

Türkiye’deki Suç Türlerinin İllere Göre İkili Kümeleme Yöntemiyle İncelenmesi

Demet Albasar¹ , Ahmet Kocatürk ^{*2} , Hacı Hasan Örkücü³ 

¹Adalet Bakanlığı, 06500, Ankara, Türkiye

²Erzurum Teknik Üniversitesi, Fen Fakültesi, Matematik Bölümü, 25100, Erzurum, Türkiye

³Gazi Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 06500, Ankara, Türkiye

Öne Çıkanlar

- Türkiye’deki 58 suç türü, 81 il temelinde ikili kümeleme yöntemi ile analiz edilmiştir.
- Suç türleri ve coğrafi bölgeler arasındaki örüntüler Cheng & Church (CC) algoritması ile belirlenmiştir.
- Elde edilen ikili kümeler, illere göre suç yoğunluğunu ve birlikte görülme yapılarını ortaya koymaktadır.
- Analiz sonuçları, suç önleme politikaları ve güvenlik stratejileri için mekânsal düzeyde veri temelli öneriler sunmaktadır.
- Bu çalışma, Türkiye’de suç verileri üzerinde ikili kümeleme yönteminin uygulanmasına yönelik özgün bir model sunmaktadır.

Makale Bilgileri

Geliş: 22/07/2025

Kabul: 01/10/2025

Anahtar Kelimeler

Suç Analizi
Suç Örüntüleri
İkili Kümeleme
Cheng & Church
Algoritması

Öz

Bu çalışma, Türkiye’de 2023 yılına ait 81 ili ve 58 farklı suç türünü kapsayan adli veriler kullanılarak, suçların mekânsal ve tematik düzeydeki örüntülerini belirlemeyi amaçlamaktadır. Geleneksel analiz yöntemlerinin sınırlı kaldığı durumlarda, özellikle yüksek boyutlu ve çok değişkenli veri yapılarında, ikili kümeleme tekniklerinin sunduğu analitik güç ön plana çıkmaktadır. Bu bağlamda, çalışmada Cheng & Church (CC) algoritmasına dayalı ikili kümeleme yöntemi kullanılarak iller ve suç türleri eşzamanlı olarak analiz edilmiştir. Uygulanan yöntem sayesinde benzer suç profiline sahip iller ile bu illerde yoğunlaşan suç türleri aynı kümelerde tanımlanabilmektedir. Toplam yedi ikili küme elde edilmiş, bunlardan beş tanesi içerdiği örüntü gücü nedeniyle harita üzerinde detaylı biçimde görselleştirilmiştir. Bulgular, suç türlerinin yalnızca bireysel değil aynı zamanda yapısal ve bölgesel örüntülerle şekillendiğini ve suçla mücadele stratejilerinin yerel düzeyde yapılandırılması gerektiğini ortaya koymaktadır. Ayrıca çalışma, veri madenciliği tekniklerinin suç analitiğinde kullanılabilirliğini ortaya koyarak hem akademik alana katkı sunmakta hem de kamu güvenliği ve stratejik planlama açısından pratik öneriler geliştirmektedir.

A Biclustering-Based Examination of Crime Types Across Provinces in Türkiye

Highlights

- A total of 58 crime types in Turkey were analyzed using the biclustering method based on 81 provinces.
- Patterns between crime types and geographic regions were identified through the Cheng & Church (CC) algorithm.
- The resulting biclusters reveal crime intensity and co-occurrence structures across provinces.
- The findings offer spatially informed, data-driven insights for crime prevention policies and security strategies.
- This study presents a novel application of the biclustering approach to crime data analysis in the Turkish context.

Article Info

Received: 22/07/2025

Accepted: 01/10/2025

Keywords

Crime Analysis,
Crime Patterns,
Biclustering,
Cheng & Church
Algorithm

Abstract

This study aims to identify the spatial and thematic patterns of crimes in Turkey by analyzing official judicial data from 2023, covering 81 provinces and 58 different crime types. In cases where traditional analysis methods fall short—particularly in high-dimensional and complex datasets—biclustering techniques offer a powerful alternative. In this context, the study employs the Cheng & Church (CC) biclustering algorithm to simultaneously analyze crime types and provinces. This method enables the identification of provinces with similar crime profiles and the co-occurrence of specific crime types within these regions. A total of seven biclusters were obtained; five of them, due to their strong pattern structure, were visualized and interpreted in detail through thematic maps. The findings reveal that crimes are not solely shaped by individual factors but also by structural and regional patterns, emphasizing the need for locally grounded crime prevention strategies. Moreover, this study demonstrates the applicability of data mining methods in crime analytics, contributing not only to the academic literature but also providing practical insights for public safety planning and policy-making.

1. GİRİŞ

Toplumsal düzenin korunmasında önemli bir yere sahip olan suç olgusu, bireylerin hukuki normlara, etik kurallara ve toplumsal değerlere aykırı davranışlarını kapsar. Bu davranışların ortaya çıkmasında demografik yapı, sosyoekonomik koşullar, bireysel psikolojik özellikler ve politik faktörler gibi çok yönlü değişkenlerin etkili olduğu bilinmektedir. Suç ve suça neden olan unsurların çok boyutlu doğası, bu olgunun yalnızca bireysel ya da mekânsal değil, aynı zamanda zamansal ve yapısal yönleriyle de ele alınmasını gerektirmektedir. Bu nedenle suç analizi, suçu oluşturan bileşenlerin bütünsel bir bakış açısıyla incelenmesini hedefleyen sistematik bir araştırma alanı haline gelmiştir [1].

Günümüzde suç oranlarında yaşanan artış ve suç türlerinin çeşitlenmesi, suç analizine olan ilgiyi artırmış ve özellikle kamu güvenliği, kaynak planlaması, polis devriye optimizasyonu ve stratejik güvenlik politikalarının geliştirilmesi gibi uygulamalı alanlarda analitik yaklaşımların önemini pekiştirmiştir. Geleneksel istatistiksel yöntemler belirli durumlarda fayda sağlasa da, büyük hacimli, karmaşık ve çok boyutlu suç verilerinin analizinde yetersiz kalabilmektedir [2]. Bu noktada veri madenciliği, makine öğrenmesi ve yapay zekâ temelli yöntemler, örüntü tanıma, sınıflandırma ve tahminleme gibi işlevleri sayesinde yeni nesil analiz yaklaşımlarının merkezinde yer almaktadır [3,4].

Veri madenciliğinin temel tekniklerinden biri olan kümeleme, benzer özellik gösteren veri noktalarının gruplanması yoluyla anlamlı yapılar ortaya çıkarır. Ancak geleneksel kümeleme teknikleri genellikle yalnızca satır veya sütun boyutlarını dikkate alarak işlem yaptığından, özellikle yüksek boyutlu ve heterojen veri kümelerinde gözlemlenen karmaşık ilişkileri ortaya çıkarmada sınırlı kalmaktadır. Bu bağlamda, ikili kümeleme yöntemi, hem satır (gözlem birimi) hem de sütun (değişken) boyutundaki örüntüleri eşzamanlı olarak analiz edebilme gücüyle öne çıkmaktadır [5].

Başlangıçta gen ekspresyon analizlerine yönelik geliştirilen ikili kümeleme algoritmaları, son yıllarda sosyal bilimlerden pazarlama analizlerine, suç örüntülerinden eğitim araştırmalarına kadar farklı disiplinlerde uygulanabilirliğini kanıtlamıştır [6,7]. Suç verileri özelinde bu yöntem, hem belirli suç türlerinin mekânsal dağılımını hem de bu suçların birlikte görülme yapılarını ortaya çıkarmada etkili bir analitik yaklaşım sunmaktadır [8]. Örneğin, Brown [1] tarafından geliştirilen ReCAP sistemi ile bölgesel suç analizleri desteklenmiş; Adderly ve Musgrove [3] ise Self-Organizing Map (SOM) algoritmasını cinsel saldırı suçlarında tahmin ve sınıflandırma amacıyla kullanmışlardır.

Bu çalışmada, Türkiye’de 2023 yılına ait Cumhuriyet Başsavcılıklarında kayıt altına alınan suç verileri temel alınarak, 81 ilde işlenen 58 farklı suç türü analiz edilmiştir [9]. Çalışmanın amacı, suç türleri ile iller arasındaki örüntüleri ikili kümeleme yöntemi aracılığıyla ortaya koymaktır. Bu amaç doğrultusunda Cheng & Church (CC) algoritması tercih edilmiş, analizler R programlama dili üzerinden gerçekleştirilmiştir. Elde edilen kümeler aracılığıyla, benzer suç profiline sahip iller ile bu illerde yoğunlaşan suç türleri eş zamanlı olarak belirlenmiş; böylece klasik analiz tekniklerinin ötesine geçilerek daha derin yapısal örüntüler ortaya çıkarılmıştır.

Bu araştırma, suçun mekânsal analizine dayalı güvenlik stratejilerinin geliştirilmesine katkı sunmayı amaçlamakta ve ikili kümeleme yöntemlerinin suç analitiği bağlamında sağlayabileceği potansiyeli göstermektedir. Bulgular, hem suç önleme politikaları hem de bölgesel güvenlik planlamaları açısından karar vericilere anlamlı bilgiler sunabilecek niteliktedir.

2. İKİLİ KÜMELEME YÖNTEMLERİNİN SUÇ ANALİZİNDE KULLANIMI

Yüksek boyutlu ve çok değişkenli suç verilerinin analizinde, yalnızca sayısal yoğunluklara değil, tematik ve mekânsal örüntülere de odaklanmak, daha anlamlı sonuçlar elde edilmesini mümkün kılar. Bu bağlamda, ikili kümeleme yöntemi, gözlem birimlerini (iller) ve değişkenleri (suç türleri) eşzamanlı olarak analiz edebilme potansiyeliyle, klasik kümeleme yöntemlerine kıyasla önemli avantajlar sunmaktadır [5,6]. Özellikle büyük veri ortamlarında, CC, Bimax, xMotif ve Plaid gibi algoritmalar sayesinde karmaşık örüntü yapıları yüksek doğrulukla tespit edilebilmektedir [7].

Suç verilerinin ikili kümeleme ile analizine ilişkin çalışmalar da giderek artmaktadır. Örneğin, Arslan ve arkadaşları [8], kaçakçılık suçları üzerine yürüttükleri çalışmada, suç türleri ile yakalanan materyaller arasında anlamlı kümeler ortaya koymuştur. Çil ve arkadaşları [10], Boston'daki suç verileri üzerinde yürüttükleri çalışmada, CC ve XMotif algoritmalarını karşılaştırmalı biçimde uygulamış ve CC algoritmasının örüntü tanımada daha başarılı sonuçlar ürettiğini göstermiştir.

Uluslararası alanda Chen ve arkadaşlarının [11] geliştirdiği COPLINK sistemi, polis birimleri arasında bilgi paylaşımı ve suç analizi açısından güçlü bir örnek teşkil etmektedir. Benzer şekilde, Nath [12] mekânsal suç örüntülerini k-ortalamalar algoritması ile analiz ederken; Bruin ve arkadaşları [13], suçlu tiplerini sınıflamak amacıyla görsel ikili karşılaştırmalar yöntemini kullanmıştır. Türkiye bağlamında ise Oğuzlar [14], Bursa iline ait ağır suç verileri üzerinde SOM algoritması ve karar ağaçlarıyla yapısal kümeler belirlemiştir. Sevüktekin ve arkadaşları [15] ise CHAID algoritması ile suçlu profillemeye çalışmaları gerçekleştirmiştir.

Bu araştırma, ikili kümeleme yöntemini Türkiye genelindeki resmi verilerle bütünleştirerek, mekânsal örüntülerin eşzamanlı ve yapısal bir bakışla analiz edilmesine olanak tanımaktadır. CC algoritmasının örüntü yoğunluğu yüksek alt kümeleri tanımlamadaki başarısı, suç türlerinin tematik olarak nasıl kümelendiği ve bu kümelerin hangi illerde yoğunlaştığı sorularına bütüncül yanıtlar sunmuştur.

Sonuç olarak, ikili kümeleme yöntemleri yalnızca istatistiksel yoğunlukları değil; aynı zamanda suçlar arası yapısal ilişkileri de açığa çıkararak, mekâna duyarlı güvenlik politikalarının geliştirilmesinde stratejik bir araç olarak değerlendirilebilir. Bu yönüyle çalışma, klasik yöntemlerin ötesine geçerek, suç analitiğine hem teorik hem de uygulamalı düzeyde katkı sunmaktadır.

3. YÖNTEM

3.1. Araştırmanın Yöntemsel Yaklaşımı

Bu araştırmanın temel amacı, Türkiye'deki suç türleri ile bu suçların işlendiği iller arasındaki örüntüleri eşzamanlı olarak analiz ederek, mekânsal ve tematik düzeyde anlamlı kümeler ortaya koymaktır. Bu hedef doğrultusunda, klasik kümeleme tekniklerinden farklı olarak, ikili kümeleme yaklaşımı tercih edilmiştir. İkili kümeleme, verinin hem satır (örneğin coğrafi birimler) hem de sütun (örneğin suç türleri) boyutundaki yapısını eşzamanlı olarak dikkate alarak, yalnızca belirli alt kümelerdeki örüntüleri yakalamayı amaçlayan bir veri madenciliği yöntemidir [6].

İkili kümeleme teknikleri başlangıçta gen ekspresyon analizlerinde kullanılmak üzere geliştirilmiş olsa da, son yıllarda sosyal bilimlerde, eğitim, pazarlama ve özellikle suç analitiği gibi çok boyutlu sosyal veri kümelerinin incelendiği alanlarda da etkili biçimde uygulanmaktadır [16,17]. Bu çalışmada, suç türleri ile iller arasındaki birlikte varyasyonları belirleyebilmek amacıyla Cheng & Church (2000) tarafından geliştirilen algoritmaya dayalı bir ikili kümeleme yöntemi uygulanmıştır. Bu yöntem, veri matrisindeki satır ve sütun alt kümeleri arasında örüntüsel olarak tutarlı bölgeleri ortaya çıkarmayı amaçlar ve özellikle yapısal benzerliklerin yüksek olduğu alt matrislerin tespiti için uygundur [5].

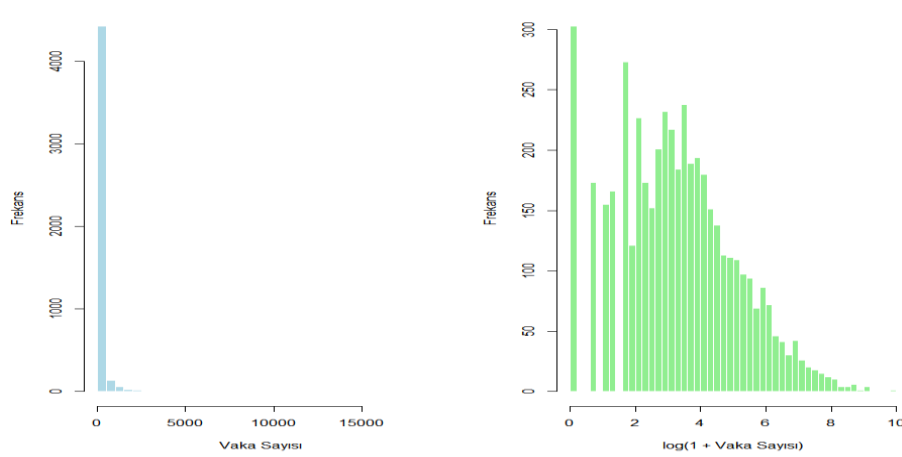
3.2. Veri Seti ve Ön İşleme Süreci

Çalışmada kullanılan veri seti, Adalet Bakanlığı Adli Sicil ve İstatistik Genel Müdürlüğü tarafından yayımlanan 2023 yılına ait resmi istatistiklere dayanmaktadır [9]. Veriler, Cumhuriyet Başsavcılıklarında kesinleşmiş suç kayıtlarını içermekte olup, 81 ili ve 58 farklı suç türünü kapsamaktadır. Bu yapı doğrultusunda oluşturulan veri matrisi, satırlarında suç türlerini, sütunlarında ise illeri içerecek şekilde 58 × 81 boyutundadır. Her bir hücre, belirli bir suç türünün ilgili ildeki vaka sayısını temsil etmektedir. Veri setine ilişkin genel bir bakış sunmak amacıyla, 81 il ve 58 suç türünü kapsayan toplam 4758 hücredeki vaka sayılarına ait tanımlayıcı istatistikler *Çizelge 1*'de verilmiştir.

Çizelge 1. Veri setine ait tanımlayıcı istatistikler

İstatistik	Değer
Ortalama	148,201
Standart Sapma	572,333
Minimum	0
1.Çeyrek (Q1)	6
Medyan (Q2)	23
3.Çeyrek (Q3)	80
Maksimum	18439
Toplam Vaka	696285

Çizelge 1'de görüldüğü gibi, vaka sayılarının ortalaması ile medyanı arasında büyük bir fark bulunmaktadır ve standart sapma değeri oldukça yüksektir. Bu durum, veri setinin sağa çarpık (pozitif çarpık) bir dağılıma sahip olduğunu ve bazı il/suç türü kombinasyonlarında aşırı yüksek vaka sayıları (uç değerler) bulunduğunu göstermektedir. Veri analizinden önce, özellikle farklı suç türleri arasındaki ölçek farklarını dengeleyebilmek amacıyla $\log(1+x)$ dönüşümü uygulanmıştır. Bu dönüşümün gerekliliği ve etkisi, Şekil 1'de sunulan dağılım grafikleri ile açıkça görülmektedir.

**Şekil 1.** Ham vaka sayısı verisinin histogramı ve $\log(1+x)$ dönüşümü sonrası verinin histogramı

Şekil 1'de ham verinin dağılımının, sıfıra yakın değerlerde yoğunlaştığı ve uzun bir sağ kuyruğa sahip olduğu görülmektedir. Bu çarpık yapı, analizde yanlılığa neden olabilir. Ayrıca $\log(1+x)$ dönüşümünün, dağılımı nasıl daha simetrik hale getirdiğini ve veriyi merkezileştirdiğini göstermektedir. Bu dönüşüm, yüksek uç değerlerin etkisini azaltarak, daha dengeli bir dağılım elde edilmesini sağlamış ve CC algoritmasının temel bileşeni olan Ortalama Kare Artığı (Mean Squared Residue – MSR) hesaplamalarının güvenilirliğini artırmıştır [5,6].

3.3. Kullanılan Algoritma ve Uygulama Süreci

Araştırmada tercih edilen Cheng & Church algoritması, ikili kümeleme yöntemleri arasında en yaygın kullanılan yaklaşımlardan biridir. Algoritmanın temel amacı, satır ve sütunlardan oluşan bir alt matrisin MSR değerini minimize ederek örüntüsel olarak homojen kümeler elde etmektir. MSR değeri, bir alt matrisin hücre değerlerinin satır ve sütun ortalamalarından sapma miktarını hesaplayan bir hata fonksiyonudur ve (3.1)'deki gibi hesaplanır

$$MSR(I, J) = \frac{1}{|I||J|} \sum_{i \in I, j \in J} (b_{ij} - b_{i.} - b_{.j} + b_{..})^2. \quad (3.1)$$

Burada, b_{ij} matrisin i . satır ve j . sütundaki gözlem değerini; $b_{i\cdot}$ satır ortalamasını; $b_{\cdot j}$ sütun ortalamasını; ve $b_{\cdot\cdot}$ ise tüm alt matrisin genel ortalamasını ifade etmektedir. $|I|$ ve $|J|$ sırası ile satır ve sütun toplam sayılarını belirtmektedir. MSR değerinin düşük olması, satır ve sütunlar arasında yüksek örüntüsel benzerliğin göstergesidir.

CC algoritması üç aşamalı iteratif bir süreçte çalışır: (1) yüksek varyans yaratan satır ve sütunlar, MSR'ye katkı düzeyine göre belirlenir ve eğer katkı belirli bir eşik değerini (δ) aşıyorsa kümelerden çıkarılır; (2) MSR değerini artırmayacak şekilde, uygun yeni satır ve sütunlar kümelere eklenerek alt matris genişletilir; (3) silme ve ekleme işlemleri belirli bir durağanlık sağlanana kadar tekrar edilir. Bu yapısı sayesinde CC algoritması, sabit bir küme sayısına ihtiyaç duymadan, veri yapısına dinamik biçimde uyum gösterebilir [5].

3.4. Yazılım ve Uygulama Ortamı

Çalışmada analizlerin gerçekleştirilmesi için açık kaynak kodlu R programlama dili kullanılmıştır. Özellikle “biclust”, “gplots”, “pheatmap” ve “stats” gibi paketlerden yararlanılarak veri görselleştirme ve analiz süreçleri yürütülmüştür. CC algoritmasının uygulanabilirliğini artırmak için, analiz öncesi veri matrisinin sayısal formatı standardize edilmiş ve parametrik ayarlamalar yapılmıştır. Uygulama sürecinde MSR için eşik değeri olarak $\delta = 0.05$ belirlenmiş, silme işlemlerinde $\alpha = 1$ olarak tanımlanmış ve çıkarılabilecek maksimum küme sayısı varsayılan değer 100 ile sınırlandırılmıştır. Elde edilen çıktılar hem sayısal hem de görsel olarak değerlendirilmiş, ısı haritaları üzerinden kümeler arasındaki yapısal benzerlikler analiz edilmiştir.

4. BULGULAR

Bu çalışmada kullanılan suç verileri, kişisel verilerin gizliliği ve bireysel hakların korunmasına özen gösterilerek, Cumhurbaşkanlığı İletişim Merkezi (CİMER) aracılığıyla resmi izin alınmak suretiyle temin edilmiştir. Veri seti, Adalet Bakanlığı Adli Sicil ve İstatistik Genel Müdürlüğü tarafından yayımlanan ve 2023 yılına ait olan, Cumhuriyet Başsavcılıkları tarafından kesinleşmiş suç kayıtlarını kapsamaktadır [9]. Analiz sürecinde, Türk Ceza Kanunu (TCK - 5237), 6136 Sayılı Ateşli Silahlar ve Bıçaklar ile Diğer Aletler Hakkında Kanun, ve 5607 Sayılı Kaçakçılıkla Mücadele Kanunu çerçevesinde tanımlanan suç türleri dikkate alınmıştır [18].

4.1. Veri Yapısı ve Değişkenler

Çalışmada kullanılan veri seti, toplam 696.285 suç kaydından oluşmaktadır. Veri seti, yalnızca suç türlerine ilişkin bilgileri değil; aynı zamanda suçluların uyruk, cinsiyet, yaş, medeni durum, eğitim düzeyi, meslek, il ve ilçe bilgilerini de içeren çok boyutlu bir yapıya sahiptir. Bu çok değişkenli yapı, yalnızca frekans bazlı değil, aynı zamanda örüntü temelli analizlerin de yapılmasına olanak tanımaktadır.

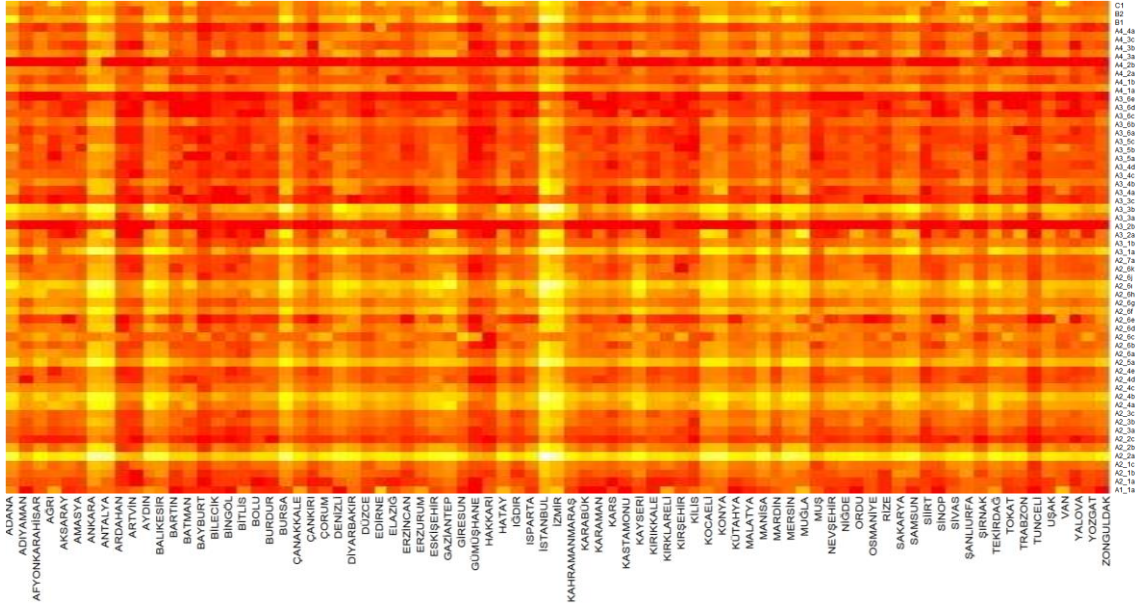
Suç türleri, içeriksel benzerliklerine göre 58 ana kategori altında sınıflandırılmış ve her biri sistematik olarak kodlanmıştır. Örneğin, “A2_2a” kodu kasten yaralama suçunu, “A2_5a” kodu ise hakaret suçunu temsil etmektedir. Söz konusu kodlamalar, analiz sürecinde algoritmalarla uyumlu bir biçimde çalışabilmesi amacıyla düzenlenmiştir. Tüm suç türlerine ilişkin kodlar ve açıklamaları *Çizelge 2*'de sunulmaktadır.

Çizelge 2. Suç türlerine ilişkin kodlamalar

A1_1a	Göçmen kaçakçılığı ve insan ticareti	A3_1b	Genel tehlike yaratan diğer suçlar
A2_1a	Kasten öldürmede nitelikli haller	A3_2a	İmar kirliliğine neden olma
A2_1b	Kasten öldürme	A3_2b	Çevreye karşı diğer suçlar
A2_1c	Taksirle öldürme	A3_3a	Uyuşturucu veya uyarıcı madde imal ve ticareti
A2_2a	Kasten yaralama	A3_3b	Kullanmak için uyuşturucu veya uyarıcı madde satın almak, kabul etmek veya bulundurmaya ya da uyuşturucu veya uyarıcı madde kullanmak
A2_2b	Taksirle yaralama	A3_3c	Kamunun sağlığına karşı diğer suçlar
A2_2c	Vücut dokunulmazlığına karşı diğer suçlar	A3_4a	Mühür bozma
A2_3a	Cinsel saldırı	A3_4b	Resmî belgede sahtecilik
A2_3b	Cinsel taciz	A3_4c	Parada sahtecilik
A2_3c	Cinsel dokunulmazlığa karşı diğer suçlar	A3_4d	Kamu güvenine karşı diğer suçlar
A2_4a	Konut dokunulmazlığının ihlali	A3_5a	Fuhuş
A2_4b	Tehdit	A3_5b	Kumar oynanması için yer ve imkân sağlama
A2_4c	Kişileri hürriyetinden yoksun kılma	A3_5c	Genel ahlaka karşı diğer suçlar
A2_4d	Kişilerin huzur ve sükununu bozma	A3_6a	Aile düzenine karşı suçlar
A2_4e	Hürriyete karşı diğer suçlar	A3_6b	Bilişim alanında suçlar
A2_5a	Hakaret	A3_6c	Ekonomi, sanayi ve ticarete ilişkin suçlar
A2_6a	Dolandırıcılık	A3_6d	Kamu barışına karşı suçlar
A2_6b	Güveni kötüye kullanma	A3_6e	Ulaşım araçlarına veya sabit platformlara karşı suçlar
A2_6c	Hakkı olmayan yere tecavüz	A4_1a	Görevini yaptırmamak için direnme
A2_6d	Hırsızlık	A4_1b	Kamu idaresinin güvenilirliğine ve işleyişine karşı diğer suçlar
A2_6e	Karşılıksız yararlanma	A4_2a	Silahlı örgüt
A2_6f	Mala zarar verme	A4_2b	Anayasal düzene ve bu düzenin işleyişine karşı diğer suçlar
A2_6g	Mala zarar vermenin nitelikli halleri	A4_3a	Hükümlü veya tutuklunun kaçması
A2_6h	Nitelikli dolandırıcılık	A4_3b	İnfaz kurumuna veya tutukevine yasak eşya sokmak
A2_6i	Nitelikli hırsızlık	A4_3c	Adliye karşı diğer suçlar
A2_6j	Nitelikli yağma	A4_4a	Topluma karşı diğer suçlar
A2_6k	Malvarlığına karşı diğer suçlar	B1	Mermileri satın alma, taşıma veya bulundurmaya ilişkin suçlar
A2_7a	Kişilere karşı diğer suçlar	B2	6136 sayılı kanuna muhalefet diğer suçlar
A3_1a	Trafik güvenliğini tehlikeye sokma	C1	Kaçakçılık suçları

İllere göre gerçekleştirilen analizde kullanılan veri matrisi, satırlarda 58 farklı suç türünü, sütunlarda ise Türkiye'nin 81 ilini içerecek şekilde yapılandırılmıştır. Matrisin her bir hücresi, belirli bir suç türünün ilgili ilde kaç kez işlendiğini niceliksel olarak göstermektedir. Bu yapı sayesinde hem yatay (suç türleri arası) hem dikey (iller arası) karşılaştırmalar yapılabilmekte ve örüntülerin eşzamanlı analizi mümkün olmaktadır.

Elde edilen 58×81 boyutundaki bu veri matrisi, suçların coğrafi dağılımını ve yoğunluklarını görselleştirmek amacıyla bir ısı haritası (heatmap) biçiminde sunulmuştur. Suç türlerine göre illerdeki frekans dağılımını yansıtan bu harita, örüntülerin ilk görsel ipuçlarını sunmakta olup, Şekil 2’de görselleştirilmiştir.



Şekil 2. Suç türlerine göre illerin frekans dağılımını gösteren ısı haritası

4.2. CC Algoritması Uygulaması

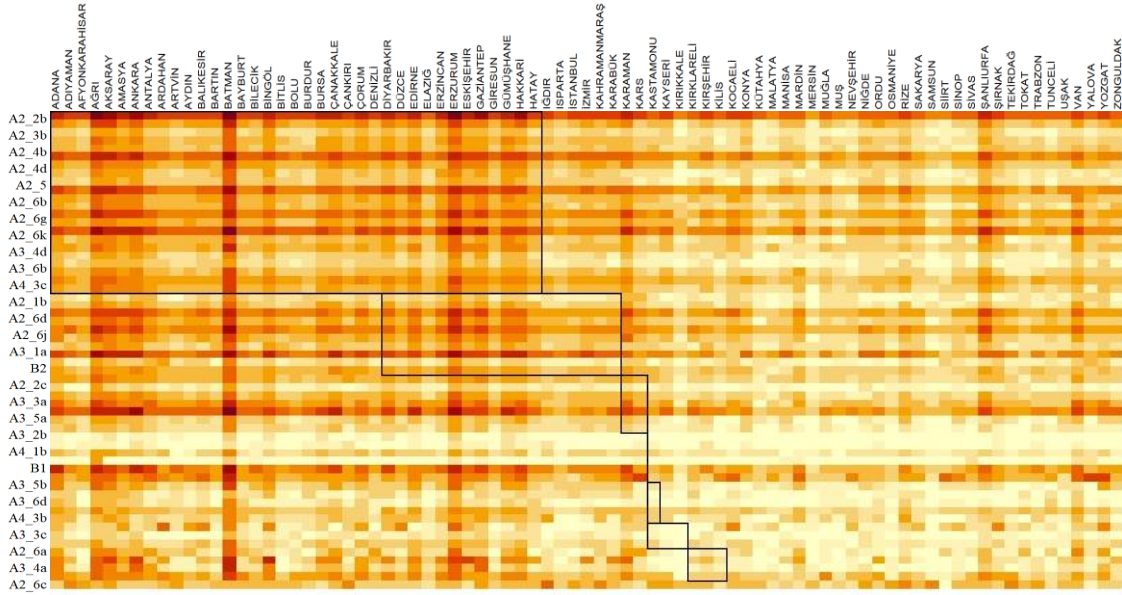
Bu çalışmada, suç türlerinin satırlarda ve illerin sütunlarda yer aldığı veri matrisine Cheng & Church (CC) algoritması uygulanmıştır. Uygulama sırasında, algoritmanın temel parametrelerinden biri olan ve küme içi varyans için tolerans sınırını belirleyen delta (δ) değeri 0,05 olarak atanmıştır. Satır ve sütun silme işlemlerinde kullanılan ölçek faktörü α , literatüre uygun şekilde 1 olarak sabitlenmiş; çıkarılabilecek maksimum ikili küme sayısı ise 100 ile sınırlandırılmıştır.

Belirlenen bu parametrelerle gerçekleştirilen analiz sonucunda, veri matrisinden toplam 7 anlamlı ikili küme elde edilmiştir. Bu kümeler, belirli suç türlerinin benzer yoğunlukta ortaya çıktığı illeri eşzamanlı olarak gruplamakta ve suç örüntülerinin hem mekânsal hem de tematik boyutlarını açığa çıkarmaktadır. Her bir kümenin kapsadığı suç türü (satır) ve il (sütun) sayıları Çizelge 3’de sunulmaktadır. En büyük küme olan BC1, 22 suç türünü ve 37 ili kapsarken; en küçük kümelerden biri olan BC6, yalnızca 3 suç türü ve 10 ili içermektedir.

Çizelge 3. CC algoritması ile elde edilen ikili kümelerin boyutları

	BC1	BC2	BC3	BC4	BC5	BC6	BC7
Satır	22	10	7	6	5	3	4
Sütun	37	18	16	9	8	10	8

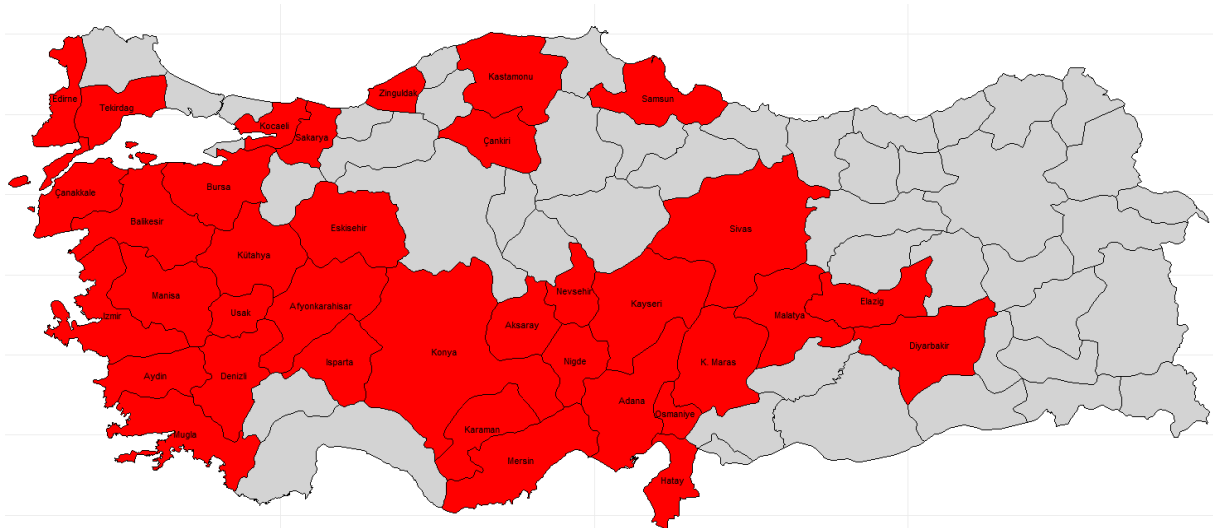
Bu yedi kümeye ilişkin genel yapılar ve yoğunluk desenleri Şekil 3’de ısı haritası biçiminde görselleştirilmiştir. Haritada, özellikle bazı illerin birden fazla kümeye dâhil olduğu dikkat çekmektedir. Bu durum, belirli illerin farklı suç gruplarında da benzer örüntüler sergilediğini ve suç profilinin çok boyutlu bir karakter taşıdığını göstermektedir. Yatay ekseninde yer alan suç türlerinde belirgin renk yoğunlukları, bazı suçların çok sayıda ilde benzer sıklıkla işlendiğine işaret etmektedir.



Şekil 3. CC algoritması ile elde edilen ikili kümelerin ısı haritası

Detaylı inceleme sonucunda, CC algoritmasının ürettiği ilk beş küme, Türkiye genelindeki baskın suç türleri ile bu suçların yoğunlaştığı iller arasında belirgin örüntüler ortaya koymuştur:

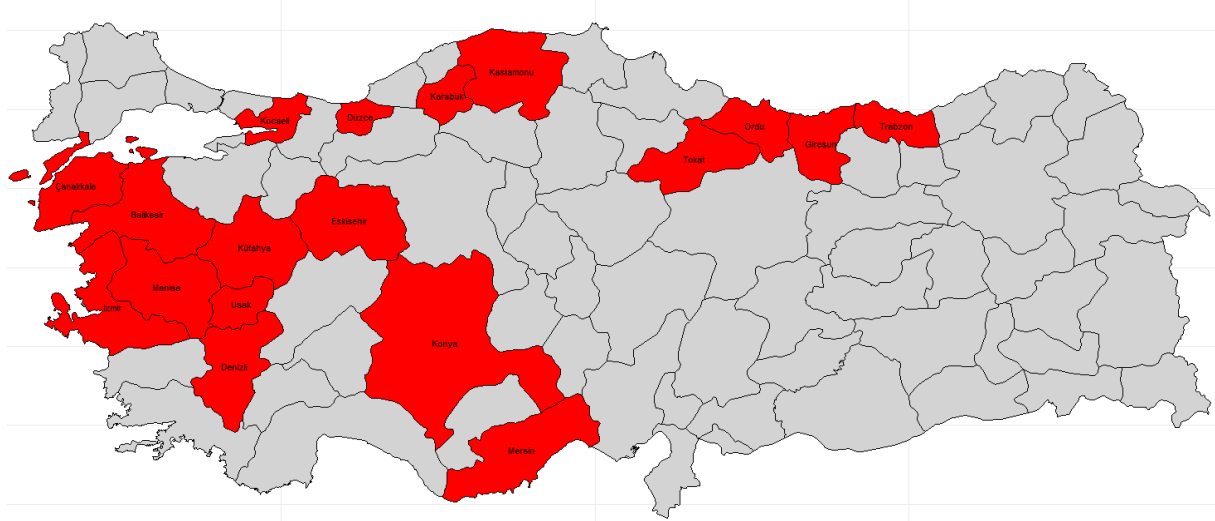
BC1 kümesinde; kasten yaralama (A2_2a), tehdit (A2_4b), hakaret (A2_5a), mala zarar verme (A2_6f) ve nitelikli hırsızlık (A2_6i) suç türlerinin belirgin biçimde öne çıktığı görülmektedir. Bu suçlar, özellikle büyükşehir statüsündeki yüksek nüfuslu illerde (başta İstanbul, Ankara, İzmir, Adana ve Antalya olmak üzere toplam 37 ilde) birlikte yoğunlaşmakta ve sosyal çatışmalar ile bireysel hak ihlallerine ilişkin örüntüleri yansıtmaktadır. Söz konusu illerin coğrafi dağılımı Şekil 4'te görselleştirilmiştir.



Şekil 4. BC1 kümesine dâhil olan 37 ilin Türkiye haritası üzerinde gösterimi

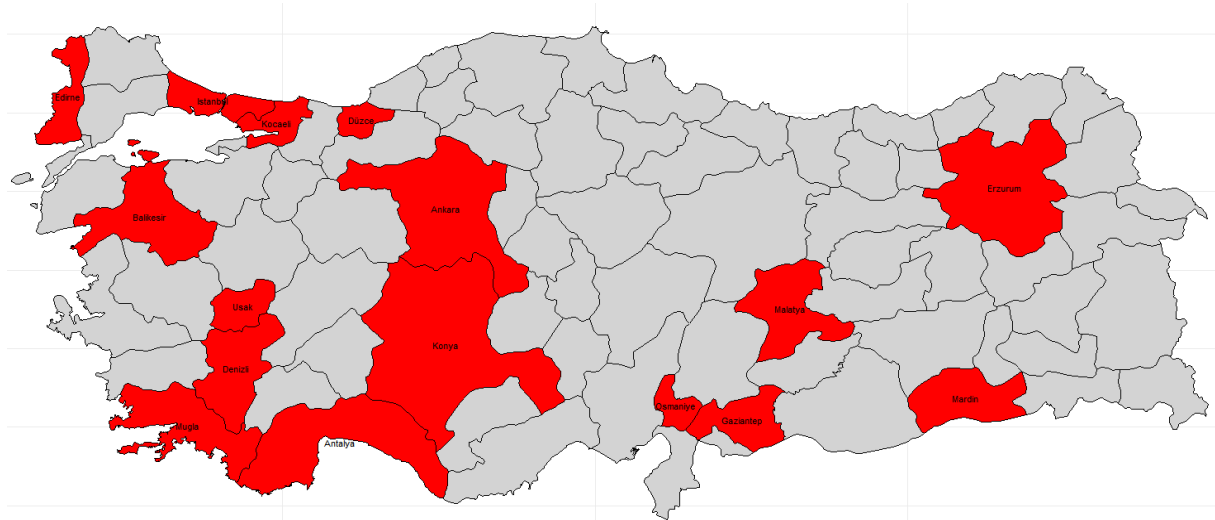
BC2 kümesi; trafik güvenliğini tehlikeye sokma (A3_1a), konut dokunulmazlığının ihlali (A2_4a) ve nitelikli dolandırıcılık (A2_6h) gibi suç türlerinin öne çıktığı bir örüntüyü temsil etmektedir. Bu suçların, özellikle İzmir, Eskişehir, Konya ve Manisa gibi kentlerde benzer düzeylerde görülmesi, kentsel yaşamın getirdiği yapısal ve sosyal risklerin ortaklığına işaret etmektedir.

Söz konusu illerdeki suç dağılımları, BC2 kümesinde yer alan diğer şehirlerle birlikte değerlendirildiğinde, bu kümelenmenin şehirleşme, göç dinamikleri, ekonomik ilişkiler ve yaşam tarzına bağlı olarak şekillenen toplumsal risk faktörlerini yansıttığı görülmektedir. Küme kapsamında yer alan iller ve bu illerin Türkiye haritasındaki coğrafi dağılımları Şekil 5’te sunulmuştur.



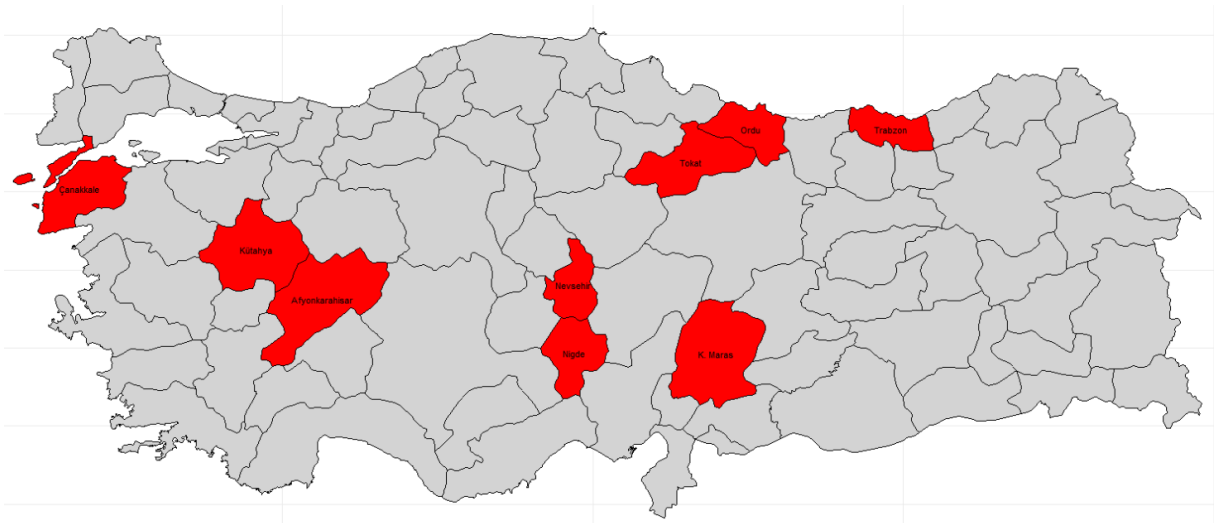
Şekil 5. BC2 kümesine dâhil olan 18 ilin Türkiye haritası üzerindeki gösterimi

BC3 kümesinde; uyuşturucu madde bulundurma (A3_3b) ve uyuşturucu madde ticareti (A3_3a) suç türlerinin yüksek düzeyde temsil edildiği gözlenmiştir. Bu suçların özellikle İstanbul, Ankara, Antalya ve Gaziantep gibi büyükşehirlerde yoğunlaşması, söz konusu kentlerin hem uyuşturucu madde tüketimi hem de dağıtımı açısından merkezî roller üstlendiğini göstermektedir. BC3 kümesindeki örüntü, büyükşehirlerin sunduğu geniş nüfus yapısı, sosyoekonomik çeşitlilik ve ulaşım ağlarının, bu tür suçların yayılımını kolaylaştıran dinamikler olarak işlev gördüğünü ortaya koymaktadır. Küme kapsamında yer alan illerin coğrafi dağılımı Şekil 6’te sunulmaktadır.



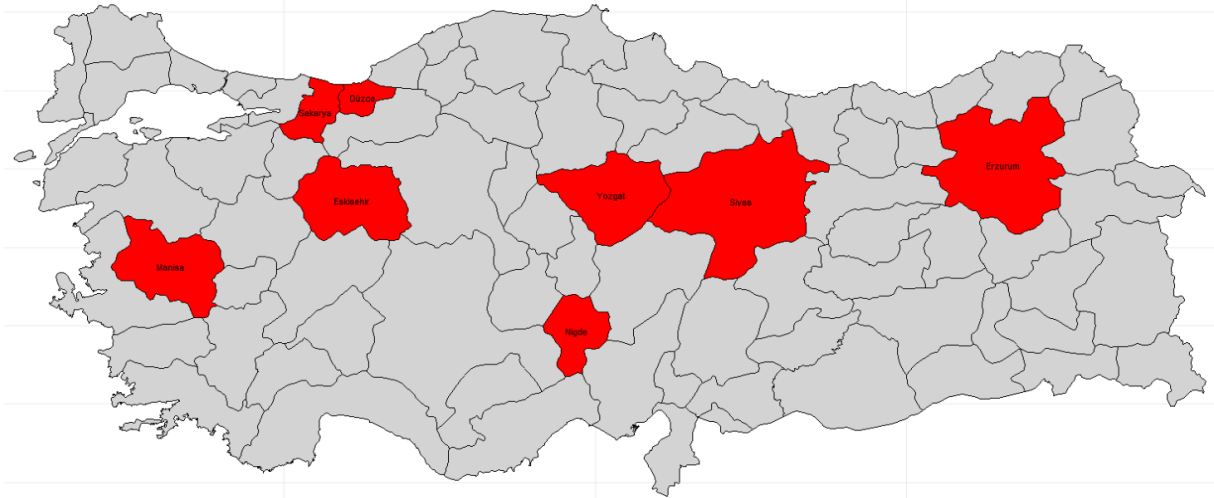
Şekil 6. BC3 kümesine dâhil olan 16 ilin Türkiye haritası üzerindeki dağılımı

BC4 kümesinde; mermi bulundurma (B1) ve kaçakçılık (C1) suç türlerinin, özellikle Kütahya, Trabzon, Ordu ve Kahramanmaraş gibi illerde yoğunlaştığı görülmektedir. Bu bulgular, ateşli silahlar ve yasa dışı ticaretle ilişkili suçların, bölgesel dinamikler doğrultusunda belirli illerde kümelenme eğilimi gösterdiğini ortaya koymaktadır. Söz konusu illerde bu tür suçların öne çıkması, sınır hattı olmayan bölgelerde dahi silah edinimi ve yasa dışı ürün dolaşımının yapısal, ekonomik ya da kültürel nedenlerle gerçekleştiğine işaret edebilir. BC4 kümesine dâhil olan illerin coğrafi dağılımı Şekil 7’de sunulmuştur.



Şekil 7. BC4 kümesine ait 9 ilin Türkiye haritası üzerindeki gösterimi

BC5 kümesinde; silahlı örgüt üyeliği (A4_2a), tutukevine yasak eşya sokma (A4_3b) ve kumar oynanması için yer temini (A3_5b) gibi suç türlerinin özellikle Manisa ve Eskişehir gibi illerde yoğunlaştığı gözlemlenmiştir. Bu suç örüntüsü, söz konusu kentlerde örgütlü suç faaliyetlerinin daha belirgin hâle geldiğini ve belirli şehirlerde güvenlik açısından yapısal risklerin bulunduğunu düşündürmektedir. Bu bulgu, suçun yalnızca bireysel davranışlarla sınırlı kalmadığını; yerleşik ve organize yapılar aracılığıyla da yayıldığını ortaya koymaktadır. BC5 kümesine ait illerin coğrafi dağılımı Şekil 8’de görselleştirilmiştir.



Şekil 8. BC5 kümesine dâhil olan 8 ilin Türkiye haritası üzerindeki gösterimi

Çalışma kapsamında toplam yedi ikili küme elde edilmiş olmakla birlikte, temsil gücü yüksek ve yorumlanabilirliği daha güçlü olan ilk beş kümeye odaklanılarak bu kümelerin mekânsal dağılımları harita üzerinde görselleştirilmiştir. Diğer iki küme ise niceliksel olarak daha sınırlı sayıda il ve suç türü içermekte olup, tematik yoğunluk açısından da görece daha düşük bir örüntü sunmaktadır. Bu nedenle, çalışmanın bütünlüğünü korumak ve harita üzerindeki görsel yükü azaltmak amacıyla yalnızca ilk beş küme haritalandırılmıştır. Ancak bu iki küme de analiz bütünlüğü içerisinde değerlendirilmiş ve genel yorumlamalara katkı sunmuştur. Bu tercih, çalışmanın hem görsel sunum hem de analitik açıklık açısından optimal düzeyde yapılandırılmasını sağlamıştır.

Elde edilen bulgular, yalnızca istatistiksel yoğunlukların değil, aynı zamanda farklı suç türlerinin birlikte ortaya çıkma eğilimlerinin de anlamlı biçimde çözümlenebildiğini göstermektedir. Bu bağlamda, ikili kümeleme yaklaşımının sunduğu eşzamanlı analiz gücü, hem mekânsal (iller) hem de tematik (suç türleri) boyutlarda örüntüleri ortaya çıkarmada etkili bir araç olarak öne çıkmaktadır. Kullanılan CC algoritması sayesinde suç türleri ve iller arasındaki çift yönlü örüntüler tanımlanabilmiş, böylece suçların sadece bireysel değil aynı zamanda yapısal ve bölgesel düzeyde nasıl kümelendiği daha açık biçimde anlaşılmıştır.

5. SONUÇ VE ÖNERİLER

Bu çalışma, Türkiye'nin 81 ili ve 58 farklı suç türünden oluşan bir veri matrisi üzerinde CC algoritmasına dayalı ikili kümeleme yöntemi uygulanarak, suçların mekânsal ve tematik düzeydeki örüntülerini ortaya koymayı amaçlamıştır. Klasik analiz tekniklerinin çoğunlukla suçun sadece toplam sayıları veya oranları üzerinden değerlendirme yaptığı mevcut literatürden farklı olarak, bu çalışma hem iller hem de suç türleri arasında eşzamanlı analiz yapılmasını sağlayarak daha rafine bir suç profili sunmuştur. İkili kümeleme yaklaşımı sayesinde, belirli suç türlerinin birlikte ortaya çıktığı ve bu suçların benzer yoğunluklarla işlendiği iller anlamlı biçimde gruplandırılmıştır.

Uygulanan CC algoritması, yüksek boyutlu ve karmaşık veri yapıları içerisinde tematik olarak ilişkili olan suç türlerini ve bu türlerin coğrafi olarak kümelendiği illeri aynı anda analiz etme imkânı tanımıştır. Elde edilen yedi ikili küme sonucundan beşi, gerek içerdiği veri yoğunluğu gerekse temsil ettikleri örüntü gücü bakımından öne çıktığından, bu kümeler detaylı olarak yorumlanmış ve harita üzerinde görselleştirilmiştir. Kalan iki küme ise, sınırlı sayıda ili ve/veya suç türünü içermesi nedeniyle haritalandırmaya dahil edilmemiş, ancak analizin bütünlüğü içinde değerlendirilmiştir. Bu tercih, çalışmanın sunum ve yorumlanabilirlik niteliğini güçlendirme amacı taşımaktadır.

Elde edilen bulgular, bazı suç türlerinin belirli bölgelerde birlikte yoğunlaştığını ve bazı illerin özgül suç profilleri açısından diğerlerinden sistematik olarak ayrıştığını göstermektedir. Örneğin, metropol illerde bireysel şiddet, hırsızlık ve tehdit suçları bir arada yoğunlaşırken; bazı kırsal veya orta ölçekli illerde silah kaçakçılığı, yasak eşya sokma ve örgütlü suç yapıları daha belirgin biçimde ortaya çıkmıştır. Bu durum, suçun yalnızca bireysel düzeydeki eğilimlerle değil; aynı zamanda sosyal, ekonomik ve coğrafi dinamiklerle şekillendiğini ortaya koymaktadır. Dolayısıyla, suç önleme ve güvenlik politikalarının tekil ve genelleyici yaklaşımlardan ziyade, yerel düzeyde farklılaşan örüntüleri dikkate alarak yapılandırılması gerekmektedir.

Çalışmanın yöntemsel katkılarından biri de, ikili kümeleme tekniklerinin suç analizi bağlamında işlevsel bir araç olarak kullanılabileceğini göstermesidir. Bu yöntem, yalnızca yüksek frekanslı verileri tanımlamakla kalmayıp; suç türlerinin birlikte ortaya çıkma eğilimlerini ve bu eğilimlerin mekânsal bağlamlarını eşzamanlı olarak analiz edebilmektedir. Böylece, hem suç türleri arası tematik yakınlık hem de bu suçların iller düzeyindeki coğrafi yayılımı bütüncül biçimde değerlendirilebilmektedir.

Gelecekte gerçekleştirilecek araştırmalarda, suç verilerinin sosyo-ekonomik, demografik, eğitim düzeyi, işsizlik oranı, göç profili ve kültürel faktörler gibi değişkenlerle entegrasyonu, suç örüntülerinin nedensel açıklamalarını daha güçlü biçimde yapmaya olanak sağlayacaktır. Ayrıca, ikili kümeleme kapsamında farklı algoritmaların (örneğin Bimax, Plaid, xMotif) karşılaştırmalı olarak uygulanması, yöntemsel duyarlılığı artırarak analizlerin güvenilirliğini destekleyebilir. Zaman serisine dayalı suç verilerinin kullanılmasıyla zamana duyarlı ikili kümeleme modellerinin geliştirilmesi de dinamik örüntülerin izlenmesine katkı sağlayacaktır.

Sonuç olarak, bu çalışma veri madenciliği tekniklerinin suç analizi bağlamında uygulanabilirliğini ortaya koyarak, akademik literatüre özgün bir katkı sunmakta; aynı zamanda kamu güvenliği, suç önleme stratejileri ve yerel politika geliştirme açısından pratik yönelimlere dayanak oluşturmaktadır. İkili kümeleme yaklaşımının sunduğu bütüncül analiz çerçevesi, suçla mücadelede daha hedefe yönelik ve etkili müdahale alanlarının belirlenmesine olanak tanımaktadır.

TEŞEKKÜR

Bu çalışma, Demet Albasar tarafından Gazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik Ana Bilim Dalı'nda hazırlanan "Suç verilerinin ikili kümeleme yöntemleri ile analizi" (2025) başlıklı yüksek lisans tezinden türetilmiştir.

YAZAR(LAR)'IN KATKISI

Demet Albasar: Araştırma, Materyal temini, İçerik analizi, Makalenin yazımı- İnceleme ve Düzenleme.
Ahmet Kocatürk: Araştırma, Materyal temini, İçerik analizi, Makalenin yazımı- İnceleme ve Düzenleme.
Hacı Hasan Örkücü: Kavramlaştırma, Metodoloji, Yazılım, Makalenin yazımı- İnceleme ve Düzenleme, Danışman/Kontrolörlük

ÇIKAR ÇATIŞMASI/ÇAKIŞMASI BİLDİRİMİ

Yazarlar arasında çıkar çatışması/çakışması bulunmamaktadır.

KAYNAKLAR

- [1] Brown, D. E. (1998). The Regional Crime Analysis Program (RECAP): A framework for mining data to catch criminals. IEEE International Conference on Systems. *Man and Cybernetics*, 2848–2853.
- [2] Orakcı, M., Ciyilan, B., Kök, İ., & Sevre, M. (2016). Suç analizinde veri madenciliği teknikleri ve makine öğrenmesi algoritmalarının kullanılması. *18. Akademik Bilişim Konferansı*, Adnan Menderes Üniversitesi, Aydın.
- [3] Adderley, R., & Musgrove, P. B. (2001). Police crime recording and investigation systems: A user's view. *Policing: An International Journal of Police Strategies and Management*, 24(1), 100–114.
- [4] Nath, S. V. (2006). Crime pattern detection using data mining. *IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology – Workshops*, 41–44.
- [5] Cheng, Y., & Church, G. M. (2000). Biclustering of expression data. Bildiri, *International Conference on Intelligent Systems for Molecular Biology*, UC San Diego, California, USA.
- [6] Madeira, S. C., & Oliveira, A. L. (2004). Biclustering algorithms for biological data analysis: A survey. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 1(1), 24–45.
- [7] Marin, J., Guerreros, L., & Calderon, D. (2024). Crime trend analysis using big data and clustering methods. *International Journal of Data Science and Analytics*, 12(1), 33–47.
- [8] Arslan, R., Örkücü, H. H., & Altunkaynak, B. (2018). Clustering of criminals according to the types of material caught in smuggling: Biclustering method. *International Journal of Economics and Administrative Studies*, 883–896.
- [9] Adalet Bakanlığı Adli Sicil ve İstatistik Genel Müdürlüğü. (2023). *Adli İstatistikler 2023*. Adalet Bakanlığı Yayınları.
- [10] Çil, İ., Çakar, S. G., Sarı, N., & Eydemir, O. (2019). Detection of crime regions with biclustering approach and comparison of methods. *Sakarya University Journal of Computer and Information Sciences*, 2(3), 144–157.
- [11] Chen, H., Chung, W., Qin, Y., Chau, M., Xu, J. J., Wang, G., Zheng, R., & Atabakhsh, H. (2003). Crime data mining: An overview and case studies. *ACM International Conference Proceeding Series*, 130, 1–9.
- [12] Nath, S. V. (2006). Crime pattern detection using data mining. *IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology – Workshops*, 41–44.
- [13] Bruin, J. S., Cox, T. K., Kusters, W. A., Laros, J., & Kok, J. N. (2006). Data mining approaches to criminal career analysis. *Proceedings of the Sixth International Conference on Data Mining (ICDM'06)*, 171–177.
- [14] Oğuzlar, A. (2005). Kümeleme analizinde yeni bir yaklaşım: Kendini düzenleyen haritalar (Kohonen ağları). *Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 19(2), 1–12.
- [15] Sevrüktekin, M., Oğuzlar, A., Aydın, B., & Nargeleşkenler, M. (2007). Karar ağacı yardımıyla suçluların özelliklerinin belirlenmesi. *Öneri Dergisi*, 7(27), 291–298.
- [16] Prelic, A., Bleuler, S., Zimmermann, P., Wille, A., Bühlmann, P., Gruissem, W., Henning, L., Thiele, L., & Zitzler, E. (2006). A systematic comparison and evaluation of biclustering methods for gene expression data. *Bioinformatics*, 22(9), 1122–1129.
- [17] Pontes, B., Giráldez, R., & Aguilar-Ruiz, J. S. (2015). Quality measures for gene expression biclusters. *Plos One*, 10(3), e0115497.
- [18] Emniyet Genel Müdürlüğü Narkotik Suçlarla Mücadele Daire Başkanlığı. (2023). *2023 Türkiye Uyuşturucu Raporu*. <https://www.narkotik.pol.tr/kurumlar/narkotik.pol.tr/TUB%C4%B0M/Ulusal%20Yay%C4%B1nlar/NARKOLOG-2023-PROFIL-ANALIZI.pdf>