

ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

PANEL VERİLERDE ÇOK DEĞİŞKENLİ İSTATİSTİKİ MODELLEME

Fatma Esra UĞURCA

İSTATİSTİK ANABİLİM DALI

ANKARA
2023

Her hakkı saklıdır

ÖZET

Doktora Tezi

PANEL VERİLERDE ÇOK DEĞİŞKENLİ İSTATİSTİKİ MODELLEME

Fatma Esra UĞURCA

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Prof. Dr. Cemal ATAKAN

Panel veri; birimler üzerinden belli bir zaman aralığında yapılan gözlemlerden oluşur. Bununla birlikte panel verinin yapısında zaman serilerinin yer alması, hem özelliklerini hem de problemlerini beraberinde getirir.. Panel veri analizi, bir birimin zaman içindeki değişimi ya da bir değişkenin birimler arası değişimini gösterir ve ilgilenilen değişkenler ile model oluşturarak bir ilişkinin tespit edebilmesini sağlar. Ayrıca, ilgilenilen değişkenler için her bir birim bakımından ilişkileri karşılaştırma imkanı da tanır. Bu çalışmada, 1975-2017 yıllarına ait kişi başına düşen karbondioksit emisyonu, kişi başına gayri safi yurtiçi hasıla ve yenilenebilir enerji için olan gözlemler kullanılmıştır. İlgilenilen değişkenlere göre kümeleme yapıldıktan sonra panel veri analizine geçiş yapılmıştır. Kümeleme analizinde çok değişkenli panel veri kümeleme yöntemleri kullanılarak beş kümeye karar verilmiştir. Elde edilen kümeler yeni birimler olarak adlandırılıp panel veri analizi uygulanmıştır. Kurulan sabit etkili model ile birimlerin etkileri tahmin edilmiş ve uzun dönem ilişkileri incelenmiştir. Ek olarak karbondioksit emisyonu, yenilenebilir enerji, GSYİH ve ülke ortalamalarına göre de kümeler oluşturularak kointegrasyon ilişkileri araştırılmıştır.

Mart 2023, 94 sayfa

Anahtar Kelimeler: Panel veri analizi, kümeleme analizi, çok değişkenli panel veride kümeleme

ABSTRACT

Ph.D. Thesis

MULTIVARIABLE STATISTICAL MODELLING ON PANEL DATA

Fatma Esra UĞURCA

Ankara University
Graduate School of Natural and Applied Sciences
Statistics Dep.

Supervisor: Prof. Dr. Cemal ATAĞAN

Panel data consists of observations over units over a certain period of time. However, the inclusion of time series in the structure of panel data brings both its features and problems. Panel data analysis shows the change of a unit over time or the change of a variable between units and enables the detection of a relationship by creating a model with the variables. It also provides the opportunity to compare the relationships in terms of each unit for the variables. In this study, observations for carbon dioxide emissions per capita, gross domestic product per capita and renewable energy for the years 1975-2017 are used. After clustering according to the variables, panel data analysis is applied. In the clustering analysis, five clusters are decided by using multivariate panel data clustering methods. The clusters obtained are named as new panel units and panel data analysis is applied. With the fixed effect model, the effects of the units are estimated, and long-term relationships are analyzed. In addition, clusters are formed according to carbon dioxide emissions, renewable energy, GDP and country averages and cointegration relationships are investigated.

March 2023, 94 pages

Key Words: Panel data analysis, clustering analysis, multivariable statistics on panel data

TEŞEKKÜR

Yüksek lisans öğrenimimden beri her sorunumda yanına gidebildiğim, zamanını ayıran ve sabırla yön veren kıymetli danışmanım Prof. Dr. Cemal ATAKAN’a teşekkür ederim. Bu süreçte yaşadığım zorluklarda samimiyetle yardım eden, hoşgörüsünü eksik etmeyen ve durumumu kırmadan anlatmaya çalışan hocam Prof. Dr. Yılmaz AKDİ’ye teşekkürlerimi sunarım. Tez çalışmalarım sırasında desteğini hiç esirgemeyen ve her zaman yapabileceğime inanan Prof. Dr. Yeliz YALÇIN’a çok teşekkür ederim.

Her zaman, her anlamda yanımda olan anneme, babama ve kardeşime teşekkür ederim. Bu süreçteki destekleri için eşim Deniz’e ve oğlum Ender’e çok teşekkür ederim.

Fatma Esra UĞURCA
Ankara, Mart 2023

İÇİNDEKİLER

TEZ ONAY SAYFASI

ETİK.....	i
ÖZET.....	ii
ABSTRACT	iii
TEŞEKKÜR	iv
SİMGELER DİZİNİ	vii
ŞEKİLLER DİZİNİ	viii
ÇİZELGELER DİZİNİ	ix
1. GİRİŞ.....	1
2. LİTERATÜR.....	5
3. PANEL VERİ MODELLERİ	9
3.1 Havuzlanmış Regresyon Modeli	11
3.2 Sabit Etkili Model	11
3.2.1 Birim etkisine sahip sabit etkili model	12
3.2.2 Zaman etkisine sahip sabit etkili model	14
3.2.3 Birim ve zaman etkisine sahip sabit etkili model	15
3.3 Rasgele Etkili Modeller.....	15
3.4 Breusch-Pagan Lagrange Çarpanı testi.....	16
3.5 Hausman Testi.....	17
3.6 Panel Veri Modellerinde Varsayımlar	18
3.6.1 Panel veri yatay kesit bağımlılığı.....	18
3.6.2 Panel veride otokorelasyon	19
3.6.3 Panel veride değişen varyans	20
3.7 Driscoll-Kraay Tahmin Edicisi	21
4. KÜMELEME ANALİZİ	23
4.1 Uzaklık Ölçüleri	23
4.1.1 Öklid uzaklığı	24
4.1.2 Minkowski uzaklığı.....	24
4.1.3 Manhattan uzaklığı.....	24
4.2 Bağlantı Yöntemleri.....	25
4.2.1 Hiyerarşik yöntemler	25

4.2.1.1 Tek bağlantı yöntemi	26
4.2.1.2 Tam bağlantı yöntemi	26
4.2.1.3 Ortalama bağlantı yöntemi	27
4.2.1.4 Ward's bağlantı yöntemi	27
4.2.1.5 Merkezi bağlantı yöntemi.....	28
4.2.2 Hiyerarşik olmayan yöntemler	29
4.2.2.1 K- Ortalama yöntemi.....	29
4.3 Küme Sayısına Karar Verme.....	30
4.4 Küme Geçerlilikleri.....	31
4.4.1 Kofenetik korelasyon katsayısı	32
4.4.2 Silhouette indeksi	35
4.4.3 Dunn indeksi.....	36
5. ÇOK DEĞİŞKENLİ PANEL VERİDE KÜMELEME ANALİZİ.....	37
6. UYGULAMA.....	42
6.1 Çok Değişkenli Panel Veride Kümeleme Analizi Uygulaması.....	46
6.2 Panel Veri Analizi Uygulaması	53
6.3 Uzun Dönem İlişkisi	61
7. SONUÇ VE TARTIŞMA.....	69
KAYNAKLAR	73
EK 1	78
EK 2	85
EK 3	90
ÖZGEÇMİŞ.....	94

SİMGELER DİZİNİ

CO ₂	Karbondioksit
ρ	Otokorelasyon
σ^2	Varyans
χ^2	Ki-Kare Dağılımı
REN	Yenilenebilir Enerji
GDP	Gayri Safi Yurt İçi Hasıla

Kısaltmalar

AB	Avrupa Birliği
GSYİH	Gayri Safi Yurt İçi Hasıla
BRICS	Brezilya, Rusya, Hindistan, Çin ve Güney Afrika Cumhuriyeti
LM	Lagrange Çarpanı
AR	Otoregresif
OECD	İktisadi İşbirliği ve Gelişme Teşkilatı
EKC	Çevresel Kuznets Eğrileri
REN	Yenilenebilir Enerji
ADF	Genişletilmiş Dickey-Fuller
AIC	Akaike Bilgi Kriteri
VAR	Vektör Otoregresif

ŞEKİLLER DİZİNİ

Şekil 4.1 Kümelerin uzaklıklarını gösteren dendogram.....	34
Şekil 6.1 CO ₂ emisyonu(kişi başına) için ülkelere ait zaman serisi grafikleri	44
Şekil 6.2 Yenilenebilir enerji için ülkelere ait zaman serisi grafikleri.....	44
Şekil 6.3 Kişi başına GSYİH için ülkelere ait zaman serisi grafikleri.....	45
Şekil 6.4 Bağımlı değişken gözlem değerleri(kırmızı) ile tahmin değerleri(yeşil)	59
Şekil 6.5 Bağımlı değişken gözlem değerleri (kırmızı) ile tahmin değerleri (yeşil) / k-ortalama.....	61



ÇİZELGELER DİZİNİ

Çizelge 4.1 Kofenetik için örnek	33
Çizelge 4.2 Kofenetik korelasyon katsayısı için hesaplama adımları.....	35
Çizelge 5.1 Çok değişkenli panel veri	38
Çizelge 5.2 Düzenlenmiş panel veri.....	40
Çizelge 6.1 Uygulamaya dahil edilen ülkeler ve kodları	43
Çizelge 6.2 Silhoutte indeks değerleri	47
Çizelge 6.3 Dunn indeks değerleri	48
Çizelge 6.4 Calinski Harabasz indeks değerleri.....	48
Çizelge 6.5 Kofenetik Korelasyon Katsayısı sonuçları	49
Çizelge 6.6 Verinin özel düzenlenmiş hali için oluşturulan kümeler	49
Çizelge 6.7 K-Ortalama için Silhoutte, Dunn ve Calinski-Harabasz indeks değerleri ...	51
Çizelge 6.8 K-ortalamaya göre kümeler	51
Çizelge 6.9 Breusch-Pagan LM testi sonuçları	54
Çizelge 6.10 Hausman Test Sonuçları	54
Çizelge 6.11 Panel veride yatay kesit bağımlılığı için Breusch-Pagan LM ve Pesaran CD test.	55
Çizelge 6.12 Panel veride otokorelasyon için Durbin-Watson testi sonuçları	56
Çizelge 6.13 Panel veride değişen varyans için Uyarlanmış Wald testi sonuçları	57
Çizelge 6.14 Parametre tahminleri.....	57
Çizelge 6.15 Breusch-Pagan LM ve Hausman testi sonuçları	59
Çizelge 6.16 Pesaran CD testi, Durbin-Watson ve Breush Pagan testi sonuçları.....	60
Çizelge 6.17 Parametre tahminleri\k-ortalama	60
Çizelge 6.18 ADF birim kök sonuçları	63
Çizelge 6.19 Johansen test sonuçları.....	64
Çizelge 6.20 Karbondioksit emisyonu için oluşturulan kümeler	66
Çizelge 6.21 Yenilenebilir enerji için oluşturulan kümeler	67
Çizelge 6.22 Kişi başı GYSİH için oluşturulan kümeler	68
Çizelge 6.23 Ülke ortalamaları için oluşturulan kümeler	58

1. GİRİŞ

Panel veri; birimler üzerinden belli bir zaman aralığında yapılan gözlemlerden oluşur. Panel veride birimler; hanehalkı, firmalar, ülkeler, bölgeler vb. için kullanılır. Bunlar, panel verinin yatay kesitini oluşturmaktadır. Aynı zamanda panel veri, zaman serileri de içerir. Bir zaman serisi, birim zaman aralıklarında yapılan gözlemlerin bir koleksiyonu olarak alınabilir. Bir ülkedeki yıllık ihracat ve ithalat miktarları, aylık enflasyon oranları, aylık faiz oranları, yıllık gayri safi milli hasıla gibi değişkenler zaman serilerine örnek olarak verilebilir (Akdi, 2010). Hem kesit, hem de zaman boyutlarından oluşan panel verinin çeşitli alanlarda uygulaması da giderek artmaktadır. Temel olarak düşünüldüğünde, birimlerin tekrarlı gözlemlerine ait yapının incelenmesinde bilinen istatistiki yöntemlerin kullanıldığı görülmektedir. Daha basit bir ifadeyle panel veri, hem kesit, hem de zaman boyutlarından meydana geldiği için yapılacak istatistiki incelemelerde kullanım alanı gün geçtikçe artmaktadır.

Panel verinin yapısındaki kesit birimler genellikle heterojen yapıdadır. Ancak bazı durumlarda, özellikle gözlem sayısının artması istendiğinde, heterojenlik göz ardı edilmektedir. Heterojenliği göz ardı etmek, varsayılan modelin eldeki durumunu en iyi şekilde ifade edememesine sebep olacaktır. Eldeki veride heterojenlik durumunu ortaya çıkarmak için kümeleme analizinden yararlanılabilir. Kümeleme analizinde amaç, aynı kümedeki birimlerin homojen kümeler arasındaki heterojen yapıyı ortaya çıkarmak olduğundan, verideki heterojenlik yapısını görebilmek için kümeleme analizinden yararlanılabilir. Böylece panel verideki birimlere kümeleme analizi uygulanarak, birimlerin heterojenliği gösterilebilir. Bu çalışmadaki amaçlardan biri de bu durumu ortaya koymaktır.

Bununla birlikte panel verinin yapısında zaman serilerinin yer alması, hem mevsimsellik, trend gibi özelliklerinin, hem de durağanlık olup olmasına sebep olur. Zaman serilerinde olduğu gibi, panel veride de durağanlığın incelenmesi gerekir. Durağanlık olmadığında uzun dönem ilişkilerinin varlığı incelenebilir. Hem zaman serisi, hem yatay kesitin birarada olması panel verinin bu iki veri yapısının özelliklerine sahip olmasını sağlar. Zaman serilerinde olduğu gibi, yatay kesit veride de kendine özgü problemler ile

karşılaşılabılır. Panel veri bu iki verinin birleşiminden oluştuğu için avantajlarının yanında bazı dezavantajları da beraberinde getirir.

Bunun yanı sıra panel veri analizi, bir çok açıdan veriyi incelememizi sağlar. Bunlardan ilki, birimin zaman içindeki değişimi ya da bir değişkenin birimler arası değişiminin görülebilmesi, ikicisi ilgilenilen değişkenler ile model oluşturarak, bir ilişkinin tespit edilebilmesidir. Üçüncü olarak, ilgilenilen değişkenler için her bir birim bakımından ilişkileri karşılaştırma imkanı sağlamasıdır.

Panel veri analizinde dikkat edilmesi gereken noktalardan biri birimler arası heterojen olma durumudur. Birimler için heterojenlik modele dahil edilebilmektedir. Panellerde yer alacak yanlış birimler parametre tahminlerde yansızlık ve tutarlılık özelliklerinin sağlanmamasına yol açar. Bu problemten kaçınmak için ilgilenilen değişkenlere göre birimleri birarada tutmak gerekir.

Panel veride birimleri kümelemek/gruplamak için çeşitli yöntemler vardır. Bunlardan biri, her bir zaman serisini bir değişken olarak alıp veriyi yeniden düzenleyerek kümeleme analizi uygulamaktır. Zheng ve Li (2014)'nin çalışmasında bu durum açıklayıcı olarak anlatılmıştır. Bu yöntem yardımıyla, panel verinin karmaşık yapısı tek boyut olarak gösterilmesi amaçlanmıştır.

Bu çalışmada, 1975-2017 yıllarına ait CO₂, kişi başına GSYİH ve yenilenebilir enerji olarak tanımlanan gözlemler kullanılmıştır. Ekonomik büyüme ve çevrenin korunması, insanlığın karşı karşıya olduğu iki önemli konudur. Sera gazı (GHG) emisyonlarından kaynaklanan iklim değişikliği ile ilgili endişeler gün geçtikçe artmaktadır. Hem gelişmiş hem de gelişmekte olan ülkelerin bugün alacağı uygun kararlar gelecek nesilleri etkileyecektir. Her ülke, ekonomik büyüme için kararlar alırken bir çok faktörü göz önünde bulundurmalıdır. Ülkeler büyüme oranını yüksek tutmayı teşvik ederken, doğal kaynaklarına dayalı olarak kalkınmayı unutmamalıdır. Büyüme, çevresel koşullar (kirlilik), doğal kaynakların aşırı kullanımı ve iklim değişikliği gibi birçok yönden çevre üzerinde olumsuz etki yaratabilir. Çevre kalitesindeki düşüş, hem mevcut hem de uzun vadeli perspektifte nüfusun yaşam koşulları için ciddi bir sorun oluşturur. Karbondioksit,

küresel ısınmanın ve çevresel bozulmanın ana nedeni olarak kabul edilen sera gazlarından biridir. Gelirin, çevre kalitesindeki rolünü incelerken CO₂ emisyonu sıklıkla kullanılmaktadır. Yenilenebilir enerji, doğal kaynakların aşırı tüketimindeki çıkmaza bir kapı olarak görülmektedir. Literatürde bu özellikleri inceleyen oldukça fazla çalışma bulunmaktadır. Bu çalışmada da ekonomik büyüme, karbondioksit emisyonu ve yenilenebilir enerji arasındaki ilişki incelenecektir. Literatürde genelde ülkelerin gelişmişlik seviyelerine göre paneller oluşturularak bu ilişki incelenmiştir. Bu çalışmada ise gelişmişlik seviyelerine göre panellemek yerine ilgilenilen değişkenler bakımından ülkeler kümelenecek paneller oluşturulmuştur. Oluşturulan bu kümelere panel veri analizi uygulanmıştır.

Bu çalışmada ilgilenilen değişkenlere göre kümeleme yapıldıktan sonra panel veri analizine geçiş yapılacaktır. Bunun sebepleri şöyle sıralanabilir;

- Ele alınan 67 ülkeyi kümeleyerek (gruplayarak) daha az birim ile çalışma imkanı sağlamak,
- Kümeleme yapıldıktan sonra panel veri analizi ile yapılan tahminler yardımıyla bu ülkelerin özelliklerini incelemek,
- Kümeleri inceleyerek ülkelerin bir araya gelmesine neden olan durumu ortaya çıkarıp isimlendirme yapmak,
- Son olarak ise ülke grupları için önerilerde bulunmak.

Ülkeler için önerilerde bulunurken, gruplar arasındaki fark dikkate alınmalıdır. Şöyle ki, her gruba aynı öneride bulunmak gerçekçi bir yaklaşım olmayabilir. Kimi ülkelerde alınan yol ileri iken, kimileri için daha başlangıç olabilir. Başlangıç seviyesindeki ülkeler için gelişmiş ülkelerin sahip olduğu olanaklar varmış gibi önerilerde bulunmak ilerleme kaydetmesini zorlaştırabilir.

Çalışma yedi bölümden oluşmaktadır. Bu bölümün ardındaki ikinci bölümde konuyla ilgili önceki çalışmalardan bahsedilmiştir. Üçüncü bölümde, panel veri ile ilgili tanımlar, modeller, varsayımlar ve tahmin yöntemleri üzerinde durulmuştur. Dördüncü bölümde, kümeleme analizi ele alınmıştır. Beşinci bölümde, çok değişkenli panel veride kümeleme

analizi açıklanmaya çalışılmıştır. Altıncı bölümde ise uygulama verilmiştir. Uygulamada; kümeleme, panel veri analizi ve uzun dönemi ilişkisi olmak üzere üç alt başlık vardır. Son olarak çalışma için sonuç ve tartışma bölümü yer almaktadır.



2. LİTERATÜR

Panel veri analizi, veriyi bir çok açıdan incelemeyi sağladığından pek çok bilim alanında uygulaması ve çeşitli gelişmeleri vardır. Bu bölümde, panel verinin kümelenmesi ve uygulamada kullanılan değişkenlerle ilgili çalışmalardan bahsedilecektir.

Banzo ve Hermosilla(2002), panel veri için hiyerarşik olmayan kümeleme analizi, yapısal koruma, hesaplama zorlukları ve bağımlılık problemlerinden dolayı zorlukları olduğunu ve bunu PASA(perturbed adaptive simulated annealing) yöntemini genetik algoritmanın içine yerleştirerek bu problemlerin üstesinden geldiklerini ifade etmişlerdir.

Nie ve ark. (2012) panel veri kümelemesi için yeni bir uzaklık ölçüsü önermişlerdir. Önerdileri uzaklık ölçüsü ağırlıklandırılmış ve ağırlıklandırılmamış olarak hesaplanabilmektedir. Uygulamalarında, önerdikleri yeni uzaklık ölçüsünün performansını incelemek için ticari bankaların kredi kartlarına ait panel veri setini kullanmışlardır.

Lu ve ark. (2014) çalışmasında iki problemi ele almışlardır. Birinci problem, birimlerin havuzlanabilirliği ve ikinci problem ise değişken seçimidir. Sosyal bilimlerde gözlem sayısının, bazı durumlarda, yeterince elde edilememesi nedeniyle birimlerin heterojen yapısı göz ardı edilerek havuzlanması yoluna başvurulduğunu aktarmışlardır. Bu durumun güvenilir bir analiz uygulamasını desteklemediğini vurgulamışlardır. İkinci problem olan değişken seçiminde ise tüm bağımsız değişkenlerin bağımlı değişken ile anlamlı bir ilişkisi olmayabileceği durumudur. Bu durumdaki değişkenlerin regresyon modelinde yer alması yanlış tahminlere yol açabilir. Bu problemleri gidermek için ‘Seçme ve Kümeleme’ olarak adlandırdıkları bir optimizasyon modelini ortaya koymuşlardır.

Baştürk ve ark. (2012), heterojenliği veri odaklı şekilde ele almışlardır. 1971-2000 yılları için Asya, Latin ve Orta Amerika ve Afrika’daki 59 ülkenin ekonomik büyümesi üzerine çalışmışlardır. Farklı büyüme özellikleri olan ülkeleri gruplandırırken bir ön bilgi (coğrafi

konum, vb.) kullanmak yerine, ülkeler arasındaki farklılıkları ve benzerlikleri büyüme belirleyicilerinin üzerindeki etkisini incelemek için sonlu karma panel modelden yararlanmışlardır.

Frühwirth-Shnatter (2011), bilinen uzaklık ölçülerinin panel veya boylamsal (longitudinal) verideki zaman serisine kolaylıkla genişletemeyeceğini aktarmıştır. Sonlu karma modele dayalı kümeleme için önerilen yaklaşımları incelemiştir. Çeşitli yaklaşımların hem sürekli hem de kategorik zaman serisine ilişkin gözlemleri için uygun olacağını aktarmıştır. Markov zinciri Monte Carlo yöntemleri aracılığıyla Bayesgil tahmini detaylı olarak tanımlanmıştır ve küme sayısına karar vermek için çeşitli kriterler incelenmiştir. Uygulamada ise bir tür keyif verici madde olan marijiana kullanan gençler üzerinden oluşturulan panel veri incelenmiştir.

Juarez ve Steel (2010), panelin içindeki birimleri kümelemek için modele dayalı bir yöntem önermişlerdir. Altta yatan model, otoregresif ve Gaussian olmayan bir modeldir. Bu model ile hem basıklık hem de kalın kuyruğu olan durumlarda kullanılabilir. Ayrıca birimler kümelenirken onların dinamik davranışları, denge seviyeleri ve ortak değişkene etkilerine göre kümelenmektedir. Sonuç çıkarımları Bayesgil bakış açısıyla irdelenmiştir ve model karşılaştırmalarında Bayes faktörleri kullanılmıştır. Ön bilgi için biraz öznel girdi yapılması tavsiye edilmiştir. Uygulama, Avrupa bölgesindeki GSYİH büyümesi ve İspanyol firmalarının istihdam büyümesi üzerine olmuştur.

Kosmelj ve Batagelj (1990), birimler, değişkenler ve zaman olmak üzere üç-modlu veri matrisi üzerinde kümeleme analizini uygulamışlardır. Zaman içinde değişen iki birim arasındaki uzaklığı hesaplamak için genel bir model sunmuşlardır. Belirli bir model geliştirilmiş ve 23 Avrupa ülkesini, 1976-1982 yılları arasındaki enerji tüketimine göre kümelemişlerdir.

Lombardo ve Falcone (2011) çalışmasında, suç göstergeleri, ekonomik performans göstergeleri ve sosyo-demografik değişkenlere göre kümeleme analizi ile panel zaman serisinin kombinasyonu olan yeni bir yöntem önererek İtalyan şehirlerinin benzerliklerine göre sınıflandırmışlardır.

Bu çalışmanın uygulama bölümü için önceki çalışmalar incelenmiş ve özet halde sunulmaya çalışılmıştır. Menyah ve Wolde-Rufael (2010) çalışmalarında, 1960–2007 yılları için ABD’de karbondioksit emisyonları, yenilenebilir ve nükleer enerji tüketimi ve reel GSYİH arasındaki nedensel ilişkiyi araştırmışlardır. Granger nedensellik testinin değiştirilmiş bir versiyonunu kullanarak, nükleer enerji tüketiminden CO₂ emisyonuna tek yönlü nedensellik bulmuşlar, ancak yenilenebilir enerjiden CO₂ emisyonlarına hiç bir nedensellik bulamamışlardır. Ekonometrik kanıtlara bakıldığında, nükleer enerji tüketiminin CO₂ emisyonlarını azaltmaya yardımcı olabileceğini öne sürüyor gibi görüldüğünü, ancak şimdiye kadar yenilenebilir enerji tüketiminin, emisyon azaltımına önemli bir katkı sağlayabilecek bir seviyeye ulaşmadığının altını çizmişlerdir.

Apergis vd (2010) çalışmalarında, panel hata düzeltme modeli kullanarak 1984–2007 yılları için toplam 19 gelişmiş ve gelişmekte olan ülke grubunu CO₂ emisyonu, nükleer enerji tüketimi, yenilenebilir enerji tüketimi ve ekonomik büyüme arasındaki nedensel ilişkiyi incelemişlerdir. Uzun vadeli tahminler, nükleer enerji tüketimi ile emisyon arasında istatistiki olarak anlamlı negatif bir ilişki olduğunu, ancak emisyon ve yenilenebilir enerji tüketimi arasında istatistiki olarak anlamlı pozitif bir ilişki olduğunu göstermektedir. Panel Granger nedensellik testlerinden elde edilen sonuçlar, kısa vadede nükleer enerji tüketiminin CO₂ emisyonlarının azaltılmasında önemli bir rol oynadığını, yenilenebilir enerji tüketiminin ise emisyonlardaki azalmaya katkıda bulunmadığını göstermektedir. Bunun nedeni, elektrik üreticilerinin en yüksek yük talebini karşılamak için emisyon üreten enerji kaynaklarına güvenmek zorunda olabileceklerini vurgulamışlardır. Problemleri durumların üstesinden gelmek için yeterli depolama teknolojisinin bulunmaması bu durumu ortaya çıkarmış olabileceğinden bahsetmişlerdir.

Bölük ve Mert (2014) çalışmalarında, 16 AB (Avrupa Birliği) ülkesi için 1990-2008 yıllarını incelemişlerdir. Değişkenler olarak fosil ve yenilenebilir enerji tüketimi, sera gazları ve geliri olarak ele almışlardır. Açıklayıcı değişkenler arasındaki çoklu bağlantı problemini çözerek panel veri analizi uygulamışlardır. Yenilenebilir enerji tüketimi, fosil enerji tüketimine göre tüketilen birim enerji başına sera gazı emisyonunu yaklaşık %50 daha az katkı sağlamasıdır. Bu durumun, enerji tüketiminde, alternatif yenilenebilir enerji teknolojilerine geçişin sera gazı emisyonlarını azaltabileceği vurgulamışlardır.

Long vd (2015) çalışmalarında, 1952 yılından 2012 yılına kadar Çin'deki enerji tüketimi, karbon emisyonları ve ekonomik büyüme arasındaki ilişkiyi incelemişlerdir. Kömürün ekonomik büyüme ve karbon emisyonları üzerinde baskın etkisi olduğunu söylemişlerdir. GSYİH, karbondioksit emisyonu, kömür, gaz ve elektrik tüketimi ile iki yönlü bir ilişki olduğunu ortaya koymuşlardır. Çin'deki enerji tüketim yapısının değiştirilmesi zorunlu olduğunu vurgulamışlardır.

Saidi ve Mbarek (2016) çalışmalarında, 1990-2013 yıllarına ilişkin dokuz gelişmiş ülke için dinamik panel kullanarak nükleer enerji tüketimi, CO₂ emisyonları, yenilenebilir enerji ve kişi başına reel GSYİH arasındaki nedensel ilişkiyi araştırmışlardır. Sermaye ve iş gücünü ek değişkenler olarak almışlardır. Kısa vadede, yenilenebilir enerji tüketimi ile GSYİH arasında tek yönlü bir nedensellik olduğunu göstermişlerdir. Bu durumu, enerji tüketimini azaltmaya yönelik politikaların ekonomik büyümeyi ve geliri geciktiremeyeceği şeklinde yorumlamışlardır. Yenilenebilir enerji tüketimi ile GSYİH uzun vadede incelediklerinde çift yönlü nedensellik bulmuşlar ve yenilenebilir enerjinin ekonomik büyüme için çok önemli bir bileşen olduğunu vurgulamışlardır.

Dong vd (2017) çalışmalarında, 1985-2016 yılları için BRICS ülkelerindeki CO₂ emisyonu, GSYİH, doğal gaz ve yenilenebilir enerji tüketimini ele almışlardır. Sonuçlarda, doğal gaz ve yenilenebilir enerji tüketiminin artması, CO₂ emisyonunu düşürdüğü görülmüştür. Doğal gaz ve CO₂ arasında kısa ve uzun vadeli çift yönlü ve yenilenebilir enerji ile CO₂ arasında da kısa ve uzun vadeli çift yönlü nedensellik olduğunu söylemişlerdir.

Yao vd (2019) çalışmalarında, yenilenebilir enerji ile GSYİH arasındaki ilişkiyi incelemişlerdir. 1990-2014 yılları arasında dünyanın altı jeo-ekonomik bölgesinin yanı sıra 17 büyük gelişmekte olan ve gelişmiş ülkenin iki panel veri setini kullanarak dinamik ilişkiyi araştırmışlardır.

3. PANEL VERİ MODELLERİ

Panel veri, ilgilenilen birimlerin belli zaman aralığında araştırılmak istenen durumla ilgili gözlemlerin toplanması ile oluşturulur. Panel veri ile daha çok iktisadi çalışmalarda karşılaşmaktadır. Hem zaman hem de kesit boyutu ele alınabildiğinden, sosyo-ekonomik çalışmalarda önemli bir yeri vardır. Aynı zamanda meteoroloji, jeoloji ve ziraat gibi diğer alanlardaki araştırmalar için de kullanılmaktadır.

Zaman serileri birçok bilim dalındaki uygulamalarda sıklıkla kullanılır. Zaman serilerinde önemli amaçlardan biri, ilgili rasgele değişkenin gelecekte alacağı değerlerin öngörülmesidir. Öngörü yapabilmek için öncelikle serinin durağan olması gerekir. Eğer seri durağan değil ise herhangi bir dönüşüm ile serinin durağanlığı sağlandıktan sonra istatistiki sonuç çıkarımlar yapılmalıdır (Akdi, 2010).

Yatay kesit verileri ise birden fazla birimin belirli bir zamanda yapılan gözlemlerinden oluşur. Bu veri türünün zaman serilerinde olduğu gibi kendine ait problemleri vardır. En önemlilerinden biri homojen olmama durumudur. Yatay kesit veriler ile zaman serilerinin birleşmesi sonucu oluşan panel veri hem yatay kesit hem de zaman serilerinin özelliklerini içermektedir (Gujarati, 2004). Yatay kesitte bulunan gözlemlerin zaman boyunca tekrarı olduğundan, panel veri analizinin tekrarlı varyans analizi olduğu da söylenebilir.

Panel veri, belli zaman periyodunda, hanehalkı, firmalar, ülkeler gibi grupların kesitinden alınan gözlemlerin birleştirilmesi ile meydana gelmektedir (Baltagi, 2005). Panel regresyon modeli aşağıda gösterilen aşamalara göre oluşturulmaktadır:

i. Y_i , $i=1, 2, \dots, n$ ile gösterilen kesitli veri, belli noktalarda bireyler, hanehalkı, ülkeler ve şehirler gibi birimler üzerinde ölçülen değişkene ait gözlemlerden oluşur.

ii. Y_{it} , $t=1, 2, \dots, T$ ile gösterilen zaman serisi verisi ise zaman içinde bir değişken üzerinden alınan gözlemlerden meydana gelir.

iii. Y_{it} , $i=1, 2, \dots, n$, $t=1, 2, \dots, T$ ile gösterilen panel verisi ise her bir kesit için zaman serisinden oluşur.

Panel regresyon modeli;

$$Y_{it} = \alpha + \beta_{1it}X_{1it} + \beta_{2it}X_{2it} + \dots + \beta_{pit}X_{pit} + u_{it}, i=1, 2, \dots, n, t=1, 2, \dots, T \quad (3.1)$$

biçiminde gösterilir. Yukarıda (3.1) eşitliğindeki i , hanehalkı, bireyler, firmalar ve ülkeleri belirtirken, t indisi de zamanı göstermektedir. Y bağımlı değişkeni, α sabit katsayıyı, $\beta_{1it}, \beta_{2it}, \dots, \beta_{pit}$ regresyon katsayılarını, $X_{1it}, X_{2it}, \dots, X_{pit}$ p tane açıklayıcı(bağımsız) değişkeni ve u_{it} sıfır ortalama ve sabit varyansa sahip hata terimini göstermektedir (Baltağı 2005, Asteriou ve Hall, 2006).

Panel veri özellikleri özetlenecek olursa;

- Panel veri birimlerin hem belli bir zaman aralığındaki hem de birimler arasındaki özelliğini ortaya koyar.
- Panel veri T gibi belli bir zaman aralığında gözlenen n tane birimi içerir.
- Tüm birimler aynı zaman aralığında gözlenmiş (tüm i 'ler için, $T_i = T$) ise dengeli panel, eğer aynı değil ise ($T_i \neq T$) dengeli olmayan/dengesiz panel veri denir.
- Bir birim için alınan gözlemlerde kendi içerisinde korelasyon görülürken, rasgele seçilen birimlere ait gözlemler arasında olabilir. Örneğin aynı bireyin yıllar itibarı ile elde ettiği gelirleri arasında bir korelasyon oluşurken, diğer bireylerin gelirleri ile bu korelasyon görülemez.

Panel regresyon modelleri oluşturulurken sabit, eğim katsayıları ve hata terimi üzerinde yapılan çeşitli varsayımlar dikkate alınır. Aşağıdaki maddelerde bir kaç durum verilmiştir:

1. Sabit ve eğim katsayılarının birimlere ve zamana göre değişmeyen durum modeli,
2. Sabit katsayının birimlere göre değiştiği, eğim katsayılarının değişmediği durum modeli,
3. Sabit katsayının hem birimlere hem de zamana göre değiştiği, eğim katsayılarının değişmediği durum modeli,

4. Tüm katsayıların birimlere göre değiştiği durum modeli ve
5. Tüm katsayıların hem birimlere hem de zamana göre değiştiği durum modelidir.

Yukarıda verilen modellere ilişkin özellikler arttıkça, daha doğru tahminler yapılabilmektedir (Gujarati 2004).

Panel veri modelleri birimlerin hem zaman içerisinde hem de birimler arasındaki davranışa göre modellenir. Havuzlanmış(pooled) regresyon modeli, sabit etkili model ve rasgele etkili model panel verinin temel üç modelini oluşturur.

3.1 Havuzlanmış Regresyon Modeli

Havuzlanmış regresyon modelinde yer alan parametrelerin birimlere ve zamana göre değişmediği varsayılmaktadır. İki boyuta göre verinin homojen olduğu durumlarda kullanılır. Panel veri modeli sadece bir sabit katsayı içeriyorsa, bilinen en küçük kareler (OLS) tahmin edicisi sabit eğim vektörü için tutarlı tahmin ediciler verir (Greene, 2002)

Bu model,

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_p X_{pit} + u_{it} \quad i=1, 2, \dots, n, t=1, 2, \dots, T \quad (3.2)$$

biçiminde gösterilir. Bu model oldukça kısıtlayıcı olduğundan uygulamalarda çok sık kullanılmamaktadır.

3.2 Sabit Etkili Model

Sabit etkili modeller tek veya iki yönlü olmaktadır. Eğer sadece birim veya zaman etkisi araştırılıyor ise tek yönlü sabit etkili model, hem birim hem de zaman etkisi inceleniyorsa iki yönlü sabit etkili model kurulmaktadır. Bunlar aşağıdaki kısımlarda verilmiştir.

3.2.1 Birim etkisine sahip sabit etkili model

Birim etkisine sahip sabit etkili modelde, sabit katsayının birimlere göre farklılık gösterdiği varsayılır. Bu durum, her bir birim için farklı katsayıların modelde yer alması anlamına gelir.

Bu model,

$$Y_{it} = \alpha_i + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_p X_{pit} + u_{it} \quad i=1, 2, \dots, n, t=1, 2, \dots, T \quad (3.3)$$

biçiminde gösterilir. Burada α_i zamana göre değişiklik göstermez, birimlere göre farklılık gösterir. Aynı zamanda grup-özel etkisi denmektedir (Greene 2002, Astreiou ve Hall 2006).

Bu modele ait parametre tahmin yöntemlerinden biri kukla(dummy, gösterge) değişkenli en küçük kareler(LSDV) yöntemidir. Bu yöntemi daha açık bir şekilde görmek için birim sayısı 3, zaman uzunluğu 5 ve açıklayıcı değişken sayısı 2 olan bir örnek üzerinden ilerlemek faydalı olacaktır. Yukarıdaki (3.3) eşitliğine göre,

$$Y_{it} = \alpha_i + \beta_1 X_{1it} + \beta_2 X_{2it} + u_{it} \quad i=1, 2, 3, t=1, 2, 3, 4, 5 \quad (3.4)$$

açık bir şekilde yazılırsa,

$$Y_{11} = \alpha_1 G_{11} + \alpha_2 G_{21} + \alpha_3 G_{31} + \beta_1 X_{111} + \beta_2 X_{211} + u_{11} \quad (3.5)$$

.

$$Y_{15} = \alpha_1 G_{11} + \alpha_2 G_{21} + \alpha_3 G_{31} + \beta_1 X_{115} + \beta_2 X_{215} + u_{15} \quad (3.6)$$

$$Y_{21} = \alpha_1 G_{12} + \alpha_2 G_{22} + \alpha_3 G_{32} + \beta_1 X_{121} + \beta_2 X_{221} + u_{21} \quad (3.7)$$

.

$$Y_{25} = \alpha_1 G_{12} + \alpha_2 G_{22} + \alpha_3 G_{32} + \beta_1 X_{125} + \beta_2 X_{225} + u_{25} \quad (3.8)$$

$$Y_{31} = \alpha_1 G_{13} + \alpha_2 G_{23} + \alpha_3 G_{33} + \beta_1 X_{131} + \beta_2 X_{231} + u_{31} \quad (3.9)$$

.

$$Y_{35} = \alpha_1 G_{13} + \alpha_2 G_{23} + \alpha_3 G_{33} + \beta_1 X_{135} + \beta_2 X_{235} + u_{35} \quad (3.10)$$

Buradan hareketle,

$$\underline{Y} = \begin{bmatrix} Y_{11} \\ \vdots \\ Y_{15} \\ Y_{21} \\ \vdots \\ Y_{25} \\ Y_{31} \\ \vdots \\ Y_{35} \end{bmatrix}_{15 \times 1}, \quad \underline{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}_{3 \times 1}, \quad G = \begin{bmatrix} \underline{1}_{5 \times 1} & \underline{0}_{5 \times 1} & \underline{0}_{5 \times 1} \\ \underline{0}_{5 \times 1} & \underline{1}_{5 \times 1} & \underline{0}_{5 \times 1} \\ \underline{0}_{5 \times 1} & \underline{0}_{5 \times 1} & \underline{1}_{5 \times 1} \end{bmatrix}_{15 \times 3}, \quad \underline{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}_{2 \times 1},$$

$$X = \begin{bmatrix} X_{111} & \dots & X_{115} & X_{121} & \dots & X_{125} & X_{131} & \dots & X_{135} \\ X_{211} & \dots & X_{215} & X_{221} & \dots & X_{225} & X_{231} & \dots & X_{235} \end{bmatrix}_{2 \times 15} \quad \text{ve} \quad \underline{u} = \begin{bmatrix} u_{11} \\ \vdots \\ u_{15} \\ u_{21} \\ \vdots \\ u_{25} \\ u_{31} \\ \vdots \\ u_{35} \end{bmatrix}_{15 \times 1}$$

biçiminde ifade edilir. Burada,

$$\underline{1}_{5 \times 1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}_{5 \times 1} \quad \text{ve} \quad \underline{0}_{5 \times 1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}_{5 \times 1}$$

vektörleridir. Bu örnekten yararlanarak genel matris notasyonları,

$$\underline{Y} = \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \vdots \\ \underline{Y}_n \end{bmatrix}_{nT \times 1}, \quad \underline{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}_{n \times 1}, \quad G = \begin{bmatrix} \underline{1}_{Tx1} & 0 & 0 & \dots & 0 \\ 0 & \underline{1}_{Tx1} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \underline{1}_{Tx1} \end{bmatrix}_{nT \times n},$$

$$\underline{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}_{px1}, \quad X = \begin{bmatrix} \underline{X}_{11} & \underline{X}_{12} & \cdots & \underline{X}_{1n} \\ \underline{X}_{21} & \underline{X}_{22} & \cdots & \underline{X}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{X}_{p1} & \underline{X}_{p2} & \cdots & \underline{X}_{pn} \end{bmatrix}_{pxnT}, \quad \underline{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}_{nTx1}$$

biçiminde elde edilmiştir. Burada, $\underline{Y}_1, \underline{Y}_2, \dots, \underline{Y}_n$ 'ler, $\underline{X}'_{11}, \underline{X}'_{12}, \dots, \underline{X}'_{pn}$ 'ler ve $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n$ 'ler $T \times 1$ boyutlu vektörlerdir. Birim etkisine sahip sabit etkili model matris notasyonu ile verildiğinde,

$$\underline{Y}_{nTx1} = G\underline{\alpha}_{nx1} + X'_{pxnT}\underline{\beta}_{px1} + \underline{u}_{nTx1} \quad (3.11)$$

biçiminde elde edilir. Bu model “Kukla Değişkenli En Küçük Kareler” olarak tanımlanır. Bu model sayesinde parametre tahminleri yapılabilir (Astreiu and Hall 2006, Greene 2002).

3.2.2 Zaman etkisine sahip sabit etkili model

Zaman etkisine sahip sabit etkili model, birim etkisine benzer şekilde düşünülebilir. Teknolojik ilerlemeler, çeşitli vergilerin düzenlenmesi ya da savaş gibi dış etkenler sebebiyle zaman etkileri görülebilir (Gujarati 2004). Bu modelde sabit katsayının zamana göre farklılık gösterdiği varsayılır.

Burada γ_t zamana göre farklılık gösterir. Model,

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \gamma_1 g_1 + \gamma_2 g_2 + \dots + \gamma_T g_T + u_{it}, i=1, 2, \dots, n, t=1, 2, \dots, T \quad (3.12)$$

biçiminde yazılabilir. Burada $\gamma_1, \dots, \gamma_T$ bilinmeyen katsayılardır Böylece kukla değişkenli en küçük kareler tahmin edicileri elde edilebilir (Stock ve Watson 2019).

3.2.3 Birim ve zaman etkisine sahip sabit etkili model

Birim ve zaman etkisine sahip modeller, iki yönlü sabit etkili modellerdir. Birim etkili modele zaman etkisini de dahil edip model,

$$Y_{it} = \beta_1 X_{1it} + \beta_2 X_{2it} + \alpha_1 G_{1i} + \alpha_2 G_{2i} + \dots + \alpha_n G_{2n} + \gamma_1 g_{1t} + \dots + \gamma_T g_{Tt} + u_{it} \quad (3.13)$$

biçiminde oluşturulabilir. Benzer şekilde n tane birime ait ve T tane zamana ait ikili göstergeler ve sabit katsayısı ile birlikte model yukarıdaki şekilde oluşturulur (Stock ve Watson 2007).

3.3 Rasgele Etkili Modeller

Rasgele etkili modellerde birim, zaman, hem birim hem de zaman etkileri rasgele değişken olarak düşünülmektedir ve modele hata teriminin bir bileşeni olarak eklenir. Rasgele etkili model,

$$Y_{it} = \alpha_i + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_p X_{pit} + u_{it} \quad i=1, 2, \dots, n, t=1, 2, \dots, T \quad (3.14)$$

biçiminde tanımlanır. Modelde her birimin ortalaması,

$$\alpha_i = \alpha + \varepsilon_i, i=1, 2, \dots, n \quad (3.15)$$

biçiminde ifade edilir. ε_i , ortalaması 0, varyansı σ_ε^2 'dir. α_i (3.14) eşitliğinde yerine yazıldığında,

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_p X_{pit} + \varepsilon_i + u_{it} \quad i=1, 2, \dots, n, t=1, 2, \dots, T \quad (3.16)$$

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_p X_{pit} + w_{it} \quad i=1, 2, \dots, n, t=1, 2, \dots, T \quad (3.17)$$

elde edilir. Burada,

$$w_{it} = \varepsilon_i + u_{it}, \quad (3.18)$$

biçimindedir. w_{it} , ε_i birim etkilerini ya da hata bileşeni ve u_{it} ise zaman serisi ve birimlerin birleşiminden meydana gelen hata bileşenidir. Bu iki hata bileşeninden oluştuğu için w_{it} , hata bileşenleri modeli olarak adlandırılır (Gujarati 2004).

Hata bileşenlerine ait varsayımlar,

$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$$

$$u_{it} \sim N(0, \sigma_u^2)$$

$$E(\varepsilon_i u_{it}) = 0, E(\varepsilon_i \varepsilon_j) = 0, \quad (i \neq j) \quad (3.19)$$

$$E(u_{it} u_{is}) = E(u_{it} u_{jt}) = E(u_{it} u_{js}) = 0, \quad (i \neq j; t \neq s) \quad (3.20)$$

olarak verilmiştir. Böylece,

$$E(w_{it}) = 0, \quad (3.21)$$

$$Var(w_{it}) = \sigma_\varepsilon^2 + \sigma_u^2 \quad (3.22)$$

olarak bulunur. (3.22) eşitliğinde yer alan $\sigma_\varepsilon^2 = 0$ ise model, havuzlanmış panel regresyon modeline dönüşür. Bu durumu kontrol etmek için çeşitli testler uygulanır (Gujarati 2004).

3.4 Breusch-Pagan Lagrange Çarpanı Testi

Breusch ve Pagan (1980) rasgele etkiler için bir Lagrange çarpanı testi geliştirmişlerdir. Havuzlanmış en küçük kareler modeline karşılık rasgele etkili modeli test etmek, yani bireysel heterojenliğin varlığını sınamak için bu testi önermişlerdir. Bu testin yokluk hipotezi, rasgele etkili modelin artıklarına ait varyansın sıfır olması üzerine kuruludur. Hipotezler;

$$H_0: \sigma_\varepsilon^2 = 0$$

$$H_1: \sigma_\varepsilon^2 > 0$$

şeklindedir ve test istatistiği,

$$BP = \frac{nT}{2(T-1)} \left[\frac{\sum_{i=1}^n [\sum_{t=1}^T \hat{u}_{it}]^2}{\sum_{i=1}^n \sum_{t=1}^T \hat{u}_{it}^2} - 1 \right]^2 \quad (3.23)$$

biçiminde verilmiştir. BP test istatistiği, 1 serbestlik dereceli χ^2 dağılımına sahiptir. Yokluk hipotezi reddedildiğinde, rasgele birim etkilerinin varlığı kabul edilir (Baltagi 2005).

3.5 Hausman Testi

Sabit etki ve rasgele etki modellerinin seçiminde en sık kullanılan testlerden biri Hausman (1978) testidir. Hausman (1978) çalışmasının başlangıcında,

$$\underline{Y} = X' \underline{\beta} + \underline{\varepsilon} \quad (3.24)$$

biçiminde basit regresyon modelini vermiştir ve bu modelden yola çıkarak açıklamaya devam etmiştir. Burada,

$$i) E(\underline{\varepsilon}|X) = 0 \quad (3.25)$$

$$ii) Var(\underline{\varepsilon}|X) = \sigma^2 I \quad (3.26)$$

varsayımları üzerinde durulmuştur. ii. varsayım, yani ortogonal olması varsayımı sağlanmadığında yanlış tahmin edici, i. varsayım sağlanmadığında ise tahmin edicide etkinlik kaybı olacağını aktarmıştır.

Yukarıda verilen sabit etkili ve rasgele etkili iki model arasındaki en önemli fark, açıklayıcı değişkenler ile hata terimi arasındaki korelasyonun varlığıdır. Eğer korelasyon var ise sabit etkili, yok ise rasgele etkili model kullanılır (Wooldridge 2002).

Hausman (1978) testi için hipotezler,

$$H_0: Cov(\alpha_i, X_{it}) = 0$$

$$H_1: Cov(\alpha_i, X_{it}) \neq 0$$

şeklinde dir. Hausman testi için test istatistiği,

$$H_{Hes} = (\underline{\hat{\beta}}_{SE} - \underline{\hat{\beta}}_{RE})' [Var(\underline{\hat{\beta}}_{SE}) - Var(\underline{\hat{\beta}}_{RE})]^{-1} (\underline{\hat{\beta}}_{SE} - \underline{\hat{\beta}}_{RE}) \quad (3.27)$$

biçimindedir. Burada $\underline{\hat{\beta}}_{SE}$ ve $\underline{\hat{\beta}}_{RE}$ sırasıyla sabit etkili ve rasgele etkili model parametre tahminlerini ifade etmektedir. $Var(\underline{\hat{\beta}}_{SE})$ ve $Var(\underline{\hat{\beta}}_{RE})$ ise sırasıyla sabit etkili ve rasgele etkili modellerin tahmininden elde edilen vektörlerin varyans-kovaryans matrisleridir Hausman testi p serbestlik dereceli χ^2 dağılımına sahiptir. Burada p değişken sayısıdır (Hausman 1978, Gujarati 2004, Baltagi 2005).

3.6 Panel Veri Modellerinde Varsayımlar

Panel veri modellerinin temelinde yatay kesit bağımlılığı, otokorelasyon ve değişen varyans problemlerinin olmadığı varsayımlarına dayanmaktadır. Bu varsayımlardan bir ya da birkaçının sağlanmaması, parametrelerin yanlış tahmin edilmesine yol açmaktadır. Bu sebeple varsayımların geçerli olup olmadığının test edilmesi gerekir. Varsayımların sağlanmadığı durumlarda uygun sağlam(robust) tahmin edicilerden yararlanılır.

3.6.1 Panel veri yatay kesit bağımlılığı

Panel veri modeli varsayımlarında, yatay kesit (birimler) birbirinden bağımsız olarak kabul edilmektedir. Ancak uygulamalarda birimler arasında bağımlılık ile karşılaşmaktadır. Panel verideki seriye belli bir şok geldiğinde tüm yatay kesit birimlerinin ilgili şoktan aynı derecede etkilenip etkilenmediği incelenir. Yatay kesit bağımlılığı, diğer adıyla birimler arası korelasyon araştırılır. Hata terimlerinde yatay kesit bağımlılığı durumunun var olmasına göre kullanılacak tahmin yöntemlerinin seçimi büyük önem taşır.

Literatürde yatay kesit bağımlılığını incelemek için bir çok yöntem bulunmaktadır. Bu testlerin uygulanabilmesi için zaman boyutu (T) ve yatay kesit boyutu (n) önemlidir. En

sık kullanılan yatay kesit bağımlılığı testlerinden birisi Breusch-Pagan (1980) LM testidir. Bu test, $T > n$ durumunda sabit etkili modelin yatay kesit bağımlılığını araştırmak için uygundur. Hem sabit hem de rasgele etkili modellerde kullanılabilen Pesaran (2004) test istatistiği, ikili korelasyon katsayılarının ortalamasına dayanır. Bu test $n \geq T$ durumunda kullanılabilmektedir ve $n \rightarrow \infty$ olduğu durumlarda tutarlı sonuçlar vermektedir (Pesaran 2004). Test istatistiği,

$$CD_{LM} = \sqrt{\frac{2T}{n(n-1)}} (\sum_{i=1}^{n-1} \sum_{j=i+1}^n \hat{\rho}_{ij}) \quad (3.28)$$

biçiminde tanımlanmıştır. Burada $\hat{\rho}_{ij}$ i 'inci birim ile j 'inci birimin artıkları arasındaki anlık korelasyonu göstermekte olup,

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^T \hat{u}_{it} \hat{u}_{jt}}{(\sum_{t=1}^T \hat{u}_{it}^2)^{1/2} (\sum_{t=1}^T \hat{u}_{jt}^2)^{1/2}} \quad (3.29)$$

şeklinde ifade edilir. Burada $\hat{u}_{it}$ panel veri modelinden tahmin edilen artıklardır. Pesaran testinde $i \neq j$ için hipotezler,

$$H_0: \rho_{ij} = \text{Corr}(u_{it}, u_{jt}) = 0,$$

$$H_1: \rho_{ij} \neq 0$$

biçiminde tanımlanır. “Yatay kesit bağımlılığı yoktur” sıfır hipotezi altında test istatistiği asimptotik olarak $n(n-1)/2$ serbestlik dereceli χ^2 dağılımına sahiptir (Pesaran 2004).

3.6.2 Panel veride otokorelasyon

Panel veri, hem kesit hem de zaman boyutunu bünyesinde barındırdığı için serilerin otokorelasyonu incelenmelidir. Bilindiği gibi zaman serileri analizinde zaman serilerin otokorelasyonları incelenerek başlanır. Burada da zaman serilerinin panel veride yer alması durumunda uygulanan otokorelasyon analizleri ele alınacaktır.

Otokorelasyon durumunda yapılan tahminler hala tutarlı olacak ancak etkin olmayacaktır. Bununla birlikte elde edilen standart hatalar da sapmalı olacaktır.

Durbin-Waston testi zaman serilerinde otokorelasyonun varlığını test etmede kullanılan en temel testlerdendir. Bhargava vd. (1982), genelleştirilmiş Durbin-Watson test istatistiğini önermişlerdir. Hipotezler,

$$H_0: \rho = 0 ,$$

$$H_1: \rho < 1$$

şeklinde oluşturulmuştur. Test istatistiği,

$$DW = \frac{\sum_{i=1}^n \sum_{t=2}^T (\hat{u}_{it} - \hat{u}_{it-1})^2}{\sum_{i=1}^n \sum_{t=1}^T \hat{u}_{it}^2} \quad (3.30)$$

biçiminde hesaplanır. Burada $\hat{u}_{it} = y_{it} - x'_{it}\hat{\beta} - \hat{u}_i$ dir. $\hat{\beta}$ ve $\hat{u}_i$ sırasıyla β ve u 'nun en küçük kareler tahminleridir (Baltagi 2005, Asteriou ve Hall 2006).

Yukarıdaki model öncelikle AR(1) kalıntıları kullanılarak sabit etkiler varsayımı ile tahmin edilmektedir. Her iki otokorelasyon test istatistiğinin değeri 2'den küçük ise birinci dereceden otokorelasyonun olmadığını belirten yokluk hipotezi reddedilmektedir. DW değeri 0 ile 4 arasında değer alır. Değer 0'a yakın ise aralarındaki korelasyonun pozitif, 4'e yakın ise negatif korelasyon göstergesidir.

3.6.3 Panel veride değişen varyans

Panel veri modelleri ile ilgili varsayımlardan biri, birimler arası hata terimlerinin varyanslarının sabit olduğudur. Uygulamada ise değişen varyans sorunu ile sıklıkla karşılaşmaktadır. Sabit etki modelinde değişen varyans sorunu doğrudan panel veri modelinin hata teriminden kaynaklanır. Birimlere göre değişen varyans gözlenir. Değişen varyans sorunu test edilirken uyarlanmış Wald testi ve Breush Pagan testlerinden yararlanılır.

Uyarlanmış Wald testinde hipotezler,

$$H_0: \sigma_i^2 = \sigma^2, i=1, 2, \dots, n$$

$$H_1: \sigma_i^2 \neq \sigma^2$$

şeklinde oluşturulmuştur. Test istatistiği,

$$W = \sum_{i=1}^n \frac{(\hat{\sigma}_i^2 - \widehat{\sigma^2})^2}{f_{ii}} \quad (3.31)$$

biçiminde ifade edilmiştir. Yukarıda yer alan $\hat{\sigma}_i^2$ birimlerin hata varyans tahmini ve f_{ii} aşağıdaki eşitliklerde verilmiştir.

$$\hat{\sigma}_i^2 = \frac{1}{T} \sum_{t=1}^{T_i} \hat{u}_{it}^2, \quad f_{ii} = \frac{1}{T} \frac{1}{T-1} \sum_{t=1}^{T_i} (\hat{u}_{it}^2 - \hat{\sigma}_i^2)^2 \quad (3.32)$$

W test istatistiği (3.32) eşitliklerinden elde edilmiştir. n serbestlik dereceli χ^2 dağılımına uymaktadır (Greene 2002).

3.7 Driscoll-Kraay Tahmin Edicisi

Panel veri modelinde bulunan parametre tahminleri için gerekli varsayımların biri ya da bir kaç gerçekleşmediğinde, diğer durumlarda olduğu gibi, burada da sağlam (robust) tahmin edici yöntemlerden yararlanılır. sağlam tahmin ediciler sayesinde bir tahmin edicide aranan özellikler sağlanır.

Panel veri uygulamasında yatay kesit bağımlılığı, otokorelasyon ve değişen varyans varsayımların sağlanmadığı durumlar oldukça sık karşılaşılmaktadır. Eğer bu varsayımlar kontrol edilmez ise tahminlerin sapmalı ve tutarsız olmasına neden olur (Hoechle 2007).

Driscoll ve Kraay (1998) tahmin edicisi ise birimlerin ortalamasına dayanan ve modeldeki birim boyutuna ve zaman boyutuna bağlı olmayan tutarlı bir yöntemdir. Bu yöntem, havuzlanmış en küçük kareler modelinde,

$$\underline{Y} = \underline{\beta}'X + \underline{u} \quad (3.33)$$

biçiminde tanımlansın. Bu modelde, hata terimi vektörü ($\underline{u}$) ile bağımsız değişkenlerin matrisi (X) aralarında korelasyonun olmadığı varsayımı altındaki tahmin edicisi,

$$\underline{\hat{\beta}} = (X'X)^{-1}X'y \quad (3.34)$$

olarak elde edilir. Driscoll ve Kraay standart hataları ise, kovaryans matrisinin sağlam(robust) diagonal elemanlarının karekökleri yardımıyla

$$V(\underline{\hat{\beta}}) = (X'X)^{-1}\hat{S}_T(X'X)^{-1} \quad (3.35)$$

olarak elde edilir. Burada yer alan $\hat{S}_T$, Newey ve West (1987) tarafından

$$\hat{S}_T = \hat{\Omega}_0 + \sum_{j=1}^{m(T)} w(j, m)[\hat{\Omega}_j + \hat{\Omega}_j'] \quad (3.36)$$

olarak tanımlanmıştır. Eşitlikte $m(T)$ otokorelasyon için gecikme uzunluğunu ifade etmektedir ve

$$w(j, m) = (1 - j)/\{m(T) + 1\} \quad (3.37)$$

olarak uyarlanmış Bartlett ağırlıklarıdır. Böylece $\hat{S}_T$ 'nin pozitif yarı tanımlı olmasını sağlar ve yüksek düzeydeki gecikmelerin ağırlığını azaltarak düzenler (Driscoll ve Kraay 1998, Hoechle 2007).

Diğer bir ifade olan

$$\hat{\Omega}_j = \sum_{t=j+1}^T h_t(\hat{\beta})h_{t-j}(\hat{\beta})' \quad (3.38)$$

$$h_t(\hat{\beta}) = \sum_{i=1}^{n(t)} h_{it}(\hat{\beta}) \quad (3.39)$$

biçiminde tanımlanır. Her bir birim için t moment koşullarının karesi,

$$h_{it}(\hat{\beta}) = x_{it}\hat{u}_{it} = x_{it}(y_{it} - \hat{\beta}'x_{it}) \quad (3.40)$$

şeklinde hesaplanır. Bu düzeltme sayesinde Driscoll-Kraay tahmin edicisi dengesiz panelde de kullanılır (Driscoll ve Kraay 1998, Hoechle 2007).

4. KÜMELEME ANALİZİ

Kümeleme analizinde gruplandırma yapısı hakkında ya çok az ya da hiç bilgi bulunmaz. Problem sıklıkla ‘doğal gruplardan’ birini belirlemek olarak tanımlanır. Kümeleme analizinin amacı, farklı gruplardaki birimler arasında düşük ve aynı gruptaki birimler arasında yüksek olan ‘doğal ilişkinin’ derecesini dikkate alarak atama yapmaktır (Anderberg, 1973).

Kümeleme analizinin aşağıdaki özel amaçlarından da söz edilebilir. Bunlar,

- Gerçek tiplerin (cinslerin-ırkların) belirlenmesi
 - Gruplar için ön tahmin
 - Hipotezlerin testi
 - Veri yapısının netleştirilmesi
 - Veri indirgemesi (veriler yerine kümelerin değerlendirilmesi)
 - Aykırı değerlerin bulunması
- şeklinde özetlenebilir (Everit 1974).

Kümeleme analizine başlarken çeşitli uzaklık ölçüleri ve bağlantı yöntemlerinden yararlanılır. Aşağıdaki bölümde bu çalışmada kullanılan uzaklık ölçüleri ile bağlantı yöntemlerine değinilmiştir.

4.1 Uzaklık Ölçüleri

Uzaklıklar veya benzerlikler kümeleme analizinin temel adımını oluşturur. Gözlemler arasındaki benzerlik ya da farklılıklar kullanılarak çeşitli bağlantı yöntemleri ile kümeler oluşturulur. Ele alınan verinin özellikleri göz önüne alınarak uzaklık ölçülerine karar verilir. Bu çalışma için kullanılan uzaklık ölçülerinin tanımları aşağıdaki bölümlerde verilmiştir.

4.1.1 Öklid uzaklığı

Öklid uzaklığı en sık kullanılan uzaklık ölçüsüdür. Basit olarak çok boyutlu uzayda geometrik uzaklıktır ve

$$d_{ij} = \sqrt{\sum_{m=1}^p (x_{im} - x_{jm})^2} \quad (4.1)$$

biçiminde hesaplanır. Burada,

d_{ij} : i . ve j . birimin birbirine olan uzaklığı

x_{im} : i . birimin m . değişken değeri

x_{jm} : j . birimin m . değişken değeri

göstermektedir. $i=1, \dots, n$; $j=1, \dots, n$ ve $m=1, \dots, p$ 'dir. n kümelenecek birim sayısı ve p değişken sayısıdır (Tatlidil 1996).

4.1.2 Minkowski uzaklığı

Uzaklıkların belirlenmesinde bir diğer ölçü Minkowski olarak bilinen uzaklık ölçüsüdür. Minkowski uzaklık ölçüsü,

$$d_{ij} = [\sum_{m=1}^p |x_{im} - x_{jm}|]^{\frac{1}{q}} \quad (4.2)$$

olarak tanımlanmıştır. x_{im} ve x_{jm} eşitlik (4.1)'de verilmiştir. Minkowski uzaklık $q=1$ için City-Block(Manhattan) uzaklık ölçüsüne, $q = 2$ için ise Öklid uzaklık ölçüsüne eşit olacaktır. Görüldüğü üzere Minkowski uzaklık ölçüsü genel bir uzaklık ölçüsüdür (Anderberg, 1973).

4.1.3 Manhattan uzaklığı

Manhattan uzaklık ölçüsü, birimler arası mutlak uzaklıkların toplamı alınarak hesaplanan bir uzaklık ölçüsüdür. Bu uzaklık aşağıdaki formül yardımı ile hesaplanmaktadır.

$$d_{ij} = \sum_{m=1}^p |x_{im} - x_{jm}| \quad (4.3)$$

biçiminde ifade edilir. x_{im} ve x_{jm} eşitlik (4.1)'de verilmiştir.

Manhattan uzaklık ölçüsü uygulamada bazı sorunlara neden olmaktadır. Bu sorunlardan en belirginini Manhattan uzaklık ölçüsünün değişkenler arasında korelasyon olmadığını varsaymasıdır. Araştırmada tanımlanan değişkenler arasında anlamlı korelasyon varsa Manhattan uzaklık ölçüsüne göre hesaplanan uzaklık değerlerine göre yapılan kümeleme anlamlı olmayacaktır. Sorunlardan bir diğeri de değişkenlerin ölçüm birimlerinin farklı olması durumunda standartlaştırılmış Karesel Öklid uzaklığıyla karşılaştırıldığında Manhattan uzaklık ölçüsünün anlamlı sonuçlar vermediği görülmüştür (Johnson ve Wichern 1998).

4.2 Bağlantı Yöntemleri

Kümeleme analizinde, uzaklıklar belirlendikten sonra, bağlantı yöntemlerine geçilir. Bu aşamada ilgilenilen soruna dayalı olarak bağlantı yöntemine karar verilir. Yönteme karar verirken küme içi benzerliklerin ve küme arası farklılıkların en yüksek olması amaçlanır. Uygulamada en çok karşılaşılan hiyerarşik (aşamalı) ve hiyerarşik olmayan yöntemler üzerinde durulacaktır. Bu çalışmada kullanılan yöntemler dikkate alınarak aşağıdaki bölümlerde açıklanacaktır.

4.2.1 Hiyerarşik yöntemler

Aşama sıralı olarak da bilinen bu yöntem, birleştirerek ya da bölerek kümeleri oluşturur. Toplamalı aşamalı kümelemede başlangıçta verideki her bir birim bir küme olarak ele alınır. Birleştirme işlemleri kümeler tek bir küme oluncaya kadar sürmektedir. Bölünmeli hiyerarşik kümeleme de ise, ilk olarak bütün gözlemlerin bir küme oluşturduğu kabul edilir, gözlemleri aşamalı olarak kümelere ayırır.

Toplamalı ya da bölmeli yöntemle oluşturulan kümeler, dendogram adı verilen iki boyutlu diyagram yardımıyla gösterilir (Everitt 1973). Toplamalı hiyerarşik yöntemler

arasında en çok tek bağlantı, tam bağlantı, ortalama bağlantı, Ward ve merkezi bağlantı yöntemleri ile karşılaşılmaktadır.

4.2.1.1 Tek bağlantı yöntemi

Tek bağlantı yöntemi en yakın komşuluk olarak da bilinir. Birbirine en yakın ya da aralarındaki uzaklık en küçük olacak şekilde birimler birleştirilir. En küçük uzaklığa sahip i ve j nci birimler birleştirilmiş ise birleştirilerek oluşturulan bu kümenin diğer k . küme ile uzaklığı,

$$d_{k(i,j)} = \text{Min}(d_{ki}, d_{kj}) \quad (4.4)$$

olarak ifadesinden elde edilir. Yukarıdaki eşitlikte;

$d_{k(i,j)}$; k .kümenin daha önce oluşan (i,j) kümesine olan uzaklığını,
 d_{kj} ; k . kümenin j . kümeye olan uzaklığını,
 d_{ki} ; k . kümenin i . küme ile olan uzaklığını

göstermektedir (Johnson ve Wichern 1998). Benzer şekilde işlemler devam ederek tüm birimler tek bir kümede toplanacak şekilde süreç tamamlanır ve 1 ve n arasında uygun küme sayısı belirlenir.

4.2.1.2 Tam bağlantı yöntemi

Tam bağlantı yöntemi, en uzak komşuluk olarak da bilinmektedir. Tek bağlantı yöntemine benzer şekilde ilerler. Burada da en küçük uzaklığa sahip i ve j nci birimler birleştirilmiş ise birleştirilen bu kümenin diğer k . küme ile uzaklığı,

$$d_{k(i,j)} = \text{Max}(d_{ki}, d_{kj}) \quad (4.5)$$

şeklinde ifade edilmektedir. Bu yönteme tam bağlantı denmesinin nedeni, bir küme içindeki tüm birimlerin birbirlerine maksimum uzaklık veya minimum yakınlığa bağlı olmasıdır (Green 1978).

4.2.1.3 Ortalama bağlantı yöntemi

Yönteme tek ve tam bağlantıda olduğu gibi işleme başlanmaktadır. Fakat kümeleme kriteri olarak, bir küme içerisindeki birim ile diğer kümedeki her bir birim ile uzaklıkların ortalaması alınarak . Ortalama bağlantı yönteminde kümeler küçük varyanslar ile birbirine bağlıdır. Bu yöntem; tam bağlantı ve tek bağlantı yöntemleri arasında sonuçlar vermesi sebebiyle bir alternatif teknik olarak önerilmektedir (Hubert 1974). Ortalama bağlantı uzaklıklar,

$$d_{k(i,j)} = \frac{\sum_i \sum_j d_{ij}}{n_k n_{(i,j)}} \quad (4.6)$$

biçiminde hesaplanır. Burada, d_{ij} (i, j) kümesindeki i . birimin arasındaki uzaklık

$d_{k(i,j)}$: (i, j) kümenin k . küme ile olan uzaklığını,

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

d_{jk} : j . kümenin k . kümeye olan uzaklığını,

n_k : k . kümedeki toplam birey sayısını,

$n_{(i,j)}$: (i, j) . kümedeki toplam birey sayısını göstermektedir (Rencher 2002, Özdamar 2004).

4.2.1.4 Ward's bağlantı yöntemi

En küçük varyans tekniği olarak da bilinen bu yöntem, bir kümenin ortasında yer alan birimin, aynı kümenin içinde bulunan birimlerinden ortalama mesafeyi ele almaktadır. Aynı zamanda yöntem hata kareler toplamından yararlanır (Ward, 1963). Yöntemin ilk basamağında her gözlem bir küme olduğundan Hata Kareler Toplamı sıfır olmaktadır (Everitt 1974).

K tane küme olduğu varsayalım. Bu durumda Hata Kareler Toplamı,

$$HKT = HKT_1 + HKT_2 + \dots + HKT_K \quad (4.7)$$

şeklinde yazılabilir. Yöntemin her bir adımında, kümelerin olabilecek her birleşimi göz önüne alınır ve Hata Kareler Toplamını en küçük arttıracak olan (minimum loss information) birimler aynı küme için seçilir. Hata Kareler Toplamı,

$$HKT = \sum_{k=1}^K \sum_{i=1}^{n_k} (\underline{x}_{ki} - \underline{\bar{x}}_k)^2 \quad (4.8)$$

biçiminde ifade edilir. Burada $\underline{x}_{ki}$ k. kümenin i. biriminin p tane değişkenine ait gözlem ve $\underline{\bar{x}}_k$ k. kümenin ortalama vektörleridir (Johnson and Wichern 1998).

k. kümenin Hata Kareler Toplamı ise,

$$HKT_k = \sum_{i=1}^{n_k} (\underline{x}_{ki} - \underline{\bar{x}}_k)^2 \quad k=1,2,\dots,K \quad (4.9)$$

biçiminde ifade edilir.

4.2.1.5 Merkezi bağlantı yöntemi

Merkezi bağlantı yönteminde, A ve B gibi iki küme arasındaki uzaklık, ikisinin ortalama vektörlerinin Öklid uzaklığı olarak tanımlanır. Bu uzaklık,

$$D(A,B) = d(\underline{\bar{x}}_A, \underline{\bar{x}}_B) \quad (4.10)$$

biçiminde gösterilir. Eşitlikteki $\underline{\bar{x}}_A$, ve $\underline{\bar{x}}_B$ sırasıyla A ve B kümesinin ortalama vektörleridir ve

$$\underline{\bar{x}}_A = \frac{1}{n_A} \sum_{i=1}^{n_A} \underline{x}_{Ai}, \quad \underline{\bar{x}}_B = \frac{1}{n_B} \sum_{i=1}^{n_B} \underline{x}_{Bi} \quad (4.11)$$

biçiminde gösterilir. n_A ve n_B sırasıyla A ve B kümelerindeki birim sayısını göstermektedir. Merkezleri arasında en küçük uzaklığa sahip olan iki küme birleştirilir. A ve B birleştirildikten sonra, AB'nin yeni merkezi,

$$\underline{\bar{x}}_{AB} = \frac{n_A \underline{\bar{x}}_A + n_B \underline{\bar{x}}_B}{n_A + n_B} \quad (4.12)$$

şeklinde oluşturulur (Rencher 2002).

4.2.2 Hiyerarşik olmayan yöntemler

Bu yöntemlerin başlıca özelliği, verinin kaç küme ile çalışılacağı önceden belirlenmesidir. Hiyerarşik olmayan yöntemlerde küme sayısı ile ilgili bir ön bilgi vardır veya küme sayısına karar verilmiştir. Bu durumda, daha uzun işlem gerektiren hiyerarşik yöntemler yerine hiyerarşik olmayan yöntemler kullanılmaktadır (Anderberg, 1973).

Hiyerarşik olmayan yöntemler birimlerin ilk parçalarından ya da kümelerin çekirdeklerinden oluşan merkez noktaların ilk kümesiyle başlar. İlk kümeler için seçim serbesttir. Başlangıçta, birimler arasındaki çekirdek noktalar rasgele seçilir veya ilk gruplardaki birimler rasgele parçalanır (Johnson ve Wichern, 1998). Bu durum, hiyerarşik kümeleme yönteminde dezavantaja sebep olmaktadır (Blashfield ve Aldenderfer 1978).

Bu çalışmada, en sık kullanılan k-ortalama yönteminden yararlanılmıştır.

4.2.2.1 K- Ortalama yöntemi

K-ortalama yöntemi, kümelere ayırırken küme merkezi olarak her kümenin ortalamasını alır.

Bu yöntemde aşağıdaki adımlar izlenir.

Adım 1: Veri seti, başlangıç için keyfi olarak k kümeye ayrılır ve her bir kümenin merkezi veya ortalamaları hesaplanır.

Adım 2: Her birim, kümelerin ortalamasına dayanarak en yakın olan kümeyle atanır. Burada yakınlık hesaplanırken birim ile küme merkezi(ortalaması) arasındaki uzaklık hesaplanır.

Adım 3: Her birim atandıktan sonra yeniden oluşan kümelerin merkezleri(ortalamaları) hesaplanır.

Adım 4: Birimlerin kümelerinde değişiklik olmayana kadar devam eder.

Uygulamada iyi sonuçlar elde etmek için, k-ortalama yöntemini farklı başlangıç küme merkezleriyle birden çok kez çalıştırmak yaygındır (Han vd 2012).

4.3 Küme Sayısına Karar Verme

Küme sayısına karar verirken baştan itibaren göz önünde bulundurulması gerekenler vardır. Analiz uygulanırken ilk adım görsel çıktılarından yararlanılır. Dendogram, oluşabilecek küme yapısını özet bir şekilde karşımıza sunar. Çeşitli uzaklık ve bağlantı yöntemleri kullanılarak oluşturulan dendogramlar incelenir. Araştırmacıya genel olarak kümelenmeyle ilgili fikir verir.

Küme sayısına karar verirken belli amaçlar vardır. Küme sayısının olabildiğince az olması istenen bir durumdur. Oluşan kümeler arasında anlamlı fark olması beklenir. Küme içi uzaklığın en az, kümeler arası en çok elde edilmeye çalışılır.

Küme sayısına karar vermede bazı kriterler:

$$1. \quad k = \sqrt{(n/2)} \quad (4.13)$$

biçiminde verilen eşitliktir. (4.13) eşitliğinde k küme sayısını, n ise tüm birim sayısını göstermektedir. Örneklemin küçük olduğu durumlar için tavsiye edilmektedir (Everit, 1974).

Marriott(1971) tarafından geliştirilen eşitlik,

$$2. \quad M = k^2 |W| \quad (4.14)$$

şeklinde verilmiştir. Yukarıdaki (4.14) eşitliğinde verilen k küme sayısını ve W grup içi kareler-çarpımlar toplamı matrisini göstermektedir. M değerini en küçük yapan k , küme sayısı olarak alınır.

Grup içi kareler-çapraz çarpımlar toplamı matrisi,

$$W = \sum_{i=1}^k \sum_{j=1}^{n_j} (\underline{x}_{ij} - \bar{\underline{x}}_i)(\underline{x}_{ij} - \bar{\underline{x}}_i)' \quad (4.15)$$

biçiminde verilmiştir. Burada k küme sayısını ve n_j j . kümedeki eleman sayısını göstermektedir. Ayrıca,

$\underline{x}_{ij}$: i . kümedeki j . birim gözlem vektörü,

$\bar{x}_i$: i . kümenin örneklem ortalama vektörü

biçiminde hesaplanır (Marriot 1971).

3. Calinsky Harabasz İndeksi

Calinsky ve Harabasz (1974) çalışmalarında önerilen yöntem

$$CH(k) = \frac{tr(B)/(k-1)}{tr(W)/(n-k)} \quad (4.32)$$

biçiminde tanımlanmıştır. Eşitlikteki B gruplar arası kareler ve çapraz çarpımlar toplamı matrisidir ve

$$B = \sum_{l=1}^k n_l (\bar{x}_l - \bar{x})(\bar{x}_l - \bar{x})' \quad (4.33)$$

biçiminde ifade edilmiştir. W ise grup içi (Küme içi) kareler ve çapraz çarpımlar toplamı matrisi (4.15) eşitliğinde verilmiştir. $CH(k)$ değerini en büyük yapan k uygun küme sayısıdır.

Küme sayısına karar verdikten sonra küme geçerliliği yöntemlerinden yararlanarak oluşan kümelerin yapısı değerlendirilir.

4.4 Küme Geçerlilikleri

Uygulanılan kümeleme yöntemlerinden elde edilen sonuçların karşılaştırılması küme belirlemede yardımcı olur. Bu sonuçlar belli yöntemler ile değerlendirilir. Küme sayısını belirleme yöntemlerinin biri ile önerilen uygun küme sayısı, başka bir kritere ile geçerli olmayabilir. Bu nedenle küme sayısının belirlenmesi, bazı geçerlilik işlemleri ile birlikte ele alınması uygun olacaktır.

Kümeleme sonuçlarının geçerliliği, kümeleme analizinin en zor ve önemli kısmını oluşturmaktadır. Dikkat edilmesi gereken noktalar ve araştırmacının amacı kümelemeyi yönlendirir. Kümeleme analizinin sonuçlarını değerlendirmek için küme geçerliliği yöntemlerine ihtiyaç vardır.

Kümeleme çözümlerinin kalitesini test etmek amacıyla geliştirilen bazı küme geçerlilik yöntemleri aşağıda verilmiştir.

4.4.1 Kofenetik korelasyon katsayısı

Hiyerarşik kümeleme yöntemleri kullanıldığında geçerli olan bu yöntem (Sokal & Sneath, 1963) tarafından önerilmiştir (Aldenderfer ve Blashfield 1984). Söz konusu korelasyon, dendogramların performansını belirlemek için kullanılır (Sokal ve Rohlf 1962).

Kofenetik korelasyon katsayısı, başlangıçta hesaplanan uzaklık matrisi(Dis) ile dendogram oluşturulduktan sonra bağlantı uzaklıklarını içeren uzaklıklar matrisinin(R), karşılıklı elemanları arasındaki ilişkiyi ölçen bir ilişki katsayısıdır ve CPCC ile gösterilir ve formül (Hawkins 1982).

CPCC ile gösterilir ve formül

$$CPCC = \frac{\sum_{i < j} (d_{ij} - \bar{d})(r_{ij} - \bar{r})}{\sqrt{\sum_{i < j} (d_{ij} - \bar{d})^2 \sum_{i < j} (r_{ij} - \bar{r})^2}} \quad (4.16)$$

Biçimindedir Burada,

d_{ij} : Dis matrisindeki i .ve j . birimler arasındaki uzaklığı,

$\bar{d}$: Dis matrisindeki tüm d_{ij} uzaklıklarının ortalamasını,

r_{ij} : R matrisinde i .ve j . birimler arasındaki uzaklığı,

$\bar{r}$: R matrisindeki r_{ij} uzaklıklarının ortalamasını

göstermektedir. Kofenetik korelasyon katsayısı $[-1,1]$ aralığında değer alır ve 1'e ne kadar yakın değer alırsa kümeleme de o kadar başarılı sayılır (Erişoğlu 2011).

Burada daha iyi anlaşılması için küçük bir örnek ile devam edilecektir. Örnekte A, B, C ve D birimlerine ait X_1 ve X_2 değişkenlerine ilişkin gözlem değerleri verilmiştir. Öklid uzaklığı ile tam bağlantı yöntemi kullanılarak kümeler oluşturulacaktır.

Çizelge 4.1 Kofenetik için örnek

Birimler	Gözlem Değerleri	
	$\underline{x}_1$	$\underline{x}_2$
<i>A</i>	5	3
<i>B</i>	-1	1
<i>C</i>	1	-2
<i>D</i>	-3	-2

Bu birimlere ait Öklid uzaklıkları (4.1) eşitliğinden yararlanarak yazılırsa,

$$d_{AB} = \sqrt{(x_{1A} - x_{1B})^2 + (x_{2A} - x_{2B})^2} = \sqrt{(5 - (-1))^2 + (3 - 1)^2} = \sqrt{40} = 6.33$$

sonucu elde edilir. Buna göre hesaplamalar benzer şekilde yapılır. Öklid uzaklıkları matrisi,

$$Dis = \begin{matrix} & \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{bmatrix} 0 & & & \\ 6.33 & 0 & & \\ 6.40 & 3.61 & 0 & \\ 9.43 & 3.61 & 4 & 0 \end{bmatrix} \end{matrix} \quad (4.18)$$

şeklinde oluşturulur. Tam bağlantı yöntemine göre uzaklık matrisindeki en küçük değer ile birleştirme işlemine başlanır. Buna göre *B* ve *C* birimleri birleştğinde diğer uzaklıklar,

$$d_{(BC)A} = \max\{d_{AB}, d_{AC}\} = \max\{6.33, 6.40\} = 6.40 \quad (4.19)$$

$$d_{(BC)D} = \max\{d_{BD}, d_{CD}\} = \max\{3.61, 4\} = 4 \quad (4.20)$$

olarak hesaplanır. Yeni uzaklık matrisi,

$$Dis = \begin{matrix} & \begin{matrix} (BC) \\ A \\ D \end{matrix} & \begin{bmatrix} 0 & & \\ 6.40 & 0 & \\ 4 & 9.43 & 0 \end{bmatrix} \end{matrix} \quad (4.21)$$

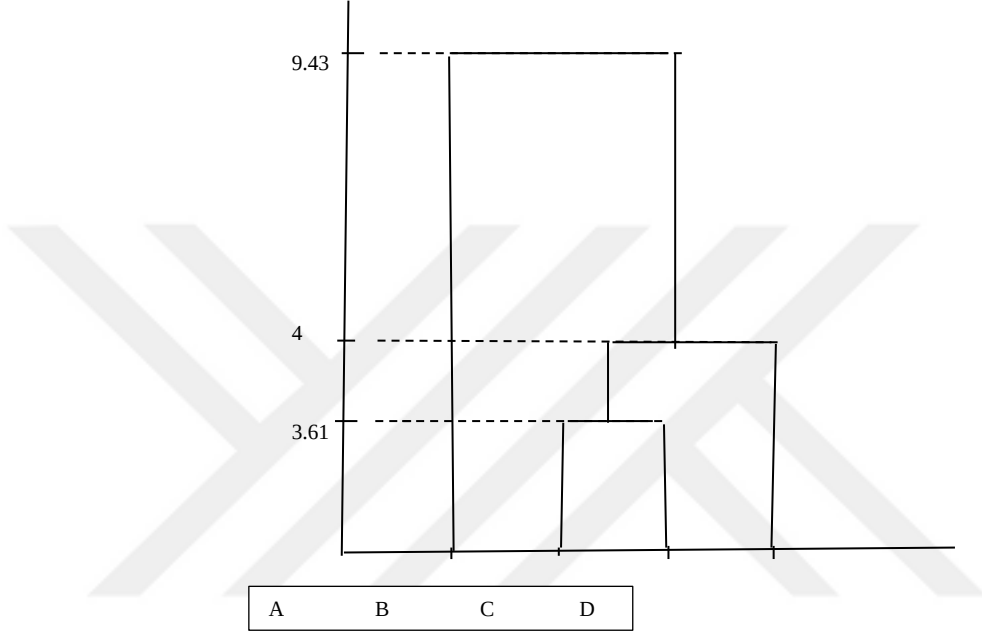
şeklinededir. Bu matriste yer alan en küçük değer 4'tür, yeni birleşmen (*BC*) kümesi ile *D* kümesi ile oluşur. Buna göre yeni uzaklıklar,

$$d_{(BCD)A} = \max\{d_{A(BC)}, d_{AD}\} = \max\{6.40, 9.43\} = 9.43 \quad (4.22)$$

ve yeni uzaklık matrisi,

$$Dis = \begin{matrix} (BCD) \\ A \end{matrix} \begin{bmatrix} 0 & \\ 9.43 & 0 \end{bmatrix} \quad (4.23)$$

biçiminde elde edilir. Dendogram oluşturulduktan sonra kofenetik matrisi elde edilir.



Şekil 4.1 Kümelerin uzaklıklarını gösteren dendogram

Kofenetik korelasyon matrisi R ,

$$R = \begin{matrix} A \\ B \\ C \\ D \end{matrix} \begin{bmatrix} 0 & & & \\ 9.43 & 0 & & \\ 9.43 & 3.61 & 0 & \\ 9.43 & 4 & 4 & 0 \end{bmatrix} \quad (4.24)$$

elde edilir. $CPCC$ değerinin hesaplamak için Dis ve R matrislerinden yararlanarak aşağıdaki Çizelge 4.2 elde edilmiştir.

Çizelge 4.2 Kofenetik korelasyon katsayısı için hesaplama adımları

	d_{ij}	r_{ij}	$d_{ij} - \bar{d}$	$r_{ij} - \bar{r}$	$(d_{ij} - \bar{d})(r_{ij} - \bar{r})$	$(d_{ij} - \bar{d})^2$	$(r_{ij} - \bar{r})^2$
AB	6.33	9.43	0.767	2.78	2.131	0.588	7.728
AC	6.40	9.43	0.837	2.78	2.326	0.700	7.7284
AD	9.43	9.43	3.867	2.78	10.749	14.951	7.7284
BC	3.61	3.61	-1.953	-3.04	5.938	3.816	9.2416
BD	3.61	4	-1.953	-2.65	5.176	3.816	7.0225
CD	4	4	-1.563	-2.65	4.143	2.444	7.0225
Toplam					30.464	26.314	46.472

Yukarıdaki Çizelge 4.2 yardımı ile

$$CPCC = \frac{\sum_{i < j} (d_{ij} - \bar{d})(r_{ij} - \bar{r})}{\sqrt{\sum_{i < j} (d_{ij} - \bar{d})^2 \sum_{i < j} (r_{ij} - \bar{r})^2}} \quad (4.25)$$

eşitliğindeki ifadelerin değerleri yerine yazıldığında,

$$CPCC = \frac{30.464}{\sqrt{26.314 \times 46.472}} = 0.871 \quad (4.26)$$

olarak hesaplanır. Bu değer 1'e yakın olduğundan kümelemenin iyi olduğu söylenebilir.

4.4.2 Silhouette indeksi

Rousseuw (1987), her bir birimin kendi kümesine uygunluğunu tanımlayacak bir Silhouette indeksi önermiştir. Silhouette değeri, kümelere atanan gözlemlerin güvenilirlik derecesini ölçer. $a(i)$; i . birimin kendi kümesindeki tüm noktalara olan ortalama uzaklıklarını (benzerliğini) ve $b(i)$; i . birimin diğer kümelerdeki tüm noktalara olan ortalama uzaklıkların minimumunu gösterebilir. Buradan i . birim için Silhouette indeksi;

$$sil(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (4.27)$$

olarak tanımlanır. Eğer $sil(i)$ değeri 1'e yaklaşırsa i . birimin atandığı kümeye ait olduğu, $sil(i)$ değeri 0'a yaklaşırsa veya negatif olursa i . birimin atandığı kümeye ait olmadığı sonucuna varılır. Negatif değerler yalnızca bir birim en uygun kümesine atanamadığında ortaya çıkar. Kümelemenin kalitesi (geçerliliği) için doğal bir ölçü olarak;

$$sil(C) = \frac{1}{n} \sum_{s_i \in S} sil(i) \quad (4.28)$$

tüm birimler için ortalama Silhouette değeri olarak tanımlanabilir. Bu kritere göre, maksimum ortalama Silhouette değerine ulaşılan küme sayısı uygun küme sayısı olarak alınır (Rousseeuw 1987).

4.4.3 Dunn indeksi

Dunn indeks değeri hesaplanırken, K_i ve K_j ile gösterilen iki küme arasındaki uzaklığı $D(K_i, K_j)$, sırasıyla K_i ve K_j kümelerinde yer alan x ve y gözlem çiftleri arasındaki minimum uzaklık olarak,

$$D(K_i, K_j) = \min_{x \in K_i, y \in K_j} (d_i, d_j) \quad (4.29)$$

tanımlanır. K_i kümesinin çapı ise bu küme içerisinde yer alan gözlem çiftleri arasındaki maksimum uzaklık olarak,

$$cap(K_i) = \max_{(x,y) \in K_i} (d_i, d_j) \quad (4.30)$$

şeklinde tanımlanır. Bu tanımlamalar altında Dunn indeksi,

$$Dunn(K) = \min_{i=1,2,\dots,k} \left(\min_{j=1+k,\dots,k} \left(\frac{D(K_i, K_j)}{\max_{l=1,2,\dots,k} cap(K_l)} \right) \right) \quad (4.31)$$

şeklinde ifade edilir. Dunn indeks değeri daha büyük olan kümeleme, geçerli kümeleme olarak kabul edilir (Xu ve Wunsch 2009).

5. ÇOK DEĞİŞKENLİ PANEL VERİDE KÜMELEME ANALİZİ

Panel veride kümeleme yapılmasının amacı birbirine benzeyen birimlerin bir araya getirerek homojen gruplar oluşturmaktır. Bu gruplar arasında heterojen yapı doğal olarak meydana geleceği için birimlerin havuzlanması doğru bir şekilde yapılmış olur. Ele alınan modelde birimlere göre farklı parametre olduğu düşünülüp ona göre tahmin edilecek ise kümeleme analizi bu aşama için yararlı olacaktır. Birbirine yakın birimlerin havuzlandığı ve ilgilenilen özelliklere göre uzak olanların ayrıldığı gruplandırmada tahmin işlemlerini yapmak daha uygun olacaktır.

Zheng ve Li (2014)'te tek değişkenli ve çok değişkenli panel veri için açıklayıcı bir kümeleme çalışması yapmışlardır. Bu çalışmada ilk olarak tek değişkenli panel veri için kümelemenin çok karmaşık olmayacağını vurgulamışlardır. Tek değişkenli panel veri setini iki boyutlu tabloya kolayca dökülebileceği ve kümeleme analizini uygulamanın kolay olacağından bahsetmişlerdir.

Zheng ve Li (2014) çalışmasında, çok değişkenli panel veri kümelemesinde dikkat edilmesi gereken noktalar olduğunu vurgulamışlardır. Çok değişkenli panel veri yapısının daha karmaşık olduğunu, doğal olarak iki boyutlu tabloya geçmenin daha kolay olmayacağını söylemişlerdir. Şöyle ki, biçiminde tablolastırma yapmışlardır. Çizelge 5.1'de yer alan ifadeler şu şekildedir. n toplam birim sayısını, p değişken sayısını ve T zaman uzunluğunu göstermektedir. Böylece X_{ijt} , t zamanındaki i . birimin j değişkenine ait değeri olarak tanımlanmıştır.

Çizelge 5.1 Çok değişkenli panel veri

	Zaman	1	...	t	...	T
	Değişken	$X_1, \dots, X_j, \dots, X_p$	...	$X_1, \dots, X_j, \dots, X_p$	...	$X_1, \dots, X_j, \dots, X_p$
Birim	1	$X_{111}, \dots, X_{1j1}, \dots, X_{1p1}$		$X_{11t}, \dots, X_{1jt}, \dots, X_{1pt}$		$X_{11T}, \dots, X_{1jT}, \dots, X_{1pT}$
	.					
	.					
	.					
	i	$X_{i11}, \dots, X_{ij1}, \dots, X_{ip1}$		$X_{i1t}, \dots, X_{ijt}, \dots, X_{ipt}$		$X_{i1T}, \dots, X_{ijT}, \dots, X_{ipT}$
	.					
	.					
	.					
	n	$X_{n11}, \dots, X_{nj1}, \dots, X_{np1}$		$X_{n1t}, \dots, X_{njt}, \dots, X_{npt}$		$X_{n1T}, \dots, X_{njT}, \dots, X_{npT}$

(Zheng ve Li 2014). Buna göre,

t. zamandaki j değişkeninin ortalaması,

$$\underline{x}_{jt} = \frac{1}{n} \sum_{i=1}^n x_{ijt} \quad (5.1)$$

j. değişkenin ortalaması,

$$\underline{x}_j = \frac{1}{n} \frac{1}{T} \sum_{i=1}^n \sum_{t=1}^T x_{ijt} \quad (5.2)$$

t. zamandaki j değişkeninin varyansı,

$$S^2(x_{jt}) = \frac{1}{n-1} \sum_{i=1}^n [x_{ijt} - \underline{x}_{jt}]^2 \quad (5.3)$$

j. değişkeninin varyansı,

$$S^2(X_j) = \frac{1}{N-1} \frac{1}{T} \sum_{i=1}^N [x_{ijt} - \underline{x}_{jt}]^2 \quad (5.4)$$

şeklinde verilmiştir (Zheng ve Li 2014).

Çok değişkenli panel veride kümeleme karmaşık yapıda olduğundan dikkat edilmesi gereken noktalar vardır. Panel veri özelliği göz ardı edilebilecek bir durum ise her değişkenin, zaman boyutunda ortalaması alınır. Böylece veriden zaman boyutu kaldırılmış olur ve panel veri özelliği kalmaz. Ancak bu durum en az iki sıkıntıyı beraberinde getirir. Bunlardan birincisi, çok fazla bilgi kaybolacaktır. Bunun sebebi, ortalama değerin birimdeki sadece ortalama değişimi göstermesidir, yani diğer dağılım özelliklerini göstermemesidir. İkincisi, tüm birimlerin aynı değişkeni aynı yönde değişir. Aksi durumda, doğru olmayan sonuçlar meydana gelir.

Zheng ve Li (2014) çalışmasında iki ana sorunu çözmeye çalışmışlardır. Bunlardan birincisi, birimlerin benzerliğini ölçmek için hangi istatistiğin kullanılacağı ve ikincisi hiyerarşik kümeleme yöntemlerinden hangi yöntemin uyum sağlayacağıdır.

Zheng ve Li (2014) çalışmasında yeni bir uzaklık ölçüsü sunmuşlardır ve bağlantı yöntemi olarak da Ward yöntemini kullanmışlardır. Panel verinin hata kareler ortalamasının kesit verinin hata kareler ortalamasından farklı olduğunu söylemişlerdir ve uzaklık fonksiyonu olarak tanımlamışlardır. Uzaklık fonksiyonu,

$$D_k = \sum_{t=1}^T \sum_{j=1}^p \sum_{i \in i^k} [x_{ijt} - \underline{x}_{jt}^k]^2 \quad (5.5)$$

biçiminde tanımlanmıştır. Eşitlikteki D_k , k grubunu kareler toplamını vermektedir. k , grubu temsil etmektedir. i^k , k grubundaki birimleri, x_{ijt} , t zamanındaki i . birimin j değişkenine ait değerini ve $\underline{x}_{jt}^k$ ise k grubundaki birimlerin t . zamanda j değişkenine ait ortalamasını göstermektedir (Zheng ve Li 2014).

Yukarıdaki Çizelge 5.1’de verilen panel veri bu durum için düzenlendiğinde,

Çizelge 5.2 Düzenlenmiş panel veri

	V-Z	X_{i11}	X_{i21}	...	X_{ip1}	X_{i12}	..	X_{ip2}	..	X_{i1T}	...	X_{ipT}
i	1	X_{111}	X_{121}		X_{1p1}	X_{112}		X_{1p2}		X_{11T}		X_{1pT}
	2	X_{211}	X_{221}		X_{2p1}	X_{212}		X_{2p2}		X_{21T}		X_{2pT}
	3	X_{311}	X_{321}		X_{3p1}	X_{312}		X_{3p2}		X_{31T}		X_{3pT}
	.	.	.		.	.		.		.		.
	.	.	.		.	.		.		.		.
	.	.	.		.	.		.		.		.
	i	X_{i11}	X_{i21}		X_{ip1}	X_{i12}		X_{ip2}		X_{i1T}		X_{ipT}
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	n	X_{n11}	X_{n21}		X_{np1}	X_{n12}		X_{np2}		X_{n1T}		X_{npT}

Çizelge 5.12 oluşturulurken Akay (2018) ve Zheng ve Li (2014) çalışmalarından yararlanılmıştır. V-Z sütunu değişken ve zaman için verilmiştir. Veri seti bu şekilde düzenlenerek kümeleme analizi aşamaları uygulanmıştır.

Öklid uzaklığı bu duruma göre genişletildiğinde,

$$d_{ir} = \sqrt{\sum_{t=1}^T \sum_{j=1}^p (x_{ijt} - x_{rjt})^2}, \quad i, r = 1, \dots, n \quad (5.6)$$

biçiminde elde edilir. Burada,

d_{ir} : i . ve r . birim arasındaki uzaklığını,

x_{ijt} : i . birimin j değişkeninin t zamanındaki değerini,

x_{rjt} : i . birimin j değişkeninin t zamanındaki değerini

göstermektedir (Zheng ve Li 2014).

İki birim arasındaki uzaklığın,

$$\bullet \quad d_{ir} > 0, \quad i \neq r, \quad \text{ve} \quad d_{ii} = 0, \quad (5.7)$$

- $d_{ir} = d_{ri},$
(5.8)

- $d_{ir} \leq d_{ik} + d_{rk}$ (5.9)

özelliklerini sağlaması gerekir(Rudin 1976).

Uzaklıklar hesaplandıktan sonra Dis uzaklık matrisi oluşturur.

$$Dis = \begin{bmatrix} 0 & & \\ d_{21} & 0 & \\ \dots & \dots & \ddots \\ d_{n1} & d_{n2} & \dots & 0 \end{bmatrix}_{n \times n} \quad (5.10)$$

Uzaklık matrisiden sonra bağlantı yöntemleri kullanılarak kümeler oluşturulur.

6. UYGULAMA

Karbondioksit emisyonu, iklim deęişikliğine sebep olan altı ana sera gazından birisidir. Bir ülkenin yapabileceęi maksimum salınım karbondioksit eşdeęeri olarak verilir (Anonymus, 2021).

Yenilenebilir enerji, doğal kaynakların sürekli olarak yenilenebildięi enerji kaynaklarıdır. Örneęin, rüzgar, güneş, su ve biyokütle gibi kaynaklar yenilenebilir enerji kaynaklarıdır. Bu kaynaklar kullanıldığında, emisyonların önemli ölçüde azaltılmasına ve iklim deęişikliğinin yavaşlatılmasına yardımcı olur.

Bu çalışmada, 1975-2017 yıllarına ait karbondioksit emisyonu, kiři başına gayri safi yurtiçi hasıla ve yenilenebilir enerjiye ilişkin gözlemler kullanılmıştır. Sera gazları, iklim deęişikliği üzerinde doğrudan etkileri olan yedi gazın toplamını ifade eder. Bunlar, karbondioksit (CO₂), metan (CH₄), azot oksit (N₂O), kloroflorokarbonlar (CFC'ler), hidroflorokarbonlar (HFC'ler), perflorokarbonlar (PFC'ler), kükürt hekzaflorür (SF₆) ve nitrojen triflorür (NF₃) gazlarıdır. Veriler, CO₂ cinsinden eşdeęeri olarak verilir ve insan faaliyetlerinden kaynaklanan brüt doğrudan emisyonu ifade eder. Veriler Uluslararası Enerji Ajansı tarafından sağlanmaktadır. Hava ve sera gazı emisyonları, CO₂ hariç, bin ton, kiři başına ton veya kiři başına kilogram olarak ölçülür. Sadece CO₂, milyon ton ve kiři başına ton olarak ölçülmektedir.

Yenilenebilir enerji, yenilenebilir kaynakların toplam birincil enerji arzına (TPES: Birincil enerji arzı; enerji üretimi artı enerji ithalatı, eksi enerji ihracatı, eksi uluslararası bunkerler, sonra artı/eksi stok deęişiklikleri olarak) katkısı olarak tanımlanır. Yenilenebilir kaynaklar hidro (pompa depolama hariç), jeotermal, güneş, rüzgar, gelgit ve dalga kaynaklarının birincil enerji eşdeęerini içerir. Katı biyoyakıtlar, biyogazolin, biyodizeller, dięer sıvı biyoyakıtlar, biyogazlar ve kentsel atıklardan elde edilen enerji de dahildir. Biyoyakıtlar, doğrudan veya dolaylı olarak biyokütleden (canlı veya yakın zamanda yaşıyan organizmalardan elde edilen malzeme) türetilen yakıtlar olarak tanımlanır. Buna odun, bitkisel atıklar (odun atıkları ve enerji üretimi için kullanılan mahsuller dahil), etanol, hayvansal malzemeler/atıklar ve sülfite kül suyu dahildir.

Belediye atıkları, konut, ticaret ve kamu hizmeti sektörleri tarafından üretilen ve yerel makamlar tarafından ısı ve/veya elektrik üretimi için merkezi bir yerde bertaraf edilmek üzere toplanan atıkları içerir. Bu gösterge, toplam birincil enerji arzının yüzdesinin yanı sıra ton petrol eşdeğeri cinsinden ölçülür (Anonymous 2019).

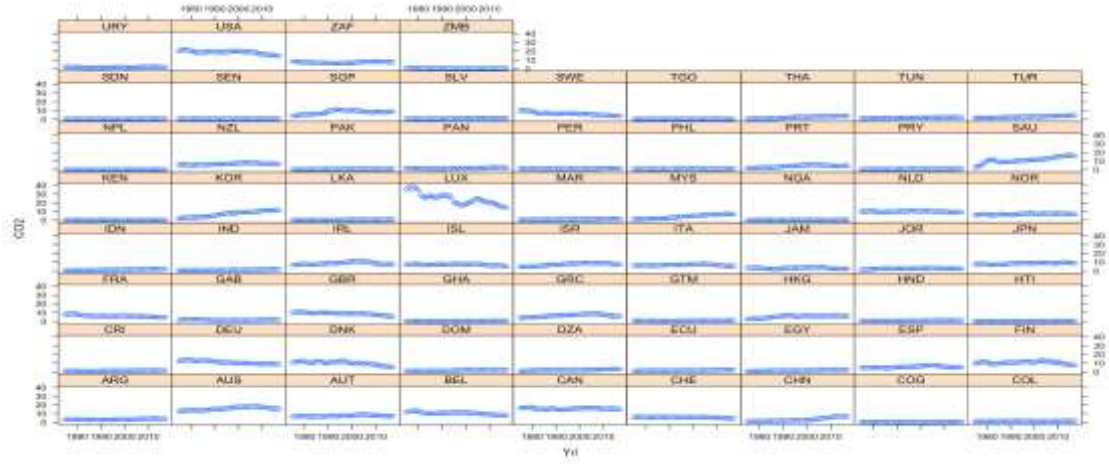
Karbondioksit emisyonu kişi başına ton ve yenilenebilir enerji petrol eşdeğeri ton olarak OECD'den elde edilmiştir.

Kişi başı GSYİH ise Dünya Bankasından elde edilmiştir. Bu üç değişkene göre ülkeler incelendiğinde 67 ülkenin verilerine ulaşılmıştır.

Çizelge 6.1 Uygulamaya dahil edilen ülkeler ve kodları

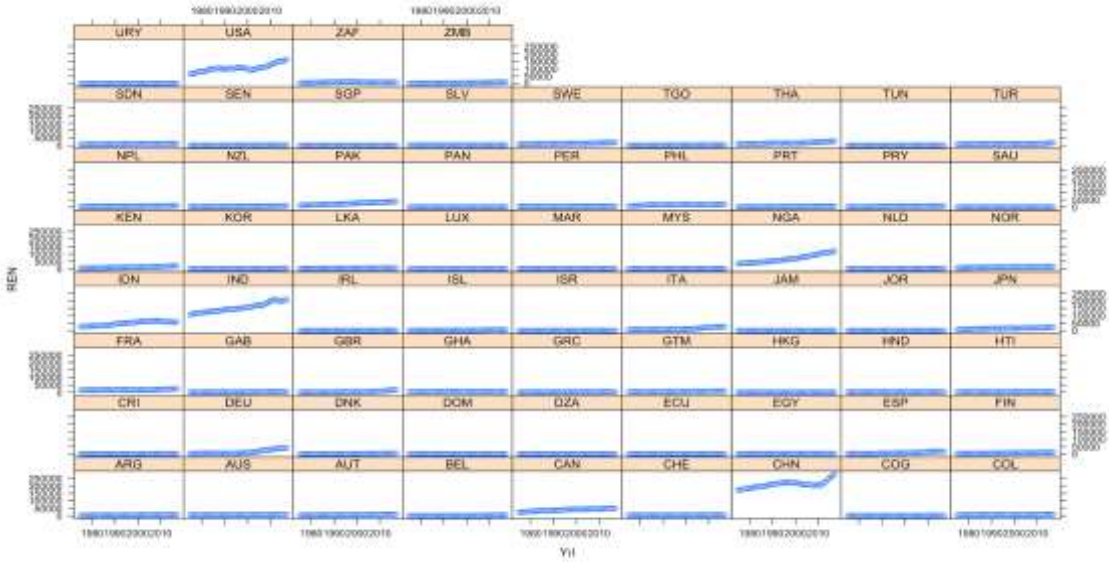
AUS : Avustralya	GBR : Büyük	GTM : Guatemala	PHL : Filipinler
AUT : Avusturya	Britanya	HTI : Haiti	SAU : Suudi
BEL : Belçika	USA : Amerika	HND : Honduras	Arabistan
CAN : Kanada	Birleşik Devletleri	HKG : Hong	SEN : Senegal
DNK : Danimarka	DZA : Cezayir	Kong	SGP : Singapur
FIN : Finlandiya	ARG : Arjantin	IND : Hindistan	ZAF : Güney
FRA : Fransa	CHN : Çin	IDN : Endonezya	Afrika
DEU : Almanya	COL : Kolombiya	ISR : İsrail	LKA : Sri Lanka
GRC : Yunanistan	COG : Kongo	JAM : Jamaika	SDN : Sudan
ISL : İzlanda	CRI : Kosta Rika	JOR : Ürdün	THA : Tayland
IRL : İrlanda	DOM : Dominik	KEN : Kenya	TGO : Togo
ITA : İtalya	Cumhuriyeti	MYS : Malezya	TUN : Tunus
JPN : Japonya	PRT : Portekiz	MAR : Fas	URY : Uruguay
KOR : Kore	ESP : İspanya	NPL : Nepal	ZMB : Zambiya
LUX : Lüksemburg	SWE : İsveç	NGA : Nijerya	NZL : Yeni
MEX : Meksika	CHE : Şili	PAK : Pakistan	Zelanda
NLD : Hollanda	ECU : Ekvator	PAN : Panama	NOR : Norveç
GHA : Gana	EGY : Mısır	GAB : Gabon	
PER : Peru	SLV : El Salvador	TUR : Türkiye	
	PRY : Paraguay		

Bu ülkelere ait zaman serisi grafikleri her üç değişken için verilmiştir. Zaman serilerini oluştururken trellis grafiklerinden yararlanılmıştır. Trellis grafikleri, her panel için aynı anda zaman serisi grafiklerini çizdirir. Böylece veriyle ilgili özellikleri görsel olarak bütün haliyle inceleme fırsatı tanır (Frees 2004, Becker vd. 1996).



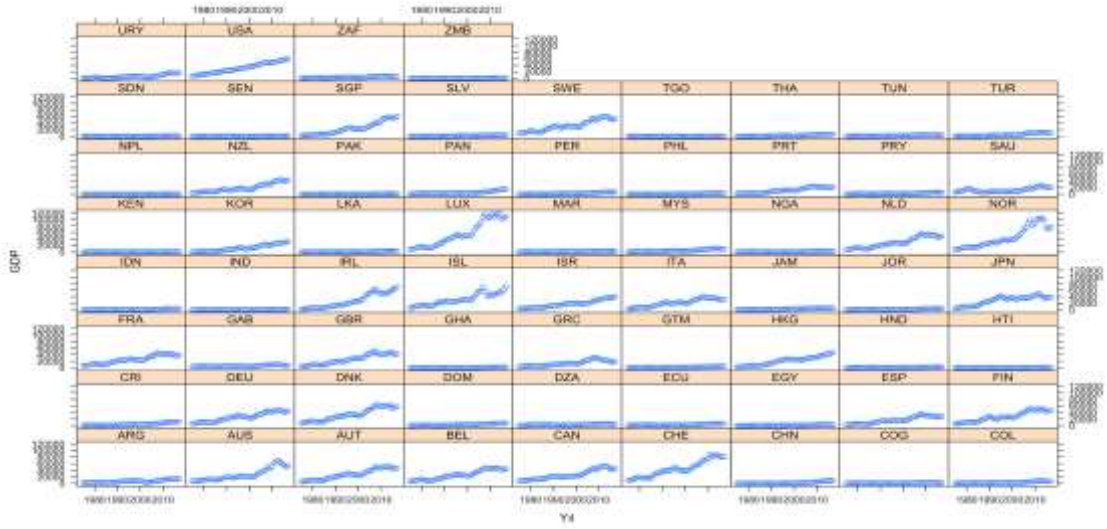
Şekil 6.1 CO₂ emisyonu(kişi başına) için ülkelere ait zaman serisi grafikleri

Şekil 6.1 incelendiğinde; Amerika Birleşik Devletleri(USA) ve Lüksemburg(LUX)'a ait grafiklerdeki dalgalanmalar hemen göze çarpmaktadır. Ayrıca Güney Kore (KOR) ve Suudi Arabistan(SAU)'da artış gözlenmektedir.



Şekil 6.2 Yenilenebilir enerji için ülkelere ait zaman serisi grafikleri

Yenilenebilir enerjide; Amerika Birleşik Devletleri(USA), Endonezya(IDN), Hindistan(IND), Nijerya(NGA) ve Çin(CHN) grafikleri artış eğilimi göstermektedir.



Şekil 6.3 Kişi başına GSYİH için ülkelere ait zaman serisi grafikleri

Şekil 6.3’de; Amerika Birleşik Devletleri(USA), Singapur(SGP), İsveç(SWE), Yeni Zelanda(NZL), Kore(KOR), Luksemburg(LUX), Hollanda(NLD), Norveç(NOR), İrlanda(IRL), İzlanda(ISL), Japonya(JPN), İsrail(ISR), Fransa(FRA), Birleşik Krallık(GBR), HongKong(HKG), Almanya(DEU), Danimarka(DNK), Finlandiya(FIN), Avustralya(AUS), Avusturya(AUT), Belçika(BEL) Kanada(CAN) ve İsviçre(CHE) grafiklerinde artış görülmektedir.

Bu üç değişken arasındaki ilişki çeşitli şekillerde olabilir. Artan GSYİH seviyeleri, ülkenin ekonomik faaliyetlerinde daha fazla enerji kullanmasına, dolayısıyla daha fazla CO₂ emisyonuna yol açabilir. Bu nedenle, GSYİH’nın artması, emisyonun artmasına sebep olabilir. Diğer yandan, emisyonları azaltmak için yenilenebilir enerji kaynaklarına yönelmek, GSYİH’nın artmasına katkıda bulunabilir. Son olarak, ülkenin GSYİH’nın artması ve CO₂ emisyonunun azaltılması arasında bir denge bulunabilir. Ülkenin ekonomik büyüme sırasında daha verimli teknolojiler kullanması ve daha çevreci üretim tekniklerini benimsemesi, GSYİH’nın artmasını sağlarken emisyonun azalmasına yardım eder.

İlgilenilen değişkenlere göre ülkelerin tanımlayıcı istatistiklerini incelemek ilk adım olarak faydalı olacaktır. Değişkenlere ait tanımlayıcı istatistik çizelgeleri Ek-2’de verilmiştir.

6.1 Çok Değişkenli Panel Veride Kümeleme Analizi Uygulaması

Çalışmanın amacında belirtildiği gibi altmış yedi ülke için panel veri modeli oluşturmak yerine, bu ülkeleri gruplayarak (kümeleme yaparak) daha az panel ile çalışılması amaçlanmıştır. Bu amaca yönelik olarak kümeleme analizinden yararlanılmıştır.

Veri seti, Zheng ve Li (2014) önerdiği gibi düzenlenmiş ve R programında işlemler yapılmıştır. Kümeleme analizine ilk olarak değişkenlerin birimleri farklı olduğundan standartlaştırma işlemi ile başlanmıştır. Kümeleme analizinde Öklid, Manhattan ve Minkowski($q=3$) uzaklıkları ile tam, tek, ortalama, Ward ve merkezi bağlantı yöntemlerinden yararlanılmıştır. Kullanılan yöntemlere göre oluşturulan dendogramlar incelenmiştir. Gerekli adımlar izlenerek kümeler elde edilmiştir.

İlk olarak eldeki üç değişken, Zheng ve Li (2014) önerdiği şekilde düzenlenmiştir. Kümeleme analizinde dendogramlar çizildikten sonra, görsel olarak oluşan kümeler incelenir. Dendogramlar ilk aşamada küme yapısı hakkında çok yardımcı olur. Çalışmaya ait dendogramlar Ek 1’de verilmiştir.

Ek-1’deki dendogramlar incelendiğinde Çin, Hindistan, Lüksemburg ve ABD’nin diğer ülkelerden ayrıştığı kolayca göze çarpmaktadır. Özellikle Çin ve Hindistan’ın bir kümede, Lüksemburg ile ABD’nin bazen aynı bazen farklı kümede yer aldığı görülmektedir.

Küme sayısına karar vermede dendogramlar incelendikten sonra fikir vermesi için ilk olarak

$k = \sqrt{(n/2)}$ eşitliğinden yararlanılır ise,

$$k = \sqrt{(67/2)} \cong 5.78 \quad (6.1)$$

elde edilir. Küme sayısının 5 ya da 6 olması beklenebilir.

Silhouette, Dunn ve Calinski-Harabazs yöntemlerinden yararlanarak küme sayısına karar verilmeye çalışılmıştır. Kofenetik korelasyon katsayısı kullanılarak küme geçerliliği bakımından incelenmiştir. İlk olarak aşağıdaki çizelgede Silhouette indeks değerleri verilmiştir.

Çizelge 6.2 Silhouette indeks değerleri

k	Tam			Tek			Ortalama			Ward			Merkezi		
	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min
2	0.48	0.55	0.49	0.59	0.58	0.58	0.54	0.52	0.69	0.49	0.59	0.5	0.52	0.62	0.69
3	0.42	0.55	0.62	0.6	0.58	0.62	0.52	0.58	0.62	0.53	0.60	0.55	0.52	0.58	0.62
4	0.59	0.56	0.48	0.6	0.56	0.58	0.56	0.56	0.58	0.55	0.61	0.45	0.56	0.56	0.56
5	0.59	0.56	0.48	0.56	0.59	0.56	0.56	0.6	0.57	0.46	0.61	0.56	0.56	0.49	0.58
6	0.51	0.53	0.47	0.23	0.39	0.27	0.54	0.58	0.55	0.46	0.51	0.44	0.54	0.59	0.54
7	0.52	0.49	0.48	0.27	0.18	0.3	0.5	0.55	0.53	0.48	0.51	0.47	0.54	0.56	0.53
8	0.49	0.47	0.5	0.26	0.19	0.29	0.49	0.52	0.51	0.5	0.48	0.49	0.51	0.52	0.48

k: küme sayısı

Çizelge 6.2 incelendiğinde, Silhouette indeks değerlerinin en yüksek olduğu küme sayısı 4 ve 5 olduğu görülmektedir. Öte yandan, Minkowski uzaklığı ile ortalama ve merkezi bağlantı yöntemleri kullanılması durumunda küme sayısı 2 önerilmektedir.

Çizelge 6.3 Dunn indeks değerleri

k	Tam			Tek			Ortalama			Ward			Merkezi		
	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min
2	0.408	0.148	0.465	0.405	0.341	0.417	0.408	0.285	0.465	0.056	0.109	0.076	0.408	0.286	0.464
3	0.522	0.125	0.206	0.478	0.285	0.554	0.150	0.405	0.516	0.066	0.111	0.101	0.522	0.286	0.560
4	0.140	0.094	0.165	0.522	0.305	0.560	0.150	0.405	0.565	0.095	0.140	0.140	0.522	0.406	0.545
5	0.171	0.238	0.565	0.481	0.406	0.566	0.501	0.565	0.182	0.095	0.187	0.140	0.481	0.340	0.564
6	0.200	0.146	0.181	0.321	0.229	0.376	0.150	0.225	0.182	0.131	0.148	0.183	0.171	0.181	0.156
7	0.223	0.199	0.227	0.318	0.227	0.359	0.151	0.225	0.196	0.171	0.188	0.127	0.171	0.226	0.156
8	0.238	0.201	0.235	0.279	0.238	0.291	0.151	0.226	0.199	0.171	0.136	0.137	0.171	0.226	0.156

Dunn indeks değerleri için Çizelge 6.3 ele alındığında, küme sayısının 3, 4 ve 5 etrafında toplandığı göze çarpmaktadır.

Çizelge 6.4 Calinski Harabasz indeks değerleri

k	Tam			Tek			Ortalama			Ward			Merkezi		
	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min	Öklid	Man	Min
2	22.67	18.85	22.67	8.699	8.69	8.698	22.67	13.68	22.67	46.32	54.34	51.31	22.67	13.68	22.67
5	5	3	5				5	9	5	6	1	1	5	9	5
3	24.56	30.16	24.56	19.19	8.398	19.19	24.56	24.56	24.56	62.36	75.47	70.33	24.56	8.398	19.19
4	4	2	4	0		0	4	4	4	6	9	6	4		0
4	71.27	71.89	47.74	18.48	18.48	18.48	18.48	18.48	18.48	67.60	84.40	79.98	18.48	18.48	18.48
		3	7	8	8	8	8	8	8	9	2	1	8	8	8
5	78.76	63.92	41.58	14.17	14.17	14.17	64.68	71.76	73.39	72.94	76.57	73.94	64.17	14.17	14.17
	5	0	1	3	2	2	7	5	3	8	9	3	2	2	2
6	69.60	58.31	73.46	11.91	12.99	11.91	54.15	68.56	61.77	64.12	76.94	75.33	60.34	62.66	58.25
	7	0	4	9	4	9	5	5	2	9	2	7	1	6	1
7	71.09	66.81	65.81	13.01	11.40	13.01	55.31	60.61	63.60	67.91	77.16	75.22	58.96	53.12	56.60
	0	9	7	7	9	7	4	4	2	3	8	2	4	4	9
8	70.68	62.72	71.57	11.10	11.10	11.10	49.77	55.84	66.55	64.23	73.66	66.23	62.73	54.52	51.52
	9		7	0	0	0	1	9	6	3	5	2	6	1	6

Calinski-Harabasz indeks sonuçlarında, 4, 5 ve 6 küme sayısı yoğunlukta olmasına rağmen küme sayısının 3 olabileceğini söylemektedir.

Çizelge 6.5 Kofenetik Korelasyon Katsayısı sonuçları

	Bağlantı Yöntemleri				
Uzaklık Ölçüleri	Tam	Tek	Ortalama	Ward	Merkezi
Öklid	0.905	0.867	0.910	0.609	0.915
Manhattan	0.822	0.842	0.898	0.658	0.896
Minkowski	0.919	0.881	0.929	0.640	0.907

Yukarıdaki kofenetik korelasyon katsayıları incelendiğinde bağlantı yöntemlerinden tam, ortalama ve merkezi yöntemlerinin öne çıktığı görülmektedir.

Buna göre oluşan kümeleme sonucu aşağıdaki Çizelge 6.6’da verilmiştir.

Çizelge 6.6 Verinin özel düzenlenmiş hali için oluşturulan kümeler

KÜME 1: ARG COG COL CRI DOM DZA ECU EGY GAB GHA GTM HND HTI IDN JAM JOR KEN LKA MAR MYS NGA NPL PAK PAN PER PHL PRT PRY SDN SEN SLV TGO THA TUN TUR URY ZMB
KÜME 2: AUS AUT BEL CAN CHE DEU DNK ESP FIN FRA GBR GRC HKG IRL ISL ISR ITA JPN KOR NLD NOR NZL SAU SGP SWE ZAF
KÜME 3: CHN IND
KÜME 4: LUX
KÜME 5: USA

Yukarıda oluşturulan kümeler ele alınan değişkenler açısından incelendiğinde, birinci kümedeki ülkeler; Arjantin, Kongo Cumhuriyeti, Kolombiya, Kosta Rika, Dominik Cumhuriyeti, Cezayir, Ekvador Cumhuriyeti, Mısır, Gabon, Gana, Guatemala, Honduras, Haiti, Endonezya, Jamaika, Ürdün, Kenya, Sri Lanka, Fas, Malezya, Nijerya, Nepal, Pakistan, Panama, Peru, Filipin, Portekiz, Paraguay, Sudan, Senegal, El Salvador, Togo, Tayland, Tunus, Türkiye, Uruguay ve Zambiya’dır.

İkinci küme ülkeleri; Avustralya, Avusturya, Belçika, Kanada, İsviçre, Almanya, Danimarka, İspanya, Finlandiya, Fransa, İngiltere, Yunanistan, Hong Kong, İrlanda,

İzlanda, İsrail, İtalya, Japonya, Güney Kore, Hollanda, Norveç, Yeni Zelanda, Suudi Arabistan, Singapur, İsveç ve Güney Afrika yer almaktadır.

Üçüncü küme ülkeleri; Çin ve Hindistan'dan oluşmaktadır. Dördüncü küme Lüksemburg ve beşinci küme ise ABD ülkeleridir.

Bu kümelerdeki ülkelerin biraraya gelme sebebi incelenmelidir. Aslında ülkelerin bu şekilde gruplanması çok şaşırtıcı değildir. İlgilenilen değişkenler CO₂ emisyonu, yenilenebilir enerji ve kişi başına GSYİH olduğundan bazı ülkelerin ayrı bir yerde yer alması normal karşılanmaktadır. Ek-1 verilen dendogramlarda da görülen ABD, Lüksemburg, Çin ve Hindistan'ın ısrarla ayrı bir kümede olması beklenen bir durumdur. Şöyle ki, bu ülkeler CO₂ emisyonu bakımından ilk sıralarda yer alır (Anonymys, 2023). Lüksemburg ise kişi başına düşen GSYİH bakımından en üst sıralardadır (Anonymys, 2023). Kişi başına GSYİH bakımından çok yüksek değer alması o ülkeyi kendi başına ayrı bir grupta yer almasına sebep olabilir.

Küme 2'deki ülkelere bakıldığında genel olarak Avrupa Birliği ülkelerinin çoğunlukla yer aldığı görülmektedir. Avrupa Birliği ülkelerinin yanı sıra Avustralya, Kanada, İsviçre, Büyük Krallık, Hong Kong, İrlanda, İsrail, Japonya, Güney Kore, Yeni Zelanda, Suudi Arabistan, Singapur ve Güney Afrika biraraya gelme sebepleri araştırılmıştır. (Guisan 2009) çalışmasında Norveç, Danimarka, İsveç, ABD, Avusturya, İrlanda, İsviçre, Kanada, Finlandiya, Hollanda ve Birleşmiş Krallık ülkelerini üç indeks üzerinden incelemiştir. Bu indekslerden biri yaşam memnuniyeti ve kişi başına gelir üzerinden hesaplanmaktadır. Bu indeks bakımından ülkelerin sıralarının en üst seviyede olduğunu akmaktadır. Aynı zamanda, Güney Afrika hariç insani gelişmişlik indeksi (HDI) bakımından aynı grupta yer almaktadır. İnsani gelişmişlik indeksi birinci grubu olan yüksek HDI'ya sahip ülkelerdir.

Küme 1'de yer alan ülkeler insani gelişmişlik indeksi(HDI) bakımından incelendiğinde farklı gruplardaki ülkelerin bir araya geldiği görülmektedir. Ancak yenilenebilir enerji bakımından ele alındığında Endonezya hariç birbirine yakın ülkeler olduğu görülmektedir. Ayrıca Portekiz bir Avrupa Birliği ülkesi olmasına rağmen bu gruba dahil

olmuştur. Li ve Lin (2013), çalışmasında 1971-2008 yılları için 110 ülkeyi CO₂ emisyonlarının dinamik değişimleri bakımından incelemişlerdir. Ülkeleri gelir gruplarına(düşük gelir, düşük-orta, yüksek-orta ve yüksek gelir grubu) göre ayırıp farklı gelir grubuna göre kişi başına GSYİH ile kişi başına CO₂ emisyonunun ilişkisinin farklı olduğunu ortaya koymuşlardır.

İkinci aşamada hiyerarşik olmayan k-ortalama yöntemi kullanılarak kümeler oluşturulmuştur. Burada da yukarıda hiyerarşik yöntemlerde olduğu gibi kümeleme adımları izlenmiştir. Sonuçlar özet halinde çizelgelerde sunulmuştur.

Çizelge 6.7 K-Ortalama için Silhoutte, Dunn ve Calinski-Harabazs indeks değerleri

k	2	3	4	5	6	7	8
Silhoutte	0.5539	0.5873	0.5935	0.5966	0.5770	0.5423	0.4859
Dunn	0.1366	0.0958	0.1395	0.1898	0.1898	0.1898	0.127
C-H	129.353	75.5825	135.235	179.325	162.43666	151.17759	143.1354

Küme değerlendirme indeksleri incelendiğinde yine 5, 6 ve 7 tane küme etrafında karar verilmesi önerilmektedir. İndeksler en çok küme sayısı olarak beşi göstermektedir. Bu aşamadan sonra K-ortalama yöntemi kullanılarak elde edilen kümeler incelenmiştir.

Çizelge 6.8 K-ortalamaya göre kümeler

1. KÜME: AUT BEL CHE DEU DNK FIN FRA GBR ISL JPN NLD NOR SWE
2. KÜME: AUS CAN LUX USA
3. KÜME: CHN IND
4. KÜME: ESP GRC HKG IRL ISR ITA KOR NZL PRT SAU SGP ZAF
5. KÜME: ARG COG COL CRI DOM DZA ECU EGY GAB GHA GTM HND HTI IDN JAM JOR KEN LKA MAR MYS NGA NPL PAK PAN PER PHL PRY SDN SEN SLV TGO THA TUN TUR URY ZMB

Çizelge 6.8’de yer alan ülkeler aşağıda açıklanmıştır.

Birinci küme; Avusturya, Belçika, İsviçre, Almanya, Danimarka, Finlandiya, Fransa, İngiltere, İzlanda, Japonya, Hollanda, Norveç ve İsveç ülkelerinden oluşmaktadır.

İkinci kümede Avustralya, Kanada, Lüksemburg, ABD bir araya gelmiştir.

Üçüncü kümede sadece Çin ile Hindistan bir aradadır.

Dördüncü küme ise İspanya, Yunanistan, Hong Kong, İrlanda, İsrail, İtalya, Güney Kore, Yeni Zelanda, Portekiz, Suudi Arabistan, Singapur, Güney Afrika yer almaktadır.

Son olarak beşinci küme; Arjantin, Kongo Cumhuriyeti, Kolombiya, Kosta Rika, Dominik Cumhuriyeti, Cezayir, Ekvador, Mısır, Gabon, Gana, Guatemala, Honduras, Haiti, Endonezya, Jamaika, Ürdün, Kenya, Sri Lanka, Fas, Malezya, Nijerya, Nepal, Pakistan, Panama, Peru, Filipinler, Paraguay, Sudan, Senegal, El Salvador, Togo, Tayland, Tunus, Türkiye, Uruguay ve Zambia'dan oluşmaktadır.

K-ortalama yöntemine göre oluşan kümeler incelendiğinde, Lüksemburg ile ABD, Kanada ve Avustralya birleştiği ve çoğunluğu Avrupa Birliği ülkelerinden oluşan ikinci bir küme oluşturulduğu görülmektedir. Beşinci kümede (hiyerarşikte) ise sadece Portekiz' in bu kümeden ayrılıp Avrupa Birliği ülkeleri grubuna dahil olduğu görülmektedir. Her iki kümelemede de üçüncü küme aynı ülkelerden oluşmuştur.

Birinci kümedeki Avrupa Birliği ülkelerine bakıldığında kişi başına GSYİH bakımından bölge içinde üst sıralarda yer aldıkları bilinmektedir. İsviçre, Avrupa Birliği'nde yer almamasına rağmen hem coğrafi hem de diğer kriterler bakımından benzerliklere sahiptir. Bunun yanı sıra Japonya için incelendiğinde Hamiltona ve Turton (2002) çalışmasında ekonomik büyüme bakımından AB ülkeleri ile benzer olduğunu ifade etmişlerdir. Bu sebepten dolayı Japonya'nın AB ülkelerinin yanında olması beklenebilir.

Aynı zamanda İspanya, Yunanistan, İtalya ve Portekiz'in AB ülkesi olmasına rağmen, başka kümede yer almasının sebepleri incelenmiştir. Balsalobre-Lorente vd. (2021)

çalışmasında İspanya, Yunanistan, İtalya ve Portekiz'in 1995-2015 yılları için Çevresel Kuznets Eğrileri (EKC) bakımından ele almışlardır. Bu Güney Avrupa ülkelerinin artan enerji ihtiyacı, şehirleşme durumu ve artan nüfusu ile karbondioksit emisyonunun nasıl etkileneceği üzerinde durmuşlardır.

6.2 Panel Veri Analizi Uygulaması

Kümeleme analizi uygulandıktan sonra çalışmanın amacında belirtildiği gibi kümelerde yer alan ülkelerin ortalamaları alınarak paneller oluşturulmuştur. Panel veri analizinde parametre tahmini için uygun yöntemi belirlemek amacıyla yatay kesit bağımlılığı, değişen varyans ve otokorelasyon testleri yapılır. Bu testler sonucunda parametre tahminine geçilir.

Oluşturulan panel veri için bireysel etkilerin olup olmadığı yani havuzlanmış model incelenmiştir. Bunun için dengeli panel veri olma durumunda kullanılan Lagrange Çarpınları Testi - (Breusch-Pagan) testinden yararlanılmıştır. Breusch-Pagan LM testi birim etkiler varlığının sınanması için kullanılır. Testte kullanılan hipotez, rasgele birim etkiler varyansının sıfır olup olmadığını sınanmaktadır. Hipotez,

$$H_0: \sigma_\varepsilon^2 = 0$$

$$H_1: \sigma_\varepsilon^2 > 0$$

biçimindedir. Test istatistiği ise,

$$BP = \frac{5*43}{2(43-1)} \left[\frac{\sum_{i=1}^5 [\sum_{t=1}^{43} \hat{u}_{it}]^2}{\sum_{i=1}^5 \sum_{t=1}^{43} \hat{u}_{it}^2} - 1 \right]^2 \quad (6.2)$$

olarak hesaplanır. Test istatistiğinde yer alan $\hat{u}_{it}$, rasgele etkili modele ait artıkların tahmin değerleridir. Testin kuralı, $BP > X_{1,0.05}^2$ ise yokluk hipotezi reddedilir ve birim etkisinden bahsedilebilir.

Çizelge 6.9 Breusch-Pagan LM testi sonuçları

Test Adı	Hesaplanan Ki-Kare İstatistik Değeri	p-değeri
Breusch-Pagan LM testi	1410.2	0.00

Yukarıdaki analiz sonucu incelendiğinde, yokluk hipotezinin $\alpha=0.05$ anlam düzeyinde reddedildiği görülmektedir ($p < \alpha$). Böylece birim etkilerin var olduğu yani havuzlanmış modelin kullanılamayacağı sonucuna varılır.

Birim etkiler, sabit ve rasgele etkili modeller için karşılaştırıldığında, Hausman testinden yararlanılır. Bu testin yokluk hipotezinde, ‘rasgele etkili model’ sunulmasına karşın alternatif hipotezde ‘sabit etkili model’ yer alır. Yokluk hipotezinde, “Birim etkiler ile açıklayıcı değişkenler arasında ilişki yoktur” iddiası, alternatif hipotezde ise “ilişki vardır” iddiası yer alır. Eğer yokluk hipotezi reddedilir ise sabit etkili model seçilir. Bu test için hesaplanan istatistik,

$$H_{Hes} = (\hat{\beta}_{SE} - \hat{\beta}_{RE})' [Var(\hat{\beta}_{SE}) - Var(\hat{\beta}_{RE})]^{-1} (\hat{\beta}_{SE} - \hat{\beta}_{RE}) \quad (6.3)$$

biçimindedir. Burada, $H_{Hes} > X_{3;0.05}^2 = 7.815$ ise yokluk hipotezi red edilir.

Analiz sonucu aşağıdaki çizelgede yer almaktadır.

Çizelge 6.10 Hausman Test Sonuçları

Test Adı	Hesaplanan Ki-Kare İstatistik değeri	p- değeri
Hausman Test	43.318	0.0000

Yukarıdaki sonuçlara bakıldığında $\alpha= 0.05$ anlam düzeyinde yokluk hipotezi reddedilir. Sabit etkili model ile devam edilir.

Bu aşamadan sonra yatay kesit bağımlılığı testi uygulanır. Kümeleme uygulandıktan sonra yatay kesit sayısı beşe düşmüştür. Buna göre sabit etkili modeli regresyon modelinde olduğu gibi düşünülür ise,

$$\hat{u}_{it} = y_{it} - \hat{u}_{it}, i=1, 2, 3, 4, 5, t= 1, 2, \dots, 43 \quad (6.4)$$

eşitliği yardımıyla hata teriminin tahmini bulunur. Tahmin edilen bağımlı değişken ($\hat{y}_{it}$) ile gözlenen bağımlı değişken (y_{it}) değerlerinin farkının alınmasıyla hata teriminin tahminleri elde edilir.

Yatay kesit testlerinin yokluk hipotezi,

$$H_0: \rho_{it} = \rho_{jt} = 0, \text{ tüm } t\text{'ler ve } i \neq j \text{ için}$$

$$H_1: \rho_{it} = \rho_{jt} \neq 0$$

biçiminde tanımlanır. “Yatay kesit bağımlılığı yoktur” sıfır hipotezi altında test istatistiği hesaplanır.

$$CD_{LM} = \sqrt{\frac{2T}{n(n-1)}} (\sum_{i=1}^{n-1} \sum_{j=i+1}^n \hat{\rho}_{ij}) \quad (6.5)$$

eşitliği yardımı ile hesaplanır. Burada $\hat{\rho}_{ij}$,

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^T \hat{u}_{it} \hat{u}_{jt}}{(\sum_{t=1}^T \hat{u}_{it}^2)^{1/2} (\sum_{t=1}^T \hat{u}_{jt}^2)^{1/2}} \quad (6.6)$$

olarak elde edilir.

Çizelge 6.11 Panel veride yatay kesit bağımlılığı için Breusch-Pagan LM ve Pesaran CD test

Test Adı	Hesaplanan Ki-Kare İstatistik Değeri	p-değeri	Test Adı	Hesaplanan Z İstatistik Değeri	p- değeri
Breusch-Pagan LM testi	200.46	0.00	Pesaran CD test	-3.9573	0.00

Çizelge 6.10 incelendiğinde her iki test sonucunda $\alpha = 0.05$ anlam düzeyinde yokluk hipotezinin reddedilir. Bu durumda yatay kesit bağımlılığı olduğu sonucuna ulaşılır.

Diğer bir varsayım, zaman serilerinde otokorelasyon üzerinedir. Durbin-Watson testi, zaman serilerinde otokorelasyonun varlığını test etmede kullanılan en temel testlerdendir. Hata terimlerinin kendi aralarında $u_{it} = \rho u_{it-1} + \varepsilon_{it}$ biçiminde otokorelasyonlu olabilir (Asteriou and Hall, 2006, Akdi, 2010).

Hipotezler,

$$H_0: \rho = 0,$$

$$H_1: |\rho| < 1$$

şeklindedir.

Çizelge 6.12 Panel veride otokorelasyon için Durbin-Watson testi sonuçları

Test Adı	Hesaplanan İstatistik Değeri
DW testi	DW = 0.18678

Yukarıdaki Çizelge 6.11’de yer alan DW = 0.18678 değeri 2’den küçük olduğundan yokluk hipotezi reddedilir. Bu durumda, otokorelasyon varlığından söz edilir. DW < 2 olduğundan aralarında pozitif bir ilişki vardır (Brocklebank & Dickey, 1986).

Son olarak değişen varyans sorunu olup olmadığı incelenir. Sabit etki modelinde değişen varyans sorunu doğrudan panel veri modelinin hata teriminden kaynaklanır. Birimlere göre değişen varyans gözlenir. Değişen varyans sorunu test edilirken Uyarlanmış Wald testinden yararlanılabilir. Her birim için varyansların eşit olduğu $H_0: \sigma_i^2 = \sigma^2, i=1, 2, \dots, 5$ hipotezi ile

$H_1: \sigma_i^2 \neq \sigma^2$ alternatif hipotezi için

$$W = \sum_{i=1}^n \frac{(\hat{\sigma}_i^2 - \hat{\sigma}^2)^2}{f_{ii}} \quad (6.7)$$

test istatistiği kullanılır. Eşitlikteki, $\hat{\sigma}_i^2 = \frac{1}{T} \sum_{t=1}^{T_i} \hat{u}_{it}^2$, $f_{ii} = \frac{1}{T} \frac{1}{T-1} \sum_{t=1}^{T_i} (\hat{u}_{it}^2 - \hat{\sigma}_i^2)^2$ ifadelerini gösterir.

Çizelge 6.13 Panel veride değişen varyans için Uyarlanmış Wald testi sonuçları

Test Adı	Hesaplanan Ki-Kare İstatistik Değeri	p- değeri
Uyarlanmış Wald testi	360.38	0.0000

Çizelge 6.12’de görüldüğü gibi $\alpha = 0.05$ anlam düzeyinde yokluk hipotezi reddedilir ve hata terimlerinin sabit varyanslı olmadığı sonucuna varılır.

Yukarıdaki test sonuçlarına göre yatay kesit bağımlılığı, durağan olmama durumu ve değişen varyans olduğu sonucuna varılmıştır. Bu durumda, panel veri analizinde parametre tahmininde kullanılan Driscoll ve Kraay(1998) tahmin edicisinden yararlanılır (Millo, 2017). Parametre tahminleri aşağıdaki çizelgede verilmiştir.

Çizelge 6.14 Parametre tahminleri

Log(CO ₂)	Tahmin Değeri	Standart Hata	t istatistik değeri	p-değeri
Log(REN)	-0.55061	0.03643	-15.114	0.00
Log(GSYİH)	0.43300	0.02300	18.824	0.00
Kukla 1	1.79647	0.23530	7.635	0.00
Kukla 2	2.67628	0.22685	11.798	0.00
Kukla 3	4.35576	0.35142	12.395	0.00
Kukla 4	0.59687	0.17830	3.348	0.00
Kukla 5	4.85928	0.30387	15.991	0.00
Adjusted R-squared: 0.993				

Çizelge 6.14'deki parametre tahminleri incelendiğinde, yenilenebilir enerji ve GSYİH değişkenlerinin model için $\alpha = 0.05$ anlam düzeyinde anlamlı oldukları görülmektedir. Yenilenebilir enerjinin modele negatif yönlü, GSYİH ise pozitif yönlü katkı sağladığı ortadadır. Bunun yanı sıra modelde birim etkiler için oluşturulan kukla değişkenlere ait katsayı tahminleri de anlamlı görülmektedir. Düzeltilmiş açıklama oranı ise 0.99 olarak elde edilmiştir.

Çizelge 6.14'ün daha iyi anlaşılması için aşağıdaki model göz önüne alınsın.

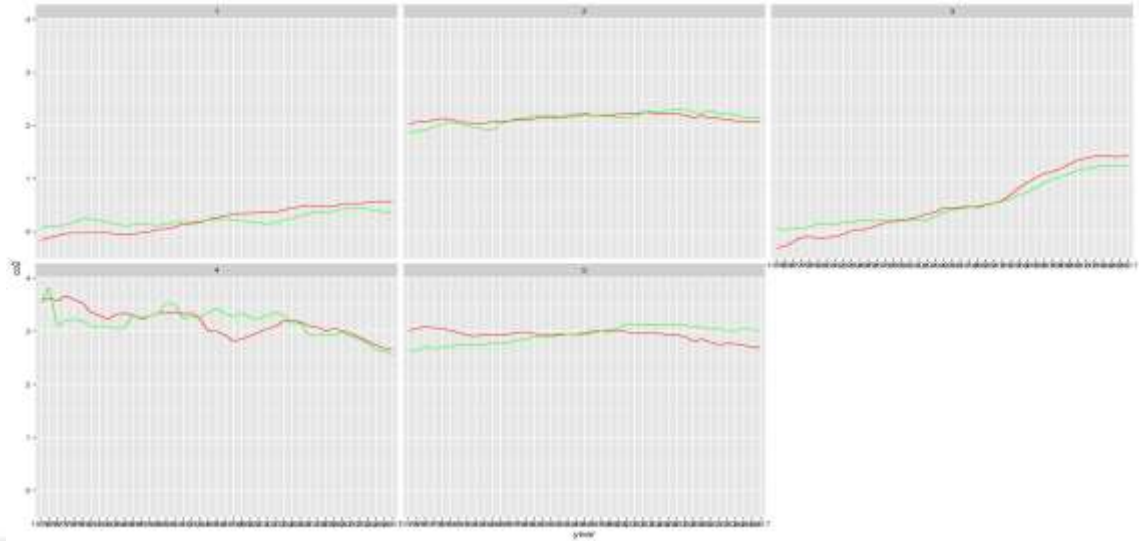
$$Y_{it} = \alpha_i + \beta_1 X_{i1t} + \beta_2 X_{i2t} + u_{it}, \quad i=1, 2, 3, 4, 5, t=1, \dots, 43 \quad (6.8)$$

Bu modele göre tahmin değerleri yerine yazıldığında,

$$\begin{aligned} \hat{y}_{1t} &= \hat{\alpha}_1 + \hat{\beta}_1 x_{11t} + \hat{\beta}_2 x_{12t} \\ (\log(\widehat{CO_2}))_{2t} &= 1.797 - 0.551(\log(REN))_{11t} + 0.433(\log(GDP))_{12t} \\ \hat{y}_{2t} &= \hat{\alpha}_2 + \hat{\beta}_1 x_{21t} + \hat{\beta}_2 x_{22t} \\ (\log(\widehat{CO_2}))_{2t} &= 2.676 - 0.551(\log(REN))_{21t} + 0.433(\log(GDP))_{22t} \\ \hat{y}_{3t} &= \hat{\alpha}_3 + \hat{\beta}_1 x_{31t} + \hat{\beta}_2 x_{32t} \\ (\log(\widehat{CO_2}))_{3t} &= 4.356 - 0.551(\log(REN))_{31t} + 0.433(\log(GDP))_{32t} \\ \hat{y}_{4t} &= \hat{\alpha}_4 + \hat{\beta}_1 x_{41t} + \hat{\beta}_2 x_{42t} \\ (\log(\widehat{CO_2}))_{4t} &= 0.597 - 0.551(\log(REN))_{41t} + 0.433(\log(GDP))_{42t} \\ \hat{y}_{5t} &= \hat{\alpha}_5 + \hat{\beta}_1 x_{51t} + \hat{\beta}_2 x_{52t} \\ (\log(\widehat{CO_2}))_{5t} &= 4.856 - 0.551(\log(REN))_{51t} + 0.433(\log(GDP))_{52t} \end{aligned}$$

eşitlikleri elde edilir. Bu eşitliklerde $(\hat{y}_{it}) = (\log(\widehat{CO_2}))_{it}$, $(x_{i1t}) = (\log(REN))_{i1t}$ ve $(x_{i2t}) = (\log(GDP))_{i2t}$ 'dir. Birim etkilerinin tahminleri incelendiğinde en yüksek 3. (CHN, IND) ve 5.(USA) kümeye ait olduğu görülmektedir. Bu durumda, Çin, Hindistan kümesinin ve Amerika Birleşik Devletleri kümesinin birim etkilerinin en yüksek olduğu yani karbondioksit emisyonuna en fazla etkili olduğu yorumu yapılabilir.

Aşağıdaki Şekil 6.4'de, gözlem değerleri ve modelden elde edilen tahmin değerlerinin grafikleri verilmiştir.



Şekil 6.4 Bağımlı değişken gözlem değerleri(kırmızı) ile tahmin değerleri(yeşil)

K- ortalama yöntemine göre oluşturulan kümeler için panel veri analizi yapılarak devam edilmiştir. Yukarıda yapılan panel veri modeli için varsayımların sınama aşamalarının aynısı burada da uygulanacaktır.

Öncelikle model seçimiyle ilgili olan test sonuçları aşağıdaki Çizelge 6.15 'te verilmiştir.

Çizelge 6.15 Breusch-Pagan LM ve Hausman testi sonuçları

Test Adı		
Breusch-Pagan LM testi	668.79	p- değeri
Hausman Test	29.641	0.0000

Sabit etkili model seçilir. Sonrasında yatay kesit bağımlılığı, otokorelasyon ve değişen varyans sorunları için testler yapılır.

Çizelge 6.16 Pesaran CD testi, Durbin-Watson ve Breush Pagan testi sonuçları

Test Adı	Hesaplanan Z İstatistik Değeri	p-değeri	Test Adı	Hesaplanan İstatistik Değeri	Test Adı	Hesaplanan Ki-Kare İstatistiği	P değeri
Pesaran CD test	-3.9573	0.00	DW testi	0.19648	Breush Pagan testi	59.381	0.0000

Çizelge 6.16 incelendiğinde varsayımların sağlanmadığı görülmektedir. Bu durum için önerilen Driscoll ve Kraay tahmin edicisi kullanılmıştır.

Çizelge 6.17 Parametre tahminleri\k-ortalama

Log(CO ₂)	Tahmin Değeri	Standart Hata	t istatistik değeri	p- değeri
Log(REN)	-0.592089	0.064603	-9.1650	0.00
Log(GSYİH)	0.456455	0.027913	16.3527	0.00
Kukla 1	1.979988	0.450452	4.3956	0.00
Kukla 2	4.704465	0.662199	7.1043	0.00
Kukla 3	2.353665	0.378460	6.2191	0.00
Kukla 4	4.470491	0.515219	8.6769	0.00
Kukla 5	2.727880	0.420825	6.482	0.00
Adjusted R-squared: 0.9909				

Bu modele göre tahmin değerleri yerine yazıldığında,

$$\hat{y}_{1t} = \hat{\alpha}_1 + \hat{\beta}_1 x_{11t} + \hat{\beta}_2 x_{12t}$$

$$(\log(\widehat{CO_2}))_{1t} = 1.980 - 0.592(\log(REN))_{11t} + 0.457(\log(GDP))_{12t}$$

$$\hat{y}_{2t} = \hat{\alpha}_2 + \hat{\beta}_1 x_{21t} + \hat{\beta}_2 x_{22t}$$

$$(\log(\widehat{CO_2}))_{2t} = 4.705 - 0.592(\log(REN))_{21t} + 0.457(\log(GDP))_{22t}$$

$$\hat{y}_{3t} = \hat{\alpha}_3 + \hat{\beta}_1 x_{31t} + \hat{\beta}_2 x_{32t}$$

$$(\log(\widehat{CO_2}))_{3t} = 2.354 - 0.592(\log(REN))_{31t} + 0.457(\log(GDP))_{32t}$$

$$\hat{y}_{4t} = \hat{\alpha}_4 + \hat{\beta}_1 x_{41t} + \hat{\beta}_2 x_{42t}$$

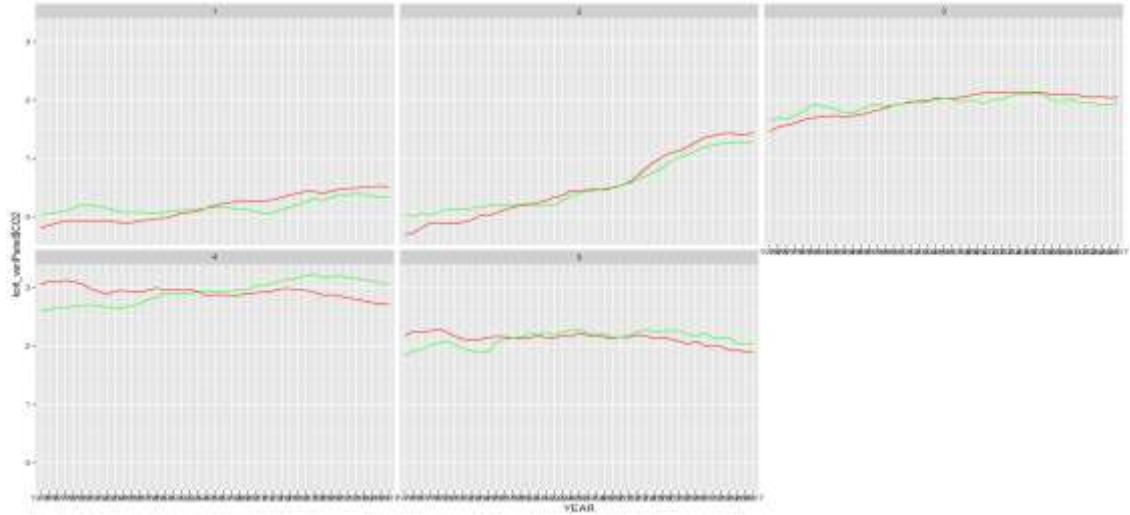
$$(\log(CO_2))_{4t} = 4.471 - 0.592(\log(REN))_{41t} + 0.457(\log(GDP))_{42t}$$

$$\hat{y}_{5t} = \hat{\alpha}_5 + \hat{\beta}_1 x_{51t} + \hat{\beta}_2 x_{52t}$$

$$(\log(\widehat{CO_2}))_{5t} = 2.728 - 0.592(\log(REN))_{51t} + 0.457(\log(GDP))_{52t}$$

eşitlikleri elde edilir. Kümelerin CO₂ emisyonuna tahmin edilen etkileri karşılaştırıldığında en düşük etkinin birinci kümeye ait olduğu söylenebilir. Birinci kümedeki ülkeler Avusturya, Belçika, İsviçre, Almanya, Danimarka, Finlandiya, Fransa, İngiltere, İzlanda, Japonya, Hollanda, Norveç ve İsveç'tir.

Şekil 6.5'de kümelere ait CO₂ emisyonu gözlem değerleri kırmızı, tahmin değerleri yeşil renkte belirtilmiştir.



Şekil 6.5 Bağımlı değişken gözlem değerleri(kırmızı) ile tahmin değerleri(yeşil)/k-ortalama

6.3 Uzun Dönem İlişkisi

Çalışmanın bu kısmında, oluşturulan kümelere göre ilgilenilen değişkenler arasındaki eşbütünleşme ilişkisi Johansen Eşbütünleşme Testi (1988) ile araştırılmıştır. Analizde değişkenlerin logaritmalarının alınmış halleri kullanılmıştır.

Çok değişkenli zaman serilerinde, rasgele vektörün bileşenleri arasındaki ilişkiyi ortaya çıkarmak en önemli amaçlardan biridir. Tek değişkenli zaman serilerinde olduğu gibi, çok değişkenli zaman serilerinde de istatistiki sonuç çıkarımları serinin durağanlık varsayımına bağlıdır. Çok değişkenli bir zaman serisi durağan olmamasına rağmen, bileşenlerinin herhangi bir lineer birleşimi durağan olabilir. Böyle serilere kointegrasyonlu (eşbütünleşik) seriler denilmektedir (Akdi 2010). Durağan olmayan çok değişkenli zaman serisinden yapılacak istatistiki sonuç çıkarımları yerine durağan bir lineer birleşimi ele alınarak yapılan çıkarımlar daha anlamlı olabilir.

Bu sebeple ilk olarak çok değişkenli zaman serilerinin durağanlıkları kontrol edilmiştir. Bunun için Dickey-Fuller (ADF) testinden yararlanılmıştır. Kointegrasyon analizine geçmeden önce serilerin bütünleşik sıraları araştırılmalıdır. Bunun için serilere DickeyFuller (ADF) testi uygulanmıştır. Bu teste ait hipotezler,

H_0 : Seri birim köke sahiptir. (Durağan değildir),

H_1 : Seri birim köke sahip değildir. (Durağan)

şeklinde. Aşağıda yer alan Çizelge 6.18’de birim kök testlerine ait sonuçlar yer almaktadır.

Çizelge 6.18. ADF birim kök sonuçları

Küme		ADF		
		Sabit	Sabit + Trend	1.Fark(p-değeri)
1	log(co ₂)	0.8385	0.6299	0.0028*
	log(ren)	0.2715	0.7399	0*
	log(gdp)	0.8187	0.2683	0.0013*
2	log(co ₂)	0.4221	0.9907	0.0001*
	log(ren)	0.9928	0.8661	0*
	log(gdp)	0.2071	0.1799	0.0023*
3	log(co ₂)	0.9251	0.2808	0.0096*
	log(ren)	0.9995	0.7834	0.0025*
	log(gdp)	0.9977	0.9143	0.0007*
4	log(co ₂)	0.5969	0.3064	0.0004*

Çizelge 6.18. ADF birim kök sonuçları (devam)

Küme		ADF		
		Sabit	Sabit + Trend	1.Fark(p-değeri)
	log(ren)	0.7898	0.1481	0*
	log(gdp)	0.4716	0.8836	0.0012*
5	log(co ₂)	0.9876	0.9448	0.001*
	log(ren)	0.8205	0.7924	0*
	log(gdp)	0.2093	0.1054	0.0043*

*: $p < \alpha = 0.05$ 'tir. 0.05 anlam düzeyine göre yokluk hipotezi(H_0) red edilir.

Yukarıdaki Çizelge 6.18, sabit ve sabit ile trend içeren model ele alınarak elde edilen Dickey -Fuller (ADF) testlerinin sonucunu göstermektedir.

Sonuçlar incelendiğinde düzey durumunda tüm değişkenler için serinin durağan olmadığını iddia eden yokluk hipotezi 0.05 anlam düzeyinde durağan olduğunu iddia eden alternatif hipotezine karşı reddedilemeyeceği, yani her kümedeki tüm değişkenlerin birim köke sahip oldukları görülmektedir. Birinci sıra farkları alınan değişkenlerin ise 0.05 anlam düzeyinde durağan oldukları görülmektedir. Dolayısıyla, genel olarak her iki test de tüm değişkenlerin bütünleşik sıralarının $I(1)$ birinci dereceden gecikmeli olduğu görülmektedir.

Johansen kointegrasyon testi kapsamında kurulması gereken VAR(Vektör Otoregresif Seriler) modelinde uygun gecikme uzunluğu belirlenmelidir. Uygun gecikme uzunluğu için maksimum gecikme 3 alınmış ve kriter olarak Akaike Bilgi Kriteri (AIC) kullanılmıştır.

Johansen yönteminde, parametre matrisinin özdeğerlerinden yararlanılır. Bileşenleri $X_{1t}, X_{2t}, \dots, X_{kt}$ olan VAR(1) modeli

$$X_t = AX_{t-1} + e_t$$

olarak verilmiş olsun. Burada A matrisi p boyutlu parametre matrisi olmak üzere, e_t 'ler varyans kovaryans matrisi V olan beyaz gürültü sürecini göstermektedir. Modelin her iki tarafından $X_t - I$ çıkartılırsa, $\Delta X_t = A - I$ olmak üzere, $\Delta X_t = \Delta X_{t-1} + e_t$ şeklinde yazılabilir. $\text{rank}(\Delta) = r$ olsun. $r = p$ ise sistem durağan, $r = 0$ ise serinin hiçbir lineer birleşeni durağan değildir, yani kointegrasyon ilişkisi bulunmaz. Ayrıca, $0 < r < p$ ise $\Delta = \alpha\beta'$ olacak şekilde öyle α ve β vektörleri (veya matrisleri) vardır ki $\beta'X_t$ durağandır. Yani, β kointegrasyon vektörüdür (Akdi 2010).

Johansen kointegrasyon analizi daha önce detaylı bir şekilde anlatılan ve elde edilen 5 küme içinde yer alan değişkenler için uygulanmıştır. İlk olarak birinci küme olarak adlandırılan ve karbon emisyonu, GSYİH ve yenilenebilir enerji değişkenlerinin Zheng ve Li(2014) önerdiği şekilde düzenlenerek elde edilen küme değişkenleri için elde edilen kointegrasyon analizi sonuçları elde edilmiştir. Birinci küme için elde edilen iz ile maksimum özdeğer istatistikleri Çizelge 6.19'da verilmektedir.

Çizelge 6.19 Johansen test sonuçları (Verinin özel düzenlenmiş hali için)

Küme	Hipotez	Özdeğer	İz İstatistik Değeri	Maks. Özdeğer İstatistik Değeri
1	$r = 0$	0.303941	28.55763 (0.069)	14.49284 (0.326)
	$r \leq 1$	0.234475	14.06479 (0.0812)	10.68774 (0.1706)
	$r \leq 2$	0.080961	3.377048 (0.0661)	3.841465 (0.0661)
2	$r = 0$	0.187887	11.87947 (0.936)	21.13162 (0.8828)
	$r \leq 1$	0.076666	3.554819 (0.9361)	14.26460 (0.9332)
	$r \leq 2$	0.009065	0.364265 (0.5461)	3.841465 (0.5461)
3	$r = 0$	0.294827	18.56126 (0.5248)	13.62316 (0.3967)
	$r \leq 1$	0.115817	4.938099 (0.8154)	4.800551 (0.7668)
	$r \leq 2$	0.003521	0.137547 (0.7107)	0.137547 (0.7107)

Çizelge 6.19 Johansen test sonuçları (Verinin özel düzenlenmiş hali için) (devam)

Küme	Hipotez	Özdeğer	İz İstatistik Değeri	Maks. Özdeğer İstatistik Değeri
4	$r = 0$	0.278967	18.76557 (0.5099)	13.08281 (0.4446)
	$r \leq 1$	0.132041	5.682755 (0.7326)	5.664454 (0.6566)
	$r \leq 2$	0.000457	0.018302 (0.8923)	0.018302 (0.8923)
5*	$r = 0$	0.546693	43.03858 (0.0009*)	31.64744 (0.0012*)
	$r \leq 1$	0.245400	11.39115 (0.1886)	11.26269 (0.1415)
	$r \leq 2$	0.003206	0.128452 (0.7200)	0.128452 (0.7200)

Not: Parantez içinde yer alan değerler, p değerini göstermektedir.

* 0.05 düzeyinde istatistik olarak anlamlıdır.

Yukarıdaki Çizelge 6.19 incelendiğinde sadece 5. küme için

$$H_0 : r = 0$$

$$H_1 : r \geq 1$$

H_0 yokluk hipotezi %5 anlamlılık düzeyinde reddedilir. Böylece kointegrasyonun yani uzun dönem bir ilişkinin olduğu söylenebilir.

Uzun dönem modeli,

$$\log(\text{CO}_2) = 0.035\log(\text{GSYİH}) - 0.597\log(\text{REN})$$

biçiminde oluşturulmuştur. Yukarıdaki model tahmini 5. küme(ABD) için ele alındığında karbondioksit emisyonu ile GSYİH arasında pozitif, yenilenebilir enerji ile negatif bir ilişki görülmektedir. Bilgili ve ark.(2016) çalışmasında 1977-2010 yıllarına ait 17 OECD ülkesi için karbondioksit emisyonu, GSYİH, GSYİH'nin karesi ve yenilenebilir enerji tüketimi değişkenleri ile çalışmışlardır. Çevresel Kuznetz eğrisi üzerinde, yenilenebilir enerjinin etkisini araştırmışlardır. Bu çalışmada, ülkelerin bireysel olarak panel kointegrasyon testlerinin sonuçları verilmiştir. ABD için karbondioksit emisyonu ile GSYİH arasında pozitif bir ilişki, yenilenebilir enerji ile negatif ilişki ortaya koymuşlardır.

Literatür incelendiğinde, Menyah ve Wolde-Rufael(2010) çalışmasında, 1960-2007 yıllarına ait ABD için karbondioksit emisyonu, yenilenebilir ve nükleer enerji tüketimi ve reel GSYİH arasındaki nedensel ilişkiyi araştırmıştır. Granger nedensellik testinin değiştirilmiş bir versiyonunu kullanarak, nükleer enerji tüketiminden karbondioksit emisyonu, tek yönlü bir nedensellik bulmuşlar, ancak yenilenebilir enerjiden karbondioksit emisyonuna uzanan bir nedensellik bulamadıklarını söylemişlerdir. Ekonometrik kanıtlar, nükleer enerji tüketiminin karbondioksit emisyonunu azaltmaya yardımcı olabileceğini gösteriyor gibi görünürken yenilenebilir enerji tüketimi, emisyonların azaltılmasına önemli bir katkı sağlayabilecek bir düzeye ulaşmadığını açıklamışlardır.

Sadece karbondioksit emisyonu alınarak oluşturulan kümeler için de benzer adımlar izlenmiştir.

Çizelge 6.20 Karbondioksit emisyonu için oluşturulan kümeler

KÜME 1: ARG CHN COG COL CRI DOM DZA ECU EGY GAB GHA GTM HND HTI IDN IND JAM JOR KEN LKA MAR MYS NGA NPL PAK PAN PER PHL PRT PRY SDN SEN SLV TGO THA TUN TUR URY ZMB

KÜME 2: AUS CAN LUX USA

KÜME 3: AUT BEL CHE DEU DNK ESP FIN FRA GBR GRC HKG IRL ISL ISR ITA JPN KOR NLD NOR NZL SAU SGP SWE ZAF

Karbondioksit emisyonu bakımından oluşturulan kümeler için ADF birim kök sonuçları Ek- 3’de verilmiştir. Bu kümeler için oluşturulan zaman serileri de durağan değildir. Birinci farkları alındığında 0.05 anlam düzeyinde durağan oldukları görülmüştür. Johansen test sonuçları Ek-3’de incelendiğinde sadece 2. Küme için maksimum özdeğer istatistiği(21.132), 0.05 anlam düzeyinde (p değeri=0.0407) yokluk hipotezi red edilir. Sonuçlar incelendiğinde ikinci küme için kointegrasyondan bahsedilebileceği görülmektedir. Ancak iz istatistiği ve maksimum özdeğer istatistiği incelendiğinde farklı sonuçlar ile karşılaşılmaktadır. Bu durumda Engle ve Granger yönteminden yararlanılabilir. Bu yöntem, artıklar serisini kullanılır. Burada bahsedilen artıklar,

değişkenleri ikili olarak alıp regresyon denklemlerinden elde edilir. Bu çalışma için $\log(\text{CO}_2)$ ile $\log(\text{REN})$ ve $\log(\text{CO}_2)$ ile $\log(\text{GSYİH})$ alınmıştır. İkinci küme için $\log(\text{CO}_2)$ serisinin $\log(\text{REN})$ üzerine regresyonu yapılmış ve bu denklemden artıklar elde edilmiştir. Bu artıkların durağanlık sonuçları Ek-3'deki çizelgede verilmiştir. Tabloya göre bu artıklar serisi durağan değil yani birim kök içermektedir. Bu durumda kointegrasyon ilişkisi gözlenmemiştir.

Sonraki adımda $\log(\text{CO}_2)$ serisinin $\log(\text{GSYİH})$ üzerine regresyonu yapılmış ve bu denklemden artıklar elde edilmiştir. $\log(\text{CO}_2)$ ile $\log(\text{GSYİH})$ için oluşturulan artıklar da durağan değildir. Bu durumda kointegrasyon ilişkisi yoktur biçiminde yorum yapılabilir.

Yenilenebilir enerji bakımından ülkeler kümelendiğinde dört küme elde edilmiştir. Bu kümeler için de uzun dönem ilişkisi incelendiğinde 0.05 anlam düzeyinde hiç bir kümede olmadığı görülmüştür.

Çizelge 6.21 Yenilenebilir enerji için oluşturulan kümeler

KÜME 1: ARG AUS AUT BEL CHE COG COL CRI DEU DNK DOM DZA ECU
EGY ESP FIN FRA GAB GBR GHA GRC GTM HKG HND HTI IRL ISL ISR ITA
JAM JOR JPN KEN KOR LKA LUX MAR MYS NLD NOR NPL NZL PAK PAN
PER PHL PRT PRY SAU SDN SEN SGP SLV SWE TGO THA TUN TUR URY
ZAF ZMB

KÜME 2: CAN IDN NGA

KÜME 3: CHN

KÜME 4: IND USA

Sonrasında kişi başı GSYİH için kümeler aşağıdaki çizelgede verilmiştir.

Çizelge 6.22 Kişi başı GSYİH için oluşturulan kümeler

KÜME 1: ARG CHN COG COL CRI DOM DZA ECU EGY GAB GHA GTM HND HTI IDN IND JAM JOR KEN LKA MAR MYS NGA NPL PAK PAN PER PHL PRY SDN SEN SLV TGO THA TUN TUR URY ZAF ZMB

KÜME 2: AUS AUT BEL CAN DEU DNK FIN FRA GBR IRL ISL JPN NLD SGP SWE USA

KÜME 3: CHE LUX NOR

KÜME 4: ESP GRC HKG ISR ITA KOR NZL PRT SAU

Bu kümelerde de uzun dönem ilişkisi bulunamamıştır. Son olarak ülke ortalamaları alınarak oluşturulan kümeler için eşbütünleşme denklemleri araştırılmıştır. Elde edilen kümeler aşağıdaki çizelgededir.

Çizelge 6.23 Ülke ortalamaları için oluşturulan kümeler

KÜME 1: ARG COG COL CRI DOM DZA ECU EGY GAB GHA GTM HND HTI IDN JAM JOR KEN LKA MAR MYS NGA NPL PAK PAN PER PHL PRT PRY SDN SEN SLV TGO THA TUN TUR URY ZAF ZMB

KÜME 2: AUS AUT BEL CAN CHE DEU DNK ESP FIN FRA GBR GRC HKG IRL ISL ISR ITA JPN KOR NLD NOR NZL SAU SGP SWE

KÜME 3: CHN IND

KÜME 4: LUX USA

Burada 4. kümede iz istatistiğine göre bir eşbütünleşme vektörü 0.05 anlam düzeyinde kurulabileceği görülmektedir. Engle-Granger nedensellik testi ile incelendiğinde nedensellik ilişkisi bulunamamıştır. Bu durumda ülke ortalamaları için oluşturulan kümelerin hiç birinde uzun dönem ilişkisi ortaya çıkarılamamıştır.

7. SONUÇ VE TARTIŞMA

Panel veri analizi, zaman serisi ve yatay kesit boyutunu birarada taşıdığı için oldukça geniş uygulama alanı bulmaktadır. Panel veri analizi sayesinde bir birimin zaman içindeki hareketlerini ve bir değişkenin zaman içinde birimler üzerinde nasıl ilerlediği görülebilir. Bu yüzden panel veri karışık bir çok olay içerebilir. Elverişli bir şekilde istatistiki analizleri uygulamak için veride düzenlemeler gerekir. Bunlardan dikkat edilmesi gereken birimlerin heterojen olma durumudur. Birimlerin heterojenliğini bir ön bilgi ile karar vermek yerine ilgilenilen değişkenlere göre oluşturulması uygun olabilir. Bu çalışmadaki amaçlardan biri, birimleri değişken özelliklerine gruplamaktır. Bunun için kümeleme analizi uygulanmıştır.

Bu çalışmanın ilk bölümünde konuyla ilgili genel çerçeve verilmiştir. Çalışmanın amacı ve sınırları üzerinde durulmuştur. İkinci bölümde, önceki çalışmalardan bazıları ele alınarak literatür incelenmesi yapılmıştır. Sonraki bölümde panel veri modeline ilişkin tanımlamalara yer verilmiştir. Bunun yanı sıra model seçimi ile ilgili test tanımları verilmiş ve tahmin edici yöntemlerden bahsedilmiştir. Dördüncü bölümde, kümeleme analizinde kullanılacak uzaklıklar ile bağlantı yöntemleri ele alınmıştır. Beşinci bölümde, çok değişkenli panel veride kümeleme için kullanılacak olan veri düzenleme açıklanmıştır. Altıncı bölümde 1975-2017 yıllarına ait karbondioksit emisyonu(kişi başına), kişi başına gayri safi yurtiçi hasıla ve yenilenebilir enerji ilişkin gözlemlere ait analizler bulunmaktadır. İlk olarak panel veriye kümeleme analizi uygulanarak ülkelere ait gruplar oluşturulmuştur. Bu kümeler için panel veri modellerinde yer alan parametre tahminleri elde edilmiştir.

Panel veriye kümeleme analizi uygulanırken Zheng ve Li (2014) çalışmasından yararlanılmıştır. Önerilen yöntemde verinin düzenlenme yöntemi panel verinin karmaşık yapısını basitleştirerek analizlerin yapılmasını sağlamaktır. Her değişkeninin her bir zaman serisi ayrı ayrı değişken olarak alındığında, çok sayıda değişken ile karşı karşıya kalınacağından zaman boyutu uzun olan durumlar için bu yöntem tekrar değerlendirilmelidir. Değişken sayısı bir anda birim sayısını geçmektedir. Böyle bir durumda, çok değişkenli varyans analizi gibi bir istatistiki analizi uygulama imkanı

ortadan kalkmaktadır. Bunun için yeniden veriyi düzenlemek gerekmektedir. Her bir değişken yıllara göre ortalaması hesaplanıp, kümeler arasında anlamlı bir farkın olup olmadığı çok değişkenli varyans analizi ile incelenir. Öte yandan her bir zaman serisi ayrı bir değişken olarak düzenlendiğinde verideki bilgi kaybı önlenir.

Hiyerarşik kümeleme yöntemleri tam, tek, ortalama, Ward ve merkezi bağlantı yöntemleri ile k-ortalama yöntemi, Öklid, Manhattan ve Minkowski uzaklıklarına uygulanmıştır. Küme sayısına karar vermede bazı tekniklerden yararlanarak en uygun olacak kümelere karar verilmiştir. Elde edilen kümelere panel veri analizi uygulamak için kümede bulunan ülkeler için gözlenen değerlerin ortalamaları alınarak veri yeniden düzenlenmiştir. Panel veri analizi aşamaları oluşturulan beş küme için uygulanmıştır.

Kümeleme analizinin başlangıcında dendogramlardan yararlanarak oluşacak kümeler hakkında bilgi edinilmiştir. Burada göze çarpan durum Çin, Hindistan, ABD ve Lüksemburg ülkelerinin hep ayrılması olmuştur. Bu ülkeler için ele alınan değişkenler bakımından literatür incelendiğinde beklenen bir durumdur. Çin, Hindistan ve ABD CO₂ emisyonunda ilk sırada yer alan ülkelerdir. Öte yandan Lüksemburg ise kişi başına GSYİH’da ilgilenilen yıllara göre ön sıralarda yer alır. Diğer kümeler incelendiğinde ise genellikle AB ülkelerinin birlikte yer aldığı görülmüştür. Ancak k-ortalama yöntemine göre kümelemede İspanya, Yunanistan, İtalya ve Portekiz’in başka bir kümede yer alması incelenmiştir. Literatüre bakıldığında bu ülkelerin AB’de yer alan ve güneyde bulunması sebebiyle ayrı bir şekilde ele alındığı görülmüştür.

Panel veri analizi uygulanırken varsayımların kontrol edilmesi gerekir. İstatistiki model oluşturulduğunda genellikle hata teriminin sıfır ortalamalı, sabit varyanslı ve birbiriyle ilişkisiz olma gibi durumları varsayılır. Ancak uygulamada verinin yapısı, bu varsayımları sağlamaya uygun olmayabilir. Böyle bir durumda parametre tahmini yapıldığında, tahmin edicide aranan özellikler elde edilemez. Tahmin edicilerin istenen özelliklere sahip olmasını sağlamak için çeşitli tahmin yöntemlerinden yararlanılır. Çalışmada, yatay kesit bağımlılığı, otokorelasyon ve değişen varyans sorunları ile karşılaşıldığından Driscoll-Kraay dirençli tahmin edicisinden yararlanılmıştır. Bu

analizlerin tamamı hem hiyerarşik yöntemlerden elde edilen hem de k-ortalama ile elde edilen iki farklı beş küme için uygulanmıştır.

Panel veri için kurulan modelde $\log(\text{REN})$ ile $\log(\text{GSYİH})$ değişkenleri için katsayı tahminleri elde edilmiştir. Elde edilen 5 kümeye göre $\log(\text{REN})$ değişkeninde meydana gelecek 10 birimlik artışın $\log(\text{CO}_2)$ 'e yaklaşık 0.60 azalmaya neden olacağı görülmüştür. Öte yandan $\log(\text{GSYİH})$ değişkenindeki 1 birimlik artışın $\log(\text{CO}_2)$ değişkeninde yaklaşık 0.47 artışa sebep olabileceği görülmektedir.

Panel veri analiz sonuçlarındaki tahminler arasında karşılaştırma yaparak birim etkileri incelenmiştir. Hiyerarşik yöntemlere göre oluşan kümelemede en yüksek Çin, Hindistan ve ABD'ye ait olduğu görülmektedir. Çin, Hindistan kümesinin ve ABD kümesinin birim etkilerinin en yüksek olduğu yani karbondioksit emisyonuna diğer kümelerle göre daha fazla etkili olduğu şeklinde yorumlanabilir.

Panel veri analiz kısmı tamamlandıktan sonra kümelerin uzun dönem ilişkileri incelenmiştir. Bunun için elde edilen beş kümenin yanında karbondioksit emisyonu, yenilenebilir enerji ve kişi başına düşen GSYİH değişkenlerine göre kümeler elde edilmiş ve bu kümeler için de uzun dönem ilişkisi araştırılmıştır. Uzun dönem ilişkisi için Johansen kointegrasyon yönteminden yararlanılmıştır. Bu yöntemde, iz istatistiği ile maksimum özdeğer istatistiği olmak üzere iki hesaplanan istatistikten yararlanılır. Bu iki istatistiğin birbirini desteklemediği durumlarda Engle-Granger nedensellik testine başvurulmuştur.

Bu sonuçlarda sadece ABD için bir tane kointegrasyon denklemi elde edilmiştir. Bu denklem incelendiğinde yenilenebilir enerji ile CO_2 emisyonu arasındaki ilişki beklenildiği gibi negatif yönlü olmuştur. Kişi başına GSYİH ile CO_2 emisyonunun aynı yönde hareket edeceği yorumu yapılabilir.

Bu çalışmanın ilerleyen aşamalarında CO_2 emisyonu ile GSYİH arasındaki ilişki hakkında ortaya konan Çevresel Kuznet Eğrisi bakımından kümelerin durumu

incelenebilir. Bu açıdan kümelerin özelliklerine göre isimlendirmeleri yapılabilir. Burada kümeler için Çevresel Kuznet Eğrisi'nin geçerliliği sınanabilir. Dönüm noktası hesaplanarak CO₂ emisyonunda azalmanın başlanacağı GSYİH değeri elde edilebilir.



KAYNAKLAR

- Akay, Ö. 2018. Panel Verilerde Kümeleme Analizi. Doktora Tezi. Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, 131, Adana.
- Akdi, Y. 2010. Zaman Serileri Analizi(Birim Kökler ve Kointegrasyon). Gazi Kitabevi, 396, Ankara
- Anderberg, M. R. 1973. Cluster Analysis for Applications. Academic Press, 359, New York.
- Anonymous. 2020. Web Sitesi: <https://data.oecd.org/energy/renewable-energy.htm>, Erişim Tarihi: 20.11.2020.
- Anonymous. 2020. Web Sitesi: <https://data.oecd.org/air/air-and-ghg-emissions.htm>, Erişim Tarihi: 20.11.2020.
- Anonymous. 2020. Web Sitesi: <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>, Erişim Tarihi: 20.11.2020.
- Apergis, N., Payne, J.E., Menyah, K., ve Wolde-Rufael, Y. 2010. On the causal dynamics between emissions, nuclear energy, renewable energy, and economic growth, Ecological Economics,69(11): 2255-2260
- Asteriou, D. ve Hall, S. G. 2006. Applied Econometrics. Palgrave Macmillan, 421, London.
- Balsalobre-Lorente, D., Leitão, N. C. ve Bekun, F. V. 2021. Fresh Validation of the Low Carbon Development Hypothesis under the EKC Scheme in Portugal, Italy, Greece and Spain. Energies, 4, 1-17.
- Baltagi, B. H. 2005. Econometric Analysis of Panel Data Third edition. John Wiley & Sons Ltd, 302, New York
- Banzo, D.C. ve Hermosilla, A. G. 2002. Clustering Panel Data via Perturbed Adaptive Simulated Annealing, Advances in Complex System, Vol 5(4), 339-360.
- Baştürk, N., Paap, R. ve van Dijk. D. 2014. Structural differences in economic growth: an endogenous clustering approach. Applied Economics 44:1, 119-134.
- Becker, R. A., Cleveland, W. S. ve Shyu, M.-J. 1996. The Visual Design and Control of Trellis Display. Journal of Computational and Graphical Statistics, 5:2, 123-155.
- Bhargava, A., Franzini, L. ve Narendranathan, W. 1982. Serial correlation and the Fixed Effects Model. The Review of Economic Studies, 49(4), 533-549.
- Bilgili, F., Koçak, E. ve Bulut, Ü. 2016. The dynamic impact of renewable energy consumption on CO2 emissions: A revisited Environmental Kuznets Curve approach. Renewable and Sustainable Energy Reviews, Vol 54, 838-845.

- Bingyun Zheng, S. L. (2014). Multivariable panel data cluster analysis and its application. *Computer Modelling & New Technologies*, 553-557.
- Blashfield, R. ve Aldenderfer, M. 1978. The Literature on Cluster Analysis. *Multivariate Behavioral Research*, 13, 271-295.
- Bölük, G. ve Mert, M. 2014. Fossil & renewable energy consumption, GHGs (greenhouse gases) and economic growth: Evidence from a panel of EU (European Union) countries. *Energy*, 74: 439-446.
- Breusch, T.S. ve Pagan, A. R. 1980. The Lagrange Multiplier Test and its Application to Model Specification in Econometrics. *Review of Econometric Studies*, Vol 47, Issue 1, 239-253.
- Brocklebank, J. C. ve Dickey, D. A. 2003. SAS system for forecasting time series, 2nd Edition. SAS Institute, 390, Cary, NC
- Dong, K., Sun, R., Hockman, G. 2017. Do natural gas and renewable energy consumption lead to less CO2 emission? Empirical evidence from a panel of BRICS countries, *Energy*, 141: 1466-1478.
- Driscoll, J. C. ve Kraay, A. C. 1998. Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *The Review of Economics and Statistics*, Vol. 80, No. 4,, 549-560.
- Erişoğlu, M. 2011. Uzaklık Ölçülerinin Kümeleme Analizine olan Etkilerinin İncelenmesi ve Geliştirilmesi. Doktora Tezi, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, 199, Adana.
- Everit, B. 1974. Cluster Analysis. Heinemann Educational Books Ltd, 122, London.
- Frees, E. W. 2004. Longitudinal and Panel Data: Analysis and Applications for the Social Sciences, Cambridge University Press, 364, Cambridge.
- Frühwirth-Schnatter, S. 2011. Panel data analysis: a survey on model-based clustering of time series. *Advances in Data Analysis and Classification* 5, 251–280.
- Green, E. P. 1978. Analysing Multivariate Data,. Dryden Press, 516, California.
- Greene, W. H. 2002. Econometric Analysis 5. Edition. Pearson Education, Inc., 828, New Jersey.
- Guisan, M.C. 2009. Government Effectiveness, Education, Economic Development and Well-Being: Analysis of European Countries in Comparison with The United States and Canada, 2000-2007. *Applied Econometrics and International Development*. Vol. 9-1 , 39-55.
- Gujarati, D. N. 2004. Basic Econometrics 4th Edition. McGraw-Hill, 979, New York.
- Hamiltona, C. ve Turton, H. 2002. Determinants of emissions growth in OECD countries. *Energy Policy* 30 , 63-71.

- Han, J., Kamber, M. ve Pei, J. 2012. Data Mining Concepts and Techniques Third Edition. Elsevier, 740, Waltham.
- Hausman, J. A. 1978. Specification Tests in Econometrics. *Econometrica*, Volume 46, 1251-1271.
- Hawkins, D. M. 1982. Topics in Applied Multivariate. Cambridge Uni.Press, 362, Cambridge.
- Hoechle, D. 2007. Robust standard errors for panel regressions with cross-sectional dependence. *The Stata Journal*, 7, 281-312.
- <https://unfccc.int/process-and-meetings/the-kyoto-protocol/what-is-the-kyoto-protocol/kyoto-protocol-targets-for-the-first-commitment-period>, Erişim Tarihi: 25.10.2021.
- Hubert, L. 1974. Approximate Evaluation Techniques for the Single-Link and Complete-Link Hierarchical Clustering Procedures. *Journal of the American Statistical Association*, 69, 698-704.
- Johansen, S. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, Vol12, 231-254.
- Johnson, R. ve Wichern, D. 1998. Applied Multivariate Statistical Methods 6th Edition. Prentice Hall, 393, New Jersey.
- Juárez, M. A. ve Steel, M. F. 2010. Model-Based Clustering of Non-Gaussian Panel Data Based on Skew-t Distributions. *Journal of Business & Economic Statistics*, 1, 52-66.
- Kosmelj, K. ve Batagelj, V. 1990. Cross-Sectional Approach for Clustering Time Varying Data. *Journal of Classification* 7, 99-109.
- Li, X. ve Lin, B. 2013. Global convergence in per capita CO2 emissions. *Renewable and Sustainable Energy Reviews*, 24, 357-363.
- Lombardoa, R. ve Falcone, M. 2011. Crime and Economic Performance. A cluster analysis of panel data on Italy's NUTS 3 regions. Università della Calabria, Dipartimento di Economia, Statistica e Finanza (Ex Dipartimento di Economia e Statistica), October, Working Papers, 1-33, Italy.
- Long, X., Naminse, E.Y., Du, J. ve Zhuang, J. 2015. Nonrenewable energy, renewable energy, carbon dioxide emissions and economic growth in China from 1952 to 2012, *Renewable and Sustainable Energy Reviews*, 52: 680-688.
- Lu, H., Huang, S., Li, Y. ve Yang, Y. 2014. Panel Data Analysis Via Variable Selection and Subject Clustering. In: Yada, K. (eds) *Data Mining for Service. Studies in Big Data*, Vol 3. Springer, Berlin.
- Marriot, F. 1971. Practical Problems in a method of Cluster Analysis . *Biometrics*, 27, 501-514.

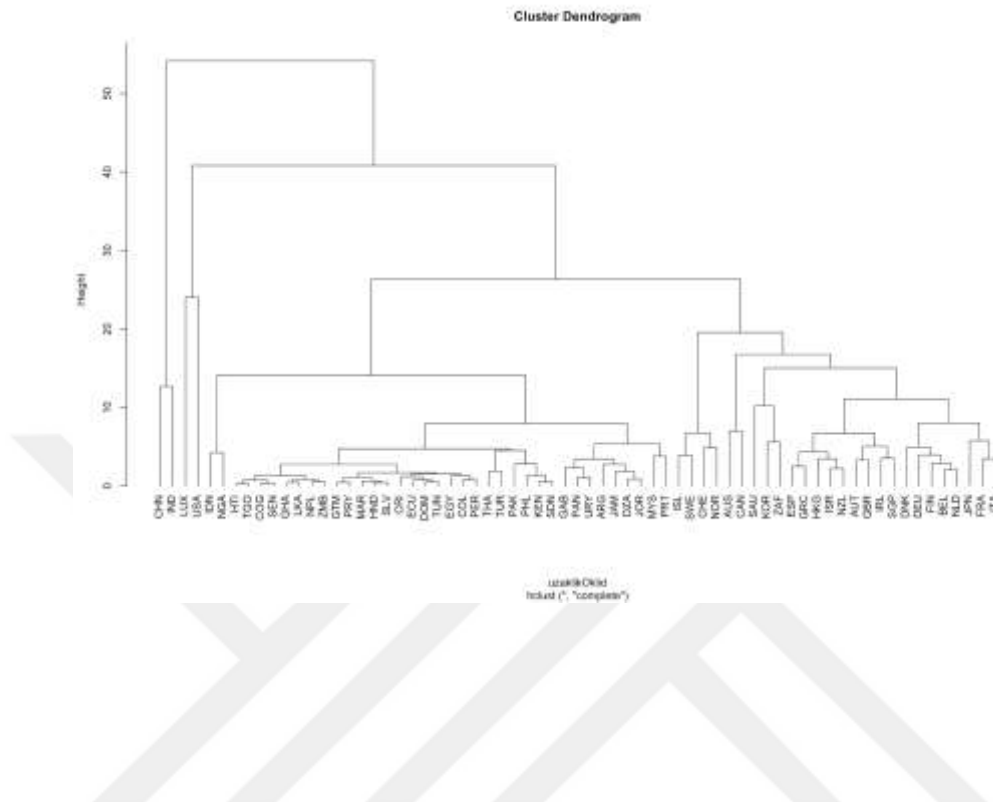
- Menyah, K. ve Wolde-Rufael, Y. 2010. CO2 emissions, nuclear energy, renewable energy and economic growth in the US, *Energy Policy*, 38(6): 2911-2915.
- Millo, G. 2017. Robust Standard Error Estimators for Panel Models: A Unifying Approach. *Journal of Statistical Software*, 82(3), 1–27.
- Mott, G., Razo, C. ve Hamwey, R. UNCTAD. <https://unctad.org/news/carbon-emissions-anywhere-threaten-development-everywhere>, Erişim Tarihi: 25.01.2023.
- Newey, W. K., and West, K. D., 1987. A Simple, Positive Semi-definite and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3): 703- 708.
- Newey, W. K., and West, K. D., 1994. Automatic Lag Selection in Covariance Matrix Estimation. *The Review of Economic Studies*, 61(4):631-653.
- Nie, G., Chen, Y., Zhang, L. ve Guo, Y. 2010. Credit card customer analysis based on panel data clustering. *Procedia Computer Science Volume 1, Issue 1*, 2489-2497.
- Özdamar, K. 2004. Paket Programlar ile İstatistiksel Veri Analizi (Çok Değişkenli Analizler).Kaan Kitabevi, 649, Eskişehir.
- Pesaran, M. H. 2004. General Diagnostic Tests for Cross Section Dependence in Panels. *IZA Discussion Paper No. 1240*, 1-41.
- Rencher, A. C. 2002. *Methods of Multivariate Analysis*, Second Edition. John Wiley & Sons, Inc., 695, New York.
- Rousseeuw, P. J. 1987. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics* 20, 53-65.
- Rudin, W. 1976. *Principles of Mathematical Analysis*. McGraw-Hill, Inc. 352. New York.
- Saidi, K. ve Mbarek, M.B. 2016. Nuclear energy, renewable energy, CO2 emissions, and economic growth for nine developed countries: Evidence from panel Granger causality tests, *Progress in Nuclear Energy*, 88: 364-374.
- Sokal, R. R. ve Sneath, P. H. 1963. *Principles of numerical taxonomy*. W. H. Freeman and Company, 963, San Francisco.
- Stock, J. H. ve Watson, M. W. 2019. *Introduction to Econometrics*, 4th edition. Pearson Addison Wesley, 726, Boston,
- Sui, B. Z., Zheng, B. ve Li, S. 2014. Multivariable panel data cluster analysis and its application. *Computer Modelling & New Technologies*, 553-557.
- Tatlıdil, H. 1996. *Uygulamalı Çok Değişkenli İstatistiksel Analiz*. Cem Ofset Ltd.Şti., 350, Ankara.
- Ward, J. 1963. Hierarchical grouping to optimise an objective function. *Journal of the American Statistics Association*. Vol. 58, 236-244.

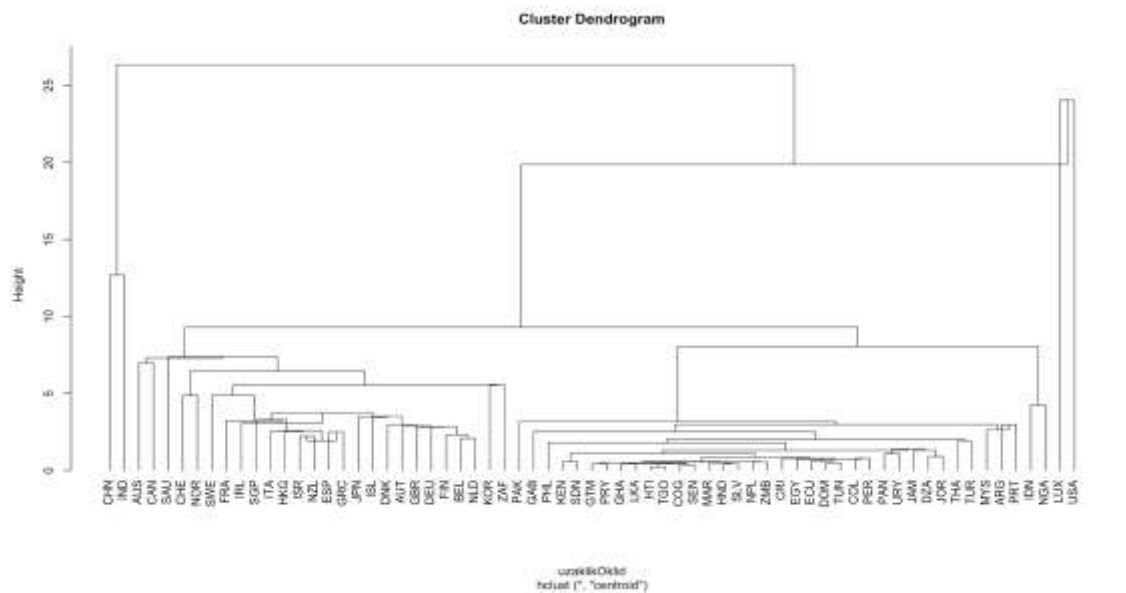
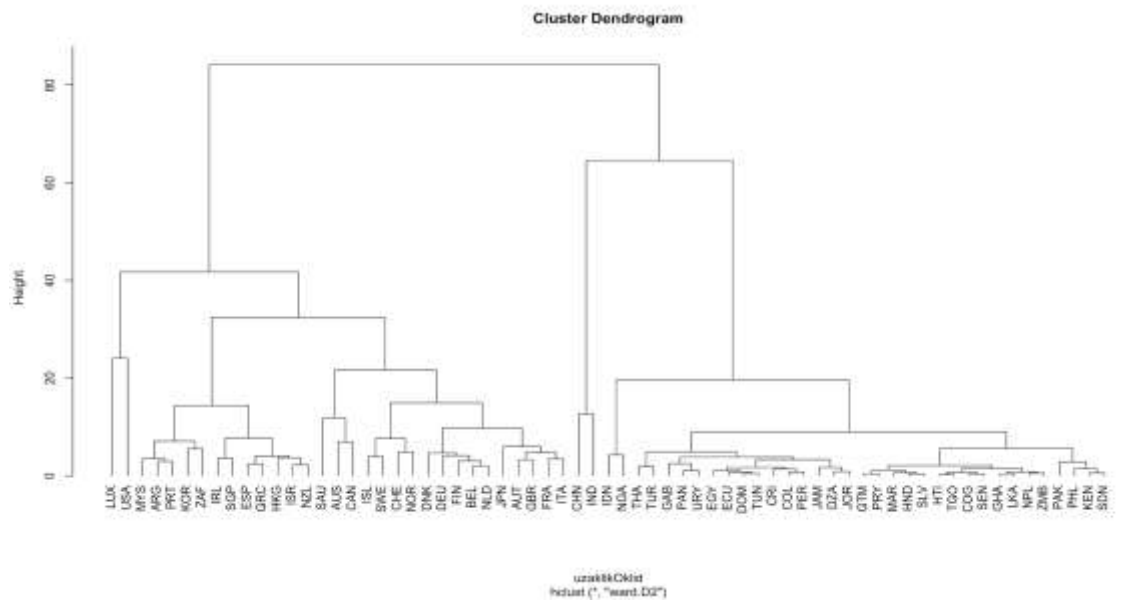
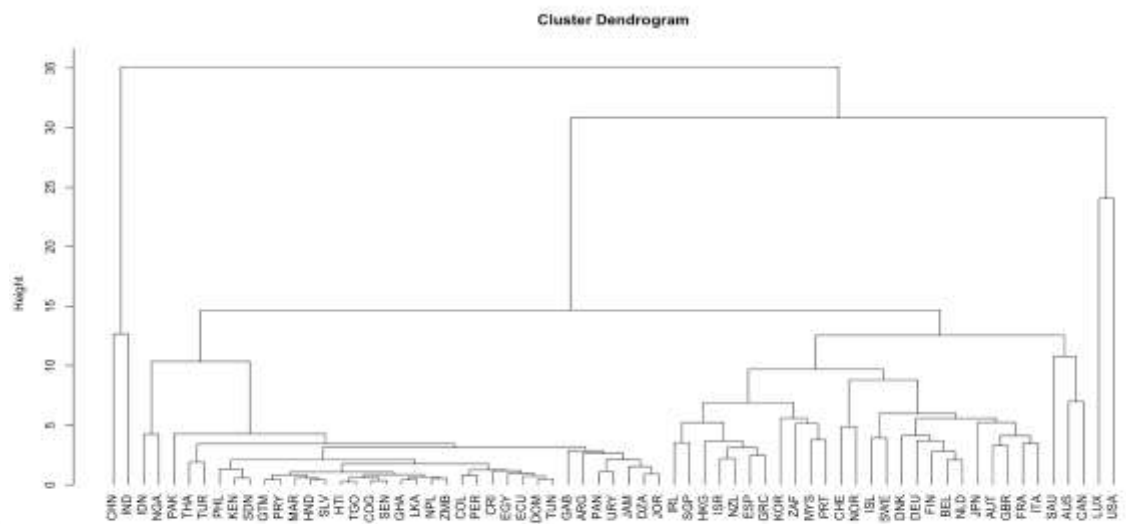
- Wooldridge, J. M. 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press, 741, Massachusetts
- Xu, R. ve Wunsch, D. C. 2009. Clustering. John Wiley&Sons, 358, New Jersey.
- <https://www.imf.org/external/datamapper/NGDPDPC@WEO/OEMDC/ADVEC/WEO/WORLD> - Eriřim Tarihi: 25.01.2023
- Yao, S., Zhang, S. ve Zhang, X. 2019. Renewable energy, carbon emission and economic growth: A revised environmental Kuznets Curve perspective, Journal of Cleaner Production, 235: 1338-1352.
- Zheng, B., and Li, S. 2014. Multivariable Panel Data Cluster Analysis and Its Application. Computer Modeling &New Technologies, 18:553-557

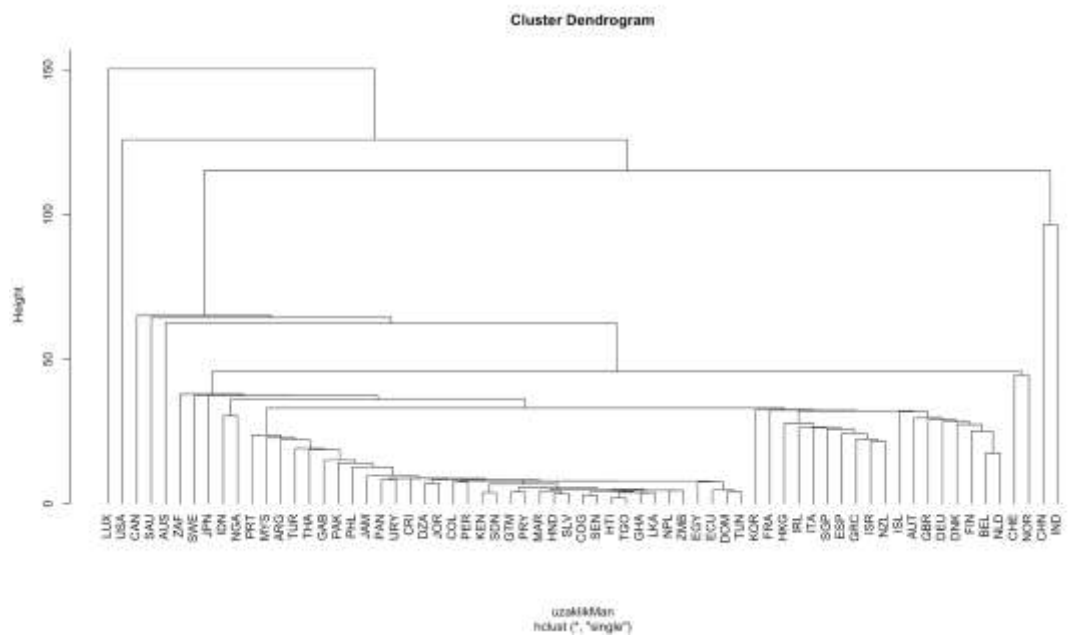
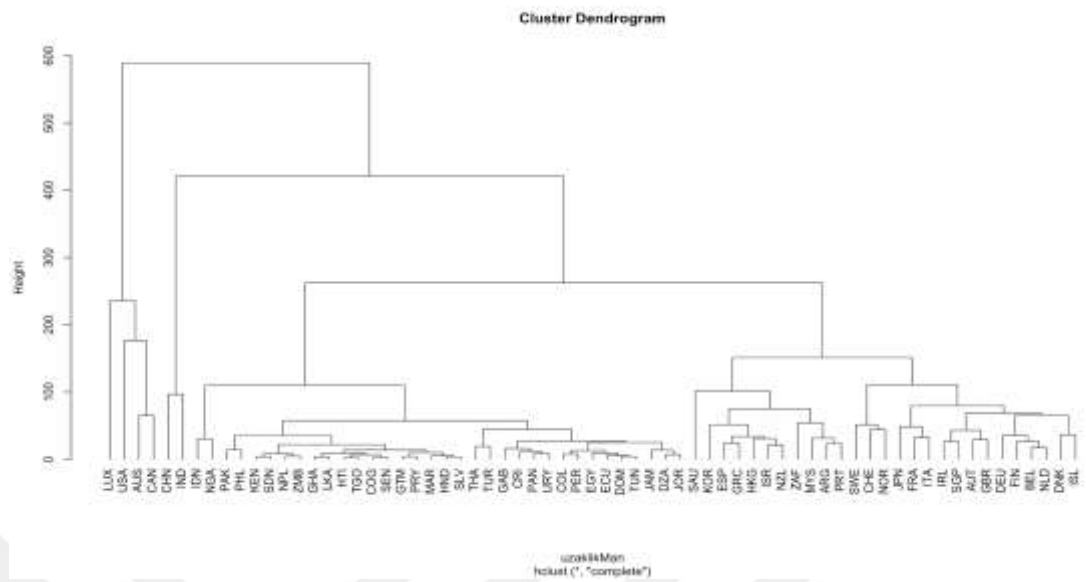


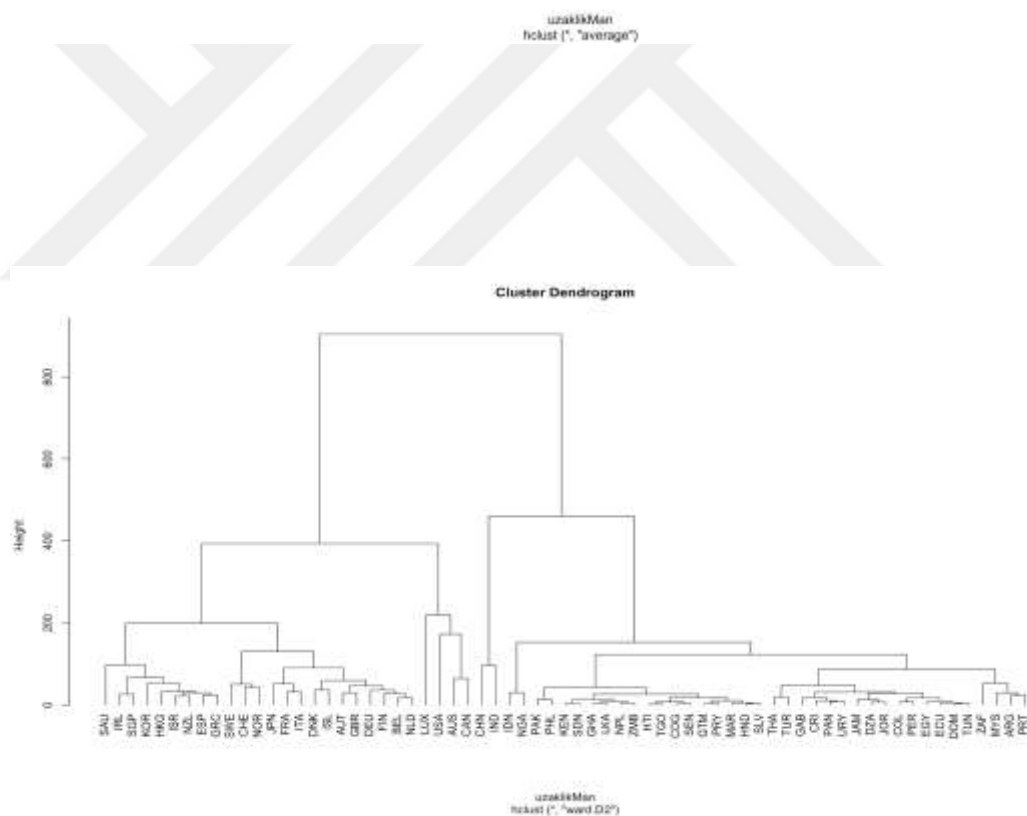
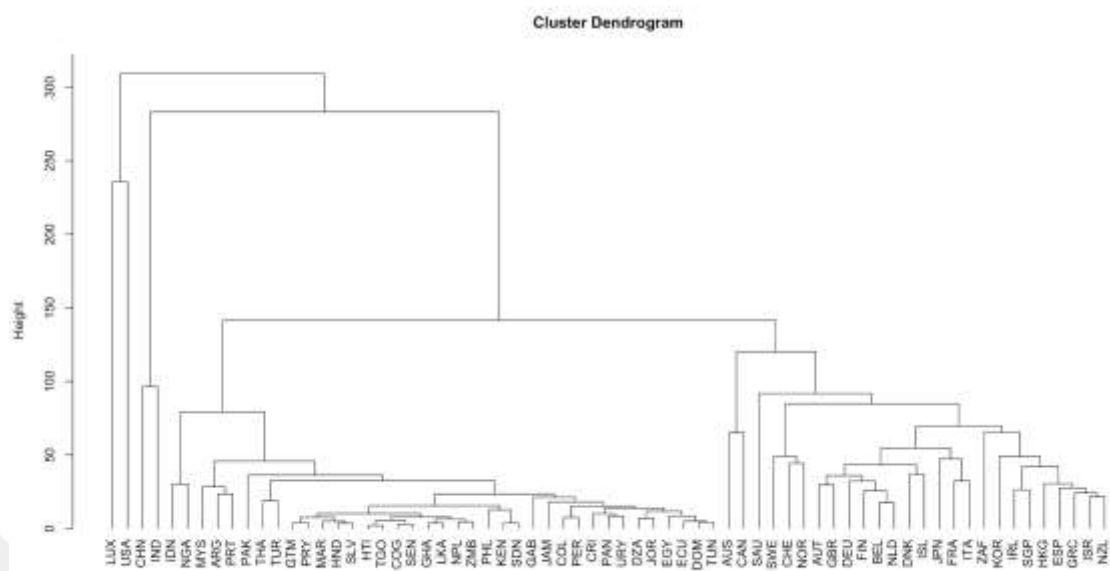
EK 1

Dendogramlar



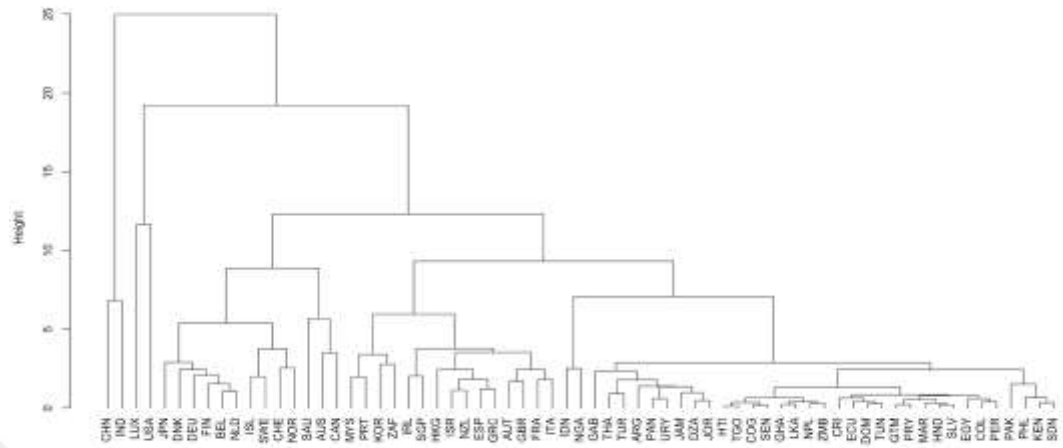






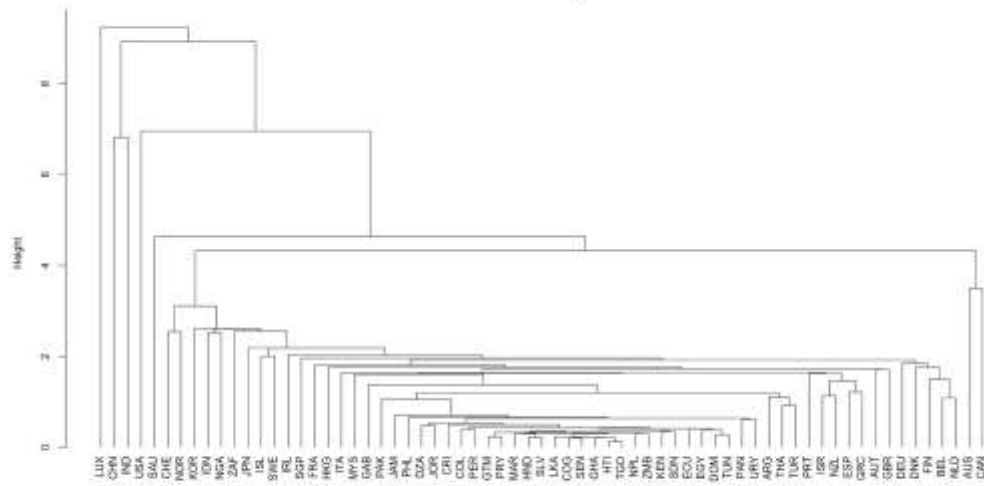
Cluster Dendrogram

Cluster Dendrogram

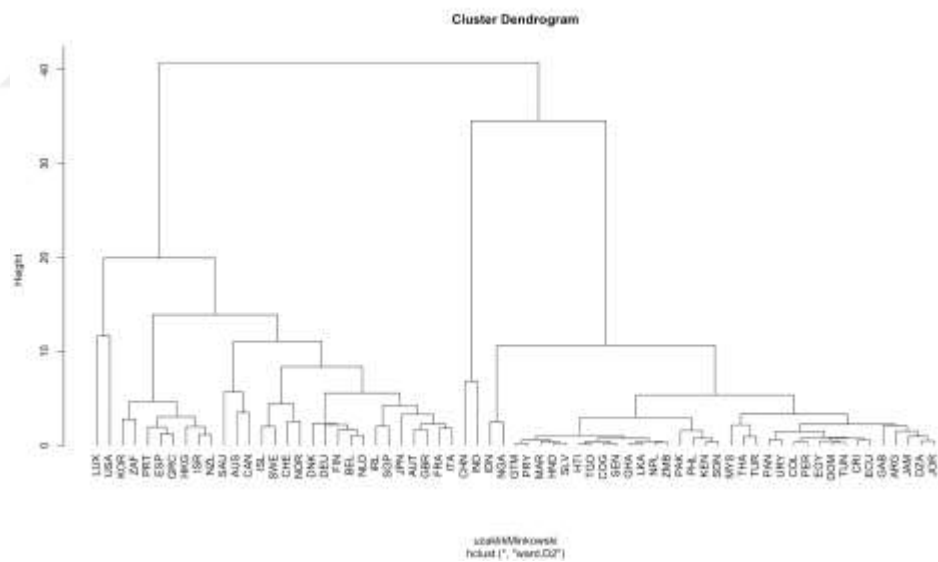
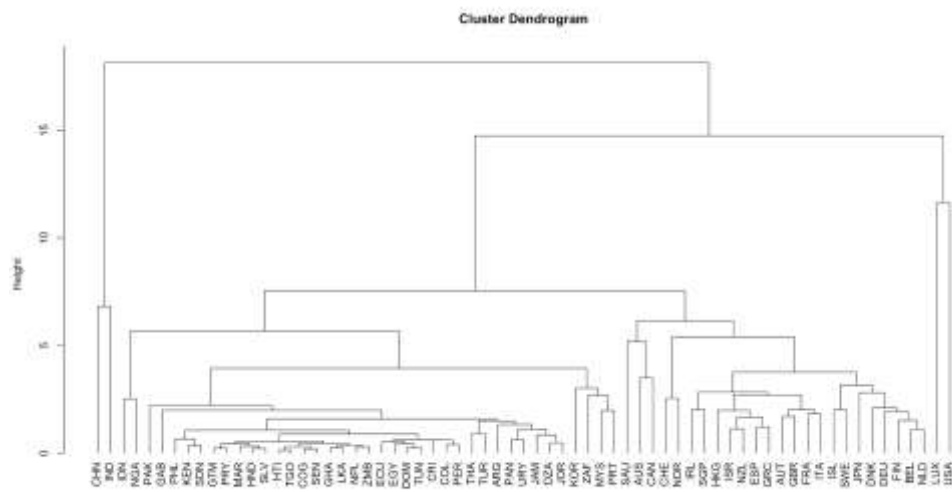


uzaklikMinkowski
hclust ("", "complete")

Cluster Dendrogram



uzaklikMinkowski
hclust ("", "single")



EK 2

CO₂ değişkeninin ülkelere göre tanımlayıcı istatistikleri

Ülke	Min(CO2)	Max(CO2)	Ort(CO2)	Med(CO2)	SD(CO2)
ARG	2.87	4.20	3.47	3.36	0.402
AUS	12.85	18.3	15.71	15.64	1.75
AUT	6.53	9.06	7.49	7.33	0.70
BEL	7.81	13.47	10.50	10.61	1.39
CAN	14.40	17.24	15.77	15.63	0.78
CHE	4.37	6.39	5.72	5.85	0.51
CHN	1.12	6.8	3.14	2.36	1.96
COG	0.11	0.65	0.33	0.33	0.14
COL	1.18	1.61	1.36	1.33	0.11
CRI	0.67	1.53	1.13	1.17	0.28
DEU	8.70	14.07	10.98	10.54	1.60
DNK	5.50	13.62	9.99	10.37	1.97
DOM	0.96	2.25	1.55	1.51	0.45
DZA	0.82	3.29	2.12	2	0.59
ECU	0.85	2.34	1.56	1.44	0.38
EGY	0.66	2.24	1.52	1.39	0.46
ESP	4.33	7.64	5.58	5.34	0.98
FIN	7.70	13.67	10.51	10.55	1.35
FRA	4.51	8.65	6.12	5.92	1.09
GAB	0.96	2.52	1.51	1.33	0.40
GBR	5.45	11.08	8.83	9.04	1.33
GHA	0.12	0.52	0.28	0.24	0.12
GRC	3.73	8.86	6.57	6.81	1.46
GTM	0.29	1.02	0.62	0.61	0.20
HKG	2.44	7.25	5.38	5.93	1.32
HND	0.39	1.01	0.67	0.6	0.23
HTI	0.03	0.3	0.16	0.14	0.06
IDN	0.29	1.84	1.04	1.05	0.51
IND	0.35	1.64	0.82	0.75	0.39
IRL	6.66	11.12	8.57	8.33	1.29
ISL	4.87	8.28	7.08	7.36	0.76
ISR	4.67	9.33	7.27	7.6	1.50
ITA	5.25	7.87	6.57	6.48	0.73
JAM	1.89	4.27	3.02	3.09	0.64
JOR	1.04	3.2	2.51	2.59	0.51
JPN	6.91	9.71	8.46	8.82	0.82
KEN	0.17	0.34	0.24	0.23	0.04
KOR	2.20	11.69	7.23	7.92	3.18
LKA	0.19	1.04	0.45	0.43	0.24
LUX	14.45	39.12	24.47	24.47	6.46

MAR	0.54	1.63	1.03	0.97	0.34
MYS	1.33	7.37	4.18	4.33	2.09
NGA	0.17	0.52	0.36	0.36	0.07
NLD	8.87	11.15	9.97	10.02	0.51
NOR	5.89	8.34	7.05	7.13	0.64
NPL	0.02	0.37	0.10	0.08	0.08
NZL	5.15	8.12	6.61	6.64	0.87
PAK	0.28	0.88	0.57	0.64	0.18
PAN	1.00	2.54	1.71	1.61	0.45
PER	0.82	1.69	1.12	1.03	0.24
PHL	0.49	1.17	0.76	0.76	0.14
PRT	1.96	5.92	4.02	4.38	1.27
PRY	0.25	1.12	0.60	0.62	0.20
SAU	3.04	16.75	11.13	11.13	3.32
SDN	0.11	0.48	0.23	0.18	0.10
SEN	0.27	0.54	0.36	0.35	0.07
SGP	3.73	11.11	7.92	8.23	2.05
SLV	0.31	1.16	0.72	0.81	0.29
SWE	3.66	10.44	6.32	6.34	1.72
TGO	0.06	0.37	0.16	0.15	0.06
THA	0.51	3.61	2.08	2.39	1.11
TUN	0.86	2.29	1.65	1.59	0.41
TUR	1.49	4.71	2.74	2.8	0.89
URY	1.00	2.42	1.61	1.59	0.34
USA	14.64	21.71	18.77	19.07	1.75
ZAF	6.09	8.47	7.20	7.17	0.72
ZMB	0.14	0.88	0.326	0.27	0.187

REN değişkeninin ülkelere göre tanımlayıcı istatistikleri

Ülke	Min(REN)	Max(REN)	Ort(REN)	Med(REN)	SD(REN)
ARG	1947.66	6633.09	4128.77	4170.24	1213.40
AUS	4700.80	8897.42	5973.55	5582.32	1116.01
AUT	2443.93	10081.40	6244.04	5902.95	2264.69
BEL	27.38	4047.57	1075.73	527.11	1279.39
CAN	23631.82	50572.53	39292.24	41478.11	7631.78
CHE	2430.48	5135.15	4024.38	4015.24	662.02
CHN	171944.49	279310.03	210042.60	208310.72	22123.28
COG	348.95	1699.64	712.94	549.30	409.23
COL	5094.69	10688.00	7304.98	7261.24	1433.46
CRI	385.33	2621.62	1288.17	984.85	797.88
DEU	4074.54	41446.85	14267.95	6281.01	12453.86
DNK	321.23	5646.99	2012.88	1374.35	1486.33

DOM	779.95	1731.75	1064.56	973.24	246.01
DZA	22.12	119.11	53.39	52.08	21.67
ECU	1016.35	2323.70	1344.63	1329.55	241.81
EGY	2098.02	3696.72	3019.53	3067.18	519.64
ESP	1843.32	17790.18	7669.79	6222.44	4972.25
FIN	3977.05	11238.11	6836.56	6162.21	2166.26
FRA	13016.21	24231.37	17136.41	16274.23	2742.82
GAB	544.73	4121.86	1639.83	914.71	1264.42
GBR	317.11	16843.28	3590.43	1835.83	4455.61
GHA	2823.57	5641.19	4041.68	3971.09	747.07
GRC	611.55	2835.14	1393.37	1362.26	676.54
GTM	2372.63	8860.41	4468.72	3409.09	2017.62
HKG	1.29	94.49	36.42	27.85	26.72
HND	1112.28	2928.87	1781.28	1670.11	466.34
HTI	1820.63	3409.03	2440.77	2301.19	526.92
IDN	27719.18	66480.96	48542.61	52203.02	13187.15
IND	107756.73	208241.09	153917.50	147396.18	30066.78
IRL	44.80	1309.80	325.79	174.12	337.71
ISL	578.61	5284.23	2384.68	1634.59	1560.27
ISR	2.56	1164.41	445.47	444.82	328.36
ITA	6472.42	26512.23	12286.21	8354.84	7007.90
JAM	199.06	312.19	261.12	261.11	28.96
JOR	0.69	372.14	74.32	62.59	80.98
JPN	6532.07	24010.94	14872.70	15405.87	4316.22
KEN	4723.78	21603.00	11128.62	10421.31	4730.41
KOR	119.78	4718.74	1081.04	678.90	1186.93
LKA	3056.63	5930.85	4337.13	4321.74	775.31
LUX	3.96	263.07	63.55	34.58	65.77
MAR	796.66	2395.06	1322.33	1274.53	404.51
MYS	1246.90	2908.80	1616.23	1595.25	303.80
NGA	34551.51	117643.10	67947.25	62360.78	25618.53
NLD	3.70	4021.81	1428.86	1038.87	1297.28
NOR	6643.67	14151.14	10835.72	11045.62	2006.49
NPL	3926.32	10175.32	6652.47	6273.80	1935.46
NZL	2753.44	8478.44	4892.23	4626.75	1703.81
PAK	12443.09	38265.18	24257.57	23752.43	7733.96
PAN	380.05	995.86	630.15	610.55	123.36
PER	3324.46	5997.75	4515.19	4330.66	788.74
PHL	8537.28	18771.24	15290.74	16044.86	2710.69
PRT	1045.41	5625.30	3256.66	3425.77	1389.85
PRY	1249.80	8559.87	5012.18	6156.78	2364.54
SAU	0.97	20.86	7.32	5.98	4.54
SDN	6182.25	12347.17	9798.34	10634.01	1977.13
SEN	852.87	2151.99	1204.90	1042.54	385.32
SGP	0.00	409.49	136.77	101.03	126.73
SLV	1348.87	2249.96	1885.78	1900.80	213.81

SWE	7985.74	19578.37	12969.80	12558.54	3154.89
TGO	667.84	2761.92	1533.37	1422.64	682.72
THA	9000.04	28862.21	16225.53	14874.77	5334.99
TUN	445.91	1282.28	835.48	792.61	278.50
TUR	7384.05	17738.49	10472.28	10020.82	2137.74
URY	520.82	3183.36	1295.23	1132.18	653.25
USA	65775.23	162450.02	105800.40	101974.20	23124.29
ZAF	5923.34	12171.55	9144.27	9101.37	1785.66
ZMB	3011.087	10138.02	5.85E+03	5275.494	1997.939

GSYİH değişkeninin ülkelere göre tanımlayıcı istatistikleri

Ülke	Min(GSYİH)	Max(GSYİH)	Ort(GSYİH)	Med(GSYİH)	SD(GSYİH)
ARG	1948.225	14591.86	6424.044	5919.012	3800.261
AUS	6994.21	68150.11	26890.31	20082.48	18520.48
AUT	5285.621	51717.5	27303.58	25646.7	15390.59
BEL	6701.377	48106.89	25897.18	24208.55	13860.49
CAN	7511.211	52678.39	26301.47	21448.36	14248.14
CHE	10028.99	88415.63	44132.27	40417.96	24224.68
CHN	156.3964	8879.438	1980.212	709.4138	2653.476
COG	469.6905	3286.695	1366.768	1058.229	805.1108
COL	544.2908	8218.034	3009.602	2284.839	2285.584
CRI	936.3359	11814.63	4271.586	3247.317	3347.126
DEU	6236.359	47959.99	26177.95	26438.23	13270.68
DNK	7999.113	64322.06	33741.2	30751.65	18418.3
DOM	699.6023	7609.339	2892.503	2293.755	2125.17
DZA	936.79	5592.257	2691.842	2237.086	1329.636
ECU	1105.42	6377.092	2804.356	2155.517	1623.803
EGY	301.7494	3598.971	1347.037	1063.004	969.4753
ESP	3209.837	35366.26	16563.95	15355.7	10066.15
FIN	6255.545	53554.04	27437.82	25484.74	14893.15
FRA	6690.578	45334.11	24798.67	23813.71	12215.54
GAB	3332.005	10809.65	5729.327	4736.046	2094.841
GBR	4138.168	50566.83	25369.27	24332.7	14691.1
GHA	258.471	2345.393	698.8428	397.1359	591.7076
GRC	3153.235	31997.28	13727.62	12538.18	8224.556
GTM	594.6518	4451.451	1834.382	1487.608	1051.201
HKG	2252.112	46165.86	20848.68	23977.02	12848.93
HND	356.4573	2453.728	1287.96	1122.064	520.4226
HTI	133.7501	831.874	439.7022	396.6338	201.5654
IDN	233.1167	3837.652	1320.662	780.1902	1168.659
IND	158.0362	1981.651	638.0158	399.9501	508.5402
IRL	2923.08	69822.35	27670.74	20860.6	21633.38

ISL	6463.879	71310.94	31905.91	27659.88	17719.39
ISR	4297.516	40541.86	18116.71	18435.89	10733.95
ITA	4033.099	40778.34	21425.77	21829.35	11278.58
JAM	899.2568	5209.852	2900.043	2879.658	1475.69
JOR	659.7758	4162.821	2157.204	1783.723	1043.456
JPN	4659.12	48603.48	28526.87	34524.47	13273.97
KEN	220.0697	1572.336	570.688	397.4827	366.454
KOR	617.4561	31616.84	12276.79	10672.42	9691.087
LKA	188.0585	4077.044	1236.386	756.6571	1225.808
LUX	9008.028	118823.6	54567.24	47584.46	37626.78
MAR	502.7567	3171.699	1619.105	1339.294	865.0186
MYS	764.5665	11319.08	4650.082	3728.112	3261.437
NGA	270.224	3222.694	1177.79	741.7475	861.6607
NLD	7335.509	57644.48	28927.23	26700.54	15996.89
NOR	8204.452	102913.5	44166.85	34875.7	30374.18
NPL	98.542	911.4443	310.5172	204.6858	227.5995
NZL	4171.77	44553.28	19277.71	14738.45	12420.34
PAK	168.0804	1464.993	614.1394	461.2167	366.0945
PAN	1395.479	15150.35	5059.071	3529.852	3741.878
PER	729.8763	6756.753	2586.369	1941.476	1982.801
PHL	360.7533	3123.234	1259.432	990.5639	837.3082
PRT	2127.638	24847.55	11791.24	11575.51	7653.549
PRY	541.487	6102.943	2312.749	1663.605	1655.894
SAU	5856.377	25243.36	12388.61	9282.648	6009.954
SDN	260.5102	1909.739	716.4043	491.5773	446.1258
SEN	545.6263	1411.929	892.1982	823.7211	288.6613
SGP	2489.912	60913.75	24082.26	21796.08	18507.15
SLV	453.4921	3910.257	1808.978	1684.786	1087.403
SWE	10117.31	61126.94	32567.81	30312.49	16203.63
TGO	228.6483	640.9342	394.3872	359.7906	120.0952
THA	351.6193	6592.915	2555.835	2033.258	1911.61
TUN	765.1894	4307.156	2319.126	2113.586	1195.322
TUR	1136.375	12519.39	4928.833	3119.566	3898.161
URY	1250.199	17322.15	6368.279	4720.991	4919.45
USA	7801.457	59957.73	31813.58	29967.71	15666.96
ZAF	1416.709	8007.413	3867.592	3285.955	1807.393
ZMB	232.539	1878.907	732.8273	511.3213	484.6215

EK 3

Karbondioksit emisyonu bakımından oluşturulan kümeler için ADF ve PP birim kök sonuçları

Küme		ADF		
		Sabit	Sabit + Trend	1.Fark
1	log(co2)	0.8751	0.5124	0.0034*
	log(ren)	0.9939	0.6479	0.3694-
	log(gdp)	0.8444	0.2980	0.0014*
2	log(co2)	0.9275	0.8756	0.0004*
	log(ren)	0.7372	0.6167	0*
	log(gdp)	0.1832	0.4421	0.0023*
3	log(co2)	0.4000	0.9920	0*
	log(ren)	0.9975	0.8693	0*
	log(gdp)	0.2072	0.1846	0.0020*

*%5 anlam düzeyine göre yokluk hipotezi(H0) red edilemez.

Johansen test sonuçları 2.çeşit (karbon)

Küme	Hipotez	Özdeğer	İz İstatistiği	Max. Özdeğer İstatistiği
1	$r = 0$	0.337901	29.18910 (0.0586)	16.49363 (0.1973)
	$r \leq 1$	0.256517	12.69547 (0.1265)	11.85638 (0.1162)
	$r \leq 2$	0.020759	0.839083 (0.3597)	0.839083 (0.3597)
2*	$r = 0$	0.411850	27.75852 (0.0844)	21.13162 (0.0407*)
	$r \leq 1$	0.094523	5.996845 (0.6959)	4.071014 (0.8516)
	$r \leq 2$	0.045885	1.925831 (0.1652)	1.925831 (0.1652)
3	$r = 0$	0.212161	15.93332 (0.7170)	9.538445 (0.7868)
	$r \leq 1$	0.131451	6.394877 (0.6488)	5.637234 (0.6601)
	$r \leq 2$	0.018763	0.757643 (0.3841)	0.757643 (0.3841)

Yenilenebilir enerji bakımından oluşturulan kümeler için ADF ve PP birim kök sonuçları

Küme		ADF		
		Sabit	Sabit + Trend	1.Fark
1	log(co2)	0.3568	0.9747	0.0004*
	log(ren)	0.9943	0.9444	0*
	log(gdp)	0.2714	0.1463	0.0017*
2	log(co2)	0.3713	0.5520	0*
	log(ren)	0.1841	0.9699	0*
	log(gdp)	0.7089	0.4100	0.0003*
3	log(co2)	0.9482	0.3043	0.0083*
	log(ren)	0.8910	0.3972	0.7240
	log(gdp)	0.9976	0.6296	0.0002*
4	log(co2)	0.9553	0.9089	0.0009*
	log(ren)	0.8544	0.7615	0*
	log(gdp)	0.0059	0.0782	0.0053*

*%5 anlam düzeyine göre yokluk hipotezi(H0) red edilemez.

Johansen test sonuçları (Yenilenebilir enerji)

Küme	Hipotez	Özdeğer	İz İstatistiği	Maks. Özdeğer İstatistiği
1	$r = 0$	0.256124	19.88089 (0.4309)	11.54197 (0.5932)
	$r \leq 1$	0.177593	8.338923 (0.4299)	7.625256 (0.4181)
	$r \leq 2$	0.018133	0.713667 (0.3982)	0.713667 (0.3982)
2* (No Coint)	$r = 0$	0.331958	28.47851 (0.0704)	16.53959 (0.1949)
	$r \leq 1$	0.144326	11.93892 (0.1599)	6.390484 (0.5636)
	$r \leq 2$	0.126570	5.548438 (0.0185*)	5.548438 (0.0185*)
3*	$r = 0$	0.386582	30.42765 (0.0423*)	19.54834 (0.0820)
		0.172984	10.87931	7.597257

	$r \leq 1$	0.078775	(0.2190) 3.282049 (0.0700)	(0.4211) 3.282049 (0.0700)
	$r \leq 2$			
4	$r = 0$	0.291082	26.91354 (0.1038)	14.10463 (0.3565)
	$r \leq 1$	0.265725	12.80890 (0.1220)	12.66371 (0.0881)
	$r \leq 2$	0.003535	0.145188 (0.7032)	0.145188 (0.7032)

Not: Parantez içinde yer alan değerler, p değeridir.

* 0.05 düzeyinde istatistiksel olarak anlamlıdır.

GSYİH bakımından oluşturulan kümeler için ADF ve PP birim kök sonuçları

Küm e		ADF		
		Sabit	Sabit + Trend	1.Fark
1	log(co2)	0.9491	0.3912	0.0058*
	log(ren)	0.9936	0.7923	0.5
	log(gdp)	0.8869	0.4193	0.0023*
2	log(co2)	0.9671	0.9639	0.0001*
	log(ren)	0.9702	0.9428	0.0002*
	log(gdp)	0.1724	0.2559	0.0022*
3	log(co2)	0.8995	0.3024	0.0002*
	log(ren)	0.3934	0.0371*	0*
	log(gdp)	0.3132	0.8294	0.0007*
4	log(co2)	0.0383*	0.9325	0.0034*
	log(ren)	0.9775	0.3885	0*
	log(gdp)	0.2161	0.1079	0.0053*

*%5 anlam düzeyine göre yokluk hipotezi(H0) red edilemez.

Johansen test sonuçları (GSYİH)

Küme	Hipotez	Özdeğer	İz İstatistiği	Maks. Özdeğer İstatistiği
1*	$r = 0$	0.368099	35.64097 (0.0095*)	18.36091 (0.1169)
	$r \leq 1$	0.311605	17.28007 (0.0267*)	14.93573 (0.0391*)
	$r \leq 2$	0.056924	2.344338 (0.1257)	2.344338 (0.1257)
2	$r = 0$	0.287981	18.61065 (0.5212)	13.58602 (0.3999)
	$r \leq 1$	0.114703	5.024631 (0.8062)	4.873293 (0.7578)
	$r \leq 2$	0.003776	0.151338 (0.6973)	0.151338 (0.6973)
3	$r = 0$	0.379280	38.19351 (0.0043*)	19.07503 (0.0947)
	$r \leq 1$	0.284199	19.11848 (0.0136*)	13.37410 (0.0688)
	$r \leq 2$	0.133774	5.744375 (0.0165*)	5.744375 (0.0165*)
4	$r = 0$	0.303048	25.67126 (0.1388)	14.44156 (0.3299)
	$r \leq 1$	0.172072	11.22969 (0.1978)	7.55184 (0.4259)
	$r \leq 2$	0.087815	3.676509 (0.0552)	3.676509 (0.0552)