

## **G20 ÜLKELERİNDE GSYH KENTLEŞME, SAĞLIK HARCAMALARI VE CO2 ARASINDAKİ İLİŞKİ**

### **1. Giriş**

G20 veya 20 grubu, dünyanın en büyük ekonomilerini oluşturan 19 ülkenin, Avrupa Birliği ve Afrika Birliği'nin de içinde bulunduğu bir fondur. Türkiye'nin de içinde bulunduğu G20 ülkeleri Almanya, Amerika Birleşik Devletleri, Arjantin, Avustralya, Birleşik Krallık, Brezilya, Çin, Endonezya, Fransa, Güney Afrika, Güney Kore, Hindistan, İtalya, Japonya, Kanada, Meksika, Rusya, Suudi Arabistan ve Avrupa Birliğinden oluşur.

Bu ülkeler küresel ticaretin %80'ini ve dünya nüfusunun 3 te 2'sini oluşturmaktadır.

G20 ülkelerinin bir araya gelme amaçları küresel ekonomik sorunlara ortak çözümler üretmek ve ülkeler arasındaki koordinasyonu sağlamaktır.

G20 ülkeleri, dünya ekonomisinin yaklaşık %85'ini temsil etmekte olup ekonomik güçlerini Gayri Safi Yurtıcı Hasıla (GSYİH) büyüklükleri üzerinden ortaya koymaktadır. 2024 yılı itibarıyla bu ülkelerin GSYİH büyüklüklerine göre sıralaması şu şekildedir: Amerika Birleşik Devletleri (29,2 trilyon \$), Çin (18,7 trilyon \$), Almanya (4,7 trilyon \$), Japonya (4 trilyon \$), Hindistan (3,9 trilyon \$), Birleşik Krallık (3,6 trilyon \$), Fransa (3,2 trilyon \$), İtalya (2,4 trilyon \$), Kanada (2,24 trilyon \$), Brezilya (2,18 trilyon \$), Rusya (2,17 trilyon \$), Meksika (1,9 trilyon \$), Avustralya (1,8 trilyon \$), Güney Kore (1,7 trilyon \$), Endonezya (1,4 trilyon \$), Türkiye (1,3 trilyon \$), Suudi Arabistan (1,2 trilyon \$), Arjantin (0,6 trilyon \$) ve Güney Afrika (0,4 trilyon \$).

Gelişmiş demokrasilere sahip G20 ülkelerinde ortalama kentleşme oranlarına bakıldığından 1960 yılında %49,5, 1970 yılında %55,4, 1980 yılında %60,2, 1990 yılında %65, 2000 yılında %68,5, 2010 yılında %72,5 ve 2018 yılında ise %75 civarında olduğu görülmektedir. Bu uzun periyod incelendiğinde yıllar itibariyle kentleşme oranında devamlı bir artış meydana gelmiştir. Ancak, küresel mali kriz birlikte vergi gelirlerinin GSYİH içindeki payında az da olsa düşüşler yaşanmış olmasına rağmen kriz sonrası süreçte tekrar 2000'li yillardaki trend yakalanmıştır. Dolayısıyla, gelişmekte olan ülkelere kıyasla G20 ülkelerinde toplanan vergi gelirlerinin GSYİH içindeki payı iki katından fazladır. Vergi gelirlerindeki bu başarı, G20 ülkelerinde gelişmekte olan ülkelere kıyasla çok sayıda vergilendirilebilecek şirketin olduğuna işaret etmektedir. (Mahir & Nazlı, 2020)

Sağlık harcamalarının ülkeler bazında artırılması sağlık çıktılarını iyileştirerek beşeri sermayeye olumlu katkı sağlamaktadır. Beşeri sermaye ekonomik kalkınmada önemli bir rol oynadığı için sağlık harcamalarına verilen önemin artırılması gerekmektedir. Yaşam kalitesini doğrudan etkileyen sağlık harcamaları bireylerin yaş, cinsiyet, ekonomik ve sosyal durumlarına bakılmadan eşit şekilde organize edilmelidir. G20 ülkeleri arasında özellikle gelişmekte olan ülkelerde sağlığa ayrılan payın artırılması gerekmektedir. Bireylere uygulanan sağlık hizmetleri sadece onların hastalıklarını iyileştirmek için değil aynı zamanda hayat kalitelerini artırmak için de sunulmalıdır. Küreselleşen dünyada bahse konu olan ülkeler ile uluslararası örgütlerin birlikte hareket etmeleri sağlığa yönelik uygulanan politikaların finansmanı konusunda da önem arz etmektedir. (Özge & Gülden, 2024)

Hızlı ekonomik büyümeye sürecinde, enerji kullanımındaki artışla birlikte karbondioksit emisyon salınınımının artması beklenen bir sonuçtır. Ancak en gelişmiş yirmi ekonomiden biri olan bir ülkede, ekonomik büyümeye ve refah artışı ile birlikte çevre politikalarına verilen önemin artması ve sürdürülebilir kalkınma sürecine girilmiş olması ve dolayısıyla bu süreçte karbondioksit emisyon salınınımının azalmış olması beklenmektedir (Oytun & Mustafa, 2018)

## 2. Literatür Taraması

(Mahir & Nazlı, 2020) çalışmasında kentleşme ve demokrasi faktörlerinin vergi gelirleri üzerindeki etkisi G20 ülkeleri açısından empirik olarak incelenmektedir. Bu amaçla, G20 ülkeleri için 2005-2018 dönemini kapsayan bir araştırma yürütülmüştür. Çalışmanın bağımlı değişkeni vergi gelirleri (%GSYİH)'dır. Bağımsız değişken olarak kentleşme oranı, demokrasi düzeyi, kişi başına GSYİH, ticari açıklık ve literatürde yer alan çalışmalarдан farklı olarak demokrasi ve kentleşmenin moderatör etkisi kullanılmıştır Analiz sonucuna göre demokrasi ve kentleşme etkisi vergi gelirlerini istatistikî olarak anlamlı ve pozitif yönde etkilemektedir. Elde edilen bu sonuç Andersson (2018) çalışmasını destekler nitelikte olup demokrasinin vergi gelirlerini üzerindeki etkisinin kentleşmeye bağlı olduğunu göstermektedir. Dolayısıyla sanayileşmenin getirdiği temel ekonomik değişiklik olarak da adlandırılan kentleşme düzeyindeki artışla birlikte demokratikleşmenin vergi gelirlerini artırmacı özelliği sahip olduğunu göstermektedir. Öte yandan, kişi başına GSYİH ile vergi gelirleri arasında istatistikî olarak anlamlı bir ilişki bulunmamasına rağmen katsayı değeri pozitiftir.

(Mahir & Nazlı, 2020) tarafından yapılan çalışmada 1995-2017 dönemi için G7 ve BRIC-T ülkeleri için vergi gelirlerinin belirleyicileri araştırılmıştır. Analiz sonucunda G7 ülkelerinde kişi başına GSYİH, küreselleşme, istihdam, sanayi üretim endeksi, ücretli çalışan sayısı, doğrudan yabancı yatırımlar ve enflasyon değişkenlerinin vergi gelirlerini artırdığı BRIC-T ülkelerinde ise küreselleşme, doğrudan yabancı yatırımlar ve enflasyonun vergi gelirlerini azalttığı tespit edilmiştir.

(Debolina, 2025) G20 ülkeleri, 2023 itibarıyla küresel kentsel nüfusun yaklaşık %67'sine ev sahipliği yaptıkları için kentleşmenin gelecekteki seyrini belirlemekte kritik bir rol oynamaya hazırlıdır. Dahası, Hindistan ve Çin, Yeni Delhi, Şanghay, Pekin ve Mumbai gibi dünyanın en kalabalık megakentlerine ev sahipliği yapan en kalabalık iki G20 üye ülkesidir. Bu ülkeler küresel GSYİH'nin yaklaşık %85'ine katkıda bulunsa da, bu megakentlerin çoğu trafik sıkışıklığı, konut sıkıntısı, kentsel hizmetlere eşitsiz erişim ve plansız büyümeye gibi zorluklarla karşı karşıyadır. Öte yandan, birçok Avrupa şehri, karbon ayak izlerinde önemli bir azalma olmaksızın nüfus azalması yaşamaktadır.

Özellikle, G20 ülkeleri, ekonomik üretim, zayıf planlama ve politika yapısı ve etkisiz çevre düzenlemeleri nedeniyle küresel karbon emisyonlarının %80'ine katkıda bulunmaktadır. Ayrıca, kentsel yayılma bu ülkelerde yaygın olup karbon emisyonlarına önemli bir katkıda bulunmaktadır. Önceki G20 zirveleri, şehirlerin "sıfıra ulaşma yarışındaki" rolünü kabul etmiş ve karbon ayak izini azaltmada "döngüsel ekonominin" kritik rolünü vurgulamıştır. BM-Habitat da sürdürülebilir yaşam alanları oluşturmak için kentsel-kırsal bağların güçlendirilmesi gerektiğini belirtmiştir.

G20 ülkeleri ekonomik büyümeye büyük önem veriyor ve bu da fosil yakıtların tamamen aşamalı olarak kullanımdan kaldırılması konusunda güçlü bir taahhüt eksikliğini açıktır. Hindistan tarafından önerilen LiFE Misyonu, sürdürülebilir ve bilinçli tüketimi teşvik ederek döngüsel ekonominin

benimsenmesini ve karbon ayak izinin azaltılmasını teşvik ediyor. Hindistan Başkanlığı tarafından ortaya konulan bu misyonun ilkelerinin benimsenmesi, daha yüksek emisyon seviyelerine sahip ülkelerin iklim değişikliğiyle mücadele stratejilerine ve önlemlerine yardımcı olabilir. Bilinçli tüketim sayesinde, gelecekteki kentleşme daha sürdürülebilir olabilir ve ülkelerin küresel sürdürülebilirlik hedeflerine ulaşmalarına yardımcı olabilir.

Ekonomide yaşanan gelişmeler, nüfus ve sanayide yaşanan artışlar çevresel sorunların yaşanmasına yol açmaktadır. Ekonomik gelişim, sağlık ve çevresel alanlarda birçok problemi de beraberinde getirmektedir. (Ceren, Aysegül, & Nergis, 2020) Tarafından yapılan bu çalışmada CO<sub>2</sub> emisyonu ve enerji tüketiminin sağlık, nüfus ve büyümeye olan ilişkisi incelenmiştir. G20 ülkelerinin ele alındığı çalışmada panel veri analizi yönteminden yararlanılmıştır. Ele alınan değişkenler için iki ayrı model kurularak analiz yürütülmüştür. CADF birim kök testi ile serilerin birinci dereceden bütünlük olduğu belirlenmiştir. Değişkenler arasındaki uzun dönemli ilişkinin incelenmesi için ikinci nesil testlerden Gengenbach, Urbain ve Westerlund Panel Eşbütnleşme Testi (2016) uygulanmış ve değişkenler arasında uzun dönemli bir ilişki tespit edilmiştir. Çalışmada son olarak ise Dumitrescu -Hurlin panel nedensellik testi yapılmıştır. Yapılan nedensellik testi için birinci modelde, CO<sub>2</sub> emisyondan, sağlık harcamalarına, nüfus artısından CO<sub>2</sub> emisyonuna ve CO<sub>2</sub> emisyondan kişi başına GSYH'ya doğru tek yönlü nedensellik saptanmıştır. İkinci modelde ise nüfustan enerji tüketimine ve kişi başına GSYH'dan enerji tüketimine doğru tek yönlü nedenselliğin olduğu bulunmuştur. Ekonomik büyümeye gerçekleştirilirken çevresel etkenlerin göz önünde bulundurulması gerekmektedir. Artan nüfus oranları ve sanayileşme gibi faktörler sürdürülebilir kalkınmanın gerçekleştirilmesini zorlaştırmaktadır. Enerji ihtiyaçlarının yenilenebilir enerji kaynaklarından sağlanması, sanayi alanında yeni ve çevreye duyarlı stratejilerin kullanılması, çevresel atıkların azaltılması ve enerji verimliliğinin sağlanması CO<sub>2</sub> emisyonunun önemli ölçüde azalmasını sağlayacaktır. Çevresel sorunlar sağlık harcamalarının da artmasına yol açmakta ve iş gücünün verimliliğinin azalmasına neden olmaktadır. Çevre üzerinde oluşturacağımız iyileşmeler ülke kalkınmasına birçok sektör üzerinde katkı sunacaktır. Üretimde verimliliğin artması, sürdürülebilir kalkınmanın gerçekleşmesi ülkelerin gelecekte hem sosyal alanda hem de ekonomik alanda gelişmesine yol açacaktır.

(Oytun & Mustafa, 2018) Tarafından yapılan bu çalışmanın temel motivasyonu, dünya ekonomisinin en büyük ekonomik payda sahip ülkeleri olarak kabul edilen G20 ülkelerinde karbondioksit emisyonları ve GSYH arasındaki Granger nedenselliğin araştırılmasıdır. Bu doğrultuda, G20 ülkeleri için 1971-2012 döneminde nedensellik ilişkisi Kónya (2006) panel nedensellik testi ile incelenmiştir. Bulgular, Birleşik Krallık, Japonya ve Hindistan'da karbondioksit emisyonundan GSYH'ye doğru bir nedensellik olduğunu, diğer G20 ülkelerinde ise karbondioksit emisyonundan GSYH'ye doğru bir nedensellik olmadığını, yine GSYH'den karbondioksit emisyonuna nedensellik incelediğinde ise Birleşik Krallık, Suudi Arabistan, Kanada ve Almanya'da GSYH'den karbondioksit emisyonuna doğru bir nedenselliğin olduğunu göstermektedir. Ayrıca daha önce de vurgulandığı gibi, Birleşik Krallık'ta değişkenler arasında çift yönlü bir ilişkinin söz konusu olduğu görülmektedir. Analizdeki değişkenler kişi başına GSYH ve karbondioksit emisyonu olarak alındığında ise kısmen benzer şekilde Birleşik Krallık, İtalya, Endonezya ve Avustralya'da karbondioksit emisyon salınımından kişi başına GSYH'ye tek yönlü bir nedensellik söz konusuyken, ABD, Kanada ve Almanya'da kişi başına GSYH'den karbondioksit emisyon salınımına doğru tek yönlü bir nedensellik gözlendiği, Japonya'da ise değişkenler arasında çift yönlü bir nedensellik tespit edildiği ifade edilebilir.

Çevresel problemler her geçen gün katlanarak artmaktadır ve bu durum doğrudan ve dolaylı olarak ülke ekonomilerini de olumsuz yönde etkilemektedir. Bu nedenle çevresel problemlerin ekonomik etkileri ve maliyeti iktisat alanında önemli bir araştırma konusu haline gelmiştir. Çevresel kirlilik göstergesi olarak literatürde genellikle karbondioksit (CO<sub>2</sub>) emisyonu kullanılmaktadır. Yenilenebilir enerji yatırımları ve çevreye yönelik inovatif yaklaşımın uzun dönemde çevresel problemlerin üstesinden gelebileceği düşünülmektedir. (Selim, 2021) Tarafından yapılan bu çalışmada çevresel inovasyonlar (ETI), yenilenebilir enerjinin (REC) ve büyümenin (GDP) CO<sub>2</sub> emisyonu üzerindeki etkisi 1993-2018 yılları arasında G-20 ülke grubunda IMF'nin sınıflandırmamasına göre gelişen statüsündeki 8 ülke için incelenmiştir. Analizde yatay kesit bağımlılığı ve heterojenliği dikkate alan Durbin-H eşbüTÜNleşme ve FMOLS testleri kullanılmıştır. Analiz sonuçlarına göre değişkenler arasında uzun dönemli bir ilişki vardır. Ele alınan değişkenlerin CO<sub>2</sub> emisyonu üzerindeki etkileri ülkeden ülkeye farklılık göstermektedir fakat panelin geneli için REC ve GDP'deki artışın CO<sub>2</sub> emisyonu üzerindeki etkisi negatifken ETI'daki artışın etkisi beklenenin aksine pozitiftir.

(Zeynep, Havva, & Gözde, 2021) tarafından yapılan çalışmada 2000-2017 yılları arasında G-20 ülkelerinde sağlık harcamaları ve ekonomik büyümeye arasındaki ilişkinin varlığı araştırılmıştır. Bu bağlamda sağlık harcamaları, özel sağlık harcamaları, kişisel sağlık harcamaları, GSYH ve doğumdan beklenen yaşam süresi değişkenleri panel ARDL testi ile analiz edilmiştir. Analiz sonucunda sağlık harcamaları, özel sağlık harcamaları ve yapılan kişisel sağlık harcamaları değişkenlerinin hem uzun hem de kısa dönem parametreleri istatistiksel olarak anlamlı bulunmuştur. Uzun dönemde sağlık harcamalarındaki %1'lük artış ekonomik büyümeye % 0,8 oranında artışa sebep olurken, özel sağlık harcamalarındaki %1'lük artış ekonomik büyümeye %0,6 oranında azalşa sebep olacaktır. Yine uzun dönemde yapılan kişisel sağlık harcamaları değişkenindeki %1'lük artış ekonomik büyümeye %0,7 oranında artışa sebep olacaktır.

Bu konuda yapılan en erken çalışmalarдан biri Bhargava vd. (2000)'e aittir. Bhargava vd. (2000) çalışmalarında 1965-1990 dönemi için gelişmiş ve gelişmekte olan ülkeleri ele alarak sağlık göstergeleri ve ekonomik büyümeye ilişkisini incelemiştir. Panel veri analizinin kullanıldığı bu çalışmada sağlık göstergelerinin ekonomik büyümeye üzerinde pozitif ancak güçlü bir etkiye sahip olmadığını ortaya koymuşlardır. (Pınar & Çebi, 2015)

(Sibel, Doğan, & Pınar, 2014) 2001-2011 döneminde 27 AB üyesi ülke ve Türkiye'de kişi başı sağlık harcaması ve ekonomik büyümeye arasındaki kısa ve uzun dönemli ilişkileri panel eşbüTÜNleşme ve hata düzeltme modelleri kullanarak incelemiştir. Çalışmadan elde edilen bulgulara göre kişi başı sağlık harcaması ve ekonomik büyümeye arasında kısa ve uzun dönemde pozitif yönlü bir ilişki olduğu belirlenmiştir.

Kamu sağlık harcamalarının sağlık çıktıları üzerindeki etkisinin G20 ülkeleri açısından değerlendirildiği (Özge & Gülden, 2024) tarafından yapılan bu çalışmada 2000-2020 yılları için Dünya Bankası'ndan alınan veriler farklı ekonometrik modeller ile analiz edilmiştir. İlk olarak yatay kesit bağımlılık varsayımlarının sınanması için yatay kesit bağımlılık testi uygulanmıştır. Uygulanan yatay kesit bağımlılık test sonuçlarına göre bebek ölüm hızı, hastane yatak sayısı, karbondioksit emisyonu, doğurganlık hızı, anne ölüm oranı, cepten ödemeler, özel sağlık harcamaları ve kırsal nüfus oranı değişkenlerinde yatay kesit bağımlılığının söz konusu olduğu sonucuna varılmıştır. Daha sonra Hadri-Kuruzami Panel Birim Kök Testi ile değişkenler için durağanlık analizi yapılmıştır. Yapılan analiz

sonucunda serinin durağan olmadığı tespit edilmiş ve serinin birinci dereceden farkı alınarak seri %5 anlam düzeyinde durağan hale getirilmiştir. Üçüncü olarak yatay kesit bağımlılığın var olduğu durumlarda kullanılan ve eşbüütünleşme ilişkisinin olup olmadığını test eden ikinci kuşak testlerden Gengenbach, Urbain ve Westerlund (2016) Eşbüütünleşme Testi kullanılmıştır. Son olarak modelde kullanılan serilerin heterojen dağılıma sahip olduğu durumlarda kullanılan Emirmahmutoğlu & Köse (2011) Nedensellik Testi kullanılmıştır. Analiz sonuçları ülkelerin sağlık harcamalarına ayırdıkları paylarda farklılık olsa dahi sağlık harcamalarında meydana gelen artışların sağlık çıktıları üzerinde aynı yönlü bir etkiye neden olduğunu göstermektedir. Nedensellik test sonuçlarına göre kamu sağlık harcamalarında meydana gelen yüzdelik bir değişimin tüm bağımlı değişkenleri etkilediği görülmektedir. Kamu sağlık harcamalarının anne ölüm oranı, bebek ölüm hızı, 5 yaş altı çocuk ölüm hızı ve doğuştan beklenen yaşam süresi üzerindeki etkisi ülkeden ülkeye farklılık göstermektedir. Kamu sağlık harcamalarının sağlık çıktılarını etkilemesinin yanı sıra bazı değişkenlerin de kamu sağlık harcamaları ile olan ilişkisi araştırılmıştır. Analizler sonucunda bebek ölüm hızı, hastane yatak sayısı, karbondioksit emisyonu, doğurganlık hızı, cepten ödemeler, özel sağlık harcamaları ve kent-kırsal nüfus oranlarında da yatay kesit bağımlılığın mevcut olduğunu sonucuna varılmış ve bu değişkenlerde meydana gelen yüzdelik bir değişimin kamu sağlık harcamalarını etkilediği sonucuna ulaşılmıştır.

(Emine, Bahar, & Aslı, 2025) Bu çalışmada, “kırsal mahalle” tanımlaması ile birlikte özellikle büyükşehirlerde süregelen yerleşim birimleri sınıflandırmalarında yaşanan kır-kent ayrımı sorunsalına yönelik olarak yerleşim birimleri sınıflandırmasında makine öğrenmesi algoritmalarından birisi olan Kmeans algoritması yöntemi kullanılması önerilmiştir. Bu amaçla, Ankara İl sınırları içerisinde yer alan 1426 yerleşim birimlerine yönelik kent-kır ayrımının yapılabilmesi için yerleşim birimi kümelenmesi üzerinde çalışılmıştır. Literatürde sıklıkla kullanımına rastlanılan kümeleme yöntemi olarak gözetimsiz makine öğrenme algoritmalarından K-means algoritması çalışmaya konu edilmiştir. OECD ve ulusal mevzuat çerçevesinde kırsal alan ve kentsel alan tanımlamaları üzerinden esas alınan nüfus, kent merkezine ve ilçe merkezine mesafe verileri kullanılarak oluşturulan veri seti üzerinden algoritma çalıştırılmıştır.

(Safa, Hatem, Durmuş, & Hasan, 2022) Bu çalışmada K-means algoritması kümeleme problemlerinde yaygın bir şekilde kullanılmasına rağmen, rastgele oluşturulmuş ilk merkez değerlerine bağlı lokal optimuma takılma problemi ve özellikle büyük boyutlu dosyalarının işlenmesi sürecinde yavaş yakınsama gibi dezavantajları vardır. Kümeleme işlemlerindeki bu tür problemlerin üstesinden gelmek için literatürdeki çalışmalarla meta sezgisel algoritmaların kullanılması önerilmiştir. Bu çalışmada, K-means algoritması ile Diferansiyel Gelişim Algoritması (DEA) algoritmaları hibrit bir yapıda kullanılarak K-means algoritmasının kümeleme bütünlüğü performansının artırılması amaçlanmıştır. Önerdiğimiz hibrit yapı ile Kalp Hastalığı veri seti kullanıldı. Hastalık tespit oranı artırdı ve yüksek başarı oranı ile kalp hastalığı teşhis edildi. Kümeleme performansını iyileştirmek için kullanılan veri setinde hastalığı belirleyen girdi parametre değerlerine öncelik vermek için optimizasyon algoritmaları tarafından belirlenen ağırlık değerlerini kullandık. Algoritmalar accuracy (doğruluk), recall (duyarlılık), specificity (belirlilik), precision (kesinlik), F1 skoru ve standart sapma gibi performans ölçütlerine dayalı olarak istatistiksel olarak değerlendirildi ve K-means algoritmasının başarısı yaklaşık %50 artırılmıştır.

Şişeci ve arkadaşları ise k-means algoritması ile resimleri alt bloklara ayırarak segmentasyon yapmışlar ve diğer yöntemlere göre daha başarılı sonuçlar almışlardır[5]. Bu alanda kullanılmak üzere çok çeşitli kümeleme algoritması geliştirilmiştir. En temel gözetimsizkümeleme algoritmaları k-means

ve fuzzy c-means algoritmasıdır[6] .Bu çalışmada küme sayısı belli değil iken verilerin k-means ile nasıl sınıflandırılacağı üzerinde durulmuş ve sınıf sayısı belirli olmayan verilerde optimum sınıf sayısının belirlenerek k-means ile sınıflandırılması amaçlanmıştır. (Samet, Seçkin, Serdar, & Erman, 2018)

Fırat ve diğerleri [10] Türkiye genelindeki toplam yağış miktarını kümelemek ve homojen bölgeleri belirlemek için K-means algoritmasını kullanmışlardır. Kümeleme analizleri için Türkiye Ulusal Meteoroloji İşleri (DMI) tarafından kullanılan 188 adet istasyondan elde edilen yıllık yağış kayıtları, boylam, enlem ve yükseklik değerlerini kullanmışlardır. Sonuç olarak, yıllık toplam yağışların kümelemesinde K-means algoritması kullanılarak elde edilen sonuçların kabul edilebilir düzeyde olduğunu söylemişlerdir. (Kemal & Ramazan, 2020)

Bu çalışmada alüminyum ve paslanmaz çelikten 200 çeşit profil, aksesuar ve komponent üretimi gerçekleştiren bir imalat işletmesindeki üretilmekte olan ürünlerde k-means algoritması ve genetik algoritma ile parça kümeleri oluşturulmuştur. Daha sonra oluşturulan hücrelere göre işletmenin yerlesimi düzenlenmiş ve toplam taşıma maliyetleri hesaplanmıştır. Son olarak ise mevcut sistem ile geliştirilen hücresel sistemler karşılaştırılmıştır. (Serap, Seda, & Reşit, 2016)

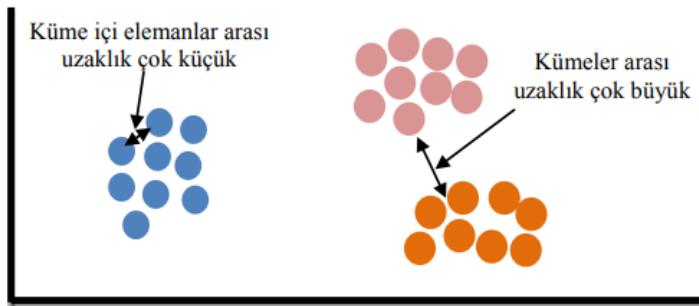
### **3. Metodoloji**

#### **3.1. K-Means Kümeleme Algoritması**

En eski kümeleme algoritmalarından biri olan K-means, 1967 yılında J. B. MacQueen tarafından geliştirilmiştir. En yaygın kullanılan gözetimsiz öğrenme yöntemlerinden biri olan K-means algoritmasının atama mekanizması, her verinin yalnızca bir kümeye ait olmasına izin verir ve bu nedenle keskin bir kümeleme algoritmasıdır. Merkez noktanın kümeyi temsil etmesi ana fikrine dayalı bir metottur. (Meltem & Yılmaz, 2007)

Kümeleme, en basit tanımlıyla benzer özellik gösteren verilerin kendi aralarında gruplara ayrılmalıdır. Kümeleme analizinde genel amaç küme içi homojenliği, kümeler arası heterojenliği sağlamaktır. Bu da benzer bireylerin aynı kümeye toplanmasıyla sağlanabilir. Kümeleme problemi de bir optimizasyon problemi olup, küme elemanlarının küme ortalamasından uzaklıklarının toplamının minimizasyonu ile optimum kümeleme gerçekleştirilmiş olur. Bireylerin benzerlikleri uzaydaki konumları ile ilgilidir. Uzaydaki konumları itibarı ile birbirine uzaklıkları daha az olan bireyler aynı kümeye toplanmış olacaklardır. Literatürde pek çok kümeleme algoritması bulunmaktadır. Kullanılacak olan kümeleme algoritmasının seçimi, amaca ve veri tipine bağlıdır. Genel olarak başlıca kümeleme yöntemleri; Bölme yöntemleri (Partitioning methods), Hiyerarşik yöntemler (Hierarchical methods), Yoğunluk tabanlı yöntemler (Density-based methods), Izgara tabanlı yöntemler (Grid-based methods), Model tabanlı yöntemler (Model-based methods) olarak sınıflandırılabilir

Birimler arasındaki uzaklıkları hesaplamak için en sık kullanılan uzaklık ölçütleri Minkowski, Öklid ( Euclide ), Pearson, Manhattan ( City-Blok ), Mahalanobis, Hotelling T2 ve Canberra Uzaklılığı'dır . Bu çalışmada uzaklık hesabı için Öklid uzaklıği kullanılmıştır.



Şekil 1. Küme Yapısı (Cluster Structure)[9]

Kümeleme işleminde küme içindeki nesneler arası uzaklık çok küçükken, kümeler arası uzaklık çok büyütür. En yaygın kullanılan gözetimsiz öğrenme yöntemlerinden biri olan K-means algoritması her verinin sadece bir kümeye ait olabilmesine izin veren keskin bir kümeleme algoritmasıdır.

K-means algoritmasının genel mantığı  $n$  adet veri nesnesinden oluşan bir veri setini,  $k$  adet giriş parametresi sayısı kadar kümeye bölmektedir. Amaç, gerçekleştirilen bölüme işlemi sonunda elde edilen kümelerin, küme içi benzerliklerinin maksimum ve kümeler arası benzerliklerinin minimum olmasını sağlamaktır. Bölümleyici kümelemeli yöntemlerden olan K-Means algoritması sürekli olarak kümelerin yenilendiği ve en uygun çözüme ulaşana kadar devam eden döngüsel bir algoritmadır.

K-means yönteminin performansını  $k$  küme sayısı, başlangıç olarak seçilen küme merkezlerinin değerleri ve benzerlik ölçümü kriterleri etkilemektedir. Küme sayısının belirlenmesi konusunda son yıllarda yoğun çalışmalar yapılmaktadır. Küme sayısının belirlenmesinde kullanılan en pratik yol (1) nolu eşitlik ile ifade edilir .Ancak veri sayısının çok büyük olması durumunda pratik değildir.

$$k = \sqrt{\frac{n}{2}} \quad (1)$$

$n$ : kümelenen birey sayısı

K-means kümeleme yönteminin değerlendirilmesinde en yaygın olarak karesel hata kriteri SSE kullanılır. En iyi sonucu en düşük SSE değerine sahip kümeleme verir. Nesnelerin bulundukları kümenin merkez noktalarına olan uzaklıklarının karelerinin toplamı (1) nolu eşitlik ile hesaplanmaktadır

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} dist^2(m_i, x) \quad (2)$$

$x$ :  $C_i$  kümelerinde bulunan bir nesne,  $m_i$ :  $C_i$  kümelerinin merkez noktası

Algoritma, karesel-hata fonksiyonunu minimize edecek şekilde k kümeyi belirlemeye gayret eder. K-means algoritması, algoritmaya kullanıcı tarafından verilen  $k$  parametresi ile  $n$  tane veriden oluşan veri setini  $k$  adet kümeye böler

K-means algoritmasının işlem basamakları şöyledir:

1.Adım: İlk olarak küme merkezleri belirlenir. Bunun için iki farklı yol vardır. Birincisinde nesneler arasında küme sayısı olan  $k$  adet rasgele nokta seçilir veya merkez noktalar tüm nesnelerin ortalaması alınarak belirlenir.

2.Adım: Her nesnenin seçilen merkez noktalara olan uzaklığı hesaplanarak tüm nesneler  $k$  adet kümeden kendilerine en yakın olan kümeye yerleştirilir.

3.Adım: Oluşan kümelerin yeni merkez noktaları o kümedeki tüm nesnelerin ortalama değeri ile değiştirilir.

4.Adım: Merkez noktalar değişmeyene kadar 2. ve 3. adımlar tekrarlanır. (Pakize, Buket, & Zehra, 2016)

k-Means, gözetimsiz öğrenme yöntemlerinden biri olup, verileri önceden tanımlanmış  $k$  sayıda kümeye ayırmayı amaçlar. Her küme, verilerin belirli özelliklerine göre gruplanması sağlar. Bu algoritma, özellikle büyük veri setlerinde verilerin doğal gruplarını belirlemek için etkilidir. k-Means algoritması, verileri  $k$  adet kümeye ayırırken, her bir veri noktasını en yakın küme merkezine atar. Bu işlem, kümeler içindeki verilerin toplam karesel hata değerini (Within-Cluster Sum of Square-WCSS) minimize etmeyi hedefler ve WCSS, her bir küme için Denklem 2.1'de görüldüğü gibi hesaplanmaktadır.

$$WCSS = \sum_{j=1}^K \sum_{i=1}^{n_j} \|x_i^{(j)} - c_j\|^2$$

Burada;  $K$ : Toplam küme sayısı,  $n_j$ :  $j$ 'inci kümedeki veri noktalarının sayısını,  $x_i^{(j)}$ :  $j$ 'inci kümedeki  $i$ 'inci veri noktasını ve  $c_j$  :  $j$ 'inci kümenin merkezini temsil etmektedir. Algoritma, başlangıçta rastgele seçilen küme merkezleriyle başlar ve tekrarlı olarak ilk önce her bir veri noktası, en yakın küme merkezine atanır. Daha sonra her küme için merkez noktası, o kümedeki veri noktalarının ortalaması olacak şekilde güncellenir. Bu süreç, küme merkezlerinin konumları sabitlenene veya belirli bir tekrar sayısına ulaşana kadar devam eder. (Tuğba & Ferhat, 2025)

Silhouette puanı ve silhouette grafiği, kümeler arasındaki ayırma mesafesini ölçmek için kullanılır. Bir kümedeki her bir noktanın komşu kümelerdeki noktalara ne kadar yakın olduğunu bir ölçüsünü görüntüler. Bu ölçüm  $[-1, 1]$  aralığına sahiptir ve kümeler içindeki benzerlikleri ve kümeler arasındaki farklılıklar görsel olarak da sunmaktadır. Silhouette katsayıları  $+1$ 'e ne kadar yakınsa, kümenin örnekleri komşu küme örneklerinden o kadar uzaktadır. 0 değeri, örneklerin iki komşu küme arasındaki karar sınırında veya çok yakınında olduğunu gösterir. Negatif değerler ise bu örneklerin yanlış kümeye atanmış olabileceğini gösterir. Silhouette katsayılarının ortalamasının alınması tüm kümenin performansının tek bir silhouette puanı ile hesaplanabilmesini sağlar. (Topaloğlu, 2024)

Silhouette İndeksi: Kümelenen verilerin bulunduğu kümedeki uygunluğunu bulmak için geliştirilen Silhouette indeksi aşağıdaki eşitlik ile hesaplanır:

$$ch = \frac{\text{tr}(B_k)x(n_E-k)}{\text{tr}(W_k)x(k-1)}$$

$\text{tr}(B_k)$ : kümeler içi kareler toplamı

$\text{tr}(W_k)$ : kümeler arası kareler toplamı

En yüksek CH değeri en iyi kümeyi ifade eder. (Murat & Mesut, 2021)

#### 4. Uygulama

Çalışma Gayri Safi Yurt İçi Hasıla (GSYİH) ile kentleşme, karbon emisyon tonu ve sağlık harcamalarını analiz etmeyi hedeflemiştir. Bu doğrultuda, 2000-2023 yılları arasındaki veriler kullanılarak, makine öğrenmesi algoritmalarından biri olan K-Means algoritması ile kapsamlı bir değerlendirme yapılmıştır.

##### 4.1. Veri Seti

Veri setinde yer alan değişkenler aşağıda sıralanmıştır:

1-Kişi Başına Düşen GSYİH: Kişi başına düşen GSYİH, gayri safi yurtiçi hasılanın yıl ortası nüfusuna bölünmesiyle elde edilmektedir. Kullanılan veriler **cari ABD doları (current US\$)** cinsindendir.

2-Kentsel Nüfus Oranı (% Toplam Nüfus): Kentsel nüfus oranı, bir ülkede kentsel alanlarda yaşayan nüfusun toplam nüfusa oranını ifade etmektedir. Değerler yüzde (%) olarak hesaplanmıştır.

3-Cari sağlık harcamaları (% GSYİH): Sağlık hizmetlerinin sunulması amacıyla yapılan ve tüketim niteliği taşıyan cari harcamaların, bir ülkenin gayri safi yurtiçi hasılası (GSYİH) içindeki payını ifade etmektedir. Değerler yüzde (%) olarak hesaplanmıştır.

4-Toplam CO<sub>2</sub> emisyonları (LULUCF hariç, Mt CO<sub>2</sub>e): Arazi kullanımı, arazi kullanım değişikliği ve ormancılık faaliyetlerinden kaynaklanan emisyonlar hariç tutulmak üzere, enerji üretimi, sanayi, ulaşım ve benzeri ekonomik faaliyetler sonucunda ortaya çıkan toplam karbondioksit emisyonlarını ifade etmektedir. Değerler milyon ton karbondioksit eşdeğeri (Mt CO<sub>2</sub>e) cinsindendir.

**Tablo 1:** Özet İstatistikler

Vars	Mean	Median	Min	Max	1.Çeyrek	3.Çeyrek
GDP	23920.6	21547.0	442.8	82304.6	7973.2	39169.5
Urban	73.73	78.52	27.67	92.46	69.63	82.08
Healthy	7.852	8.080	1.850	18.810	5.160	9.943
CO2	1442.1	519.8	124.1	13259.6	397.1	1274.1

Ortalama GSYH değeri 23920.6 ABD Doları iken, medyan değerinin (21547.0) ortalamadan düşük olması, veride yüksek gelirli bazı üç değerlerin ortalamayı yükselttiğini göstermektedir. Verilerin minimum ve maksimum değerleri arasındaki yüksek fark ve 1. çeyrek değeri ile 3. çeyrek değeri

arasındaki farkın fazla olması ülkeler arasında ciddi bir gelir farklılığı olduğunu kanıtlar. Bazı ülkeler çok yüksek gelir değerlerine sahipken bazı ülkelerin gelir değerleri oldukça düşüktür.

Ülkelerin kentleşme oranlarının ortalama değeri %73.73 iken, medyan değeri %78.52'dir. Medyan değerinin ortalama değerden daha yüksek olması çoğu ülkede kentleşme oranının yüksek olduğunu ancak bu oranın çok düşük olduğu bazı ülkelerin **genel** ortalama değeri aşağı çektiğini göstermektedir.

Ülkelerin yaptığı sağlık harcamalarının GDP içindeki genel ortalaması %7.85 iken, medyan değeri %8.08 olarak bulunmuştur. Ortalama ve medyan değerleri birbirine oldukça yakındır. Her ne kadar minimum ve maksimum arasındaki fark oldukça yüksek olsa da, verinin 1.çeyreği ve 3.çeyreği arasındaki farkın az olması verinin belirli bir aralıkta yoğunlaştığını göstermektedir. Bu durum veri setinde sağlık harcaması çok düşük ve çok yüksek üç değerler bulunduğuunu ancak ülkelerin yarısının %5 ile %10 arasında dağıldığını gösterir.

Ülkelerin karbon emisyonunun ortalama değerinin (**1.442,1 Mt CO<sub>2</sub>e**), medyandan (**519,8 Mt CO<sub>2</sub>e**) çok daha yüksek olması; bazı ülkelerin aşırı derecede fazla karbon salınımı yaptığını ve bu üç değerlerin ortalamayı yükselttiğini göstermektedir. Minimum ve maksimum değerler arasındaki farkın yüksek olması, veri setinde çok düşük ve çok yüksek karbon emisyonuna sahip ülkelerin bir arada bulunduğu kanıtlamaktadır. 1. ve 3. çeyrek arasındaki farkın fazla olması da verilerin geniş bir alana yayıldığını desteklemektedir.

#### 4.2. Veri Hazırlama ve Veri Ön İşleme Adımları

Analiz sonuçlarının daha doğru olması için veri seti üzerinde bazı hazırlık işlemleri yapılmıştır. Kullanılan K-Means algoritması, başlangıç noktalarına ve aykırı değerlere karşı oldukça hassas olduğu için veri setinde bulunan aşırı üç değerler sonuçları kolayca bozabilmektedir. Bu nedenle genel dağılımdan çok uzak olan ve algoritmanın performansını düşürebilecek aykırı noktalar analizden önce tespit edilip temizlenmiştir. Böylece K-Means algoritmasının daha dengeli çalışması sağlanmış ve verilerin birbirine en benzer olacak şekilde daha doğru gruplandırılması mümkün olmuştur.

#### 4.3. Standardizasyon İşlemi

Veri setinde bulunan değişkenlerin birimlerinin farklı olması ve değer aralıklarının çok geniş olması, yapılan analizde istenmeyen sonuçlar alınmasına; büyük sayıların analize daha fazla etki etmesine ve küçük değerli değişkenlerin etkisiz kalmasına sebep olabilmektedir. Bu dengesizliği gidermek için **Standardizasyon (Z-Score)** işlemi uygulanmıştır. Bu işlemle tüm değişkenlerin ortalaması 0 ve standart sapması 1 olacak şekilde dönüşüm yapılmıştır. Böylece her bir değişkenin analiz üzerinde eşit ağırlığa sahip olması ve sonuçların daha tarafsız çıkması sağlanmıştır

$$z = \frac{x - \mu}{\sigma}$$

$z$ : Standart puan (Z-skoru)

$x$ : Gözlem değeri (Verinin ham hali)

$\mu$ : Veri setinin ortalaması

$\sigma$ : Veri setinin standart sapması

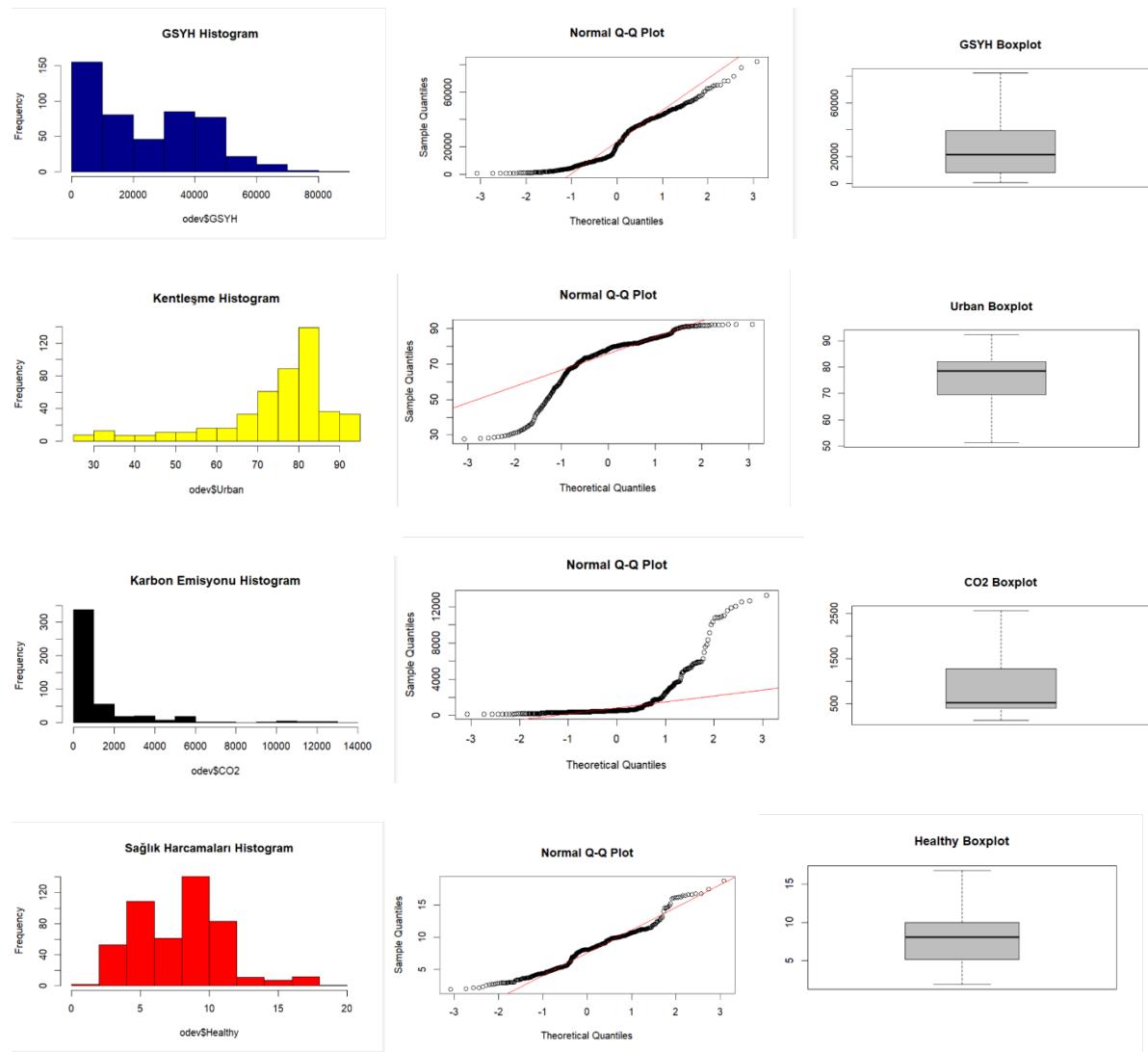
#### **4.4. Ön İşleme ve Yorumlama**

Veri analizinin en önemli adımlarından biri olan ön işleme adımı , veri setindeki yer alan değişkenlerin temizlenmesini sağlar ve veri analizini düzgün bir şekilde yapılabilmesi için uygun hale getirir. Öncelikle , veri setindeki eksik değerlerin olup olmadığı kontrol edilir . Eksik verilerin kontrol edilmesi verilerin güvenilir sonuçlar vermesi açısından önemlidir . Verilerin normal dağılıma uygun olup olmadığını değerlendirmek amacıyla histogram grafikleri incelenir. Histogramlar veri setinin genel dağılım yapısını görsel olarak değerlendirilmesini sağlar.Bu sayede verilerin simetri durumunu , çarpıklık düzeyini ve olası uç değerlerin varlığını daha kolay bir şekilde analizinin yapılabilmesini sağlar.

Verilerin normal dağılıma uygun olup olmadığını daha detaylı bir şekilde analiz edebilmek amacıyla , QQ Plot (Quantile-Quantile Plot) yöntemi kullanır. QQ Plot yöntemiyle ortaya çıkan görsele bakarak verilerin normal dağılıma ne kadar uyumlu olduğunu değerlendirilir. Bununla birlikte verilerdeki aykırı değerleri görmek ve verilerin nerede yoğunlaştığını da incelemek amacıyla Box Plot (kutu grafiği ) kullanılmıştır.

Veri setinde yer alan aykırı değerler, yapılan analizin sonuçlarını olumsuz yönde etkileyerek yanlış çıkarımlara sebep olabileceğiinden dolayı bu değerlere titizlikle dikkat edilmesi gerekiyor. Yaşanabilecek olumsuz sonuçları en aza indirebilmek ve verilerimizin daha sağlıklı sonuçlar verebilmesi sağlamak amacıyla IQR (Interquartile Range – Çeyrekler Açıklığı) yöntemi kullanılır. Bu yöntemle belirlenen aykırı değerler verilerin merkezi eğilimini dikkate alınarak medyan ile değiştirilmiştir.

Burada kullanılan ön işleme adımları sayesinde veri seti daha sağlıklı bir yapıya dönüştürü.



**Şekil 2.** Değişkenlere Ait Histogram, QQ Plot, Kutu Grafikleri

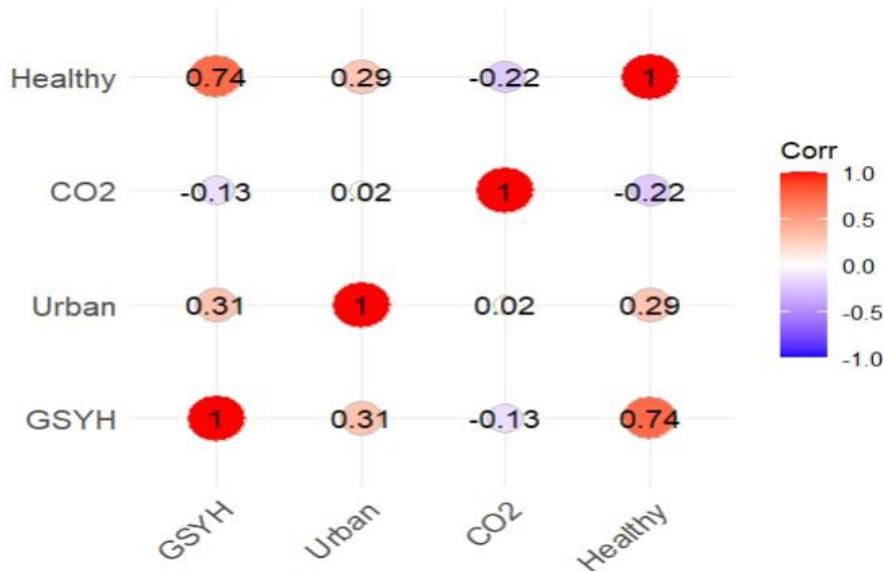
Şekil 2'de yer alan grafikler, verilerin normal dağılıma yakınlık derecesini anlamak açısından önemli bir rol oynamaktadır.

İlk olarak GSYH değişkenine baktığımızda histogramında verilerin sağa çarpık bir dağılım sergilediğini görülmektedir. Gözlem değerlerinin düşük ve orta değer aralığında yoğunlaşıırken yüksek değerlerde ise az sayıda gözlem vardır. Q-Q grafiğine baktığımızda ise verilerin yüksek değerlere doğru çıktıığı zaman doğrusal bir çizgi üzerinden belirgin bir şekilde saplığı görülmektedir. Bu durum verinin normal dağılmadığını göstermektedir. Kutu grafiğine baktığımızda ise medyan çizgisinin kutunun alt yarısına yakın olduğu görülmektedir ve aykırı değerlerin üst kısmada olduğu gözlemlenmektedir.

Kentleşme değişkeninin histogramına baktığımızda ise verilerin yüksek değerlerde toplandığını görürüz bu durum bize histogramın sola çarpık olduğunu göstermektedir. Q-Q grafiğine baktığımızda ise verilerin sadece orta bölgede normal dağılıma yaklaştığını görülmektedir. Bu durum kentleşme oranının normal dağılmadığını göstermektedir. Kutu grafiğine baktığımızda ise dağılımin düzenli olmadığını görülmektedir.

Karbon Emisyonu (CO<sub>2</sub>) değişkeninin histogramına baktığımızda diğer değişkenlere göre çarpıklığı en yüksek olan değişken olduğunu görürüz. Verilerin bir çoğu 0-2000 arasında kümelenmiştir. Bu histogram sağa çarpaktır. QQ Plot grafiğine baktığımız verilerin normal dağılım çizgisinden saplığını açık olarak görürüz. Kutu grafiğine baktığımızda ise aykırı değerlerin üst tarafta çok sayıda olduğunu ve veri setindeki dağılımin dengesiz olduğunu görürüz.

Sağlık harcamaları değişkeninin histogramına baktığımızda ise diğer 3 değişkene göre normal dağılıma en yakın duran değişken olduğunu görürüz. Veriler çoğunlukla 5 ile 15 arasında dağılmıştır. QQ Plot incelemesinde veri noktalarının büyük oranda normal dağılım çizgisine uyumlu olduğunu görmekteyiz. Kutu grafiğine baktığımızda aykırı değer sayısının az olduğunu ve dağılımin simetriye yakın olduğunu görülmektedir .



**Şekil 3.** Korelasyon Matrisi

Bu çalışmada Gayri Safi Yurt İçi Hasıla (GSYH), şehirleşme oranı (Urban), karbondioksit emisyonu (CO<sub>2</sub>) ve genel sağlık göstergeleri (Healthy) arasındaki etkileşimi korelasyon matrisi üzerinden incelemektedir. Korelasyon katsayıları -1 ile +1 arasında değer almaktadır katsayının mutlak değeri arttıkça değişkenler arasındaki ilişkinin gücü artmaktadır.

Verile baktığımızda GSYH ile sağlık göstergesi arasında güçlü bir ilişki tespit edilmiştir ( $r = 0.74$ ). Ortaya çıkan bu yüksek pozitif analizde ekonomik büyümeyenin; sağlık altyapısına erişim, beslenme kalitesi ve tıbbi teknoloji yatırımları aracılığıyla toplum sağlığını doğrudan iyileştirdiğini kanıtlamaktadır. Ekonomik büyümeyenin artmasıyla kişilerin daha nitelikli gıdaya ve sağlık hizmetlerine ulaşabilmesinin sonucu olmuştur.

GSYH ile kentleşme oranı arasında pozitif yönlü ancak orta-düşük düzeyde bir korelasyona sahip olduğu görülmektedir.Yani özetle ekonomi gelişmişçe şehirleşme de bir miktar artı.Benzer şekilde, kentleşme oranı ile sağlık göstergesi arasında zayıf düzeyde pozitif bir ilişki ( $r = 0.29$ ) bulunmaktadır. Bu sonuç, kentleşmenin sağlık üzerinde sınırlı fakat olumlu bir etkiye sahip olabileceği işaret etmektedir.

$\text{CO}_2$  emisyonu ile sağlık göstergesi arasındaki ilişki -0.22 değerindeki negatif korelasyondur. Bu bulgu, endüstriyel faaliyetlerin veya yoğun enerji tüketiminin bir çıktıları olan hava kirliliğinin halk sağlığı üzerindeki olumsuz etkisini göstermektedir.

Analiz bulgularından elde edilen sonuca göre, refah düzeyindeki artışın sağlık göstergelerini olumlu yönde etkilediğini açıklamaktadır. GSYH ile sağlık arasındaki güçlü ilişkiye karşın,  $\text{CO}_2$  emisyonlarının oluşturduğu negatif etki, günümüz ekonomilerinin karşı karşıya olduğu temel yapısal sorunlardan birini yansıtmaktadır. Toplum sağlığının korunması sadece sağlık alt yapısının geliştirilmesi ve gelir düzeylerinin artması ile sınırlı kalmamalı aynı zamanda çevresel kalitenin iyileştirilmesi ve kentleşmenin beraberinde getirdiği riskleri yönetmekte mümkündür.

Kümelere ayırmaya işlemeye geçilmeden önce, analiz için küme sayısı üç olarak belirlenmiştir. Bu doğrultuda veri seti K-means algoritması kullanılarak kümelere ayrılmış ve elde edilen sonuçlara göre birinci, ikinci ve üçüncü kümelerde yer alan gözlem sayıları sırasıyla 248, 189 ve 43 olarak tespit edilmiştir.

Kümeleme işleminin ardından, her bir küme için analiz edilen değişkenlerin ortalama değerleri hesaplanmış ve elde edilen bulgular raporlanmıştır:

**Tablo 2.** Küme bazlı değişkenlerin ortalama değerleri

Küme	GDP	Urban	CO2	Healthy
1	0.7684328	0.3992644	-0.1687308	0.7328186
2	-0.7948630	-0.4851571	-0.3992015	-0.7336962
3	-0.9381914	-0.1702996	2.7277749	-1.0016381

Kümeleme analizi sonucunda üç farklı küme belirlenmiştir. **Küme 1;** GDP, Urban ve Healthy değişkenlerinde ortalamanın üzerinde pozitif değerlere sahip tek gruptur. Özellikle GDP ve Healthy değerlerinin yüksek olması, bu ülkelerin hem ekonomik refah hem de sağlık standartları açısından en avantajlı grup olduklarını gösterir. Bu grubun CO2 değerinin negatif olması, söz konusu ülkelerin ekonomik büyümelerini çevresel maliyeti düşük tutarak gerçekleştirdiklerine işaret etmektedir. Küme 2; tüm değerlerin negatif olduğu, dezavantajlı bir kümedir. Düşük ekonomik gelir ve düşük sağlık göstergelerine sahip, iyileştirici politikalara en çok ihtiyaç duyan ülkelerin oluşturduğu bir gruptur. Küme 3, diğer kümelerden oldukça farklı bir profil sergilemektedir. Bu kümeyi dikkat çeken özelliği düşük gelir ve düşük sağlık standartlarına rağmen çok yüksek emisyon değerlerine sahip olmasıdır. Bu durum bu gruptaki ülkelerin verimsiz sanayileşme sürecinde olduklarına işaret etmektedir.

Kümeleme analizinin performansını ve kümelerin homojenliğini değerlendirmek için önemli metrikler Tablo 3'de yer almıştır. Bu tablo, bir kümeleme analizinin sonuçlarını göstermekte ve küme içi hata kareleri toplamını (Within\_SS) ve toplam açıklanan varyansı (Between\_SS / Total\_SS) sunmaktadır.

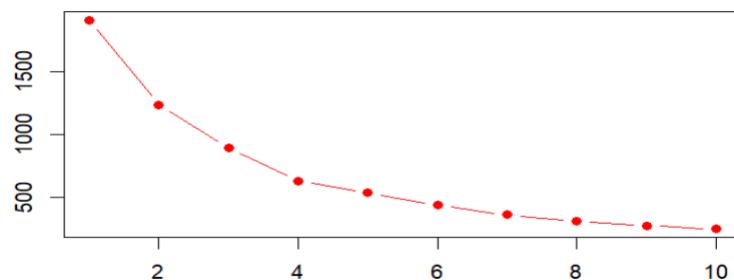
**Tablo 3.** Kümeler Arası ve Kümeler İçi Toplam Kareler (SS) Dağılımı

Cluster	Within_SS
<b>1</b>	470.32504
<b>2</b>	377.25070
<b>3</b>	44.25929
<b>Total (Between_SS / Total_SS)</b>	53.5 %

Küme 1'in **Within\_SS** değeri oldukça yüksektir (**470.32504**). Bu Küme 1'in içindeki gözlem değerlerinin dağılıminin daha geniş olduğunu ve bu kümenin diğerlerine göre daha heterojen bir yapıda olduğunu göstermektedir. Küme 2'nin **Within\_SS** değeri (**377.25070**), Küme 1'e göre daha düşük bir değerdir. Bu küme, Küme 1'e göre daha homojen olsa da hala yüksek bir varyansa sahiptir. Küme 3 en düşük **Within\_SS** değerine sahip kümedir(**44.25929**). Bu durum, bu küme içindeki gözlemlerin birbirine çok yakın olduğunu, kümenin kendi içinde homojen ve sıkı bir yapıda olduğunu göstermektedir. Tablodaki **53.5%**, toplam verinin %53.5'inin kümeler arasındaki farklılıklarla açıklanabildiğini göstermektedir. Bu, kümeleme modelinin açıklayıcı gücünün orta düzeyde olduğunu ifade eder. Ancak geriye kalan %46.5'lik varyans, kümeler içindeki farklılıklarla açıklanmaktadır, bu da modelin daha iyi iyileştirilebileceğini düşündürmektedir.

Bu çalışmada, en uygun küme sayısının (k) belirlenmesi için Dirsek (Elbow) yöntemi kullanılmıştır. Dirsek Yöntemi, her bir verinin ait olduğu küme merkezlerine olan uzaklıklarının karelerinin toplamını kullanarak k değeri belirler. Grafikte, WCSS değerindeki değişimin azaldığı nokta, yani "dirsek" noktası, verilerin en iyi şekilde modellenmesini sağlayacak k küme sayısını ifade eder.

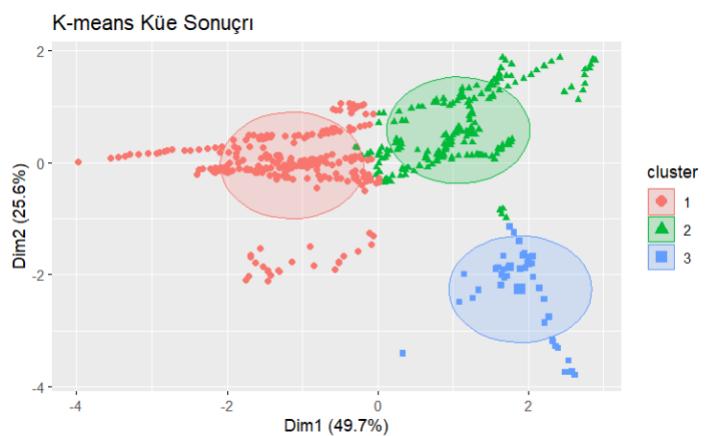
### ELBOW YÖNTEMİ



**Şekil 4.** Elbow (Dirsek) Yöntemi

Dirsek Yöntemi ile optimal küme sayısının 3 olduğu tespit edildikten sonra, bu kümeleme sonuçlarını görselleştirme işlemi yapılmıştır. Kümeleme sonuçlarının görsel olarak sunulması, elde edilen kümelerin yapısının daha iyi anlaşılmasına ve yorumlanmasına katkı sağlamaktadır. Bu görselleştirme sayesinde kümeler arasındaki ayırmalar daha net bir biçimde ortaya konulurken, her bir gözlemin hangi kümeye dahil olduğu açıkça gösterilmektedir.

Görselleştirme işlemiyle oluşan kümelerin her birini farklı renkler ve sembollerle görürüz, kümelerin birbirlerine olan uzaklıklarını ve verinin genel yapısını analiz etmemiz daha kolay hale gelir.



**Şekil 5.** Kümeleme Sonucunun Görselleştirilmesi

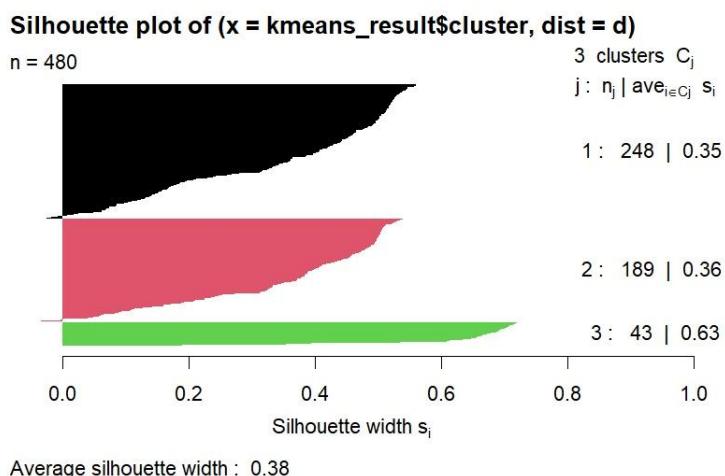
Grafikteki veriler, k-means kümeleme algoritması kullanılarak birbirine en yakın özellikler gösteren 3 farklı kümede (Cluster) toplanmıştır. Her bir kümenin merkez noktaları ve dağılım yapıları farklı renkler ile semboller kullanılarak ayrı edilmiştir. Grafiğe bakıldığından Küme 1 kırmızı renkle, Küme 2 yeşil renkle, Küme 3 ise mavi renkle gösterilmiştir. Bu kümeleme yapısı, çok sayıda değişkene sahip verilerin daha kolay anlaşılabilmesi için PCA yöntemi kullanılarak iki boyuta indirgenmiştir.

Grafik üzerinde yatay eksen Dim1'i, dikey eksen ise Dim2'yi temsil etmektedir. Boyut indirgeme sonucunda elde edilen Dim1 ve Dim2 toplam varyansın sırasıyla %49.7 ve %25.6'sını açıklayan boyutlar olarak belirlenmiştir. Bu iki boyutun birleşimi, verideki toplam varyansın %75.3'ünü temsil ederek genel veri yapısını ve gruplandırmaları anlamada oldukça yeterli olduğunu göstermektedir.

Grafikte yer alan üç kümeye bakıldığından mavi kümenin kırmızı ve yeşil kümeden belirgin bir şekilde ayrıldığı görülmektedir. Kırmızı kümenin dim1 eksen boyunca geniş bir alana yayıldığı görülür. Bu durum dim1 boyuttunda daha çeşitli olduğunu göstermektedir. Mavi kümeninde diğer kümelere göre daha düzenli bir yapıda olduğu görülmektedir. Genel olarak bakıldığından kümeler arasında belirgin bir ayrışma olduğu görülür ve elipslerin büyük ölçüde birbirine temas etmediği gözlemlenir. Noktaların büyük kısmının kendi küme renkleri içinde toplanması, K-means algoritmasının verileri ayırt edici özelliklerine göre başarılı bir şekilde sınıflandırdığını ortaya koymaktadır.

Kümeleme görselleştirmeleri, K-Means algoritmasını yorumlamak açısından önemli bir araçtır. Bu görselleştirmeler sayesinde her bir küme içerisindeki dağılımı net bir şekilde gözleme imkanı sunar.

Kümeleme başarısını değerlendirmek için Silhouette analizi kullanılır. Mevcut çalışmanın Silhouette skoru ve analizi aşağıdaki gibidir:



**Şekil 6.** Silhouette Skoru ve Analizi

Yapılan analiz sonucunda genel ortalama siluet genişliği 0.38 olarak hesaplanmıştır. Literatürde 0.25 ile 0.50 arasındaki değerler genellikle "zayıf ile orta düzeyde yapısal veri grubu" olarak tanımlanır. Bu durum, veri setinde doğal bir kümeleme yapısının mevcut olduğunu ancak kümeler arasındaki sınırların tam olarak kesinleşmediğini göstermektedir.

Silhouette analizi sonuçları, çalışmanın kümeleme başarısını değerlendirmek için önemli bilgiler sunmaktadır. Küme 1, 0.35 ortalama Silhouette skoru ile kümeler arasında en zayıf ayrışmayı göstermektedir. Bu durum, küme 1 içerisinde yer alan gözlemler arasında daha yüksek düzeyde çeşitlilik olduğu ve özellikle çok sayıda ülke ve yılı kapsaması nedeniyle, bu kümeye ait homojenliğin diğer kümelere kıyasla daha düşük olabileceği anlamına gelmektedir. Küme 2 ortalama silhouette değeri 0.36 olarak hesaplanmıştır. Bu kümeye de orta düzeyde bir ayrışma göstermektedir. Küme 3 ise 0.63 gibi yüksek bir ortalama silhouette değerine sahiptir. Bu durum, üçüncü kümeyi diğer kümelerden belirgin bir şekilde ayırttığını ve iç tutarlılığının güçlü olduğunu ortaya koymaktadır.

Genel olarak, tüm kümeler için ortalama Silhouette skoru 0.38 olarak hesaplanmıştır. İdeal Silhouette skoru literatürde genellikle 0.50 olarak kabul edilmektedir. Bu değer, oluşturulan kümeyi yapısının orta düzeyde bir ayırt ediciliğe sahip olduğunu ve kümelerin tamamen net sınırlarla ayrılmadığını göstermektedir. Bununla birlikte, veri yapısının belirli alt gruplar içerdığı ancak bu gruplar arasında kısmi örtüşmelerin bulunduğu söylenebilir.

Son olarak da Tablo 4'te çalışmaya konu olan ülkelerin hangi kümede yer aldığı bilgisi sunulmuştur.

**Tablo 4.** Kümeleme Sonuçları

Küme	Ülkeler
Küme 1	Arjantin, Avustralya, Kanada, Fransa, Almanya, İtalya, Japonya, İngiltere, ABD, Avrupa Birliği
Küme 2	Brezilya, Çin, Endonezya, Güney Kore, Meksika, Suudi Arabistan, Güney Afrika, Türkiye
Küme 3	Hindistan, Rusya

Kümeleme analizi sonuçları; ülkelerin GSYH (Ekonomik Güç), Urban (Şehirleşme Oranı), CO2 (Karbon Emisyonu/Endüstriyel Etki) ve Healthy (Sağlık/Yaşam Kalitesi) değişkenleri baz alındığında üç temel yapısal grupta toplandığını göstermektedir. Birinci kümede 248 veri, ikinci kümede 189 veri, üçüncü kümede ise 43 veri yer almaktadır. Ülkelerin 2000-2023 dönemine ait verileri incelenmiş ve ülkeler en uzun süre istikrar gösterdikleri küme içerisinde sınıflandırılmıştır. Arjantin, Brezilya, Hindistan, İtalya, Güney Kore, Suudi Arabistan ve Avrupa Birliği belli yıllarda farklı kümelerin özelliklerini sergilemişlerdir. Bu durum, söz konusu ülkelerin analiz dönemi boyunca yaşadıkları dönemsel değişimlerden geçici veya kalıcı şekilde etkilendiklerini göstermektedir. Küme 1'deki çoğu ülke daha önce Küme 2'de yer almış ülkeler olduğundan, elde edilen küme sonuçlarında bu iki kümenin bazı birimleri birbirine oldukça yakındır.

Birinci grup (ABD, Almanya, Japonya, Kanada vb.), küresel ekonominin merkezinde yer alan ülkeleri temsil eder. Bu grup; gelir düzeyi ve sağlık/yasam kalitesi göstergeleri bakımından diğer gruplara göre oldukça iyi durumdadır. İkinci grupta bulunan ülkeler ise (Türkiye, Çin, Brezilya vb.) gelişmekte olan ülkelerin genelini kapsamaktadır. Bu grubun gelir seviyesi, kentleşme oranları ve sağlık/yasam kalitesi göstergeleri düşüktür; karbon emisyonları ise birinci gruba kıyasla daha düşük seviyededir. Bu durum, söz konusu ülkelerin sanayi yapısı ve kişi başına üretim düzeyleri ile ilişkilidir. Üçüncü grup diğer gruplara göre farklı özellikler sergilemektedir. Bu grubun en belirgin özelliği CO2 emisyonu tarafındaki ağırlığıdır. Rusya'nın enerji yoğun endüstrisi ve Hindistan'ın devasa nüfusuna rağmen hızla artan kömür bazlı enerji üretimi, bu iki ülkeyi karbon ayak izi bakımından diğerlerinden net bir şekilde ayırmaktadır.

## 5.Sonuç

Yapılan bu çalışmada G20 ülkelerini 2000-2023 yılları arasında (GSYH) Gayri Safi Yurtiçi Hasıla ile Kentleşme, Sağlık harcamaları ve son olarak karbon emisyonu arasındaki ilişkiyi makine öğrenmesi algoritmalarından olan K-Means yöntemi kullanılarak bir analiz gerçekleştirilmiştir. Çalışma boyunca veri hazırlama, aykırı değerleri bulma ve analizini yapma, standardizasyon işlemi uygulanarak daha sağlıklı sonuçlar bulmayı hedeflenmiştir.

Elbow yöntemi kullanılarak bu çalışmaya uygun olan küme sayısının 3 olduğu belirlenmiştir. Birinci kümede (ABD, Almanya, Japonya, Kanada, İtalya, Fransa vb.) yer alan ülkelere bakıldığından küresel ekonominin merkezinde yer alan ülkeler görülmektedir. Bu küme diğer iki kümeye göre gelir düzeyi ve sağlık/yaşam kalitesi bakımından daha iyi durumdadır. İkinci kümeye bakıldığından ise (Brezilya, Çin, Meksika vb.) bu kümede yer alan ülkelerin gelişmekte olan ülkeler olduğunu görülür. Bu kümeyi gelir seviyesi, kentleşme oranları ve sağlık/yaşam kalitesi ve karbon emisyonları oranı daha düşüktür. Bu durumun nedeni bu ülkelerin sanayi yapısının ve kişi başına düşen üretim düzeyinden kaynaklanmaktadır. Son olarak üçüncü kümeye bakıldığından ise (Hindistan ve Rusya) burada yer alan ülkelerin karbon emisyonu oranının daha yüksek olduğu görülmektedir. Bu durumun nedeni ise bu ülkelerde artmakta olan kömür bazlı enerji üretimidir, bu iki ülkenin karbon ayak izi diğer kümelerde yer alan ülkelere göre belirgin bir şekilde yüksektir.

Kümeleme performansını Silhouette skoru ile değerlendirildiğinde yapılan analizin genel ortalama silüet genişliği 0,38 olarak belirlenmiştir. Bu durum veri setinde doğal bir kümeleme olduğu ama kümeler arasındaki sınırların net olmadığını gösterir.

Kümeler arasındaki farklılıklara bakıldığından ise birinci kümeyi GSYH, kentleşme ve sağlık değişkenlerindeki ortalamasının pozitif olduğu ama CO<sub>2</sub> değişkeninin ortalamasının negatif olduğu görülmektedir. Bu duruma bakıldığından Küme 1'deki ülkelerin ekonomik, refah ve sağlık standartlarının yüksek olduğu gözlemlenir. Karbon emisyon miktarının düşük olması bu ülkelerin çevreye duyarlı olduğunu gösterir. Küme 3' e baktığımızda ise düşük gelir ve düşük refah düzeyine oranla daha yüksek karbon emisyonu değerlerine sahip olmasıdır. Bu durum bize bu ülkelerin sanayileşme sürecini verimsiz olduğunu gösterir.

Sonuç olarak bu çalışma G20 Ülkelerinin GSYH ile kentleşme, sağlık ve karbon emisyonları arasındaki ilişkiyi kümeleme yöntemiyle analiz ederek literatüre katkı sağlamayı hedeflemiştir. Elde edilen sonuçlar G20 ülkelерinin ekonomik, refah ve sağlık standartları ilişkisini anlatmaktadır. Bu araştırmmanın gelecekteki araştırmalarda katkı sağlama beklenmektedir.

## Kaynakça

- Ceren, P., Aysegül, H., & Nergis, B. (2020). G20 ÜLKELERİNDE CO<sub>2</sub> EMİSYONU VE ENERJİ TÜKETİMİNİN SOSYAL VE EKONOMİK DEĞİŞKENLER ÜZERİNDEKİ ETKİSİ. *Beykoz Akademi Dergisi*, 8(1), 334-348.
- Debolina, K. (2025). Future Urbanisation and Habitat: Transitioning Towards Climate Responsive and Sustainable Development. 251-266.
- Emine, B., Bahar, Y. S., & Aslı, G. Z. (2025). Kır-kent ikileminde yeni arayışlar: K-means algoritması ile yerleşim birimlerinin kümelendirilmesi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 40(3), 1467-1478.
- Kemal, S., & Ramazan, A. (2020). K-Means Kümeleme Algoritması Kullanılarak Oluşturulan Yapay Zekâ Modelleri ile Sediment Taşınımının Tespiti. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 9(1), 306-322.
- Mahir, Ö., & Nazlı, K. (2020). VERGİ GELİRLERİ, KENTLEŞME VE DEMOKRASİ: G20 ÜLKELERİ ÜZERİNE EKONOMETRİK BİR ÇALIŞMA. *EKEV AKADEMİ DERGİSİ*, 24(84).
- Meltem, I., & Yılmaz, Ç. A. (2007). K-Means, K-Medoids Ve Bulanık C-Means Algoritmalarının Uygulamalı Olarak Performanslarının Tespiti. *İstanbul Commerce University Journal of Science*, 6(11), 31-45.
- Murat, A., & Mesut, A. (2021). Kümeleme Performansını Ölçmek için Yeni Bir Yöntem ve Metin Kümeleme için Değerlendirmesi. *Avrupa Bilim ve Teknoloji Dergisi*(27), 53 - 65.
- Oytun, M., & Mustafa, K. (2018). G20 ÜLKELERİNDE EMİSYONLAR VE GAYRİ SAFİ YURTİÇİ. *Uluslararası Yönetim İktisat ve İşletme Dergisi*, 14(3).
- Özge, E., & Gülden, Ü. (2024). G20 ÜLKELERİNDE SAĞLIK HARCAMALARININ SAĞLIK ÇIKTILARI ÜZERİNDEKİ ETKİSİ. *Dicle Akademi Dergisi*, 4(1), 28-44.
- Pakize, E., Buket, Ç., & Zehra, D. (2016). K\_Means Algoritması İle Otomatik Kümeleme. *El-Cezeri*, 3(2).
- Pınar, H., & Çebi, B. H. (2015). Üst Orta Gelirli Ülkelerde Sağlık Harcamaları ve Ekonomik Büyüme İlişkisi. *İşletme ve İktisat Çalışmaları Dergisi*, 3(2), 35-44.
- Safa, D., Hatem, D., Durmuş, Ö., & Hasan, T. (2022). Melezlenmiş K-means ve Diferansiyel Gelişim Algoritmaları ile Kalp Hastalığının Teşhisini. *1st International Conference on Engineering and Applied Natural Sciences*.
- Samet, G., Seçkin, C. K., Serdar, G. H., & Erman, C. (2018). TRAFİK KAZA DESENLERİİNİN TANIMLANMASINDA K-MEANS KÜMELEME ALGORİTMASININ KULLANILMASI: SAKARYA İLİ UYGULAMASI. *İşletme Bilimi Dergisi*, 6(3), 89-105.
- Selim, D. (2021). Çevresel İnovasyonlar, Yenilenebilir Enerji Tüketimi ve Ekonomik Büyümenin CO<sub>2</sub> Emisyonu Üzerine Etkileri: Seçilmiş G-20 Ülkeleri için Panel Veri Analizi. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 9(4), 1007-1017.

Serap, C. E., Seda, G. H., & Reşit, Y. H. (2016). Hücresel İmalat Sistemlerinin K-Means Algoritması ve Genetik Algoritma. *Academic Platform - Journal of Engineering and Science*, 4(3).

Sibel, S., Doğan, U., & Pınar, E. (2014). Türkiye'de Sağlık Harcamalarının Ekonomik Büyüme Üzerindeki Etkisinin Ekonometrik Analizi. *Niğde Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 7(3), 13-24.

Topaloğlu, F. (2024). Saldırı Tespit Sistemlerinde K-Means Algoritması ve Silhouette Metriği ile Optimum Küme Sayısının Belirlenmesi. *Journal of Information Technologies*, 17(2), 71 - 79.

Tuğba, T., & Ferhat, B. (2025). Üniversite Öğrencilerinin Eleştirel Düşünme Profillerinin k-Means Kümeleme Algoritması İle Analizi. *International Journal of Pure And Applied Sciences* , 11(1), 269-280.

Zeynep, K., Havva, G., & Gözde, M. (2021). G20 Ülkelerinde Sağlık Harcamaları, Yaşam Beklentisi ve Ekonomik Büyüme İlişki Üzerine Bir İnceleme. *GAZİANTEP UNIVERSITY JOURNAL OF SOCIAL SCIENCES*, 20(4), 1600-1616.

## **GRUP ÜYELERİNİN ROLÜ**

**NURGÜL YILDIZ:** Bu çalışmada giriş bölümünü ekip arkadaşlarımla birlikte hazırladım, literatür taraması kapsamında sekiz bilimsel makale alıntısını çalışmaya ekledim ve metodoloji bölümünü oluşturdum. Analiz sürecinde kullanılan kod dosyalarını hazırladım, geliştirilen analiz uygulamasını tanıttım, veri setindeki değişkenleri ve özet istatistikleri yorumladım. Değişkenlerin karşılaştırılabilir hale getirilmesi amacıyla standardizasyon işlemini tanımladım. Kümeleme analizi sonucunda elde edilen kümeler için değişken ortalamalarını inceledim, kümeler arası ve kümeler içi toplam kareler dağılımını yorumladım ve oluşan kümelerde yer alan ülkeleri değişkenlere göre sınıflandırarak sonuçları değerlendirdim.

**EBRAR TOPAL:** Makale için giriş bölümü birlikte hazırlanmış, literatür taraması kısmı ise görev dağılımı yapılarak tamamlanmıştır. Bu kapsamda beş adet literatür taraması çalışmaya ekledim. Ayrıca korelasyon matrisi analiz edilerek, matriste yer alan değişkenler arasındaki ilişkiler her bir değişken için ayrı ayrı yorumladım. Kümeleme analizinde kullanılan Elbow yöntemi görseli değerlendirdim ve uygun küme sayısını yorumladım. Bununla birlikte, Silhouette skoru tanımlanmış ve elde edilen sonuçlar doğrultusunda kümeleme başarısını yorumladım.

**NEBAHAT AKARSU:** Bu çalışmada kullanılan verileri ben topladım ve analiz sürecine uygun şekilde düzenledim. Literatür taraması ekip arkadaşlarımla birlikte gerçekleştirilmiştir. Veri ön işleme aşamasında uygulanan adımları ben yazdım; değişkenlerin dağılım yapıları histogram, Q-Q plot ve kutu grafikleri aracılığıyla incelenmiş, her bir grafik ayrı ayrı yorumlanarak değişkenlerin dağılım özellikleri ayrıntılı biçimde açıklanmıştır. K-means kümeleme analizi sonucunda elde edilen küme görselleştirmesini ben yorumladım ve Dim1 ile Dim2 boyutlarının veri seti üzerindeki anlamlarını açıkladım. Ayrıca çalışmanın sonuç bölümünü tarafımdan yazılarak elde edilen bulgular bütüncül bir bakış açısıyla değerlendirilmiştir

## GÖREV VE SORUMLULUKLAR

	Giriş	Literatür Taraması	Veri Anlama	Veri Hazırlama	Modelleme ve R Kodlama	Değerlendirme ve Sonuç
Nurgül Yıldız	P	P	S	S	P	P
Ebrar Topal	P	P	P	S	S	P
Nebahat Akarsu	P	P	S	P	S	P

**P:** Birinci derecede sorumluluk

**S:** İkinci derecede sorumluluk

No	Değerlendirme Ölçütleri Kontrol Listesi	Yapıldı (E/H)
1	Projenin amacı açık ve net olarak açıklanmış mı?	E
2	Verinin tam olarak anlaşılmasına ilişkin içerdeği nitelikler (Değişkenler) anlaşılmış mı? (veya açıklanmış mı?)	E
3	Verinin anlaşılmasını kolaylaştıran veri özetleme ve/veya gruplama yapılmış ve yorumlanmış mı? Verinin anlaşılmasına ilişkin ilgili grafikler alınmış ve yorumlanmış mı?	E
4	Veri üzerinde ön işlemler olarak bilinen düzeltme, eksik verilerin giderilmesi, tekrarlayan verilerin silinmesi, dengesizliği oluşturan verilerin düzgünleştirilmesi veya gerek duyulan normalizasyon/standardizasyon işlemleri dikkate alınmış mı?	E
5	Modelin performansını ölçmek için ayrılan eğitim ve test veri setleri alternatif yöntemler ile yapılarak performans sonuçları yorumlanmış mı?	H
6	Seçilen algoritmanın fonksiyonu üzerinde argümanları ve ilgili parametreler değiştirilerek alternatif sonuçların karşılaştırılması yapılmış mı?	E
7	Alternatif algoritmalar kullanılarak geliştirilen modelin performansları karşılaştırılmış mı?	E
8	Modelin performans raporları açıklanmış ve yorumlarınız katılmış mı?	E
9	Yeni veri tanımlanarak seçilen modelin tahmini veya öngörüsü yapılmış mı?	H
10	Elde edilen modelin sonuçlarının uygulanabilirliği ve faydalari tartışılmış mı?	E
11	Proje raporu istenen başlıklara göre ve formata uygun olarak yazılmış mı?	E
12	Projenizle ilgili çalışma ve uygulamalar araştırılmış mı? Kullandığınız kaynaklar belirtilmiş mi?	E
13	Projenizin sonuna görev dağılım tablosu eklenmiş mi?	E
14	Projenizi ekibiniz ile birlikte yaptığınız özgün bir çalışma olduğunu onaylar misiniz?	E

