyüklüğü ve sosyal destek olarak belirlenmiştir. 2021 yılı için Türkiye'nin kümesinde yer alan ülkeler Yunanistan, Macaristan, Meksika, Polonya ve Slovakya olmakla birlikte 2021 yılı için kümelerin belirlenmesinde en büyük rolü oynayan faktörler sırasıyla devlet büyüklüğü, serbest piyasa, yolsuzluk algısı, mevzuat verimliliği ve sosyal destek olarak belirlenmiştir. Analize 2017-2021 yılları aralığında bakıldığında Meksika ve Slovakya'nın 5 kez, Macaristan ve Polonya'nın 4 kez, Çekya, Litvanya, Letonya ve Güney Kore'nin 3 kez, İsrail ve Yunanistan'ın ise 1 kez Türkiye ile aynı kümede yer aldığı tespit edilmiştir.

SONUÇ

Literatürde yapılan çalışmalar incelendiğinde ekonomik özgürlük endeksi veriseti için genellikle Ward ve K-ortalamalar yöntemleri kullanılırken, dünya mutluluk raporu veriseti için genellikle regresyon analizi, korelasyon analizi ve lojistik regresyon analizi kullanılmıştır. Bu çalışmada ekonomik özgürlük endeksi ve dünya mutluluk raporu veri setleri birleştirilip ülkelere hem ekonomik özgürlük hem de mutluluk çerçevesinden bakmak amaçlanmıştır. Çalışmada kullanılan veri seti için ekonomik özgürlük endeksinden 4 faktör(Hukukun Üstünlüğü, Devlet Büyüklüğü, Mevzuat Verimliliği, Serbest Piyasa) ve dünya mutluluk raporundan 7 faktör Log(Kişi Başına GSYİH), Sosyal Destek, Doğumda Sağlıklı Yaşam Beklentisi, Yaşam Seçimleri Yapma Özgürlüğü, Cömertlik, Yolsuzluk Algısı, Distopya + Residual) kullanılmıştır. İki veri setinde de birçok ülke verisi bulunmaktadır ancak analizlerin sağlıklı olabilmesi amacıyla, veri seti OECD ülkelerine indirgenmiştir. OECD ülkelerindeki veri eksikliğinden kaynaklı bazı ülkeler veri setinden çıkartılmıştır. Türkiye'nin de içinde bulunduğu 31 OECD ülkesi 2017-2021 aralığında 5 yıl için ayrı ayrı analiz edilmiştir. Çalışmada küme sayısının belirlenmesi için, hiyerarşik kümeleme analizi yöntemlerinden Ward yöntemi sonucunda oluşturulan dendrogram ve dirsek yönteminden faydalanılmıştır. Küme sayısına karar verilmesinin ardından oluşturulacak kümelerin homojen olması için hiyerarşik olmayan kümeleme analizi yöntemlerinden K-ortalamalar yöntemi ile kümeler oluşturulmuştur. Kümelerin oluşumunda en büyük rolü oynayan faktörlerin belirlenebilmesi için karar ağacı yöntemlerinden literatürde çokça kullanılmış olan CART algoritması ile gini endeksi kullanılarak bir karar ağacı oluşturulmuştur. 5 yıllık analizlerin sonucunda Meksika, Slovakya, Macaristan ve Polonya'nın, Türkiye ile en az 4 yılda aynı kümede yer aldığı tespit edilmiştir.

KAYNAKÇA

Acar, M. (2010). Serbest Ticaret, Ekonomik Özgürlükler ve Refah. *Türk Dünyası Sosyal Bilimler Dergisi*, 53,1-28.

Alpar, R. (2017). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*. Ankara: Detay Yayıncılık.

Altın, H. Ekonomik Özgürlük Endeksinin Çok Kriterli Karar Verme Yöntemleriyle Analizi/Analysis Of The Economic Freedom Index With Multi-Criteria Decision-Making Methods. *Uluslararası Ekonomi İşletme ve Politika Dergisi*, 4(2), 441-460.

Bradshaw, J., & Rees, G. (2017). Exploring national variations in child subjective well-being. *Children and Youth Services Review*, 80, 3-14.

Bulut, H., (2018). *R Uygulamaları ile Çok Değişkenli İstatistiksel Yöntemler*. Nobel Akademik Yayıncılık, Ankara, Türkiye.

Cantril H. (1965). *The Pattern of Human Concern*. New Brunswick: Rutdgers University Press, New Brunswick.

Coşkun, A. & Rençber, Ö. F. & Bağcı, H. (2021). Ülkelerin girişimcilik ve ekonomik özgürlük özelliklerinin karşılaştırılması: bulanık k-ortalamlar yöntemi ile örnek bir uygulama. *Turkish Studies - Economy*, 16(1), 141-157.

De Neve, J. E., & Ward, G. (2017). *Does Work Make You Happy? Evidence from the World Happiness Report*. Harvard Business Review.

Deveci, E. (2020). *K-En Yakın Komşu Algoritması ile Ekonomik Özgürlük Endeksinin Tahmin Edilmesi: OECD Ülkelerinde Örnek Uygulama* (Yayınlanmamış Yüksek Lisans Tezi), İstanbul: Aydın Üniversitesi Lisansüstü Eğitim Enstitüsü, İstanbul, 2020.

Dünya Ülkelerinde Mutluluk Eşitsizliği-Mutluluğu Belirleyen Faktörlerin Analizi(Doktora Tezi) Olcay SERVET

Everitt, B. 1974. Cluster Analysis. Heinmann.London, p.122

Everitt, B. S. Landau, S., Leese, M. ve Stahl, D. (2011). *Cluster Analysis*. 5th Edition. United Kingdom: John Wiley and Sons. Ltd. 352

George, D. ve Mallery, P. (2019). *IBM SPSS Statistics 26 Step by Step: A Simple Guide and Reference*. Routledge.

Georgescu, I., Kinnunen, J., Androniceanu, A. ve, & Androniceanu, A. M. (2020). A Computational Approach To Economic İnequality, Happiness And Human Development. *Informatica Economica*, 24(4),: 16-28.

Grimm, L. G. ve Yarnold, P. R. (1995). *Reading and Understanding Multivariate Statistics*. USA: American Psychological Association.

Grimm, L. G., & Yarnold, P. R. (2000). *Reading and Understanding MORE Multivariate Statistics*. American Psychological Association.

Gülden, Tansu. *OECD Ülkelerinin Ekonomik Özgürlüklerine Göre Kümeleme Analizi İle Sınıflandırılması*, (Yüksek Lisans Tezi), Sivas Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü, Sivas, 2019.

Güzel, S. (2018). Gelir eşitsizliği, Refah ve Mutluluk, *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi* 6(3):, 389-394. <https://doi.org/10.18506/anemon.322631>

Hair, J. F. Black, W. C., Babin, B. J., Anderson, R. E. ve Tatham, R. L. (2018). *Multivariate Data Analysis*. USA: Pearson.

Hartigan, J. A. (1981). Consistency of Single Linkage for High-Density Clusters. *Journal of the American Statistical Association*, 76(374), 388-394.

Hennig, C., Meila, M., Murtagh, F., & Rocci, R. (Eds.). (2015). *Handbook of Cluster Analysis*. CRC Press.

Heritage, *Free Economies are Clean Economies*, https://www.heritage.org/index/pdf/2021/book/02_2021_IndexOfEconomicFreedom_METHODODOLOGY.pdf

Hickerson, S. R. (1984). Complexity and The Meaning of Freedom: The Classical Liberal View. *American Journal of Economics and Sociology*, 43(1), 91-101.

Irani, J., Pise, N. ve Phatak, M. (2016). Clustering Techniques and the Similarity Measures Used in Clustering: A Survey. *International Journal of Computer Applications*, 134(7), 9-14.

Iyer, L. S., & Aronson, J. E. (1999). A parallel branch-and-bound method for cluster analysis. *Annals of Operations Research*, 90, 65-86.

Johnson, D. E. (1998). Applied Multivariate Methods for Data Analysts. Duxbury Resource Center.

Kalaycı, Şeref (2010). *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri*. Ankara: Asil Yayıncılık.

Kandemir, O. (2021). Dünya Ülkelerinde Sosyal Eşitsizliklerin Mutluluk Düzeyine Etkisi: Lojistik Regresyon Analizi, *International Journal of Eurasia Social Sciences (IJOESS)*, 12(45), 612-625.

Kangallı, S. & Uyar, U. & Buyrukoğlu, S. (2014). OECD ülkelerinde ekonomik özgürlük: bir kümeleme analizi. *Uluslararası Alanya İşletme Fakültesi Dergisi*, 6(3).

Kass, G. V. (1980). "An Exploratory Technique for Investigating Large Quantities of Categorical Data". *Applied Statistics*, 29(2): , 119-127.

Kayri, M., & Boysan, M. (2007). "Araştırmalarda Chaid Analizinin Kullanımı ve Baş Etme Stratejileri İle İlgili Bir Uygulama". *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi.*, 40(2), 133-149.

Kılıcı, E. N. (2019). Analysis Of The Relationship Between Economic Freedom Index And Stock Market Indices; Evidence From Turkey. *Maliye ve Finans Yazıları*, (111), 117-134.

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, 1(6), 90-95.

Kutbay, H. (2020). Ekonomik Özgürlük Ve Vergi Gelir Performansı Arasındaki İlişki: Yükselen Piyasa Ekonomileri İçin Panel Veri Analizi. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (40), 303-318.

Markov, I., Arampatzis, A. ve Crestani, F. (2012, August). Unsupervised Linear Score Normalization Revisited. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Pp. 1161-1162).

Menteş, N . (2020). Mutluluk, Gelir ve Demokrasi: Dünya Ülkeleri İçin Yapısal Eşitlik Modeli Önerisi. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 9(3),2138-2153. <https://doi.org/10.15869/itobiad.734193>

Milligan, G. W. ve M. C. Cooper. (1988). A Study of Standardization of Variables in Cluster Analysis. *Journal of Classification*. 5.2, 181-204.

Okagbue, H. I., Oguntunde, P. E., Bishop, S. A., Adamu, P. I., Akhmetshin, E. M., & Iroham, C. O. (2019). Significant Predictors of Henley Passport Index. *Journal of International Migration and Integration*, 1-12.

Özcan, M., & Akar, G. (2020). E7 Ülkelerinde Ekonomik Özgürlük ve Doğrudan Yabancı Sermaye Yatırımları: Panel Veri Analizi. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (36), 542-570.

Saraç, M. (2021). *İktisadi Büyüme ve Ekonomik Özgürlükler Arasındaki İlişki: OECD Ülkeleri İçin Ekonometrik Veri Analizi (2000-2021)* (Yüksek Lisans Tezi, Niğde Ömer Halisdemir Üniversitesi/Sosyal Bilimler Enstitüsü).

SERVET, O. (2018). *Dünya Ülkelerinde Mutluluk Eşitsizliği-Mutluluğu Belirleyen Faktörlerin Analizi. Yayınlanmamış* (Doktora Tezi. Kocaeli: Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü.) Olcay SERVET

Sharma, S. (1996). *Applied Multivariate Techniques*. USA: John Wiley ve Sons, Inc.

Smith, A. (1937). *The wealth of nations [1776]* (Vol. 11937). na.

Şahin, Dilek (2017). Kümeleme Analizi İle Doğu Avrupa Ülkelerinin Ekonomik Özgürlükler Açısından Değerlendirilmesi, *Hitit Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, Yıl 10, Sayı 2, Aralık 2017, ss.1299-1314

Şeker, M. (2009). *Mutluluk ekonomisi. Sosyoloji Konferansları*, 39, 115-134.

TAŞAR, M. O. (2007). Ekonomik Özgürlük Kavramı Ve Türkiye’de Ekonomik Özgürlüklerin Gelişimi. *Selçuk Üniversitesi Hukuk Fakültesi Dergisi*, 15(1), 143-167.

Taşcı, Melike. (2019). *OECD Ülkelerinin Ekonomik Özgürlük Göstergelerinin K-Ortalamalar Kümeleme Yöntemi Ve Gri İlişkisel Yöntem İle Analizi.*, (Yayınlanmamış Yüksek Lisans Tezi), İstanbul: İstanbul Aydın Üniversitesi Sosyal Bilimler Enstitüsü., İstanbul, 2019.

Teague, J. ve, & Locker, C. (2019). Analyzing The Relationship Between Happiness And European State-Mandated Parental Leave: Lessons For Human Resource Policymakers. *International Journal of Business & Public Administration*, 16(1).

Ulkhaq, M. M. (2020). Clustering Countries According to The World Happiness Report. *Statistica & Applicazioni*, 18(2).

World Happiness Report, What is the theme of this year’s report?, <https://worldhappiness.report/faq/>, 12.06.2021.

Yayla, A., & Hayek, F. A. (2000). *Özgürlük Yolu: Hayek'in sosyal teorisi*. Liberte.

Zhu, X., & Lu, Y., & Xiao, L. ve& Hao, Z. (2020, October). Application of CART algorithm and parametric modeling method in the design of sit-skis. In 2020 6th International Conference on Mechanical Engineering and Automation Science (ICMEAS) (pp. 100-105). IEEE.



EKLER

EK 1: Dirsek Metodu İçin Python Kodu

Aşağıda verilen Python kodları “Google Colaboratory” programı üzerinden çalıştırılmıştır. Çalıştırılan kodun bir kopyası da https://github.com/huseyinkayaa/Thesis_dataset adresinde yer almaktadır.

```
from google.colab import files
uploaded = files.upload()
import io
df = pd.read_csv(io.BytesIO(uploaded['veriseti.csv']))

df2 = pd.DataFrame(df, columns=['EF1_Rule of law','EF2_Government
size','EF3_Regulatory efficiency','EF4_Market openness','WHI_11_Log GDP per
capita','WHI_12_Social support','WHI_13_Healthy life
expectancy','WHI_14_Freedom to make life choices',
'WHI_15_Generosity','WHI_16_Perceptions of corruption','WHI_17_Dystopia +
residual'])
X = df.iloc[:,3:].values

from yellowbrick.cluster import KElbowVisualizer
model = KMeans()
visualizer = KElbowVisualizer(model, k=(1,12)).fit(df2)
visualizer.show()
```

EK 2: Kümeleme Analizi İçin Python Kodu

Aşağıda verilen Python kodları “Google Colaboratory” programı üzerinden çalıştırılmıştır. Çalıştırılan kodun bir kopyası da

https://github.com/huseyinkayaa/Thesis_dataset adresinde yer almaktadır.

```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from sklearn.cluster import KMeans

from google.colab import files
uploaded = files.upload()
import io
df = pd.read_csv(io.BytesIO(uploaded['veriseti.csv']))

df2 = pd.DataFrame(df, columns=['EF1_Rule of law','EF2_Government
size','EF3_Regulatory efficiency','EF4_Market openness','WHI_11_Log GDP per
capita','WHI_12_Social support','WHI_13_Healthy life
expectancy','WHI_14_Freedom to make life choices',
'WHI_15_Generosity','WHI_16_Perceptions of corruption','WHI_17_Dystopia +
residual'])
X = df.iloc[:,3:].values

from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=4, init='k-means++', random_state=0).fit(df2)

kmeans.inertia_
kmeans.n_iter_
kmeans.cluster_centers_
from collections import Counter
Counter(kmeans.labels_)
```

EK 3: Karar Ağacı Oluşturulması İçin Python Kodları

Aşağıda verilen Python kodları “Google Colaboratory” programı üzerinden çalıştırılmıştır. Çalıştırılan kodun bir kopyası da https://github.com/huseyinkayaa/Thesis_dataset adresinde yer almaktadır.

```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
import graphviz
import pydotplus

from google.colab import files
uploaded = files.upload()
import io
df = pd.read_csv(io.BytesIO(uploaded['veriseti.csv']))

feature_cols = ['EF1_Ruleoflaw', 'EF2_Governmentsize',
'EF3_Regulatoryefficiency',
'EF4_Marketopenness', 'WHI_11_LogGDPpercapita', 'WHI_12_Socialsupport', 'WHI_
13_Healthylifeexpectancy', 'WHI_14_Freedomtomakelifechoices', 'WHI_15_Generosi
ty', 'WHI_16_Perceptionsofcorruption', 'WHI_17_Dystopiaresidual']

X = df[feature_cols] # Features
y = df.QCL_1 # Target variable

clf = DecisionTreeClassifier(random_state=0)
clf = clf.fit(X,y)

from sklearn.tree import export_graphviz
from six import StringIO
from IPython.display import Image
import pydotplus
```

```
dot_data = StringIO()
export_graphviz(clf, out_file=dot_data,
                filled=True, rounded=True,
                special_characters=True, feature_names =
feature_cols, class_names=['1','2','3','4'])
graph = pydotplus.graph_from_dot_data(dot_data.getvalue())
graph.write_png('tree.png')
Image(graph.create_png())
```



EK 4: Veri setleri İçin Github Linki

https://github.com/huseyinkayaa/Thesis_dataset



EK 5: Değişkenler İçin Dağılım Matrisi

