

ANALYSIS OF COVID-19 EPIDEMIC DATA ACCORDING TO THE COUNTRIES INCLUDING IN THE HUMAN DEVELOPMENT INDEX BY USING CLUSTERING ANALYSIS K-AVERAGE METHOD

Zafer Kanberođlu

Van Yüzüncü Yıl Üniversitesi, Turkey
zkanberoglu@yyu.edu.tr

Çetin Görür

İnönü Üniversitesi, Turkey
gorurcetin@hotmail.com

İbrahim Yıldırımçakar

Van Yüzüncü Yıl Üniversitesi, Turkey
ibrahimvanli.2156@gmail.com

Abstract:

One of the important disasters of the 21st century is the Covid-19 epidemic that started in China in December 2019 and spread all over the world in a short time. The Covid-19 epidemic, which started as a health problem, continues to deeply affect all structures of the countries with its economic and social dimensions. Although there is a parallelism between the development level of the country and its success in combating the epidemic, differentiation is observed from time to time. Within the scope of this study, Covid-19 epidemic data will be obtained by country and will be analyzed using the Cluster analysis K-Average Method. The sample of the study will be forty countries in total, including the first ten countries in each category, in different development categories in the United Nations human development index. In the first stage of the cluster analysis to be used in the study, the variables of total cases, daily cases, total deaths and daily deaths belonging to countries; In the second phase, total cases per million, daily cases per million, total deaths per million and deaths per million will be used. It will be analyzed whether the variables used in the said stages make a meaningful clustering. In this context, it will be revealed whether countries are in the same cluster according to their level of development in the fight against the Covid-19 epidemic.

Keywords: Cluster Analysis, K-Mean, Covid-19, ANOVA

JEL Codes:

KÜMELEME ANALİZİ K-ORTALAMA YÖNTEMİ KULLANILARAK COVID-19 SALGIN VERİLERİNİN İNSANİ GELİŞME ENDEKSİNDE YER ALAN ÜLKELERE GÖRE İNCELENMESİ

Özet:

21. yüzyılın önemli felaketlerinden birisi Aralık 2019'da Çin'de başlayan ve kısa sürede tüm dünyaya yayılan Covid-19 salgınıdır. Bir sağlık problemi olarak başlayan Covid-19 salgını, gelinen noktada iktisadi ve sosyal boyutlarıyla ülkelerin tüm yapılarını derinden etkilemeye devam etmektedir. Salgınla mücadelede ülkenin gelişmişlik düzeyi ve salgınla mücadele başarısı arasında paralellik görülse de dönem dönem farklılaşma görülmektedir. Bu çalışma kapsamında Covid-19 salgın verileri ülkelere göre elde edilerek Kümeleme analizi K-Ortalama Yöntemi kullanılarak incelenecektir. Çalışmanın örneklemini Birleşmiş Milletler insani gelişme endeksinde farklı gelişmişlik kategorilerinde yer alan ve her kategoride ilk on ülke olmak üzere toplam kırk ülke oluşturacaktır. Çalışmada kullanılacak kümeleme analizinin ilk aşamasında ülkelere ait toplam vaka, günlük vaka, toplam ölüm ve günlük ölüm değişkenleri; ikinci aşamada ise, milyonda toplam vaka, milyonda günlük vaka, milyonda toplam ölüm ve milyonda günlük ölüm değişkenleri kullanılacaktır. Söz konusu aşamalarda kullanılan değişkenlerin anlamlı bir kümeleme yapıp yapmadığı analiz edilecektir. Bu bağlamda Covid-19 salgını ile mücadelede ülkelerin gelişmişlik düzeylerine göre aynı kümede yer alıp almadıkları ortaya konulacaktır.

Anahtar Kelimeler: Kümeleme Analizi, K-Ortalama, Covid-19, ANOVA

JEL Kodları:

Giriş

İnsanoğlu var olduğu günden bugüne kadar pek çok iktisadi, toplumsal ve çevresel felaketlere maruz kalmıştır. Yakın dönemin en önemli felaketlerinden birisi Covid-19 olarak bilinen salgın hastalıktır. Söz konusu salgın 2019'un Aralık ayından bu tarafa tüm dünya ülkelerini belirgin bir şekilde etkilemektedir. Çin'de başladığı kabul edilen ve kısa sürede tüm dünyayı etkileyen salgında bugüne kadar 170 milyon civarında vaka, 3 milyon civarında ölüm kaydedilmiştir.

Covid-19 salgını ile birlikte ülkelerin sağlık sistemleri de sorgulanır olmuştur. Bugüne kadar tüm teknolojik alt yapısı ile çok güvenilen ülkelerin aynı güveni sağlık sektöründe ortaya koyamadığı da görülmüştür. Dünya üzerinde bunun pek çok örneği mevcuttur. Bu bağlamda ülkenin sağlık sektöründe diğer ülkelerle benzer yönlerinin ve farklılıklarının belirli bir sınıflandırmaya göre incelenmesi toplumların sağlığı açısından önem arz etmektedir.

Ülkelerin sağlık sektörleri ve toplumsal sağlığın hem kendi toplumları adına hem de dünyanın diğer toplumları adına önem arz etmektedir. Dünyanın daha önce yaşanan salgın hastalıklarında (Avrupa'da 1347-1351 arası yaşanan veba, Hindistan'da 1817-1824 arası yaşanan kolera, İspanya'da 1918-1920 arası yaşanan grip, ABD'de 1981 de görülen aids) görülen bu durum Covid-19 salgını ile bir kez daha toplum sağlığının ne kadar hayati bir öneme sahip olduğu ortaya koymuştur.

Toplum sağlığını yakından ilgilendiren salgın hastalıkların ülkeler bazında benzerlikleri ve sınıflandırılması salgınla mücadelede ülkelerin durumlarını bilmeleri ve bu durumdan kurtulmanın yollarını aramaları açısından önemli görülmektedir. Bu çalışma kapsamında insani gelişme endeksinde yer alan dört kategoride yer alan kırk ülke örneğinde salgının gösterge ve değerleri Kümeleme analizi yöntemi ile analiz edilmiştir.

Yazın Taraması

Covid 19 salgın verilerinin ülkeler bazında incelenmesine ilişkin teorik ve ampirik çalışmalar yazında mevcuttur. Bu çalışmalardan biri olan Demircioğlu ve Eşiyok (2020) Covid-19 salgın verilerini ülkelerin sağlık göstergeleri doğrultusunda analiz etmiş, 36 OECD ve AB ülkesi örneğinde K-ortalamar yöntemini kullanmıştır. Çalışmada sağlık göstergeleri açısından ülkeler benzerliklerine göre kümelenecek ve 36 ülkenin konumları birbirlerine göre değerlendirilmiştir.

Tekin (2020) Covid-19'un ülkelere etkilerini kümeleme analizinin hiyerarşik kümeleme ve Ward's yöntemi kullanılarak karşılaştırmalı olarak incelemiştir. OECD ülkeleri üzerinde yapılan bu çalışmada salgının her kıtadan her ülkeyi önemli derecede etkilediği, buna karşın bazı ülkelerin gösterdikleri refleksin etkisinin salgının yayılımını ve aynı zamanda etkin tedavi yöntemleri ve sağlık sistemlerinin yapısı ile ölüm sayılarının artmasını engelleyerek daha başarılı bir süreç izledikleri görülmüştür. Ve bu durumun bir sonucu olarak bu ülkelerde, salgının ülkenin bazı ekonomik ve finansal göstergelerine etkisinin sınırlı olduğunu tespit etmiştir.

Sığırlı (2006) sağlık düzeyi ölçütlerini 1998-2004 dönemi ve 25 ülke örneğinde çok boyutlu örnekleme analizi ile incelemiştir. Ülkelerin sağlık göstergeleri bakımından birbirlerine göre konumlarının, aralarındaki benzerlik ve farklılıkları ortaya konulmuştur. Çalışmada Türkiye'nin, Slovakya, Macaristan, Çek Cumhuriyeti dışındaki diğer ülkelerden birinci boyutta sağlık göstergeleri bakımından, ikinci boyutta sağlık harcamaları ve milli gelirden sağlığa ayrılan pay bakımından farklılık gösterdiğini tespit etmiştir.

Ersöz (2009) 2004 yılı için OECD üye ülkeleri örneğinde seçili sağlık göstergelerinin karşılaştırılmasında kümeleme ve ayırma analizi yardımıyla ülkelerin benzerliklerini incelemiştir. Yapılan kümeleme analizi sonucunda Türkiye'nin, OECD ülkeleri içerisinde aşamalı kümeleme yönteminde Polonya, Slovakya, Çek Cumhuriyeti, Macaristan, Meksika,

Kore Cumhuriyeti ile aşamalı olmayan kümeleme yöntemi k-ortalamar yönteminde Portekiz, Polonya, Slovakya, Macaristan, Çek Cumhuriyeti, Meksika ve Kore Cumhuriyeti ile medoid kümeleme yönteminde ise Meksika ile aynı kümede yer aldığı tespit edilmiştir.

Alptekin ve Yeşilaydın (2015) OECD'ye üye 34 ülkenin sağlık göstergelerini bulanık kümeleme analizi ile incelemiştir. Analiz sonucu en uygun küme sayısının beş olduğu; birinci kümede üç, ikinci kümede dokuz, üçüncü kümede dokuz, dördüncü kümede altı ve beşinci kümede yedi ülkenin yer aldığı tespit etmiştir. Türkiye'nin Estonya, Macaristan, Meksika, Polonya ve Şili ile aynı kümede yer aldığı belirlemiştir.

Mut ve Akyürek (2017) OECD ülkelerini sağlık göstergelerine göre kümeleme analizi ile incelemiştir. Çalışmada ülkelerin kümelenmesinde etkili olan göstergeler incelenmiş ve kümeler arasındaki farklılıklar değerlendirilmiştir. Analiz sonuçlarına göre Türkiye, Meksika ve Şili'nin aynı kümede yer aldığı tespit edilmiştir.

Küçükefe (2020) OECD ve Çin örneğinde GSYH düşüşlerini ve milyon kişi başına düşen ölümleri karşılaştırarak Covid-19 nedeniyle sağlık ekonomi mübadelesini araştırmıştır. Amprik veriler en yüksek ölüm oranlarına sahip ülkelerin en büyük ekonomik gerilemeler yaşadığını ortaya koymuştur. Yapılan kümeleme analizi cari hesap dengesi, GSYH büyümesi, bir milyon kişi başına düşen ölüm sayısı açısından ülkelerin üç bölüme ayrıldığını bulmuştur.

Neuburger ve Egger (2020) Avrupada Covid-19 dan etkilenen Almanya, İsviçre, Avusturya bölgesinde seyahat riski algısı ve seyahat davranışı arasındaki ilişkiyi incelemiştir. Kümeleme analizi yapılan çalışmada üç farklı küme tanımlanmıştır. Ayrıca analiz bulguları Covid-19 risk algısı, seyahat riski algısı ve seyahat davranışı arasında anlamlı bir artış olduğunu ortaya koymuştur.

Larco ve Cara (2020) ülke düzeyinde değişkenleri kullanarak Covid-19 vakalarının sayısına göre ülkeleri sınıflandırmıştır. 155 ülke örneğinde yapılan çalışmada k-ortalamar yöntemi kullanılmıştır. Vaka sayısına göre ülkeler beş veya altı küme oluşturmuştur. Ve model ülkeleri ölüm sayısı ile vaka ölüm oranı açısından sınıflandırmıştır. Çalışma özellikle salgın karşısında güçlü ve güçsüz ülkelerin belirlenmesini sağlamıştır.

Materyal ve Yöntem

Elde edilen çok sayıda veriyi istatistik olarak değerlendirmek amacıyla kullanılan klasik yöntemler her değişkene göre önemli bilgiler vermesine rağmen iki ya da daha fazla farklı özellik arasında bir ilişkinin varlığı hakkında gerçek bir bilgi sunmada yetersiz kalmakta ve homojen yapıya sahip örneklerin gruplandırılmasına da imkân vermemektedir. Bu nedenle orijinal değişkenlerin daha kolay yorumlanabilmesini sağlamak amacıyla çok değişkenli istatistiksel yöntemlere ihtiyaç duyulmaktadır (Demir, Esenbuğa ve Bilgin, 2016). Kümeleme analizi, bir veri matrisinde yer alan, doğal gruplamaları net bir şekilde bilinmeyen değişkenleri ya da birimleri birbirleriyle benzer olan kümelerle ayırma yani gruplama veya sınıflandırma yöntemidir (Özdamar, 2004). Kümeleme analizinin amacı, bireyleri ya da nesnelere temel özelliklerine göre gruplamaktır (Uçar, 2010). Bu sayede gruplanmamış verileri benzerlikleri ya da farklılıklarına göre gruplandırarak özet bilgiler elde edilir. Kümeleme analizi sonucu elde edilen kümelerdeki birimler kendi içlerinde homojen, birbirine çok yakın olmalı, diğer kümelerle heterojen birbirinden uzak olmalıdır. Kümeleme analizi verileri analiz etmede kullanılan

oldukça yararlı çok değişkenli istatistiksel analiz yöntemidir. Özellikle anket çalışmalarında elde edilen çok sayıda veriyi gruplandırmak ve anlamlandırmak zordur. Kümeleme analizi yöntemiyle, belirlenen kriterlere göre elde edilen tüm verileri kümelendirmek ve özet bilgiler elde etmek mümkündür. Kümeleme analizi ile faktör analizi bazı yönlerden benzerlik göstermektedir. Faktör analizi değişkenleri bağımlı ve bağımsız olmak üzere ikiye ayırmamaktadır. Kümeleme analizinde de durum böyledir. Ayrıca faktör analizi ve kümeleme analizi bireyleri ya da nesnelere aralarında bulunan benzerliklerden yola çıkarak sınıflandırma yapar. Yani iki analiz yönteminin ortak özelliklerinden biri de sınıflandırma yapmalarıdır (Uçar, 2010). Kümeleme ve faktör analizi yöntemleri değişkenler arasındaki ilişkileri araştıran çok değişkenli istatistik yöntemlerindedir. Faktör analizinde elimizdeki verileri değişkenlerin varyans ve kovaryansına bakarak gruplandırma yapılırken kümeleme analizinde ise değişkenlerin uzaklık ölçütlerine bakarak gruplama yapılır (Hair vd., 2014). Kümeleme ve açıklayıcı faktör analizinin ikisinin kullanıldığı çalışmalar mevcuttur (Şimşek,2006;Doğan ve Başokçu, 2010; Ertürk, 2016).

Çalışmada, ülkelerin korona virüs vaka ve ölüm verileri Kümeleme Analizi kullanılarak SPSS 22 programında analiz edilmiştir. Analizde Hiyerarşik Olmayan Kümeleme Yöntemi ve K-Ortalamlar Tekniği kullanılmıştır. Ünelere ait Kovid-19 verileri Github.com resmi web adresinden temin edilmiş analiz için 01 Haziran 2020 – 25 Şubat 2021 (270 gün) tarihinde gerçekleşen ölüm ve vaka sayıları ile milyonda toplam vaka ve ölüm sayıları ele alınmıştır. Analiz edilen ülkeler 2020 İnsani Gelişim İndeksi raporuna göre; Çok Yüksek İnsani Gelişme, Yüksek İnsani Gelişme, Orta İnsani Gelişme ve Düşük İnsani Gelişme kategorilerine göre ele alınmıştır. Her kategoriden seçilmiş 10 ülke analize dahil edilmiştir. Bu ülkelere ait değişkenler ise; Toplam Vaka, Günlük Vaka, Toplam Ölüm, Günlük Ölüm, Milyonda Toplam Vaka, Milyonda Günlük Vaka, Milyonda Toplam Ölüm ve Milyonda Günlük Ölüm şeklindedir. Değişkenler iki farklı grup olarak ele alınmış ve iki aşamada analiz edilmiştir. İlk aşamada; Toplam Vaka, Günlük Vaka, Toplam Ölüm ve Günlük Ölüm değişkenleri yer alırken, ikinci aşamada ise Milyonda Toplam Vaka, Milyonda Günlük Vaka, Milyonda Toplam Ölüm ve Milyonda Günlük Ölüm değişkenleri yer almaktadır. Analiz sonuçları aşağıda tablolar şeklinde detaylı olarak verilmiştir.

Analiz ve Bulgular

Çalışmada, Kümeleme analizinde kullanılan Hiyerarşik Olmayan Kümeleme Yönteminde K-ortalama tekniği kullanılarak analizler yapılmıştır. Analizler dört kümeleme ve

10 iterasyon kullanılarak elde edilmiştir. Aşağıda, kümeleme merkezleri, küme merkezleri arasındaki uzaklık, ülkelerin yer aldığı kümeler, değişkenlerin anlamlı kümeleme gösterip göstermediğini analiz etmek için ANOVA testi, değişkenlerin kümelerdeki ortalamaları ve kümelerde yer alan sayısal değerler şeklinde tablolar olarak sıralanmıştır.

Tablo 1’de değişkenlerin her iki aşama için ele alınan dört kümede nasıl dağılım gösterdikleri yer almaktadır.

Tablo 1. Kümeleme Merkezleri

	1	2	3	4
1. Aşama				
Toplam Vaka	167677,53	1360938,85	3131973,82	1402128,91
Günlük Vaka	17996,58	14222,19	37654,51	16739,72
Toplam Ölüm	72161,59	41096,35	87511,06	48424,97
Günlük Ölüm	463,53	275,36	621,04	452,34
2. Aşama				
Mil. Top. Vaka	2260,69	17238,39	1964,11	16185,50
Mil. Gün. Vaka	319,12	435,94	299,52	369,60
Mil. Top. Ölüm	190,61	421,95	131,26	323,39
Mil. Gün. Ölüm	2,63	3,79	2,34	3,00

Tablo 1 incelendiğinde, birinci aşama için Kümeleme merkezlerinin tüm değişkenler için yoğunluk merkezleri sırasıyla; Küme 3, Küme 1, Küme 4 ve Küme 2 şeklindedir. İkinci aşama için ise nüfusun belirgin bir etken olmasından kaynaklı yoğunluk; Küme 2, Küme 4, Küme 1 ve Küme 3 şeklindedir.

Aşağıda yer alan Tablo 2’de her iki aşama için oluşturulan kümeler arasında bulunan mesafeler yer almaktadır.

Tablo 2. Küme Merkezleri Arasındaki Mesafe

	Küme	1	2	3
1. Aşama	2	23445735,002	-	-

	3	2965496,331	20480409,634	-
	4	9036352,770	14409427,854	6070995,567
	2	13981,646	-	-
2. Aşama	3	61706,782	47726,909	-
	4	33935,467	19953,866	27780,446

Tablo 2 incelendiğinde, 1. aşamaya ait en uzun mesafenin Küme 1 ile Küme 2 arasında olduğu, en kısa mesafenin ise Küme 1 ile Küme 3 arasında olduğu gözlemlenmektedir. İkinci aşamaya ait en uzun mesafenin Küme 1 ile Küme 3 arasında olduğu, en kısa mesafenin ise Küme 1 ile Küme 2 arasında olduğu gözlemlenmektedir.

Tablo 3'e göre ülkelerin her iki aşama için kümelere göre nasıl bir dağılım gösterdiği ve 01 Haziran 2020 – 25 Şubat 2021 (270 gün) tarihleri arasında hangi kümede kaç gün yer aldığı belirtilmiştir.

Tablo 3. Ülkelerin Kümelenmesi

Küme	1. Aşama
1	Afganistan (1-270), Cezayir (1-270), Azerbaycan (1-270), Bangladeş (1-270), Brezilya(1-39), Çad (1-270), Çin (1-270), Küba(1-270), Cibuti (1-270), Mısır (1-270), Fransa (1-159), Gambiya (1-270), Almanya (1-212), Gana (1-270), Guyana (1-270), Hindistan (1-61), İran (1-270), Irak (1-270), İtalya (1-189), Kenya (1-270), Kırgızistan (1-270), Madagaskar (1-270), Meksika (1-236), Nepal (1-270), Yeni Zelanda (1-270), Nijerya (1-270), Norveç (1-270), Pakistan (1-270), Rusya (1-159), Senegal (1-270), Singapur (1-270), Sri Lanka (1-270), Sudan (1-270), Tacikistan (1-270), Tayland (1-270), Türkiye (1-193), Uganda (1-270), İngiltere (1-189), Yemen (1-270)
2	ABD (197-270)
3	Brezilya (40-181), Fransa (160-270), Almanya (213-270), Hindistan (62-121), İtalya (190-270), Meksika (237-270), Rusya (160-270), Türkiye (194-270), İngiltere (190-270), ABD (1-98)
4	Brezilya (182-270), Hindistan (122-270), ABD (99-196)
Küme	2. Aşama
1	Afganistan (1-270), Cezayir (1-270), Azerbaycan (1-177), Bangladeş (1-270), Brezilya (1-49), Çad (1-270), Çin (1-270), Küba (1-270), Cibuti (1-270), Mısır (1-270), Fransa (1-124), Gambiya (1-270), Almanya (1-169), Gana (1-270), Guyana (1-248), Hindistan (1-270), İran (1-171), Irak (1-132), İtalya (1-151), Kenya (1-270), Kırgızistan (-164), Madagaskar (1-270), Meksika (1-189), Nepal (1-270), Yeni Zelanda (1-270), Nijerya (1-270), Norveç (1-221), Pakistan (1-270), Rusya (1-143), Senegal (1-270), Singapur (1-162), , Sri Lanka (1-270), Sudan (1-270),

	Tacikistan (1-270), Tayland (1-270), Türkiye (1-189), Uganda (1-270), İngiltere (1-136), ABD (1-41), Yemen (1-270)
2	Azerbaycan (178-270), Brezilya (50-165), Fransa (125-159),Almanya (170-251), Guyana (249-270), İran (172-270), Irak (133-270), İtalya (152-185), Kırgızistan (165-270), Meksika (190-270), Norveç (222-270), Rusya (144-255), Singapur (162-270), Türkiye (190-219), İngiltere (137-195), ABD (42-151)
3	Fransa (246-270), İngiltere (231-270), ABD (196-270)
4	Brezilya (166-270), Fransa (160-245), Almanya (252-270), İtalya (186-270), Rusya (256-270), Türkiye (220-270), İngiltere (196-230), ABD (152-195)

Tablo 3, ele alınan ülkelerin belirlenen tarihler arasında hangi kümede kaç gün yer aldıklarını vermektedir. Birinci aşama için tablo incelendiğinde, ülke sayısının en fazla Küme 1’de yer aldığı görülmektedir. Küme 1’i sırasıyla Küme 3, Küme 4 ve Küme 2 izlemektedir. Küme 2 ‘de son 73 günlük bir süre ile sadece ABD yer almaktadır. İnsani gelişmişlik indeksine göre, düşük insani gelişmişlik indeksine sahip ülkelerin genel olarak ele alınan 270 günün tamamında Küme 1’de yer almaktadır. Aynı şekilde salgın süresince gerek izledikleri sağlık politikaları gerek çok yüksek insani gelişmişlik indeksine sahip Singapur ve Yeni Zelanda gibi ülkeleri de bu kümede yer almaktadırlar. Bu durumun temel sebepleri arasında düşük insani gelişmişlik indeksine sahip ülkelerin vaka bildirme konusunda şeffaf davranmamaları ve virüs kapmış bireylerin takip edilmemesi gibi nedenler yatmaktadır. DSÖ yaptığı son açıklamasında salgın merkez üstü olarak Brezilya, ABD ve Avrupa ülkelerinin olduğunu belirtmiştir. Tablo 1’de birinci aşama için küme merkezlerinin genelde Küme 3 olduğu sonucu elde edilmiştir. Bu durum Tablo 3’te desteklenmiş ve DSÖ’nün açıklamasında yer alan ülkeler bu kümede dağılım göstermiştir. Türkiye son 76 günde Küme 3’te yer almış bu durumun sebebi, son dönemde yapılan vaka paylaşım güncellemesi olduğu söylenebilir.

İkinci aşama için tablo incelendiğinde, ülke sayısının en fazla Küme 1’de yer aldığı görülmektedir. Küme 1’i sırasıyla Küme 2, Küme 4 ve Küme 3 izlemektedir. Kümelerin birinci aşamaya göre farklılık göstermesinin temel sebebi nüfustur. Tablo 1’de ikinci aşamaya göre küme merkezlerinin Küme 2’de olduğu sonucu elde edilmiştir. Bu duruma göre Tablo 3’te Küme 1’de ülke sayısının fazla olmasına rağmen merkezin Küme 2 olmasının temel sebebi küme 1’de sayısal olarak vakaların az görülmesidir. Bunun temel sebebi de yukarıda belirtildiği üzere şeffaf vaka sayısı paylaşımının olmaması gösterilebilir. Aynı şekilde DSÖ tarafından salgın merkez üstü olarak belirlenen ülkeler Küme 2’de büyük bir dağılım göstermişlerdir.

Tablo 4’te her iki aşama için K-Ortalamlar yönteminin kullanılması sonucu değişkenlere ilişkin oluşan istatistikî bilgileri yer almaktadır. Aynı zamanda değişkenlerin kümeleme dağılımında anlamlı olup olmadığı tabloda yer almaktadır.

Tablo 4. ANOVA

		Küme		Hata		F	Sig.
		Ortalama Kare	df	Ortalama Kare	df		
1. Aşama	Toplam Vaka	23573171668468580,000	3	406530831599,907	10793	57986,184	,000
	Günlük Vaka	9079100230774,080	3	762733125,035	10793	6631,104	,000
	Toplam Ölüm	8603676680613,208	3	500324305,615	10793	17196,200	,000
	Günlük Ölüm	304460266,081	3	46802,697	10793	6505,186	,000
2. Aşama	Mil.Top.Vaka	373501684513,424	3	12041008,252	10791	31019,137	,000
	Mil.Gün.Vaka	18603629,039	3	5918,405	10791	3143,352	,000
	Mil.Top.Ölüm	128942626,478	3	63310,624	10791	2036,667	,000
	Mil.Gün.Ölüm	5482,177	3	5,989	10791	915,418	,000

Tablo 4, her iki aşama için K-Ortalamlar yönteminin kullanılması sonucu değişkenlere ilişkin oluşan istatistikî bilgileri sunmaktadır. ANOVA tablosu değişkenlerin kümelere ayırmada anlamlı olup olmadıklarının görülmesi amacıyla kullanılmaktadır. Tablo incelendiğinde tüm değişkenlerin kümelere ayırmada anlamlı oldukları görülmektedir (Sig.≤0.05). Bunun yanı sıra birinci aşama için F değerlerine bakıldığında, kümelere ayırmada en etkili değişkenin Toplam Vaka olduğu görülmektedir. En az etkili değişkenler ise sırasıyla Günlük Ölüm ve Toplam Ölüm değişkenleridir. İkinci aşama için F değerlerine bakıldığında, kümelere ayırmada en etkili değişkenin Milyonda Toplam Vaka olduğu görülmektedir. En az etkili değişkenler ise sırasıyla Milyonda Günlük Ölüm ve Milyonda Günlük Vaka değişkenleridir. Bu sonucun ortaya çıkmasının nedeni kümeleme analizi ile kümeler arası farkın maksimum yapılmasıdır. Sonuç olarak kümelerdeki gözlemlerin dağılımı tesadüfî değildir. Tablo 5 her iki aşama için kümelerde yer alan satır sayılarını göstermektedir.

Tablo 5. Kümelerdeki Sayısal Değerler

	Kümeler	Satır Sayısı	Toplam
1. Aşama	Küme 1	9531,000	10791,000
	Küme 2	73,000	
	Küme 3	852,000	
	Küme 4	335,000	
2. Aşama	Küme 1	8951,000	10791,000
	Küme 2	1265,000	
	Küme 3	140,000	
	Küme 4	435,000	

Tablo 5 incelendiğinde, birinci aşama için satır sayısının en çok Küme 1’de yer aldığı gözlemlenmektedir. Satır sayısının en az olduğu küme ise ikinci kümedir. Küme üçte ise toplam 852 ve küme 4’te 335 satır sayısı yer almaktadır. İkinci aşamada ise satır sayısının en çok Küme 1’de yer aldığı gözlemlenmektedir. Satır sayısının en az olduğu küme ise üçüncü kümedir. Küme ikide toplam 1265, küme 4’te ise 435 satır sayısı yer almaktadır. Toplamda 10791 satırın her iki aşama için ayrı ayrı küme dağılımı tabloda detaylı şekilde yer almaktadır.

Sonuç

Covid -19 salgını 2019 Aralık ayında Çin’de başlayıp ve Mart 2020 tarihinden itibaren tüm dünyaya yayılmıştır. Salgının ülkeler arasındaki etkilerinin benzerlik ve farklılıklarının tespiti bu anlamda önem arz etmiş ve ülkelerin salgın verileri ışığında sınıflandırılması önemli hale gelmiştir. Ve bu sınıflamada ülkelerin sosyo-ekonomik gelişmişliği salgın sürecinde içinde bulunulan küme arasında bir uyum olup olmadığı belirlenmeye çalışılmıştır.

Çalışmanın örneklemini İnsani Gelişme Endeksinde dört farklı kategoride yer alan ve her kategoriden on ülke olmak üzere toplam kırk ülke oluşturmaktadır. Çalışmanın ekonometrik bulguları iki aşamada değerlendirilmiş, birinci aşamada tüm değişkenler (toplam vaka, günlük vaka, toplam ölüm, günlük ölüm) için kümelemenin yoğunluk merkezlerinin sırasıyla; Küme 3, Küme 1, Küme 4 ve Küme 2 şeklindedir. İkinci aşama için ise nüfusun belirgin bir etken olmasından kaynaklı yoğunluk; Küme 2, Küme 4, Küme 1 ve Küme 3 şeklindedir.

Ele alınan ülkelerin belirlenen tarihler arasında hangi kümede kaç gün yer aldıklarına bakıldığında, ülke sayısının en fazla Küme 1’de yer aldığı görülmektedir. Küme 1’i sırasıyla Küme 3, Küme 4 ve Küme 2 izlemektedir. Küme 2 ‘de son 73 günlük bir süre ile sadece ABD yer almaktadır. İnsani gelişmişlik indeksine göre, düşük insani gelişmişlik indeksine sahip ülkelerin genel olarak ele alınan 270 günün tamamında Küme 1’de yer almaktadır. Aynı şekilde salgın süresince gerek izledikleri sağlık politikaları gerek çok yüksek insani gelişmişlik indeksine sahip Singapur ve Yeni Zelanda gibi ülkeleri de bu kümede yer almaktadırlar. Bu durumun temel sebepleri arasında düşük insani gelişmişlik indeksine sahip ülkelerin vaka bildirme konusunda şeffaf davranmamaları ve virüs kapmış bireylerin takip edilmemesi gibi nedenler yatmaktadır. Tablo 1’de birinci aşama için küme merkezlerinin genelde Küme 3 (Brezilya, Fransa, Almanya, Hindistan, İtalya, Meksika, Rusya, Türkiye, İngiltere, ABD) olduğu sonucu elde edilmiştir. Bu durum Tablo 3’te desteklenmiş ve DSÖ’nün açıklamasında yer alan ülkeler bu kümede dağılım göstermiştir. Türkiye son 76 günde Küme 3’te yer almış bu durumun sebebi, son dönemde yapılan vaka paylaşım güncellemesi olduğu söylenebilir.

İkinci aşama için, ülke sayısının en fazla Küme 1’de yer aldığı görülmektedir. Küme 1’i sırasıyla Küme 2, Küme 4 ve Küme 3 izlemektedir. Kümelerin birinci aşamaya göre farklılık göstermesinin temel sebebi nüfustur. Tablo 1’de ikinci aşamaya göre küme merkezlerinin Küme 2’de olduğu sonucu elde edilmiştir. Bu duruma göre Tablo 3’te Küme 1’de ülke sayısının fazla olmasına rağmen merkezin Küme 2 (Azerbaycan, Brezilya, Fransa, Almanya, Guyana, İran, Irak, İtalya, Kırgızistan, Meksika, Norveç, Rusya, Singapur, Türkiye, İngiltere, ABD) olmasının temel sebebi küme 1’de sayısal olarak vakaların az görülmesidir. Bunun temel sebebi de yukarıda belirtildiği üzere şeffaf vaka sayısı paylaşımının olmaması gösterilebilir.

K-Ortalamlar yönteminin kullanılması sonucu değişkenlere ilişkin oluşan istatistikî bilgilere bakıldığında, değişkenlerin kümeleme dağılımında anlamlı olduğu görülmüştür. Kümelere ayırmada en etkili değişkenin Toplam Vaka olduğu görülmektedir. En az etkili değişkenler ise sırasıyla Günlük Ölüm ve Toplam Ölüm değişkenleridir. İkinci aşama için F değerlerine bakıldığında, kümelere ayırmada en etkili değişkenin Milyonda Toplam Vaka olduğu görülmektedir. En az etkili değişkenler ise sırasıyla Milyonda Günlük Ölüm ve Milyonda Günlük Vaka değişkenleridir.

Nihai olarak ülkelerin İnsani gelişme endeksinde sıralaması ile Covid 19 salgınında yer aldığı grup arasında bir uyum olmadığı tespit edilmiştir. Örneğin Cibuti, Türkiye, İngiltere, Sudan, Pakistan ve Yeni Zelanda gibi farklı gelişmişlik seviyesine sahip ülkeler uzun bir süre aynı kümede yer almıştır. Tabii bu durumun şeffaf veri paylaşımı, vaka sayısını bilinçli olarak düşük göstermek gibi nedenleri olabilese de bu analizin sonucu insani gelişmişlik endeksinin de gelişmişlik ölçülerini yeterince ortaya koyamaması gibi bir sonucu da gözden kaçırmamaktadır.

Kaynaklar

Alptekin, N. ve Yeşilaydın, G. (2015). OECD Ülkelerinin Sağlık Göstergelerine Göre Bulanık Kümeleme Analizi ile Sınıflandırılması. İşletme Araştırmaları Dergisi, 7(4), 137- 155.

Demir, Y., Esenbuğa, N. ve Ö., C. (2016). İvesi Koyunlarının Et Kalitesini Değerlendirmede Temel Bileşenler Analizinin (PCA) Kullanılması. Süleyman Demirel University Journal of Natural and Applied Sciences Volume 20, Issue 3, 536-541, 2016

Demircioğlu, M. ve Eşiyok, S. (2020). Covid -19 Salgını İle Mücadelede Kümeleme Analizi İle Ülkelerin Sınıflandırılması. İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi Covid-19 Sosyal Bilimler Özel Sayısı, 37, s.369-389.

Doğan, N. ve Başokçu, T.O. (2010). İstatistik Tutum Ölçeği için Uygulanan Faktör Analizi ve Aşamalı Kümeleme Analizi Sonuçlarının Karşılaştırılması. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 1(2): 65-71.

Ersöz, F. (2009). OECD'ye Üye Ülkelerin Seçilmiş Sağlık Göstergelerinin Kümeleme ve Ayırma Analizi ile Karşılaştırılması. Türkiye Klinikleri Journal of Medical Sciences, 29(6), pp.1650-1659.

Hair, J. F., Black, W. C., Babin, B. J. and Anderson, R. E. (2014). Multivariate Data Analysis: Harlow. UK: Pearson Education Limited, pp. 415-474.

Küçükefe, B. (2020). Clustering Macroeconomic Impact Of Covid -19 in OECD Countries And China. Ekonomi Politika Finans Araştırmaları Dergisi, 5, 280-291.

Larco, R. M. C. and Cara, M. C. (2020). Using country-level variables to classify countries according to the number of confirmed Covid-19 cases: An unsupervised machine learning approach. Wellcome Open Research, erişim adresi: <https://doi.org/10.12688/wellcomeopenres.15819.3>

Mut, S. ve Akyürek, Ç. E. (2017). OECD Ülkelerinin Sağlık Göstergelerine Göre Kümeleme Analizi ile Sınıflandırılması. International Journal of Academic Value Studies, 3(12), 411-422

Neuburger, L. and Egger, R. (2020). Travel risk perception and travel behaviour during the Covid-19 pandemic 2020: a case study of the DACH region. Current Issues in Tourism, erişim adresi: <https://doi.org/10.1080/13683500.2020.1803807>

Özdamar Kazım (2004). Paket Programlar İle İstatistiksel Veri Analizi 5. Baskı, Kaan Kitabevi, Eskişehir.

Sığırlı, D., Ediz, B., Cangür, Ş., Ercan, İ. ve Kan, İ. (2006). Türkiye ve Avrupa Birliği'ne Üye Ülkelerin Sağlık Düzeyi Ölçütlerinin Çok Boyutlu Ölçekleme Analizi ile İncelenmesi. İnönü Üniversitesi Tıp Fakültesi Dergisi, 13(2), 81-85.

Şimşek, D. (2006). Kümeleme Analizi, Çok Boyutlu Ölçekleme, Doğrulayıcı ve Açıklayıcı Faktör Analizi İle Elde Edilen Yapı Geçerliği Kanıtlarının Karşılaştırılması. Yüksek lisans Tezi, Hacettepe Üniversitesi, Ankara, 106.

Tekin, B. (2020). Covid-19 Pandemisi Döneminde Ülkelerin Covid-19 Sağlık Ve Finansal Göstergeler Bağlamında Sınıflandırılması: Hiyerarşik Kümeleme Analizi. Finans Ekonomi Ve Sosyal Araştırmalar Dergisi, 5 (2), s.336-349.

Uçar, N. (2010). SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri. Kalaycı, S.(Ed.), Dinamik Akademi Yayın Dağıtım, Ankara, s. 321-331.

