

# Güncel Ekonometrik ve İstatistiksel Uygulamalar ile Akademik Çalışmalar

Editörler: Doç. Dr. Elif Tuna • Doç. Dr. Serkan Akoğul



# Güncel Ekonometrik ve İstatistiksel Uygulamalar ile Akademik Çalışmalar

**Editörler:**

Doç. Dr. Elif TUNA

Doç. Dr. Serkan AKOĞUL



Published by

**Özgür Yayın-Dağıtım Co. Ltd.**

Certificate Number: 45503

📍 15 Temmuz Mah. 148136. Sk. No: 9 Şehitkamil/Gaziantep

☎ +90.850 260 09 97

📞 +90.532 289 82 15

🌐 www.ozgurayinlari.com

✉ info@ozgurayinlari.com

---

## Güncel Ekonometrik ve İstatistiksel Uygulamalar ile Akademik Çalışmalar

*Academic Studies with Current Econometric and Statistical Applications*

Editörler: Doç. Dr. Elif TUNA • Doç. Dr. Serkan AKOĞUL

---

Language: Turkish-English

Publication Date: 2024

Cover design by Mehmet Çakır

Cover design and image licensed under CC BY-NC 4.0

Print and digital versions typeset by Çizgi Medya Co. Ltd.

**ISBN (PDF):** 978-975-447-951-5

**DOI:** <https://doi.org/10.58830/ozgur.pub518>

---



This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/>  
This license allows for copying any part of the work for personal use, not commercial use, providing author attribution is clearly stated.

---

Suggested citation:

Tuna, E. (ed), Akoğul, S. (ed) (2024). *Güncel Ekonometrik ve İstatistiksel Uygulamalar ile Akademik Çalışmalar*.

Özgür Publications. DOI: <https://doi.org/10.58830/ozgur.pub518>. License: CC-BY-NC 4.0

---

*The full text of this book has been peer-reviewed to ensure high academic standards. For full review policies, see <https://www.ozgurayinlari.com/>*

---



## Ön Söz

Günümüz dünyası, ekonomik belirsizlikler, sosyal değişimler ve teknolojik gelişmelerle şekillenmektedir. Bu dinamik ortamda, doğru ve güncel verilerin analizi, yalnızca akademik bir gereklilik değil, aynı zamanda toplumsal ve ekonomik karar alma süreçlerinin de temelini oluşturmaktadır. “Güncel Ekonometrik ve İstatistiksel Uygulamalar ile Akademik Çalışmalar” isimli bu eser, bu ihtiyaçları karşılamak amacıyla hazırlanmıştır.

Kitap, farklı disiplinlerden gelen akademisyenlerin ve araştırmacıların katkılarıyla oluşturulmuş olup, ekonometri ve istatistik alanındaki gelişmeleri ve uygulamaları ele almaktadır. İçerisinde yer alan bölümler hem teorik altyapıyı hem de pratik uygulamaları bir araya getirerek okuyuculara geniş bir perspektif sunmayı hedeflemektedir. Bu bölümleri iki ana başlık olan ekonometrik ve istatistiksel uygulamalar olarak gruplayabiliriz. İlk grup, ekonomi, sosyal yapı ve istatistiksel analizlerle ilgili konuları içeren ekonomik değişkenler, sosyal durum analizleri ve makroekonomik etkiler üzerine odaklanmaktadır. İkinci grupta ise istatistiksel yöntemlerle veri analizi, makine öğrenimi, görüntü işleme gibi konuları içeren yöntemsel ve teknolojik yaklaşımlar ön plandadır. Ayrıca, sağlık, eğitim ve sosyal durum analizi gibi çeşitli konulara odaklanan araştırmalar disiplinler arası bir yaklaşım sergilemektedir.

Akademik çalışmaların yanı sıra, bu kitap aynı zamanda uygulamalı istatistik ve ekonometrik modelleme konularında çalışan araştırmacılar için bir başvuru kaynağı olmayı amaçlamaktadır. Gelişen veri analizi teknikleri ve makine öğrenimi yöntemleri, günümüz araştırmalarında vazgeçilmez araçlar haline gelmiştir. Bu bağlamda, istatistiksel ve ekonometrik modelleme, veri madenciliği ve görüntü işleme gibi konular, kitabın önemli bölümlerini oluşturmaktadır.

Sonuç olarak, “Güncel Ekonometrik ve İstatistiksel Uygulamalar ile Akademik Çalışmalar” adlı bu eser, akademisyenler, öğrenciler ve araştırmacılar için değerli bir kaynak olmayı hedeflemekte ve güncel ekonomik ve sosyal sorunlara ışık tutmaktadır.

Bu kitabın hazırlanmasında katkı sunan tüm yazarlara ve yayınlanmasında emeği geçen herkese içten teşekkürlerimizi sunarız.

**Editörler**

Doç. Dr. Elif TUNA  
Doç. Dr. Serkan AKOĞUL



# İçindekiler

Ön Söz

iii

## Bölüm 1

---

Tüketici Güven Endeksi ile Makroekonomik Değişkenler Arasındaki İlişkinin  
Ekonometrik Analizi 1

*Eda Yalçın Kayacan*

*Tuncay Morah*

*Serkan Akoğul*

## Bölüm 2

---

Cinsiyete Göre Yaşam Beklentisinin Şekillenmesinde Etkili Faktörler: OECD  
Daha İyi Yaşam İndeksi Analizi 19

*Mert Veznikli*

## Bölüm 3

---

Avrupa Birliği Ülkelerinde Temel Gelir Sorgulamasında Çoklu Uyum İyiliği  
Analizi Kullanımı 49

*Zehra Zeynep Şahinbaşoğlu*

## Bölüm 4

---

Gerçek Yaşam Temelli Matematik Etkinliklerine Katılan Dezavantajlı Ortaokul  
Öğrencilerin Matematiğe Yönelik Bakış Açılarının İstatistiksel Olarak  
İncelenmesi 69

*Elif Bahadır*

*Şervval Gökçen*

*Elif Tuna*

*Mehmet Samil Güneş*

## Bölüm 5

---

Faktör Analizi Çözümlemesine Ortak Entropi Yaklaşımı 89

*Nida Oruç*

*Doğan Yıldız*

## Bölüm 6

---

Banka Sektöründe Otomatik Makine Öğrenme Yöntemi (DataBlender) ile  
Vadeli Mevduat Talep Tahmini 109

*Gizem Aydın*

*Mehmet Yalçın*

## Bölüm 7

---

Türkiye’de Teknolojik Ürünlerin Hedef Pazarlarının İstatistiksel Yöntemlerle  
Değerlendirilmesi 127

*Aylin Uçan*

*Nuran Çakır Yıldız*

## Bölüm 8

---

Calculation of Household-Based Socioeconomic Status in Türkiye 161

*İbrahim Demir*

*Furkan Metin*

*Metin Aytaç*

*Turgay Altun*

*Mehmet Şaban Ucarı*

*Onur Şentürk*

## Bölüm 9

---

Producing CPI During Covid-19 Period in Türkiye 175

*Furkan Metin*

*İbrahim Demir*

*Cem Baş*

## Bölüm 10

---

Use of Relative Entropy Statistics in Contingency Tables 187

*Atıf Evren*

*Erhan Ustaoglu*

*Elif Tuna*

*Büşra Şahin*

## Bölüm 11

---

OpenCV and DeepFace Approach on Image Processing: Applications on Street Interviews 201

*Muzaffer Göztaş*

*Doğan Yıldız*

## Bölüm 12

---

Bi-Rads Classification in Mammography using Deep Learning 229

*Arda Tekin*

*Boran Toktay*

*Ahmet Can Günay*

*Harun Yazgan*

*Neslihan G. İnan*

*Ozan Kocadağlı*

## Bölüm 13

---

LSTM-Based Multivariate Deep Neural Networks for Stock Price Forecasting 241

*Selim Serin*

*Güldeir Kemalbay*

## Bölüm 14

---

Clustering Analysis of European Health Status Pre-Covid-19 261

*Ümmü Şahin Şener*

*Ersin Şener*





# Tüketici Güven Endeksi ile Makroekonomik Değişkenler Arasındaki İlişkinin Ekonometrik Analizi<sup>1</sup>

Eda Yalçın Kayacan<sup>2</sup>

Tuncay Moralı<sup>3</sup>

Serkan Akoğul<sup>4</sup>

## Özet

Türkiye İstatistik Kurumu (TÜİK) tarafından her ay açıklanan Tüketici Güven Endeksi (TGE), tüketicilerin bireysel finansal durumlarına ve gelecek beklentilerine odaklanarak harcama ve tasarruf eğilimlerini ölçmeyi amaçlamaktadır. Bu çalışmanın temel motivasyonu, tüketiciler için belirleyici olduğu düşünülen kredi faiz oranı, enflasyon, dolar kuru değişkenleri ile tüketici güven endeksi arasındaki ilişkiyi ortaya koymak ve belirleyici değişkenlerin tüketici güvenini nasıl etkilediğini incelemektir.

Tüketici güven endeksi ile tüketici kredileri, konut kredileri, taşıt kredileri, enflasyon ve döviz kurları arasında uzun dönemli anlamlı bir ilişkinin test edilmesinin amaçlandığı çalışmada Mayıs 2012-Eylül 2023 dönemlerine ait aylık veriler kullanılmıştır. Bu amaçla değişkenler arasındaki ilişki yapısal kırılmayı dikkate alan Gregory-Hansen Eşbütünleşme Testi ile test edilmiştir. Anlamlı ilişkilerin tespit edilmesinin ardından değişkenlerin uzun dönem katsayıları Tam Değiştirilmiş En Küçük Kareler (FMOLS), Dinamik En Küçük Kareler (DOLS) ve Kanonik Eşbütünleşme Regresyonu (CCR) yöntemleri kullanılarak tahmin edilmiştir. Ayrıca tüketici güven endeksi ve alt endeksleri ile tüketici kredileri, konut kredileri, taşıt kredileri, enflasyon ve döviz kurları arasındaki ilişkiler korelasyon analizi ile incelenmiş, göstergeler arasındaki ilişkiler hem istatistiksel hem de ekonometrik olarak detaylı bir şekilde değerlendirilmiştir.

- 1 Bu çalışmada ifade edilen bulgu ve görüşler tamamen yazarlara ait olup, Türkiye İstatistik Kurumu ile ilişkilendirilemez.
- 2 Doç. Dr., Pamukkale Üniversitesi Fen Fakültesi İstatistik Bölümü, eyalcin@pau.edu.tr, 0000-0002-1616-9121
- 3 Dr., TÜİK Denizli Bölge Müdürlüğü, tuncaymoral@gmail.com, 0000-0003-2567-2447
- 4 Doç. Dr., Pamukkale Üniversitesi Fen Fakültesi İstatistik Bölümü, sakogul@pau.edu.tr, 0000-0002-0346-4308

## 1. Giriş

Bireylerin harcama ve tasarruf etme davranışları ekonomi ve finans bilimlerinin temel inceleme konuları arasında olagelmıştır. Bireyler hayatları boyunca farklı rollere sahip olmuşlardır. Bireylerin ekonomik yaşam döngüsü içerisinde sahip oldukları roller arasında tüketici ve yatırımcı rolleri de bulunmaktadır. Bu rollere bağlı olarak bireyler satın alma, tasarruf etme ve finansal yatırım yapma kararları verebilmektedirler (Moralı, Coşkun 2021). Bireylerin günlük yaşamlarında ihtiyaç duydukları mal ve hizmetleri tüketmek için satın almaları esnasında yaptıkları parasal ödemeler harcamalarını oluşturmaktadır. Bu harcamalar anlık tüketime konu olabileceği gibi dayanıklı, yarı-dayanıklı malların alımı ile uzun süreli ihtiyaçlarını da karşılamak için kullanılabilirlerdir. Ancak bireyler ihtiyaç analizlerini yaparlarken çeşitli faktörlerden etkilenmektedirler. Ciumara (2014) çalışmasında bireylerin ekonomik ve finansal davranışlarına ilişkin karar verirken cinsiyet, eğitim seviyesi, gelir durumu gibi sosyal ve ekonomik faktörler ile birlikte, kendilerinin içinde buldukları psikolojik durum ve piyasa algılamalarının da etkili olabileceğinden bahsetmektedir.

Bireylerin piyasa algılamaları, bugün ve gelecek hakkındaki beklentilerini şekillendirmektedir. Bireylerin finansal durumlarına ilişkin doğru öngörüler yapabilmeleri, beklentilerini doğru yönetebilmeleri için konjonktürel etkileri doğru değerlendirmeleri gerekmektedir. Hem siyasi hem de ekonomik etkilerin ışığında bireylerin şekillenen beklentilerinin ve güveninin doğru olarak tayin edilmesi, karar alıcıların bireyler açısından doğru politikaları hayata geçirmesini sağlamaktadır. Karar alıcıların bilgisine sunulan ve bu değerlendirmelerin objektif olarak yapılmasını sağlayan bazı ekonomik göstergeler bulunmaktadır.

Bireylerin piyasa algılamalarını ve gelecek beklentilerini ölçmek amacıyla kullanılan endekslerden bir tanesi Tüketici Güven Endeksidir (TGE). TGE'nin veri kaynağı Türkiye genelinde aylık olarak uygulanan Tüketici Eğilim Anketi'dir ([www.tuik.gov.tr](http://www.tuik.gov.tr)). Bu endeks ülkemizin resmi istatistik otoritesi konumunda bulunan Türkiye İstatistik Kurumu tarafından 2004 yılından bu yana her ay düzenli olarak ölçülmekte ve kamuoyu ile sonuçları paylaşılmaktadır. TGE'nin alt bileşenlerinde bireylerin harcama eğilimlerinin ne yönde hareket ettiği başta olmak üzere, bireylerin enflasyon algısı ve geleceğe dönük beklentileri de ölçümlenmektedir. Ancak TGE sonuçlarının bireylerin satın alma niyetlerini nasıl etkilediği konusu tam olarak irdelenen bir mevzu olamamıştır. Yani sonuçları elde edilen endekslerin değerlendirmesini yapan tüketiciler, ileriye dönük harcama davranışını gerçekleştirirken bu endekslerden ne kadar etkilenmektedir?

Bu alıřma, TİK'in aylık Tetici Eđilim Anketi sonularından elde edilen Tetici Gven Endeksi ile bu endeksi etkilediđi dřnlen bazı makroekonomik deđiřkenler arasındaki iliřkiyi incelemektedir. Analizde Dolar Kuru, Tetici Fiyat Endeksi ve ihtiya kredisi faiz oranlarının TGE ile olan iliřkisi ele alınmıřtır. İlk olarak deđiřkenlerin birim kk testleri yapılmıř ve birim kke sahip olanlar iin yapısal kırılmalı birim kk sınaması uygulanmıřtır. Yapısal kırılmanın varlıđı durumunda Gregory-Hansen Eřtmleřme testi ile eřtmleřme sınaması gerekleřtirilmiřtir. Konuyla ilgili literatr taraması ve yapılan analizlerin bulguları sunulurak sonu kısmında deđerlendirme yapılmıřtır.

## 2. Literatr İncelemesi

Tetici gven endeksinin eřitli deđiřkenler ile olan iliřkisi literatrde farklı alıřmalarda inceleme konusu olmuřtur. Genelde TGE'nin alt endekslerinin iliřkisi incelenirken, bazı alıřmalarda TGE endeksinin diđer deđiřkenler ile arasındaki iliřki incelenmiřtir. Bu alıřmalara iliřkin literatr ařađıda verilmiřtir.

Kasapođlu (2022) alıřmasında makroekonomik deđiřkenler ile tetici gveni arasındaki iliřkileri incelemiřtir. alıřmasında makroekonomik deđiřkenlerin tetici gveni zerinde etkili olduđundan hareket ederek, 2005:01-2021:12 yılları arasındaki aylık seriyi kullanarak TGE ile eřitli makroekonomik deđiřkenler arasındaki iliřkiyi incelemiřtir. İki farklı trdeki tetici gveni olan Blomberg Tetici Gven Endeksi ile TİK tarafından aıklanan Tetici Gven endeksinin Enflasyon oranı, Sanayi retim endeksi, Borsa İstanbul-100 Endeksi, Bankalararası gecelik Faiz oranı, Reel Efektif Dviz Kuru deđiřkenleri ile aralarındaki iliřkiyi VAR modeli ile incelemiřtir. Kurulan modeller aısından sonulardaki farklılıkları ele almıřtır. alıřmasında TİK tarafından aıklanan tetici gven endeksi ile efektif dolar satıř kuru ve borsa endeksi arasında uzun dnemde anlamlı bir iliřki tespit etmiřtir.

nem (2022) alıřmasında, Trkiye'de Tetici Gven Endeksi ile bazı BIST sektrel endeksler arasındaki iliřkiyi eřbtnleřme ve nedensellik testleri ile incelemiřtir. Bu alıřmada ilk olarak arařtırma modelini kurmuř daha sonra ise modele alınan deđiřkenler arasındaki nedensellik iliřkisini lmek amacıyla ARDL Sınır Testi ve Granger Nedensellik analizini kullanmıřtır. alıřmadan elde edilen sonulara gre TGE ile BIST 100, BIST Banka, BIST Kurumsal Ynetim, BIST Sigorta ve BIST Sınai endeksleri arasında uzun dnemli eřbtnleřme iliřkisi tespit edilememiřtir. TGE ile belirtilen seriler arasında Granger Nedensellik Testine gre ise BIST banka endeksi ile tek ynl nedensellik iliřkisi tespit edilmiřtir.

Gedik ve Özbek (2021) çalışmalarında harcamaların ve yatırımların yönünü belirleyen göstergelerden birisi olan TGE'nin, makroekonomik değişkenler ile birlikte ele alınması gerektiğini belirterek; TGE ile işsizlik oranı, enflasyon oranı arasındaki ilişkiyi Ocak 2007- Ocak 2021 dönemindeki aylık verileri için incelemişlerdir. Değişkenler arasında eşbütünlüğün bulunduğu sonucuna ulaşmışlardır. Nedensellik analizleri kapsamında ise TGE'den işsizliğe doğru tek yönlü nedensellik ilişkisi bulunduğunu tespit etmişlerdir.

Prirayani (2016) çalışmasında Endonezya için Tüketici güveni ile makroekonomik değişkenler arasındaki ilişkiyi incelemiştir. Bu endeksin Endonezya ekonomisinin bugün ve gelecekteki durumu hakkında makroekonomik göstergelerle açıklanabilme durumunu araştırmıştır. Çalışmada Tüketici güven endeksi, enflasyon, büyüme oranı, döviz kuru, faiz oranı ve borsa endeksi Ocak 2001- Aralık 2015 dönemi için analiz edilmiştir. VAR yaklaşımı ile oluşturulan modelde, diğer değişkenler sabitken TGE ile enflasyon ve borsa endeksinin uzun dönemde eşbütünlüğe sağladığı, TGE ile enflasyon, faiz oranı ve borsa endeksi arasında nedensellik ilişkisi bulunduğu sonucuna ulaşmıştır.

Beşel ve Yardımoğlu (2016) çalışmasında Türkiye'de tüketici güven endeksi ile döviz kuru, petrol fiyatları ve işsizlik arasındaki ilişkiyi 2005-2014 dönemleri arasındaki aylık veriler ile analiz etmiştir. Gregory-Hansen eşbütünlük testi sonuçlarına göre değişkenler arasında eşbütünlük ilişkisi olduğu tespit edilmiştir. Toda-Yamamoto nedensellik testi sonuçlarına göre döviz kurundan TGE'ye tek yönlü ve petrol fiyatlarından işsizlik oranlarına doğru tek yönlü nedensellik ilişkisi tespit etmişlerdir.

Işık, Şahin vd. (2017) çalışmasında tüketici güven endeksi ile büyüme oranı ve enflasyon oranları arasındaki ilişkiyi 2004-2013 yılları için çeyrek dönemlik veriler üzerinden incelemiştir. Kullanılan model OECD'nin dönemlik verileri üzerinden oluşturulmuş ve birim kök testleri, nedensellik analizleri yapıldıktan sonra büyüme oranı ile tüketici güven endeksi arasında çift yönlü bir nedensellik ilişkisi olduğu, üretici ve tüketici fiyat endekslerinin büyüme oranının Granger nedeni olduğu sonucuna ulaşmıştır. Johansen eşbütünlük testi sonuçlarına göre ise değişkenlerin uzun dönemde anlamlı ilişki içerisinde olduğu sonucu ortaya çıkmıştır.

İslam ve Mumtaz (2016) çalışmalarında karar alıcılar ve ekonomi oyuncularını için bugün ve gelecek adına önemli role sahip olan tüketici güven endeksi ile seçilmiş Avrupa ülkeleri için büyüme oranı ile arasındaki ilişkiyi incelemişlerdir. Bu ülkelerin büyüme oranları ile tüketici güven endeksleri arasındaki uzun dönemli ilişkiyi 1996-2012 yılları arasındaki dönem için

analiz etmişlerdir. Analiz sonuçları tüketici güven endeksi ile büyüme oranı arasında uzun dönemli anlamlı ilişki olduğunu göstermiştir.

Sergeant, Lugay vd. (2011) çalışmalarında Jamaika ve Tobago için tüketici güveni ile gayrisafi yurt içi hasıla (GDP) arasındaki ilişkiyi Vektör otoregresyon modeli (VAR) çerçevesinde incelemişlerdir. Modele döviz kuru ve faiz oranı değişkenleri de eklenmiştir. 2005-2011 dönemleri arasındaki aylar için gerçekleştirdikleri çalışmada çeyrek dönemler için uzun dönemde GDP'nin tahmini için tüketici güveninin anlamlı olmadığı sonucuna ulaşmışlardır.

### 3. Veri Seti ve Yöntem

#### 3.1. Veri Seti

Çalışmada bireylerin tüketim eğilimi ve davranışlarının göstergesi niteliğinde olan ve her ay TÜİK tarafından açıklanan Tüketici Güven Endeksini (TGE) etkileyen bazı makroekonomik değişkenler ve bunlar ile aralarındaki ilişkiler ölçülmeye çalışılmıştır. TGE ile ilişkili olduğu değerlendirilen dolar kuru, ihtiyaç kredisi faiz oranı ve tüketici fiyat endeksi değişkenleri analize dahil edilmiştir. Bu değişkenlere ait veri setleri TÜİK ve Türkiye Cumhuriyet Merkez Bankası (TCMB)'nın veri sunucularından elde edilmiştir. Analize Mayıs 2012 (2012:05)- Eylül 2023 (2023:09) dönemlerine ait aylık veriler dahil edilmiştir. Analizlerde kullanılan değişkenlere ait bilgiler Tablo 1'de verilmiştir.

*Tablo 1. Değişkenlere Ait Bilgiler*

Değişken Adı	Değişken Tanımı	Veri Kaynağı
TGE	Tüketici Güven Endeksi	TÜİK
İhtiyac	Türk Lirası Üzerinden Açılan İhtiyac Kredisi Faiz Oranı	TCMB
Dolar	Aylık Ortalama Amerikan Doları Satış Kuru	TCMB
Enflasyon	Tüketici Fiyat Endeks Düzeyi	TÜİK

TGE; 2004-2012 yılları arasında Hane halkı İşgücü Anketine modül olarak uygulanan Tüketici Eğilim Anketi, 2012 yılından bu yana “Avrupa Birliği Uyumlulaştırılmış İşyeri ve Tüketici Anketleri Programı” kapsamında uygulanmaktadır. 2020 yılı Eylül ayında ise gerçekleştirilen ana revizyon kapsamında “mevcut dönemde hanenin maddi durumu” alt endeksi için FARK yöntemi ve “dayanıklı tüketim mallarına harcama yapma düşüncesi” alt endeksi için ARIMAX yöntemi kullanılarak 2004-2011 yılları için geri

çekilmiş seriler üretilmiştir. Söz konusu seriler, 2020 Eylül ayındaki ana revizyon kapsamında 2004-2011 yılları arasındaki tüketici güven endeksi serisinin hesaplamalarında kullanılmıştır (www.tuik.gov.tr).

### 3.2. Yöntem

Ekonometrik zaman serileriyle çalışılırken pek çok yöntemin serilerin durağanlığı varsayımına dayanması sebebiyle ilk olarak serilerin birim kök içerip içermedikleri yani durağanlık derecelerinin araştırılması gerekmektedir. Serilerin durağanlıklarının incelenmesinde, serilere ait grafiklerin ve korelogramların incelenmesi fikir verici olsa da kesin sonuçlara ulaşabilmek için birim kök testleri kullanılmaktadır.

#### 3.2.1. Birim Kök Testleri

Durağanlığın sınanmasında klasik birim kök testlerinden Augmented Dickey-Fuller (ADF) (1981) ve Phillips-Perron (PP) (1988) birim kök testleri sıklıkla tercih edilmektedir. ADF testi, hata teriminde otokorelasyon sorununu ortadan kaldırmak için bağımlı değişkene ait gecikmeli değerleri bağımsız değişken olarak modele dahil etmektedir. Dolayısıyla uygun gecikme uzunluğunun belirlenmesi için Akaike Bilgi Kriteri (AIC) ve Schwarz ya da Bayesian Bilgi Kriteri (SIC) gibi seçim kriterlerinin kullanılması gerekmektedir.

Birim kök testinde kullanılan ADF testine ait sabit ve trend içeren ekonometrik model (1) nolu denklemde yer almaktadır. Denklemde  $y_t$  serisinin birim kök içerdiğini ve durağan olmadığını ifaden eden temel hipotez sınanmaktadır.

$$\Delta y_t = \beta_0 + \beta_1 T + \theta y_{t-1} + \sum_{i=1}^k \delta_i \Delta y_{t-i} + u_i \quad (1)$$

ADF birim kök testine kıyasla, hata teriminin zayıf bağımsız ve heterojen olmasına imkân vererek daha esnek bir varsayım dayandırılarak geliştirilmiş PP testine ait sabit ve trend içeren ekonometrik model ise (2) nolu denklemde yer almaktadır.

$$\Delta y_t = \beta_0 + \beta_1 T + \theta y_{t-1} + u_i \quad (2)$$

#### 3.2.2. Yapısal Kırılmalı Birim Kök Testleri

ADF, PP gibi birim kök testlerinden serilerin durağan olmadığı bulgusu elde edildiğinde, yapısal kırılmaları dikkate alan birim kök testleriyle sınamaların tekrarlanması gerekmektedir. Çünkü zaman serilerinin durağan olmadığı bulgusunun elde edilmesinin nedenlerinden biri yapısal kırılmaların varlığı olabilmektedir. Yani seriler ADF, PP testinde durağan olmadığı halde

durađan olabilir ya da eđer kırılma varsa ve dikkate alınmadıysa durađanlık dereceleri farklı tespit edilebilmektedir.

Analizlerde kullanılan yapısal kırılmayı dikkate alan testlerden biri olan Zivot-Andrews (1992) testi, yapısal kırılmayı içsel olarak belirleyen (3), (4) ve (5) nolu denklemlerde yer alan modellerle tahminlenmektedir.

$$y_t = \mu + \theta_1 DU_t(\lambda) + \beta t + \alpha_1 y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-j} + e_t \quad (3)$$

$$y_t = \mu + \gamma_1 DT_t(\lambda) + \beta t + \gamma DT_t + \alpha_2 y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-j} + e_t \quad (4)$$

$$y_t = \mu + \theta_2 DU_t(\lambda) + \gamma_2 DT_t(\lambda) + \beta t + \alpha_3 y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-j} + e_t \quad (5)$$

Eşitlik (3), (4) ve (5) nolu denklemlerde  $t = 1, 2, \dots, T$  olmak üzere trend deđişkenini;  $T_b$  kırılma zamanı olmak üzere  $\lambda = T_b / T$  göreli kırılma yılını ifade etmektedir. (3) nolu model ortalamada görölen bir kırılmayı, (4) nolu model eğimde görölen bir kırılmayı, (5) nolu modelde hem ortalamada hem de eğimde görölen kırılmayı temsil etmektedir.

Analizlerde Zivot-Andrews testinin yanı sıra Lee-Strazicich (2003) ve Lee-Strazicich (2013) Birim Kök Testleri de kullanılmıştır. Lee-Strazicich (2003) ve Lee-Strazicich (2013) yapısal kırılmalı birim kökün sınanmasında kullanılan model denklem (6) ve (7)'de yer alan veri oluşturma sürecine dayanmaktadır (Mert ve Çađlar, 2019).

$$y_t = \delta Z_t + e_t \quad (6)$$

$$e_t = \beta e_{t-1} + \varepsilon_t \quad (7)$$

LM prensibine göre tahmin edilen denklem (8) nolu modelde yer almakla birlikte;  $\hat{S}_t = y_t - \tilde{\psi}_x - Z_t \tilde{\delta}_t$  ve  $t = 2, \dots, T$ 'dir. Ayrıca  $\tilde{\psi}_x$ ,  $y_1 - Z_1 \tilde{\delta}$  ile gösterilmektedir.  $y_1$  ve  $Z_1$  matrislerin başlangıç deđerlerini ve  $\tilde{\delta}$  katsayılar matrisini ifade etmektedir.  $Z_t$  dışsal deđişkenler vektörünü ifade etmekte olup bu vektör deđiştirilerek tek ve çift kırılmalı testler ortaya çıkmaktadır.

$$\Delta y_t = \hat{\delta} Z_t + \phi \hat{S}_{t-1} + \varepsilon_t \quad (8)$$

Lee-Strazicich (2013) Birim Kök Testinde,  $Z_t$  dışsal deđişkenler vektörü tek kırılmayı dikkate alarak oluşturulmaktadır. Kırılma modelleri düzeyde kırılma ve düzeyde-trendde kırılma olarak iki farklı yapıda ifade edilmektedir.  $T_b$  kırılma zamanını ifade etmekle birlikte düzeyde ve düzeyde & trendde kırılma modellerine ait bilgiler sırasıyla denklem (9) ve denklem (10)'da yer almaktadır.



$$Z_t = [1, t, D_t] \quad t \geq T_B + 1 \text{ için } D_t = 1, dd=0 \quad (9)$$

$$Z_t = [1, t, D_t, DT_t] \quad t \geq T_B + 1 \text{ için } DT_t = t - T_B, dd=0 \quad (10)$$

### 3.2.3. Yapısal Kırılmalı Eştleşme Testleri

Yapısal kırılmalı eştleşme testlerinden Gregory-Hansen (1996) Eştleşme Testi ve Hatemi-J (2008) Eştleşme Testi yaygın kullanılmaktadır. Gregory-Hansen (1996) eştleşme testini, Zivot ve Andrews'ın içsel kırılma tek değişkenli testlerinin çok değişkenli bir uzantısı olarak ifade etmek mümkündür. Gregory-Hansen Eştleşme Testinde, yapısal kırılma ile eştleşmenin olduğu alternatifine karşı, yapısal kırılma ile eştleşmenin olmadığı hipotezi sınanmaktadır. Gregory-Hansen (1996) eştleşme testi eştleşik vektörde yapısal kırılma tarihinin içsel olarak belirlendiği üç model ile sınanmaktadır. (12) nolu denklemde yer alan model düzeyde kırılmayı, (13) nolu denklemde yer alan model trendle birlikte düzeyde kırılmayı, (14) nolu denklemde yer alan model ise hem sabit hem de eğimde kırılmayı (rejimde) ifade etmektedir. Gregory-Hansen Eştleşme Testi, (11) nolu denklemde yer alan yapısal değişimleri modellerde temsil eden kukla değişkenler dikkate alınarak (12), (13) ve (14) nolu denklemlerde yer almaktadır. (11) nolu denklemde  $n$  gözlem sayısını,  $\tau \in (0,1)$  olmak üzere kırılma sayısını ve  $(n\tau)$  kırılma noktasını ifade etmektedir.

$$\varphi_{1\tau} = \begin{cases} 0, & t \leq [n\tau] \\ 1, & t > [n\tau] \end{cases} \quad (11)$$

$$y_{1t} = \mu_1 + \mu_2 \varphi_{1\tau} + \alpha^T y_{2t} + \varepsilon_t \quad t = 1, 2, \dots, n \quad (12)$$

(12) nolu denklem düzeyde kırılmayı (C) ifade etmekte olup  $\mu_1$  kırılma öncesi sabit terimi,  $\mu_2$  kırılmanın sabit terimdeki etkisini ve  $\alpha^T$  bağımlı değişkenlere ait katsayı vektörünü temsil etmektedir.

$$y_{1t} = \mu_1 + \mu_2 \varphi_{1\tau} + \beta t + \alpha^T y_{2t} + \varepsilon_t \quad t = 1, 2, \dots, n \quad (13)$$

(13) nolu denklemde trend varlığında sabitteki kırılmayı ifade etmektedir. Eğimde kırılma (C/T) modeli olarak da tanımlanmaktadır.

$$y_{1t} = \mu_1 + \mu_2 \varphi_{1\tau} + \alpha_1^T y_{2t} + \alpha_2^T y_{2t} \varphi_{1\tau} + \varepsilon_t \quad t = 1, 2, \dots, n \quad (14)$$

(14) nolu denklem rejim değişikliği (C/S) modelini ifade etmektedir. Bu modelde düzeyde kırılmadan farklı olarak  $\alpha_1^T$  rejim öncesi eştleşme

katsayısını,  $\alpha_2^T$  rejim değişikliği sonrası eğim katsayısındaki değişimi temsil etmektedir.

(12), (13) ve (14) nolu modelleriyle eştümleşmeyi tespit etmede  $ADF^*$ ,  $Z_a^*$  ve  $Z_t^*$  istatistikleri kullanılmaktadır. İstatistiklerin hesaplanması (15) denklemlerdeki gibi yapılmaktadır.

$$ADF^* = \min_{\tau \in T} ADF(\tau) ; Z_t^* = \min_{\tau \in T} Z_t(\tau) ; Z_a^* = \min_{\tau \in T} Z_a(\tau) \quad (15)$$

Eştümleşme testlerinde anlamlı bulguların elde edilmesinin ardından, birinci dereceden durağan I(1) ve eştümleşik anlamlı ilişkinin olduğunu tespit edilen değişkenler arasındaki ilişkiyi modellemek için Phillips ve Hansen (1990) Tam Düzeltilmiş En Küçük Kareler Yöntemini-FMOLS kullanılarak uzun dönem katsayılarına ulaşmak mümkündür. FMOLS yönteminde tutarlı, asimptotik sapmasız sonuçlar elde edilmekle birlikte küçük örneklerde de başarılı olduğu bilinmektedir.

( $n+1$ ) boyutlu zaman serisi vektörüyle FMOLS yöntemi (16) nolu denklemdeki gibi ifade edilmektedir.

$$Y_t = X_t' \beta + D_{1t}' \gamma_1 + u_{1t} \quad (16)$$

(16) nolu denklemde  $D_t = (D_{1t}', D_{2t}')$  deterministik trend değişkenleri ve (17) nolu sistemde yer alan  $n$  stokastik değişkenler  $X_t$  denklem sistemi tarafından belirlenmektedir.

$$X_t = \Gamma_{21}' D_{1t} + \Gamma_{22}' D_{2t} + \varepsilon_{2t} \quad \text{ve} \quad \Delta \varepsilon_{2t} = u_{2t} \quad (17)$$

FMOLS tahmincisi, (16) nolu denklemde hata terimlerinin simetrik ve tek taraflı uzun dönem kovaryans matrisinden  $\hat{u}_{1t}$  ve (17) nolu denklemde  $\hat{u}_{2t}$  elde edilerek uzun dönem kovaryans matrisleri  $\hat{\Lambda}$  ve  $\hat{\Omega}$  tahmin edilmektedir. Tam Düzeltilmiş En küçük Kareler (FMOLS) tahmincisi (18) nolu denklemde yer almaktadır. (18) nolu denklemde  $Z_t = (X_t', D_t')$ 'i ifade etmektedir.

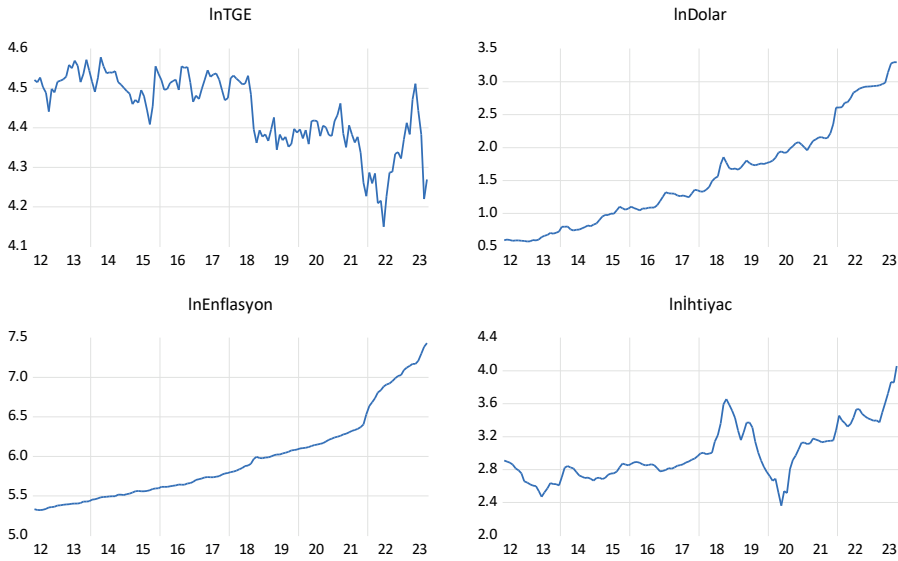
$$\hat{\theta} = \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \left( \sum_{t=1}^T Z_t Z_t' \right)^{-1} \left( \sum_{t=1}^T Z_t Y_t^+ - T \begin{bmatrix} \hat{\lambda}_{12}^+ \\ 0 \end{bmatrix} \right) \quad (18)$$

Kanonik Koentegrasyon Regresyonu (CCR) (Park 1992) ve Dinamik En Küçük Kareler Yöntemi (DOLS) (Stock & Watson 1993), FMOLS yöntemine benzer şekilde edilen uzun dönem katsayılarının yorumlanması amacıyla kullanılmaktadır. FMOLS, DOLS ve CCR yöntemlerinin kullanılabilmesi için serilerin farkta durağan olmaları koşulunun sağlanması gerekmektedir. FMOLS yönteminden farklı olarak; DOLS yönteminde

asimptotik açıdan etkin bir tahminci geliştirilmiştir. CCR yöntemi ise asimptotik olarak Ki-Kare testinin kullanılmasına imkân sağlamaktadır.

#### 4. Bulgular

Analize dahil edilen değişkenlere ve yukarıda bahsedilen analizlere ilişkin bulgular aşağıda detaylıca açıklanmıştır. Çalışmada kullanılan seriler düzey ve yalın halleri ile analizde kullanılmamış, serilerin doğal logaritması alınarak durağanlaştırma işlemi gerçekleştirilmiştir. Bu serilere ilişkin grafikler Şekil 1'de verilmiştir.



Şekil 1. Değişkenlere ait grafik

Tablo 2. Değişkenlere ait tanımlayıcı istatistikler

Değişken	Ortalama	Medyan	Mak	Min	Std. Sapma	Çarpıklık	Basıklık	Jarque-Bera
LNTGE	4.4444	4.4719	4.5785	4.1496	0.0935	-0.8313	3.0566	15.7983*
LNDOLAR	1.5553	1.3508	3.2973	0.5738	0.7633	0.6015	2.3489	10.6818*
LNENFLASYON	5.9640	5.8014	7.4331	5.3195	0.5433	1.0079	3.1160	23.2730*
LNIHTIYAC	2.9925	2.8828	4.0586	2.3616	0.3322	0.7590	3.0227	13.1567*

\*: %10 önem seviyelerini göstermektedir.

Tablo 2'de değişkenlerin doğal logaritmalarına ait yer alan tanımlayıcı istatistikler incelendiğine dolar, enflasyon ve ihtiyaç kredisi faizi serileri pozitif

asimetriye sahipken, tüketici güven endeksinin negatif asimetriye sahip olduğu görülmüştür. Tüketici güven endeksi, enflasyon ve ihtiyaç kredisi faizine ait serinin normale göre hafif sivri iken, dolar serisinin basık olduğu izlenmiştir. Ayrıca tüm serilerin normal dağılıma uymadığı sonucuna ulaşılmıştır.

Serilere ait uzun dönem ilişkilerin incelenmesinden önce, serilere ait durağanlık incelenmiş ve bulgulara Tablo 3'te yer verilmiştir. Elde edilen bulgular %1 önem seviyesinde tüm serilerin durağan olmadığını ve birinci farklarında durağan hale geldiklerini göstermiştir. Serilerin durağan olmamasının serilerde meydana gelen kırılmadan kaynaklanabileceğini düşündürmüştü ve Tablo 4'te yer alan yapısal kırılmalı birim kök testi elde edilmiştir.

*Tablo 3. ADF ve PP Birim Kök Test Sonuçları*

Değişkenler	ADF		PP	
	Sabit	Sabit&Trend	Sabit	Sabit&Trend
LNTGE	-2.0579	-3.7996**	-1.8487	-3.7911**
LNDOLAR	1.9952	-0.8967	2.3607	-0.8473
LNENFLASYON	3.7192	1.5918	5.6718	2.2787
LNIHTIYAC	-0.4504	-2.1897	-0.2403	-1.9802
$\Delta$ LNTGE	-11.7601***	-11.7301***	-13.0102***	-12.9714***
$\Delta$ LNDOLAR	-8.2864***	-8.7181***	-7.1181***	-7.2674***
$\Delta$ LNENFLASYON	-2.8125*	-6.3823***	-5.0746***	-6.3823***
$\Delta$ LNIHTIYAC	-6.8756***	-7.0352***	-6.8117***	-6.9431***

\*\*\*: %1, \*\*: %5 , \*: %10 önem seviyelerini göstermektedir.

Tablo 4'te yer alan Zivot-Andrews Birim Kök Testi sonuçları sabitte (model A), trendde (model B) ve sabit & trendde (model C) kırılmalar dikkate alındığında %5 ve %1 anlam düzeylerinde temel hipotezin reddedilemediğini ve serilerin durağan olmadıklarını göstermiştir. %10 önem düzeyinde LNTGE ve LNDOLAR serilerinde temel hipotezin reddedilmesi ise bu serilerin kırılmalar ile durağan olabileceğini ve ilgili kırılma tarihlerinin anlamlı etkilerinin olabileceğini düşündürmüştür.

Tablo 4: Yapısal Kırılmalı Birim Kök Test Sonuçları

Zivot_Andrews Analizi	Model	Kırılma Tarihi	Test İstatistiği
LNTGE	A	2018:08	-4.6789*
	B	2016:09	-4.1008
	C	2018:08	-4.7471
LNDOLAR	A	2021:11	-4.8075*
	B	2021:01	-4.1744*
	C	2021:11	-4.8112
LNENFLASYON	A	2021:12	-4.3973
	B	2021:02	-3.6167
	C	2020:12	-3.4251
LNIHTIYAC	A	2019:06	-3.0349
	B	2021:10	-3.1042
	C	2019:08	-4.6364

A (sabit), B (trend), C (sabit + trend) modellerini ve \*: %10 önem seviyelerini göstermektedir.

Elde edilen birim kök test sonuçlarıyla, seriler arasındaki uzun dönem ilişkilerinin sınanmasında kırılmayı dikkate alan eştümleşme testinin kullanılması uygun görülmüştür. Gregory-Hansen Yapısal Kırılmalı Eştümleşme testine ait sonuçlara Tablo 5'te yer verilmiştir. Elde edilen bulgular seriler arasında trendde kırılmayla anlamlı uzun dönem ilişkisi olduğunu göstermiştir. Sabitte ve rejimde kırılmayla ise uzun dönem ilişkisi olmadığı bulgusu elde edilmiştir. Trendde kırılmayla anlamlı uzun dönem ilişkisinin olduğunu ifade eden model dikkate alınması gereken kırılma tarihinin 2018 yılının 10. ayı olduğunu göstermiştir. 2018:10 kırılma tarihine ait kukla değişken oluşturularak trendde kırılma dikkate alınmış ve değişkenlere ait uzun dönem katsayıları FMOLS, DOLS ve CCR yöntemleri ile tahminlenmiştir. Elde edilen tahmin sonuçlarına Tablo 6'da yer verilmiştir.

Tablo 5: Gregory-Hansen Eşitümleşme Test Sonuçları

Test	Model	ADF	TB	$Z_t$	TB	$Z_a$	TB
Model	C	-4.7619	2015:08	-4.8173	2015:08	-40.9876	2015:08
	C/T	-5.4779*	2018:10	-5.4981*	2018:10	-50.6998	2018:10
	C/S	-5.4644	2018:10	-5.5330	2018:10	-51.7443	2018:10

\*: %10 önem seviyelerini göstermektedir. Gregory-Hansen için kritik değerler model C için  $ADF \& Z_t = -5.28$  (%5),  $-5.02$  (%10),  $Z_a = -48.65$  (%10); model C/T için  $ADF \& Z_t = -5.57$  (%5),  $-5.33$  (%10),  $Z_a = -54.94$  (%10); model C/S için  $ADF \& Z_t = -6.00$  (%5),  $-5.75$  (%10),  $Z_a = -63.42$  (%10).

Tablo 6’da yer alan bulgular incelendiğinde; FMOLS, DOLS ve CCR modellerinden benzer bulguların elde edildiği izlenmektedir. FMOLS ve CCR modellerinde dolar, enflasyon değişkenlerine ait katsayılar ile kırılma kuklasıyla dolar ve enflasyon katsayıları ve sabit terimin istatistiki olarak anlamlı olduğu bulgusu elde edilmiştir. DOLS modelinde farklı olarak sadece kırılma kuklasıyla enflasyon değişkenine ait katsayının anlamlı olmaması söz konusudur. Elde edilen bulgular tüketici güven endeksi üzerinde doların negatif, enflasyonun pozitif yönde etkiliyken ihtiyaç kredisi faizinin etkili olmadığını göstermiştir. FMOLS ve CCR modelleri için 2018 yılının 10. ayındaki kırılmanın anlamlı olduğu ve doların negatif, enflasyonun pozitif katkısını artırarak etkilediği izlenmiştir. Modellerden elde edilen katsayılar her bir model için yorumlandığında FMOLS modeli için; 2018:10 tarihinden önce dolar kurundaki %1’lik artışın tüketici güven endeksini %0.3248 azaltırken; enflasyondaki %1’lik artışın tüketici güven endeksini %0.5685 arttırmaktadır. 2018:10 tarihinden sonra ise dolar kurundaki %1’lik artışın tüketici güven endeksini %0.5973 azaltırken; enflasyondaki %1’lik artışın tüketici güven endeksini %0.6499 arttırdığını ifade etmek mümkündür.

DOLS modeli için; 2018:10 tarihinden önce dolar kurundaki %1’lik artışın tüketici güven endeksini %0.3367 azaltırken; enflasyondaki %1’lik artışın tüketici güven endeksini %0.6173 arttırmaktadır. 2018:10 tarihinden sonra ise dolar kurundaki %1’lik artışın tüketici güven endeksini %0.6125 azalttığını ifade etmek mümkündür.

CCR modeli için; 2018:10 tarihinden önce dolar kurundaki %1’lik artışın tüketici güven endeksini %0.3315 azaltırken; enflasyondaki %1’lik artışın tüketici güven endeksini %0.5788 arttırmaktadır. 2018:10 tarihinden sonra ise dolar kurundaki %1’lik artışın tüketici güven endeksini %0.6102 azaltırken; enflasyondaki %1’lik artışın tüketici güven endeksini %0.6620 arttırdığını ifade etmek mümkündür.

FMOLS modeli (19) nolu eşitlikte yer almaktadır.

$$LNTGE = 1.6909 - 0.3248 * LNDOLAR + 0.5685 * LNENFLASYON - 0.0090 * LNIHTIYAC - 0.2726 * LNDOLAR * D2018:10 + 0.0814 * LNENFLASYON * D2018:10 - 0.0417 * LNIHTIYAC * D2018:10 \quad (19)$$

2018:10 tarihi sonrasında değişen eğim katsayılarıyla FMOLS modelini (20) nolu eşitlikteki gibi ifade etmek mümkündür.

$$LNTGE = 1.6909 - 0.5973 * LNDOLAR + 0.6499 * LNENFLASYON - 0.0507 * LNIHTIYAC \quad (20)$$

Tablo 6: Uzun Dönem Model Tahmin Sonuçları

Bağımlı Değişken: LNTGE	FMOLS	DOLS	CCR
Bağımsız Değişken	Katsayı	Katsayı	Katsayı
LNDOLAR	-0.3248***	-0.3367***	-0.3315***
LNENFLASYON	0.5685***	0.6173***	0.5788***
LNIHTIYAC	-0.0090	-0.0303	-0.0080
LNDOLAR*D2018:10	-0.2726***	-0.2758***	-0.2787***
LNENFLASYON*D2018:10	0.0814**	0.0693	0.0832**
LNIHTIYAC*D2018:10	-0.0417	-0.0183	-0.0413
C	1.6909***	1.4956***	1.6365***
R <sup>2</sup>	0.7841	0.8606	0.7839
Hata Kareler Toplamı	0.2551	0.1598	0.2554
Uzun Dönem Varyans	0.0036	0.0029	0.0036

\*\*\*: %1, \*\*: %5, \*: %10 önem seviyelerini göstermektedir.

Değişkenlere ait kısa dönem ilişkilerini tespit etmek için VECM yapısına uygun olarak serilerin birinci farkları alınmış ve uzun dönem modellerinden elde edilen hata terimleri kullanılarak FMOLS, DOLS ve CCR modelleri için kısa dönem katsayıları tahminlenmiştir. Elde edilen bulgulara Tablo 7'de yer verilmiştir.

Tablo 7: Kısa Dönem Model Tahmin Sonuçları

Bağımlı Değişken: $\Delta$ LNTGE	FMOLS	DOLS	CCR
Bağımsız Değişken	Katsayı	Katsayı	Katsayı
$\Delta$ LNDOLAR	-0.2841***	-0.4493***	-0.2964**
$\Delta$ LNENFLASYON	-0.3516	0.5269***	-0.2430
$\Delta$ LNIHTIYAC	-0.0657	-0.0983***	-0.0430
$\Delta$ LNDOLAR*D2018M10	-0.0589	-0.1509*	-0.1912
$\Delta$ LNENFLASYON* D2018M10	0.5035	0.1191	0.6767
$\Delta$ LNIHTIYAC* D2018M10	0.0392	0.0599*	-0.0006
C	0.0068**	0.0010	0.0045
ECT(-1)	0.1949***	-0.0102	0.0367
R <sup>2</sup>	0.2551	0.9668	0.1274
Hata Kareler Toplamı	0.1403	0.0053	0.1644
Uzun Dönem Varyans	0.0006	0.0001	0.0006

\*\*\*: %1, \*\*: %5, \*: %10 önem seviyelerini göstermektedir.

Tablo 7'de yer alan kısa dönem model sonuçları incelendiğinde ECT (hata düzeltme teriminin) hiçbir modelde iktisadi beklentiye uygun olarak negatif ve aynı zamanda istatistiki olarak anlamlı olmaması sebebiyle kısa dönem modelinin anlamlı olmadığı sonucuna ulaşılmıştır. İstatistiki olarak anlamlı katsayılar olmasına rağmen ECT(-1) katsayısının ifade edilen özelliklere sahip olmaması sebebiyle katsayılar yorumlanamamıştır.

## 5. Sonuç

Ekonomik konjonktür incelendiğinde Covid-19 salgın hastalığı sonrasında tüm dünya ile birlikte ülkemizde de fiyat hareketlerinde yükseliş trendi görülmüştür. Bu fiyat hareketlerinin etkilerini azaltmak amacı ile fiyat istikrarını sağlama görevine sahip olan TCMB faiz politikasına ilişkin kararlar almış ve bu kararların piyasaya yansımaları olmuştur. Tüm bu dinamik yapı içerisinde gerçekleşen hareketler, bir taraftan yatırımcı diğer taraftan ise tüketici olan bireyleri etkilemiştir. Bahsedilen hareketlerin tüketici güveni üzerindeki etkisi merak konusu haline gelmiştir.

Bu çalışmada TÜİK tarafından üretilen Tüketici Güven Endeksi ile bazı makroekonomik değişkenler arasındaki ilişki araştırılmıştır. Bu amaçla TGE'yi etkilediği düşünülen tüketici fiyat endeksi, dolar kuru ve ihtiyaç kredisi faiz oranları analize dahil edilmiştir. İlk etapta 2012:05 ile 2023:09 dönemleri arasındaki 137 gözleme sahip olan serilerin düzeydeki durağanlıkları birim kök testleri ile sınanmıştır. Bu testler sonucunda serilerin birim köke sahip oldukları, %1 önem seviyesinde durağan olmadıkları sonucuna ulaşılmıştır. Durağan olmadıkları belirlenen serilerin bir yapısal kırılmaya sahip olup olmadıkları incelenerek, Zivot-Andrews testi ile yapısal kırılmalı birim kök testi gerçekleştirilmiştir. Bu test sonucunda sabit model, trend model, sabit ve trend modelde serilerin %1 ile %5 önem düzeyinde durağan olmadıkları tespit edilmiştir. %10 önem düzeyinde ise TGE ve Dolar serilerinin kırılmalar ile durağan olabileceği sonucuna ulaşılmıştır. Yapılan birim kök testi sonuçlarında yapısal kırılmaların serilerde anlamlı olduğu, bu nedenle yapısal kırılmalı eştümleşme testinin kullanılmasının uygun olacağı değerlendirilerek, Gregory-Hansen Eştümleşme Testi uygulanmıştır. Üç farklı model ile uzun dönem ilişkinin varlığının araştırıldığı yöntemde seriler arasında sabitte ve rejimde kırılmayla anlamlı uzun dönem ilişkinin bulunmadığı, trendde kırılmayla anlamlı uzun dönemli ilişkinin bulunduğu sonucuna ulaşılmıştır. Trendde kırılmayla anlamlı uzun dönem ilişkinin bulunduğu sonuçlarda kırılma tarihi olarak 2018 yılı 10. ay olduğu tespit edilmiştir. 2018:10 dönemi Türkiye ekonomisi açısından önemli olumsuz sonuçlara sahip olan ve ulusal basında ve medyada Rahip Brunson olayı olarak bilinen olayın gerçekleşme dönemidir. Bu dönemde Rahip Brunson'a



ilişkin devam eden mahkemenin karara bağlanmış, kararın neticesinde ise ABD tarafından ülkemize uygulanan ekonomik ve siyasi yaptırımlar söz konusu olmuştur. Bu yaptırımlar sonucunda Türkiye ekonomisinde dolar kurunda, bazı hammadde fiyatlarında yükselişler gerçekleşmiştir.

Burada 2018:10 kırılma tarihi kukla değişken olarak modele eklenmiş ve değişkenlere ait uzun dönemli katsayılar FMOLS, DOLS ve CCR yöntemleri ile tahmin edilmiştir. FMOLS, DOLS ve CCR modellerine göre seriler arasında Dolar kuru ve Enflasyon ile TGE arasında anlamlı uzun dönemli ilişki olduğu, Dolar kuru ile TGE arasında negatif yönlü, enflasyon ile TGE arasında ise pozitif yönlü ilişki olduğu tespit edilmiştir. Modele dahil edilen ihtiyaç kredisi faiz oranı ile TGE arasında uzun dönemli anlamlı ilişki tespit edilememiştir. Modele dahil edilen serilerin kısa dönemli hareketlerini uzun dönemde modelde yaptığı dengelemeyi görmek için yapılan kısa dönemli model tahmin sonuçlarında ise ECT (-1) katsayısı istatistiki olarak anlamlı bir sonuç ortaya koyamamıştır.

Özetle, Türkiye’de dolar kuru ve enflasyon makroekonomik değişkenleri tüketici güveni üzerinde uzun dönemli ilişki içerisinde olduğu, serilerde gözlenen tarihi kırılmalara Türkiye ile diğer ülkeler arasındaki yaşanan siyasi gelişmelerin neden olduğu, tüketici kredisi faiz oranının modelde etkisinin olmadığı tespit edilmiştir.

## Kaynakça

- Beşel, F. & Yardımcıoğlu, F., (2016), Tüketici Güven Endeksi ile Makro Değişkenler Arasındaki İlişki, International Congress on Political, Economic and Social Studies (ICPESS) Bildiriler Kitabı, İstanbul
- Dickey, D. A. & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: journal of the Econometric Society*, 1057-1072.
- Gedik, A., Özbek, Ö. (2021). Türkiye’de Tüketici Güven Endeksi, İşsizlik Oranı ve Enflasyon Oranı Arasındaki İlişki, *New Era International Journal Of Interdisciplinary Social Researches*, Year 6(2021) Vol:10, ISSN 2757-5608, Doi Number:<http://dx.doi.org/10.51296/newera.119>
- Gregory, A. W. & Hansen, B. E. (1996). Residual-based tests for cointegration in models with regime shifts. *Journal of econometrics*, 70(1), 99-126.
- Hatemi-j, A. (2008). Tests for cointegration with two unknown regime shifts with an application to financial market integration. *Empirical economics*, 35(3), 497-505.
- Islam, T. & Mumtaz, M. N. (2016), Consumer Confidence Index and Economic Growth: An Empirical Analysis of EU Countries, *Euroeconomica*, Issue 2 (35) ISSN: 1582-8859
- Işık, T. & Sahin, İ., Aydınkaya, İ.Y., (2016) Consumer Confidence Index, The Relationship Between Growth and Inflation, A case of Turkey, *International Journal Of Economics, Commerce and Management*, Vol. IV, Issue 4, April 2016, United Kingdom, <http://ijem.co.uk/>.
- Lee, J. & Strazicich, M. C. (2003). “Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks”, *The Review of Economics and Statistics*. 85(4), ss. 1082-1089.
- Lee J. & Strazicich M. C. (2013). Minimum LM Unit Root Test With One Structural Break. *Economic Bulletin*, 33(4), pp. 2483-2492.
- Kasapoğlu, S., (2022). The Relationship Between Consumer Confidence Index and Macroeconomic Indicators In Türkiye, Master In Art Thesis, Yasar University, İzmir
- Mert, M. & Çağlar, A.E. (2019). Eviews ve Gauss Uygulamalı Zaman Serileri Analizi. Detay Yayıncılık: Ankara.
- Moralı, T. & Coşkun, Ö. (2021). Türkiye’de Tüketicilerin Enflasyon Algısı ve Beklentileri Üzerine Bir İnceleme, *ASEAD 7. Uluslararası Sosyal Bilimler Sempozyumu, Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi*, 2021, 631-650
- Önem, H.B., (2022), Tüketici Güven Endeksi ve Bazı BİST Endeksleri Arasındaki Eşbütünleşme ve Nedensellik İlişkisi, *Finans Ekonomi ve Sosyal Araştırmalar Dergisi*, Cilt.7 Sayı.1, Mart 2022, 12-22

Phillips, P. C. & Hansen, B. E. (1990). Statistical inference in instrumental variables regression with I (1) processes. *The review of economic studies*, 57(1), 99-125.

Phillips, P. C. & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-346.

Prirayani, J. (2016). Macroeconomic Linkage of Consumer Confidence In Indonesia, *Jurnal BPPK*, Volume 9 Nomor 1, 2016, Halaman 67-88

Sergeant, K.A., Lugay,B.& Dokkie,M. (2011), Consumer Confidence and Economic Growth: Case Studies of Jamaica and Triniad and Tobago; Economic Commission for Latin America and The Caribbean (ECLAC) Subregional headquarters for the Caribbean, <https://repositoro.cepal.org/>

Türkiye İstatistik Kurumu (TÜİK). (Erişim Tarihi: 10.10.2023, Erişim adresi: <http://www.tuik.gov.tr/>).

Türkiye Cumhuriyet Merkez Bankası (TCMB). (Erişim Tarihi: 10.10.2023, Erişim adresi: <https://evds2.tcmb.gov.tr/>)

## Cinsiyete Göre Yaşam Beklentisinin Şekillenmesinde Etkili Faktörler: OECD Daha İyi Yaşam İndeksi Analizi

Mert Veznikli<sup>1</sup>

### Özet

Yaşam beklentisi, toplumun sağlık hizmetleri, yaşam kalitesi ve genel refah düzeyi hakkında kritik bir göstergedir. Cinsiyet, yaşam beklentisi üzerinde etkili olan birçok faktörün temelini oluşturan bir sosyal değişkendir. Son yıllarda, cinsiyet ve yaşam beklentisi arasındaki ilişki, pek çok araştırmanın odak noktası olmuştur.

Cinsiyete dayalı yaşam beklentisi farklılıklarını anlamak, toplumsal cinsiyet eşitsizliğinin anlaşılmasına katkıda bulunabilir. Genellikle kadınların erkeklere kıyasla daha uzun yaşam beklentisine sahip olduğu gözlenir. Bu farkın nedenleri karmaşıktır; kadınlar genellikle daha sağlıklı yaşam tarzları benimserler, bu da sigara içmeme alışkanlığı, dengeli beslenme ve düzenli tıbbi muayeneler gibi faktörlere dayalı daha sağlıklı yaşam biçimlerini içerir.

Ancak, yaşam beklentisi sadece kişisel sağlık davranışlarına dayalı bir olgu değildir. Cinsiyet rolleri, kadınların ve erkeklerin sosyal ve ekonomik hayattaki farklılıkları nedeniyle yaşam beklentisini etkileyebilir. Bu çalışmanın temel amacı, yaşam beklentisinin cinsiyete göre farklılaşmasına etki eden sosyal ve ekonomik faktörleri incelemek ve bu faktörler arasındaki ilişkileri açıklamak, böylece cinsiyet eşitsizliği konusunda bilimsel bir katkı sağlamaktır. Bu amaç doğrultusunda, Ekonomik Kalkınma ve İş birliği Örgütü'nün (OECD) "Better Life Index" verileri kullanılarak Yapısal Eşitlik Modellemesi (PLS-SEM) yöntemi uygulanmıştır. Çalışmada, cinsiyetin yaşam beklentisi üzerindeki etkileri çeşitli değişkenler aracılığıyla analiz edilmiş, bu değişkenler arasındaki ilişkiler ve etkiler incelenmiştir. Araştırma, yaşam beklentisi ve cinsiyet arasındaki karmaşık ilişkileri anlama amacı taşımaktadır.

1 Öğr. Gör. Dr., Koç Üniversitesi Tıp Fakültesi Tıbbi İstatistik Anabilim Dalı, mveznikli@ku.edu.tr, ORCID: 0000-0003-0713-1313

Elde edilen sonuçlar, kadınlar arasındaki cinayet oranı erkeklerle kıyasla daha düşük olduğundan, bu durum yaşam beklentisini olumlu yönde etkilediğini göstermektedir. Ancak, kadınlar içinde çalışma oranının erkeklerle kıyasla daha düşük olması, kadın bireylerin yaşam memnuniyetinin düşük olmasına ve yaşam beklentisinin kısılmasına yol açmaktadır. Ayrıca, kadınların gece sokakta tek başına dolaşırken kendilerini daha güvensiz hissetmeleri, yaşam beklentisini olumsuz etkileyen diğer bir faktördür.

## 1. Giriş

Yaşam beklentisi, toplumların sağlık hizmetleri, yaşam kalitesi ve genel refah düzeyi hakkında önemli bir gösterge olarak kabul edilmektedir. Cinsiyet ise, yaşam beklentisini etkileyen birçok faktörün temelini oluşturan sosyal bir değişkendir. Günümüzde cinsiyet temelli yaşam beklentisi farklılıklarının anlaşılması, toplumsal cinsiyet eşitsizliğinin azaltılmasına yönelik önemli bir gereklilik haline gelmiştir. Araştırmalar, genellikle kadınların erkeklerle kıyasla daha uzun yaşam beklentisine sahip olduğunu göstermektedir. Bu farklılık, bir dizi karmaşık etken tarafından şekillendirilir.

Bu bağlamda, bu çalışmanın temel amacı, yaşam beklentisinin cinsiyete göre değişmesine etki eden faktörleri incelemek ve bu faktörler arasındaki ilişkileri açıklamaktır. Cinsiyet temelinde yaşam beklentisindeki bu farklılıkları anlamak, toplumsal cinsiyet eşitsizliği konusundaki farkındalığı artırarak cinsiyet eşitsizliğini azaltmamıza yardımcı olabilir.

Araştırma, özellikle toplumsal cinsiyet eşitsizliği konusunda hala önemli soruların bulunduğu bir dönemde gerçekleştirilmiştir. Kadınların yaşam beklentisinde avantajlı bir konumda olmalarına rağmen, kadınlar arasında iş bulma oranının erkeklerle kıyasla daha düşük olduğu görülmektedir. Ayrıca, kadınların gece sokakta tek başına dolaşırken kendilerini daha güvensiz hissetmeleri, yaşam beklentisi üzerinde olumsuz bir etki yaratabilir. Bu çalışmanın hedefi, bu karmaşıklığı anlamak ve toplumsal cinsiyet eşitsizliğine odaklanan politika önlemlerine katkıda bulunmaktır. Bu amaç doğrultusunda, OECD'nin veri arşivinden (OECD, 2023) elde edilen "Daha İyi Yaşam" (Better Life Index) verileri kullanılarak Yapısal Eşitlik Modellemesi (PLS-SEM) yöntemi uygulanmıştır. Bu modellemenin sonuçları, yaşam beklentisi ve cinsiyet arasındaki ilişkileri daha ayrıntılı bir şekilde açığa çıkarmıştır.

Bu çalışmanın anahtar katkısı, cinsiyet temelinde yaşam beklentisini etkileyen faktörleri ve bu faktörlerin toplumsal cinsiyet eşitsizliği üzerindeki etkilerini daha iyi anlama fırsatı sunmasıdır. Elde edilen sonuçlar, cinsiyet eşitsizliğinin azaltılmasına yönelik politika önlemlerinin geliştirilmesine katkı sağlayabilir ve toplumsal cinsiyet eşitsizliği konusundaki boşlukları doldurabilir.

## 2. Literatür Taraması

Çalışmada cinsiyet ile yaşam beklentisi arasında aracılık rolü araştırılan değişkenler şu şekildedir:

- İstihdam oranı
- Uzun vadeli işsizlik oranı
- Uzun saatler boyu çalışma oranı
- Destek Ağı Kalitesi
- Eğitime Katılım
- Öğrenci becerileri
- Eğitimde geçirilen yıllar
- Bildirilen sağlık durumu
- Yaşam Memnuniyeti
- Geceleri yalnız yürürken güvende hissetme
- Cinayet oranı
- Boş ve kişisel bakıma ayrılan zaman

### 2.1. Yaşam Beklentisi

Yaşam Beklentisi değişkeni, bir toplumun sağlık durumu ve bireylerin uzun ömürlülüğü hakkında önemli bilgiler sağlayan temel göstergelerden biridir. Bu değişken, bir bireyin doğum anından itibaren kaç yıl yaşamasının beklendiğini ifade eder ve toplumların refah düzeyi, sağlık hizmetlerine erişim, çevresel faktörler ve yaşam koşulları gibi birçok faktörün birleşiminden etkilenir (OECD, 2020).

Yaşam beklentisi kavramı, tarih boyunca toplumların gelişmişlik seviyelerini belirlemek için sıklıkla kullanılmıştır. Sanayi Devrimi'nden önce, dünya genelinde yaşam beklentisi oldukça düşüktü. Ancak 20. yüzyılda modern tıbbın ve halk sağlığı önlemlerinin gelişmesiyle birlikte bu rakamlar hızla artmıştır (Cutler & Miller, 2005). Yaşam beklentisindeki bu artışın en önemli nedenleri arasında aşılama programları, antibiyotiklerin keşfi, çocuk ölümlerindeki azalma ve sağlık hizmetlerine erişimin artması gösterilmektedir.

Sağlık sistemlerinin etkinliği, yaşam beklentisinin artmasında kritik bir rol oynamaktadır. Sağlık harcamalarının artması ve bireylerin daha kaliteli sağlık hizmetlerine erişimi, gelişmiş ülkelerde yaşam beklentisinin yükselmesine

katkı sağlamıştır. İskandinav ülkelerinde kamu tarafından finanse edilen sağlık hizmetleri, yüksek yaşam beklentisi oranlarına katkıda bulunmuştur (Wilkinson & Pickett, 2010).

Yaşam beklentisi, ülkeler ve bölgeler arasında büyük farklılıklar göstermektedir. Gelişmiş ülkelerde yaşam beklentisi daha uzunken, düşük ve orta gelirli ülkelerde yaşam beklentisi daha kısadır. Bu fark, sağlık hizmetlerine erişim, ekonomik gelişmişlik, beslenme düzeyi ve çevresel faktörler gibi birçok unsurdan kaynaklanmaktadır. Örneğin, 2020 yılında Japonya, ortalama 84 yıl yaşam beklentisi ile dünyadaki en yüksek rakamlardan birine sahipken, bazı Sahra Altı Afrika ülkelerinde bu rakam 60 yılın altında kalmaktadır (*Life expectancy at birth*, 2020). AIDS, sıtma gibi bulaşıcı hastalıkların yaygınlığı ve beslenme yetersizlikleri bu bölgelerde düşük yaşam beklentisinin temel nedenleridir (Canning, 2012).

Yaşam beklentisi üzerinde cinsiyetin de belirgin bir etkisi bulunmaktadır. Çoğu ülkede kadınlar, erkeklere göre daha uzun yaşamaktadır. Bu durum, biyolojik farklılıklar, riskli davranışlara eğilim, mesleki tehlikeler ve sağlık hizmetlerine erişim farklılıklarıyla açıklanabilir (Riley, 2001). Erkekler arasında daha yüksek oranda sigara, alkol kullanımı ve tehlikeli işlerde çalışma gibi risk faktörlerinin yaygın olması, yaşam sürelerinin kısalmasına yol açmaktadır. Bununla birlikte, son yıllarda bazı gelişmiş ülkelerde kadın ve erkek yaşam beklentisi arasındaki farkın azaldığı da gözlemlenmiştir (Glei & Horiuchi, 2007).

Yaşam beklentisi sadece bireylerin sağlık hizmetlerine erişimiyle değil, aynı zamanda çevresel ve sosyal faktörlerle de yakından ilişkilidir. Özellikle temiz suya erişim, sağlıklı beslenme, iyi yaşam koşulları ve çevresel kirliliğin azaltılması, yaşam beklentisini artıran unsurlar arasındadır (Marmot & Wilkinson, 2006). Ayrıca, bireylerin sosyal bağları, stres düzeyleri ve yaşam kaliteleri de yaşam beklentisini doğrudan etkilemektedir. Örneğin, yüksek stres düzeyi ve yetersiz sosyal destek ağlarına sahip bireylerin yaşam beklentisi daha düşük olabilmektedir (Berkman & Glass, 2000).

Eğitim ve ekonomik durum da yaşam beklentisi üzerinde büyük bir etkiye sahiptir. Daha eğitilmiş ve yüksek gelir düzeyine sahip bireyler, genellikle daha uzun yaşamaktadır. Bu durum, eğitilmiş bireylerin daha sağlıklı yaşam tarzları benimseme olasılıklarının daha yüksek olması ve sağlık hizmetlerine daha kolay erişebilmeleriyle açıklanabilir (Lleras-Muney, 2005). Öte yandan, düşük gelirli gruplar arasında yaşam beklentisi düşüktür, çünkü bu bireyler sağlıklı gıdaya erişim, düzenli sağlık kontrolleri ve güvenli yaşam koşullarından mahrum kalabilmektedirler (Case & Deaton, 2015).

## 2.2. İstihdam Oranı

Bu değişken, bireylerin iş gücü piyasasına katılımını ve istihdam seviyesini ölçen önemli bir göstergedir. İstihdam oranı, 15-64 yaş arasındaki bireylerin çalışmakta olan yüzdesini ifade eder ve bir toplumun ekonomik sağlığı, refah düzeyi ve sosyal istikrarı hakkında önemli bilgiler sağlar (OECD, 2020). Bu gösterge, aynı zamanda bireylerin gelir, sosyal güvence ve yaşam standartları açısından ne durumda olduklarını yansıtmaktadır.

İstihdam oranı, ekonomik büyümenin ve refahın önemli bir bileşenidir. İstihdamdaki artış, ekonomik üretkenliği artırır, sosyal güvenlik sistemlerine katkıda bulunur ve bireylerin hane gelirlerini artırarak genel refah düzeyini yükseltir. Bu bağlamda, (Blanchard & Katz, 1992) istihdam oranının ekonomik büyüme üzerindeki pozitif etkilerini incelemiş ve yüksek istihdam seviyelerinin uzun vadeli ekonomik kalkınma için kritik olduğunu belirtmiştir. Özellikle tam istihdam seviyesine ulaşmak hem devlet bütçesindeki açıkları azaltmakta hem de bireylerin daha fazla tüketim yapmasına olanak sağlayarak iç talebi artırmaktadır (Krueger & Meyer, 2002).

İstihdam oranı, cinsiyet ve yaş gruplarına göre önemli farklılıklar göstermektedir. Kadınların iş gücüne katılım oranı, birçok ülkede erkeklere göre daha düşüktür. Bu fark, toplumsal cinsiyet rolleri, bakım sorumlulukları ve işyerindeki ayrımcılık gibi faktörlere dayanmaktadır (Goldin, 2006).

Yaş gruplarına göre incelendiğinde, genç nüfusun (15-24 yaş) iş gücüne katılım oranının genellikle düşük olduğu gözlemlenmektedir. Bu durum, gençlerin eğitim hayatına devam etmeleri ve iş bulma süreçlerinde karşılaştıkları zorluklarla ilişkilidir. Buna karşın, orta yaşlı bireyler (25-54 yaş) genellikle daha yüksek istihdam oranlarına sahiptir. Yaşlı nüfusun (55-64 yaş) iş gücüne katılımı ise birçok ülkede erken emeklilik politikaları ve işgücü piyasasındaki dezavantajlı konumları nedeniyle düşüktür (Gruber & Wise, 2002).

Dünya genelinde istihdam oranları, ülkelerin ekonomik gelişmişlik düzeylerine göre farklılık göstermektedir. Gelişmiş ülkelerde istihdam oranları daha yüksekken, gelişmekte olan ülkelerde işsizlik oranı daha yüksek ve kayıt dışı istihdam yaygındır. Özellikle Sahra Altı Afrika ve Güney Asya gibi bölgelerde, kayıt dışı ekonominin yaygınlığı, istihdam oranlarını tam olarak yansıtmayan bir diğer önemli sorundur (Gomis, Kapsos, & Kuhn, 2020). Buna karşın, Kuzey Avrupa ülkelerinde istihdam oranları yüksek olup, bu ülkeler sosyal güvenlik sistemleri ve aktif iş gücü politikalarıyla istihdamı artırmayı başarmışlardır (Esping-Andersen, 1990).



Eğitim düzeyi, istihdam oranlarını belirleyen en önemli faktörlerden biridir. Eğitim düzeyi yüksek bireylerin iş gücüne katılım oranı daha yüksek olmakta ve daha iyi ücretli, güvenli işlerde çalışma olasılıkları artmaktadır (Heckman, Lochner, & Todd, 2006). Özellikle üniversite mezunları arasında işsizlik oranları daha düşük olup, bu bireyler iş piyasasında daha kolay istihdam edilebilmektedirler. Buna karşın, düşük eğitim düzeyine sahip bireyler daha düşük ücretli ve güvencesiz işlerde çalışmakta, bu da gelir eşitsizliği ve sosyal sorunlara yol açabilmektedir (Autor, 2019).

COVID-19 pandemisi, dünya genelinde istihdam oranlarını önemli ölçüde etkilemiştir. Birçok ülkede karantina ve kısıtlamalar nedeniyle iş yerleri kapanmış, özellikle hizmet sektörü büyük darbe almıştır. Bu süreçte, özellikle kadınlar ve gençler gibi kırılgan grupların iş gücünden çekilme oranları daha yüksek olmuştur (Gomis, Kapsos, & Kuhn, 2020).

### **2.3. Uzun Vadeli İşsizlik Oranı**

Uzun Vadeli İşsizlik Oranı değişkeni, işgücü piyasasında işsizliğin süresini ve bunun toplumsal etkilerini ölçen önemli bir göstergedir. Uzun vadeli işsizlik oranı, genellikle 12 ay ve daha uzun süre işsiz kalan bireylerin toplam iş gücüne oranını ifade eder (OECD, 2020). Bu oran, bir ekonominin işgücü piyasasının esnekliği ve iş yaratma kapasitesi hakkında önemli ipuçları vermektedir. Uzun vadeli işsizlik, bireyler için gelir kaybı, beceri erozyonu ve sosyal dışlanma gibi olumsuz sonuçlar doğururken, aynı zamanda ekonomik kalkınmayı ve sosyal uyumu da olumsuz etkileyebilir.

Uzun vadeli işsizlik, genellikle ekonomik durgunluklar, yapısal değişiklikler ve iş gücü piyasasındaki dengesizliklerle ilişkilidir. Örneğin, (Ball, 2009) uzun vadeli işsizliğin ekonomik kriz dönemlerinde artış gösterdiğini ve kriz sonrası toparlanmaların yavaş olduğu dönemlerde bu işsizlik türünün yaygınlaştığını vurgulamaktadır. Ekonomik durgunluk dönemlerinde, iş fırsatlarının azalması ve iş gücü talebinin düşmesi, bireylerin uzun süre iş bulamamasına neden olmaktadır (Nickell, 1997).

Yapısal işsizlik de uzun vadeli işsizliğin önemli bir nedenidir. Yapısal işsizlik, bir ekonomideki üretim yapısının değişmesi, yeni teknolojilerin benimsenmesi ve bazı sektörlerin gerilemesi nedeniyle bazı işçilerin becerilerinin piyasada talep görmemesi durumunda ortaya çıkar (Layard, Nickell, & Jackman, 2005). Bu durumda, bireyler iş bulmakta zorlanabilir ve uzun süre işsiz kalabilirler. Özellikle dijitalleşme ve otomasyon gibi teknolojik değişimlerin iş gücü piyasasında beceri gereksinimlerini değiştirdiği göz önünde bulundurulduğunda, bu tür yapısal dönüşümler uzun vadeli işsizliği artırabilir (Brynjolfsson & McAfee, 2014).

Uzun vadeli işsizlik hem bireyler hem de toplumlar için ciddi sonuçlar doğurur. Bireysel düzeyde, işsizliğin uzun sürmesi, gelir kaybına, yaşam standartlarının düşmesine ve bireylerin psikolojik sağlığının bozulmasına neden olabilir. Ailelerde işsiz bir bireyin bulunması, hane gelirini olumsuz etkileyebilir ve bu durum, sosyal dışlanma, yoksulluk ve sağlık sorunları gibi çeşitli zorluklara yol açabilir (Clark, Georgellis, & Sanfey, 2001). Uzun vadeli işsiz kalan bireyler, iş gücüne geri döndüklerinde bile daha düşük ücretlerle çalışmak zorunda kalabilirler ve bu da kariyerlerinde kalıcı gelir kayıplarına neden olabilir (Arulampalam, 2001).

Ekonomik düzeyde ise uzun vadeli işsizlik, üretkenlik kaybına ve ekonomik büyümenin yavaşlamasına neden olabilir. Uzun süre işsiz kalan bireylerin becerilerinde erozyon meydana gelir ve bu bireylerin iş gücüne geri dönme şansı azalır (Pissarides, 1992). Ayrıca, işsizlik yardımlarına uzun süre bağımlı olan bireyler, sosyal güvenlik sistemleri üzerinde finansal baskı oluşturur ve hükümetlerin bütçe açıklarını artırır (Blanchard & Wolfers, 2000).

Uzun vadeli işsizliği azaltmak için uygulanan politikalar genellikle işsizlik sigortası, iş eğitimi programları ve aktif işgücü piyasası politikalarını içerir. Özellikle iş eğitimi programları, işsizlerin piyasada talep edilen becerilere sahip olmalarını sağlar ve iş bulma şanslarını artırır. Ayrıca, işsizlik sigortası sistemleri, uzun vadeli işsizler için gelir desteği sunarak yoksulluğu azaltabilir, ancak bu sistemlerin iş arama teşviklerini olumsuz etkileyebileceği de göz önünde bulundurulmalıdır (Card, Kluve, & Weber, 2010).

OECD ülkeleri arasında uzun vadeli işsizliği azaltmaya yönelik farklı stratejiler uygulanmaktadır. Örneğin, Danimarka'nın "Flexicurity" modeli, iş gücü piyasasında esneklik ve güvenceyi bir araya getirerek işsizlik dönemlerinde bireylerin iş bulma süreçlerini hızlandırmayı hedeflemektedir (Madsen, 2005). Bu model, iş güvencesi sağlarken aynı zamanda bireylerin hızlı bir şekilde yeniden istihdama katılmasını teşvik etmektedir.

COVID-19 pandemisi, küresel çapta iş gücü piyasasını derinden etkileyerek uzun vadeli işsizlik oranlarında artışa neden olmuştur. Pandemi süresince birçok sektör kapanmış ve iş fırsatları önemli ölçüde azalmıştır. Bu durum, özellikle hizmet sektörü ve düşük vasıflı işlerde çalışan bireyleri olumsuz etkilemiş ve uzun süre iş bulamayan bireylerin sayısında artışa neden olmuştur (OECD., 2021).

#### **2.4. Uzun Saatler Boyu Çalışma Oranı**

"Uzun Saatler Boyu Çalışma Oranı", iş-yaşam dengesinin bir göstergesi olarak bireylerin sağlığı, refahı ve yaşam memnuniyeti açısından önemli bir

faktör olarak değerlendirilir. Bu gösterge, haftada 50 saatten fazla çalışan bireylerin yüzdesini ifade eder ve ülkeler arası iş yükü, çalışma şartları ve çalışma kültürünü karşılaştırmaya olanak tanır. OECD, uzun saatler boyunca çalışmanın bireylerin sağlığı üzerinde olumsuz etkileri olduğunu ve bunun hem iş performansını hem de kişisel yaşam kalitesini düşürdüğünü vurgulamaktadır.

Uzun saatler boyunca çalışma, özellikle çalışanların fiziksel ve zihinsel sağlığı üzerinde olumsuz etkiler yaratabilir. İşyerinde geçirilen uzun saatler, bireylerin stres düzeyini artırabilir, tükenmişlik sendromuna yol açabilir ve iş-yaşam dengesini bozabilir. Ayrıca, uzun süreli çalışma temposu, sosyal ilişkilere ayrılan zamanı azaltarak bireylerin sosyal destek ağlarını zayıflatabilir (Sparks, Faragher, & Cooper, 2001). Bu durum, çalışanların genel yaşam memnuniyetinin düşmesine ve psikolojik rahatsızlıkların artmasına neden olabilir. OECD verilerine göre, uzun saatler çalışan kişilerin oranı, ülkeden ülkeye önemli ölçüde farklılık göstermektedir ve özellikle düşük iş güvencesine sahip işlerde çalışan bireyler bu durumdan daha fazla etkilenmektedir (OECD, 2020).

Uzun saatler boyunca çalışma ile ilgili literatür, iş sağlığı ve güvenliği, iş-yaşam dengesi ve üretkenlik gibi konuları kapsamaktadır. Bu alandaki erken çalışmalar, özellikle tükenmişlik sendromu ve iş stresi konularına odaklanmıştır. (Sparks, Faragher, & Cooper, 2001) uzun çalışma saatlerinin sağlık üzerindeki etkilerini inceleyen bir meta-analiz çalışmasında, uzun süreli çalışma temposunun kardiyovasküler hastalıklar, stres ve tükenmişlik gibi sağlık sorunlarına neden olabileceğini belirtmişlerdir.

Sonraki çalışmalarda ise uzun çalışma saatlerinin hem bireysel hem de toplumsal refah üzerindeki etkileri daha geniş bir perspektiften ele alınmıştır. Uzun çalışma saatlerinin, bireylerin psikolojik sağlığı kadar sosyal yaşamları üzerinde de olumsuz etkiler yarattığı bulunmuştur (Kivimäki et al., 2015). Bu çalışmada, uzun çalışma saatlerinin depresyon ve anksiyete gibi ruh sağlığı sorunlarını artırdığı ve bireylerin sosyal ilişkilerinin zayıflamasına yol açtığı belirtilmiştir.

Ayrıca, uzun süreli çalışma saatlerinin iş verimliliğini düşürebileceğine dair önemli bulgular da mevcuttur. OECD'nin raporlarında da belirtildiği gibi, aşırı çalışma saatleri iş verimliliğini düşürmekte ve bireylerin işlerine olan motivasyonlarını azaltmaktadır (OECD, 2020). Bu da iş-yaşam dengesinin korunmasının sadece bireyler için değil, aynı zamanda işverenler ve ekonomik sistemler için de önemli olduğunu göstermektedir.

## 2.5. Destek Ağı Kalitesi

Destek Ağı Kalitesi değişkeni, bireylerin yaşam memnuniyeti ve refahı açısından sosyal desteğin önemine vurgu yapan bir göstergedir. Bu gösterge, kişilerin ihtiyaç duyduklarında güvenebilecekleri bir aile üyesi, arkadaş veya komşu gibi sosyal bağlantılara sahip olup olmadıklarını ölçer. Sosyal destek ağları, bireylerin stresle başa çıkmasına, psikolojik iyi oluşlarının artmasına ve hatta yaşam beklentisinin uzamasına yardımcı olan önemli faktörler arasında yer alır.

Sosyal destek ağları hem duygusal hem de maddi destek sağlayan kişilerden oluşur ve bireylerin yaşam kalitesine doğrudan etki eder. OECD'nin bu endekste sosyal destek ağlarını ölçmesinin temel nedeni, bireylerin yaşamlarında karşılaştıkları zorlukları aşmada güçlü sosyal bağların hayati bir rol oynamasıdır. Bir kişinin sosyal ağının kalitesi, kriz zamanlarında duygusal rahatlama sağlamanın ötesine geçerek, sağlık hizmetlerine erişimi artırabilir ve kişilerin sosyal izolasyon yaşamalarını önleyebilir (OECD, 2020). Örneğin, güçlü sosyal destek ağları, depresyon ve anksiyete gibi ruh sağlığı sorunlarının azaltılmasına katkıda bulunur (Berkman & Glass, 2000). Ayrıca, sosyal destek bireylerin yalnızlık hissini azaltarak genel yaşam memnuniyetini artırır (Holt-Lunstad, Smith, & Layton, 2010).

Sosyal destek ağları üzerine yapılan çalışmalar, bu ağların bireylerin psikolojik ve fizyolojik sağlığında ne kadar etkili olduğunu göstermektedir. (Cohen & Wills, 1985), sosyal destek ile stres arasındaki ilişkiyi inceleyen meta-analiz çalışmalarında, sosyal destek ağlarının strese karşı tampon etkisi sağladığını bulmuşlardır. Özellikle aile, arkadaşlar ve iş arkadaşları gibi güvenilir sosyal bağlara sahip bireyler, yaşam zorlukları karşısında daha dirençli olabilmektedir.

Ek olarak, (Holt-Lunstad, Smith, & Layton, 2010) yürüttüğü geniş çaplı bir meta-analiz çalışması, sosyal ilişkilerin sağlık üzerindeki etkilerini araştırmış ve zayıf sosyal bağlantıların mortalite riskini artırdığını ortaya koymuştur. Bu çalışma, sosyal izolasyonun sağlığa olan olumsuz etkilerinin sigara içmek ve obezite kadar ciddi olduğunu vurgulamaktadır.

Son olarak, (Berkman & Glass, 2000), sosyal ağların sağlık üzerindeki etkilerini inceledikleri çalışmalarında, geniş sosyal destek ağlarına sahip bireylerin daha uzun yaşadıklarını ve daha düşük kronik hastalık riskine sahip olduklarını bulmuşlardır. Bu sonuçlar, sosyal desteğin yalnızca psikolojik bir fayda sağlamakla kalmayıp, aynı zamanda somut fiziksel sağlık yararları sunduğunu göstermektedir.

## 2.6. Eğitime Katılım

Eğitime katılım değişkeni, bireylerin eğitim sistemine ne ölçüde dahil olduklarını ve eğitimden aldıkları faydayı ölçen bir göstergedir. Bu değişken hem eğitime erişim hem de eğitimin tamamlanma oranlarını kapsar. Eğitime katılım oranları, bireylerin uzun vadede iş gücü piyasasındaki yerlerini, ekonomik fırsatlarını ve yaşam kalitelerini etkileyen temel faktörler arasında yer alır (OECD, 2020). Eğitimin sosyal, ekonomik ve kültürel açıdan bireylere sağladığı faydalar, uzun süreli refahı artırma açısından kritik bir öneme sahiptir.

Eğitime katılım, bireylerin okul öncesi eğitimden yükseköğretime kadar olan süreçte ne kadar süreyle eğitim aldığı ve hangi seviyede eğitimi tamamladığı ile ilgilidir. Eğitim, bireylerin iş gücü piyasasında daha iyi pozisyonlara ulaşmasına, gelir düzeylerini artırmaya ve yaşam standartlarını yükseltmesine olanak tanır. Bu bağlamda, eğitime erişim ve katılım oranları bireylerin sosyal mobilitesi için bir göstergedir (OECD, 2020). Literatürde, eğitim seviyesi yüksek olan bireylerin yaşam kalitesinin daha yüksek olduğu, sağlık ve refah açısından da daha avantajlı oldukları sıklıkla vurgulanmıştır (Barro, 1991). Ayrıca, eğitime katılım, demografik ve sosyoekonomik gruplar arasındaki eşitsizlikleri azaltma konusunda da kritik bir role sahiptir (Hanushek & Woessmann, 2015).

Eğitime katılım ve eğitimin tamamlanma oranları üzerine yapılan araştırmalar, eğitim seviyesinin bireylerin ekonomik refahı üzerindeki etkilerini açıkça ortaya koymuştur. (Barro, 1991) yaptığı çalışmada, eğitim ve ekonomik büyüme arasındaki ilişkiyi incelemiş ve eğitime katılım oranlarının yüksek olduğu ülkelerde ekonomik büyümenin daha hızlı gerçekleştiğini bulmuştur. Eğitim, bireylerin üretkenliğini artırmakla kalmayıp, aynı zamanda ekonomik eşitsizliklerin azaltılmasına da katkıda bulunmaktadır.

(Hanushek & Woessmann, 2015), eğitime katılımın bireylerin gelecekteki ekonomik başarılarını nasıl etkilediğini inceleyen çalışmalarında, eğitimin kalitesinin ve katılım oranlarının ulusal gelir düzeyi üzerindeki belirleyici etkisine vurgu yapmıştır. Bu araştırma, eğitim sistemlerinin yalnızca niceliksel değil, aynı zamanda niteliksel olarak da geliştirilmesinin ekonomik kalkınma için önemli olduğunu göstermektedir. Ayrıca, eğitimde fırsat eşitliğinin sağlanmasının, toplumlar arasındaki gelir eşitsizliğini azaltmada etkili olduğu da belirtilmiştir.

Eğitime katılımın bireylerin sağlık ve refah üzerindeki etkileri de literatürde sıklıkla ele alınmaktadır. Daha yüksek eğitim seviyelerine ulaşan bireylerin, sağlık hizmetlerine erişimlerinin daha iyi olduğu, sağlık bilgisine daha fazla

sahip oldukları ve yaşam beklentilerinin daha yüksek olduğu bulunmuştur (Cutler & Lleras-Muney, 2006). Eğitim aynı zamanda bireylerin toplumsal hayata katılımını ve vatandaşlık bilincini geliştiren önemli bir araç olarak da değerlendirilmektedir.

## 2.7. Öğrenci Becerileri

Bu değişken, bireylerin eğitim süreçlerinde kazandıkları bilişsel becerileri ve bu becerilerin iş gücü piyasasındaki etkinliğini ölçen bir göstergedir. Bu gösterge, öğrencilerin okul yıllarında edindikleri bilgi ve becerilerin kalitesini, özellikle okuma, matematik ve fen alanlarındaki performanslarını değerlendirir. Öğrenci becerileri, bireylerin iş yaşamına hazırlıklı olup olmadıklarını, gelecekteki gelir düzeylerini ve sosyal hareketliliklerini belirleyen temel faktörler arasında yer alır (OECD, 2020).

Öğrenci becerileri, yalnızca akademik başarıyı değil, aynı zamanda eleştirel düşünme, problem çözme, iletişim kurma ve analitik düşünme gibi becerileri de kapsar. Bu beceriler, öğrencilerin eğitim sonrasında iş gücüne dahil olmaları ve iş piyasasında rekabet edebilirlikleri açısından önemlidir (Hanushek & Woessmann, 2011). PISA (Programme for International Student Assessment) gibi uluslararası ölçüm araçları, ülkeler arasındaki öğrenci beceri farklarını ortaya koymakta ve bu becerilerin ekonomik ve sosyal gelişim üzerindeki etkilerini göstermektedir. Öğrenci becerileri, yalnızca bireylerin kariyer başarısı için değil, aynı zamanda toplumların uzun vadeli ekonomik büyümesi ve refahı için de kritik bir unsur olarak görülmektedir (OECD, 2019)

Öğrenci becerileri üzerine yapılan araştırmalar, bu becerilerin hem bireysel hem de toplumsal düzeyde önemli etkiler yarattığını göstermektedir. (Hanushek & Woessmann, 2011) eğitimde beceri kazanımının ekonomik büyüme üzerindeki etkilerini ele aldıkları çalışmalarında, ülkelerin bilişsel beceri düzeylerini artırarak uzun vadede ekonomik kalkınmalarını güçlendirebileceklerini bulmuşlardır. Özellikle matematik ve fen bilimlerindeki yüksek başarı, bireylerin teknoloji yoğun işlerde daha etkin çalışabilmelerine olanak tanır ve ulusal düzeyde inovasyon kapasitesini artırır (Hanushek & Woessmann, 2011).

Ayrıca, öğrenci becerileri ile yaşam memnuniyeti arasındaki ilişkiyi inceleyen çalışmalarda, yüksek bilişsel becerilere sahip öğrencilerin, eğitim sonrasında daha yüksek gelir seviyelerine ulaştıkları ve daha uzun süre iş gücünde kalabildikleri bulunmuştur (Heckman, Stixrud, & Urzua, 2006). Bu becerilerin, yalnızca ekonomik faydalar sağlamakla kalmayıp, aynı

zamanda bireylerin yaşam tatmini, sağlık ve sosyal katılım gibi çeşitli yaşam sonuçlarını da iyileştirdiği gösterilmiştir.

PISA gibi uluslararası öğrenci değerlendirmeleri, ülkeler arasındaki öğrenci beceri farklarını ortaya koyarak eğitimin kalitesini ve etkisini ölçmekte önemli bir rol oynamaktadır. Bu tür ölçümler, hükümetlerin eğitim politikalarını şekillendirmelerine ve eğitim sistemlerindeki eksiklikleri gidermelerine yardımcı olmaktadır (OECD, 2019). Örneğin, yüksek PISA skorlarına sahip ülkelerde iş gücü piyasasına daha donanımlı bireyler dahil edilmekte ve bu da uzun vadede ekonomik büyümeyi hızlandırmaktadır.

## 2.8. Eğitimde Geçirilen Yıllar

Bu değişken, bireylerin formal eğitim kurumlarında geçirdikleri ortalama süreyi ölçen bir göstergedir. Bu değişken, bir toplumdaki bireylerin eğitim süresini ve bu sürenin ekonomik, sosyal ve bireysel refah üzerindeki etkilerini değerlendirmek amacıyla kullanılır. Eğitimin uzunluğu, bireylerin iş gücüne katılımında, iş bulma fırsatlarında, gelir düzeyinde ve yaşam kalitesinde önemli bir rol oynamaktadır (OECD, 2020).

Eğitimde geçirilen yıllar, okul öncesi eğitimden üniversite eğitiminin sonuna kadar bireylerin formal eğitimde geçirdiği sürenin toplamını ifade eder. Daha fazla eğitim, bireylerin iş gücüne katılımında avantajlar sağlamanın yanı sıra sosyal ve kültürel hayata katılımlarını da olumlu yönde etkiler. Eğitim, bireylerin sosyal mobilitesini artırır ve onlara daha fazla ekonomik fırsat sunar. (Hanushek & Woessmann, 2008), eğitimde geçirilen sürenin bireylerin üretkenliğini artırarak ülkelerin ekonomik büyümesine katkıda bulunduğunu belirtmişlerdir.

Eğitimde geçirilen yılların bireylerin ekonomik ve sosyal refahı üzerindeki etkileri, birçok çalışmada geniş bir şekilde ele alınmıştır. (Barro & Lee, 2013) tarafından yapılan çalışmada, eğitimde geçirilen yılların bireylerin ekonomik üretkenliğine doğrudan etkisi olduğu vurgulanmıştır. Çalışmaya göre, eğitim süresi arttıkça bireylerin hem iş gücüne katılım oranları hem de gelir seviyeleri artmaktadır. Ayrıca, daha fazla eğitim gören bireyler, düşük eğitim seviyesine sahip bireylere kıyasla iş güvencesine daha fazla sahip olmakta ve işten çıkarılma riskleri daha düşük olmaktadır.

Ayrıca, eğitim süresinin artması, bireylerin iş dünyasında karşılaştıkları zorluklarla başa çıkma yeteneklerini ve problem çözme becerilerini geliştirir, bu da ekonomik verimliliği artırır (OECD, 2020).

Eğitimde geçirilen yılların sağlık üzerindeki etkilerine de dikkat çekilmektedir. (Cutler & Lleras-Muney, 2006), eğitimde geçirilen sürenin



bireylerin sađlığı üzerinde olumlu etkileri olduđunu ve eđitimin, bireylerin sađlık hizmetlerine eriřimini artırdığını belirtmişlerdir. Daha uzun süre eđitim gören bireyler, sađlık bilgisine daha fazla sahip olup, sađlık açısından daha bilinçli kararlar verebilmektedir. Bu durum, yaşam beklentisini artıran faktörlerden biridir.

## 2.9. Bildirilen Sađlık Durumu

Bildirilen Sađlık Durumu” deđişkeni, bireylerin kendi sađlık durumlarına iliřkin öznel deđerlendirmelerini temel alır. Bu deđerşken, bireylerin genel sađlık durumlarını “iyi” veya “çok iyi” olarak bildirme oranını ölçer. Bireylerin kendi sađlık durumlarını deđerlendirmeleri, sađlık sonuçları ve yaşam memnuniyeti gibi birçok alanda önemli bir gösterge olarak kabul edilir (OECD, 2020) Öznel sađlık deđerlendirmeleri, fiziksel sađlık kadar ruhsal sađlığı ve genel refahı da yansıtır.

Özellikle içermesine rağmen, bildirilen sađlık durumu birçok çalışma tarafından bireylerin sađlık risklerini ve yaşam beklentilerini öngörmeye etkili bir araç olarak kabul edilmiştir (Idler & Benyamini, 1997). Sađlık durumu hakkındaki bu öznel algılar, kişilerin sađlık hizmetlerine eriřimini ve sađlık davranışlarını da etkileyebilir.

Bildirilen sađlık durumu üzerine yapılan arařtırmalar, bu deđerşkenin bireylerin gelecekteki sađlık sonuçlarını tahmin etmek için güçlü bir araç olduđunu ortaya koymaktadır. (Idler & Benyamini, 1997) geniş çaplı meta-analiz çalışması, bildirilen sađlık durumunun bireylerin yaşam beklentisi üzerinde anlamlı bir etkisi olduđunu göstermiştir. Arařtırmaya göre, kendini “kötü” sađlık durumunda bildiren bireylerin ölüm riski, sađlıklı bildirenlere kıyasla daha yüksektir. Bu, öznel sađlık deđerlendirmelerinin hem fiziksel hem de ruhsal sađlıkla iliřkili olduđunu göstermektedir.

Özellikle, bireylerin sosyoekonomik durumu, bildirilen sađlık durumu üzerinde önemli bir etkiye sahiptir. Eđitim seviyesi ve gelir düzeyi gibi faktörler, bireylerin sađlıklarını nasıl algıladıklarını etkiler. OECD de bu göstergenin sosyoekonomik eşitsizliklerle güçlü bir şekilde iliřkili olduđunu ve toplumların sađlık sistemleri ile sosyal yapılarındaki dengesizlikleri yansıtabileceđini belirtmektedir (OECD, 2020).

Bildirilen sađlık durumu aynı zamanda yaşam memnuniyeti ve genel refah ile de iliřkilidir. Sađlığını iyi veya çok iyi olarak deđerlendiren bireyler, yaşam memnuniyetini de daha yüksek düzeyde bildirirler (Jylhä, 2009). Bu, sađlığın yalnızca fiziksel bir durum deđil, aynı zamanda genel refahın önemli bir bileşeni olduđuna işareti etmektedir.



## 2.10. Yaşam Memnuniyeti

“Yaşam Memnuniyeti” değişkeni, bireylerin yaşamlarından genel olarak ne kadar memnun olduklarını ölçen öznel bir göstergedir. Yaşam memnuniyeti, bireylerin hayatlarının çeşitli yönlerinden ne kadar tatmin olduklarını ifade etmelerine dayanan bir ölçüttür ve bireylerin genel refahını değerlendirmek için kritik bir unsur olarak kabul edilir. Bu değişken, ekonomik durum, sosyal ilişkiler, sağlık durumu, iş tatmini ve boş zaman gibi birçok faktörden etkilenir (Diener, 2000).

Yaşam memnuniyeti, bireylerin yaşam kalitesini öznel olarak nasıl değerlendirdiklerini ortaya koyan bir kavramdır. Bu gösterge, bireylerin maddi refahları kadar, psikolojik ve duygusal durumlarını da yansıtır. Örneğin, gelir seviyesi yüksek bireylerin yaşam memnuniyetinin daha yüksek olma eğiliminde olduğu gözlemlenmiştir, ancak bu ilişki doğrusal değildir; belirli bir gelir düzeyinden sonra, yaşam memnuniyeti üzerindeki marjinal etkiler azalmaya (Kahneman & Deaton, 2010). Ayrıca, sosyal ilişkiler, güçlü toplumsal bağlar ve sağlıklı bir yaşam tarzı, yaşam memnuniyetini olumlu yönde etkileyen diğer önemli faktörlerdir (Helliwell, Layard, & Sachs, 2019).

Araştırmalar, bu değişkenin bireylerin genel refahını tahmin etmek için güçlü bir gösterge olduğunu ortaya koymaktadır. (Diener, 2000) yaşam memnuniyetini değerlendiren bir model önererek, bireylerin mutluluklarını ve yaşamlarından memnuniyetlerini belirleyen faktörlerin hem içsel (psikolojik) hem de dışsal (maddi refah) olduğunu belirtmişlerdir. Özellikle, gelir ve yaşam memnuniyeti arasındaki ilişki üzerine yapılan çalışmalarda, maddi refahın yaşam memnuniyetini artırdığı, ancak bunun tek belirleyici faktör olmadığı vurgulanmıştır.

Ekonomik refah ile yaşam memnuniyeti arasındaki ilişkiyi inceleyen (Kahneman & Deaton, 2010), bireylerin duygusal refahının ve genel yaşam memnuniyetinin belirli bir gelir seviyesine kadar artış gösterdiğini, ancak bu gelir eşliğinden sonra artışın duraksadığını belirtmişlerdir. Bu bulgular, sadece maddi kazançların değil, aynı zamanda sosyal ilişkiler, sağlık durumu ve kişisel gelişim gibi diğer faktörlerin de yaşam memnuniyeti üzerinde önemli bir etkiye sahip olduğunu göstermektedir.

Sosyal ilişkiler ve yaşam memnuniyeti arasındaki güçlü ilişki, bireylerin topluma katılım düzeyleri, arkadaşlık ilişkileri ve aile bağları ile yaşamdan aldıkları tatminin arttığını göstermektedir (Helliwell, Layard, & Sachs, 2019). Özellikle güçlü sosyal bağlara sahip olan bireylerin yaşam memnuniyeti düzeylerinin daha yüksek olduğu, daha uzun süre sağlıklı ve mutlu bir yaşam sürdükleri gözlemlenmiştir (Putnam, 2000).

### 2.11. Geceleri Yalnız Yürürken Güvende Hissetme

Bu değişken, bireylerin yaşadıkları mahallede gece yalnız yürürken kendilerini ne kadar güvende hissettiklerini ölçen öznel bir göstergedir. Bu değişken, toplumsal güvenlik algısının ve bireylerin yaşam kalitesinin önemli bir göstergesidir. Bireylerin geceleri yalnızken kendilerini güvende hissetmeleri, sadece suç oranları ile değil, aynı zamanda toplumun sosyal uyumu, güvenlik önlemleri ve kamu hizmetlerinin etkinliği ile de ilişkilidir (Hale, 1996).

Bireylerin geceleri yalnız yürürken kendilerini güvende hissetmeleri hem psikolojik hem de sosyal refahları üzerinde doğrudan bir etkiye sahiptir. Korku ve güvensizlik hissi, bireylerin toplumsal yaşama katılımını azaltabilir, hareket özgürlüğünü kısıtlayabilir ve genel yaşam kalitelerini düşürebilir (Ferraro, 1995). Özellikle kadınlar ve yaşlılar gibi savunmasız grupların güvenlik algıları, toplumsal normlar ve kişisel deneyimlerle şekillenebilir ve bu gruplar arasında güvenlik hissi genellikle daha düşük olabilir (Pain, 2001).

Toplumsal güvenlik algısı ve bireylerin kendilerini güvende hissetmeleri üzerine yapılan araştırmalar, bu değişkenin bireylerin toplumsal yaşama katılımı ve genel mutluluğu üzerinde önemli bir etkisi olduğunu ortaya koymaktadır. (Ferraro, 1995), bireylerin suç korkusunun, maruz kaldıkları suç olaylarının ötesinde sosyal çevrelerinde algıladıkları tehditlerle şekillendiğini ve bu algının toplumsal normlar ve medya gibi dış faktörlerden de etkilendiğini belirtmiştir.

(Hale, 1996) ise korku ve suç arasındaki ilişkiyi incelediği çalışmasında, bireylerin kendilerini güvende hissetme düzeylerinin, yalnızca suç oranları ile değil, aynı zamanda toplumun genel sosyal yapısı ve güvenlik hizmetlerine olan güven ile şekillendiğini göstermiştir. Güvende hissetme algısı, bireylerin kendilerini toplumun bir parçası olarak hissetmeleri ve toplumsal hayata daha aktif bir şekilde katılmaları ile doğrudan ilişkilidir.

Özellikle kadınlar, yaşlılar ve çocuklar gibi savunmasız gruplar arasında güvende hissetme algısı daha düşük olabilir. (Pain, 2001), toplumsal cinsiyetin, bireylerin güvende hissetme algısı üzerinde önemli bir etkisi olduğunu ve kadınların gece dışarıda yalnızken daha fazla güvensizlik hissi yaşadığını belirtmiştir. Bu durum, sosyal yapıdaki eşitsizlikler ve toplumsal normlarla da desteklenmektedir.

## 2.12. Cinayet Oranı

Cinayet Oranı değişkeni, her 100.000 kişi başına düşen kasıtlı öldürme vakalarının sayısını ölçen bir göstergedir. Bu değişken, bir toplumun genel güvenlik seviyesini değerlendirmek için önemli bir kriter olarak kabul edilir. Cinayet oranı, sadece bireysel güvenlik ve suçla ilgili değil, aynı zamanda toplumun sosyal yapısı, eşitsizlik düzeyleri ve toplumsal barış ile de ilişkilidir (LaFree, 1999). Yüksek cinayet oranları, genellikle sosyal huzursuzluk, düşük sosyal uyum ve adalet sistemine olan güven eksikliği ile bağlantılıdır (Fajnzylber, Lederman, & Loayza, 2002).

Cinayet oranı, bireylerin yaşadıkları toplumda karşılaştıkları en ciddi şiddet biçimlerinden biridir. Bu oran, hem bireylerin güvenlik algılarını etkiler hem de toplumun genel yaşam kalitesine doğrudan katkıda bulunur. Cinayet oranlarının yüksek olduğu toplumlarda, sadece bireysel korku ve güvensizlik artmakla kalmaz, aynı zamanda sosyal ilişkiler de zayıflar ve ekonomik büyüme olumsuz etkilenir (LaFree, 1999). Araştırmalara göre, cinayet oranları genellikle düşük gelir düzeyi, yüksek işsizlik oranları ve toplumsal eşitsizliklerle bağlantılıdır (Fajnzylber, Lederman, & Loayza, 2002).

Cinayet oranlarının sosyoekonomik faktörlerle nasıl ilişkili olduğuna dair literatür, bu göstergesi toplumsal yapı ve refahın bir yansıması olarak ele alır. (LaFree, 1999) cinayet oranlarının genellikle sosyal eşitsizlikler ve zayıf hukuk sistemleri ile ilişkili olduğunu belirtmiştir. Yoksulluk, işsizlik ve düşük eğitim seviyeleri, cinayet oranlarını artıran temel faktörlerdir. Özellikle gelişmekte olan ülkelerde bu oranlar daha yüksektir ve bu durum, toplumsal istikrarsızlığın bir göstergesi olarak kabul edilir.

(Fajnzylber, Lederman, & Loayza, 2002) cinayet oranları ile gelir dağılımı arasındaki ilişkiyi araştırarak, gelir eşitsizliğinin yüksek olduğu toplumlarda cinayet oranlarının da daha yüksek olduğunu bulmuşlardır. Bu durum, sosyal eşitsizliklerin şiddet ve suç oranlarını artırdığına dair yaygın bir teoriyi desteklemektedir. Ayrıca, bu tür toplumlarda bireylerin devlet ve adalet sistemine olan güvenlerinin düşük olması, suç oranlarının kontrol altına alınmasını zorlaştırmaktadır.

## 2.13. Boş ve Kişisel Bakıma Ayrılan Zaman

Bu gösterge, bireylerin iş dışında kalan zamanlarını nasıl geçirdiklerini ve kişisel bakımlarına ne kadar zaman ayırdıklarını ölçer. Bu zaman, sosyal ilişkiler kurma, hobilerle ilgilenme, dinlenme ve kişisel sağlık bakımı gibi faaliyetleri kapsar (OECD, 2020).

Boş zamana ayrılan süre, bireylerin refahını etkileyen temel faktörlerden biri olarak kabul edilir. (Veblen, 1934)'in “boş zaman sınıfı” teorisine göre, boş zaman sosyal statüyle ilişkilendirilmiş ve belirli sınıfların bir ayrıcalığı olarak değerlendirilmiştir. Modern literatürde ise boş zamanın hem zihinsel hem de fiziksel sağlığı desteklediği ve genel yaşam doyumuna katkıda bulunduğu vurgulanmaktadır (De Bloom et al., 2009). Buna ek olarak, kişisel bakım da bireylerin stres yönetimi ve fiziksel sağlığını koruma açısından kritik öneme sahiptir (Henderson & Bialeschki, 2005).

Cinsiyet, kişisel bakım ve boş zamana ayrılan süre üzerinde belirleyici bir faktördür. Araştırmalar, kadınların boş zamanlarının erkeklere kıyasla daha sınırlı olduğunu ve bu zamanın genellikle ev işleriyle meşgul olma şeklinde harcandığını göstermektedir (Mattingly & Blanchi, 2003). Bu da kadınların iş-yaşam dengesi kurma noktasında erkeklere göre daha fazla zorluk yaşadığını ortaya koymaktadır.

Boş zaman ve kişisel bakım, bireylerin sağlığı üzerinde doğrudan bir etkiye sahiptir. Özellikle serbest zaman aktivitelerinin bireylerin stres düzeyini azaltmaya yardımcı olduğu ve bu sayede hem mental hem de fiziksel sağlıklarının iyileştiği belirtilmektedir (Sonntag, 2001). Ayrıca, kişisel bakım faaliyetlerinin de bireylerin refahını artırdığı ve yaşam kalitesini yükselttiği vurgulanmaktadır (Stebbins, 2016).

### 3. Yöntem

#### 3.1. Veri Toplama Araçları

Bu çalışmada, OECD'nin ülkelerin yaşam kalitesini ve refahını değerlendirmek için kullanılan ‘Daha İyi Yaşam Endeksi’ verileri kullanılmıştır. Bu endeks, ekonomik, sosyal ve çevresel faktörleri dikkate alarak bir ülkenin iyi yaşam düzeyini ölçer. Önceden toplanan ve ulusal istatistik ofisleri tarafından sağlanan verileri içermektedir.

Verisi analiz edilen 41 ülke şu şekildedir: Avustralya, Avusturya, Belçika, Kanada, Şili, Kolombiya, Kosta Rika, Çek Cumhuriyeti, Danimarka, Estonya, Finlandiya, Fransa, Almanya, Yunanistan, Macaristan, İzlanda, İrlanda, İsrail, İtalya, Japonya, Kore, Letonya, Litvanya, Lüksemburg, Meksika, Hollanda, Yeni Zelanda, Norveç, Polonya, Portekiz, Slovak Cumhuriyeti, Slovenya, İspanya, İsveç, İsviçre, Türkiye, Birleşik Krallık, Amerika Birleşik Devletleri, Brezilya, Rusya, Güney Afrika. Veri seti anketler, hükümet raporları ve resmi kaynaklardan toplanmıştır ve 41 OECD ülkesine ait kapsamlı bir bilgi kaynağını içermektedir. Örneklem, bu 41 ülkenin farklı alt gruplarına dayalı olarak oluşturulmuştur. Veri seti, ekonomi, eğitim, sağlık, güvenlik, yaşam memnuniyeti gibi ana temalara ait faktörleri içerir.

### 3.2. Veri Analizi

Verilerin analizi için Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesi (PLS-SEM) yaklaşımı, SmartPLS 3 yazılımı aracılığıyla gerçekleştirilmiştir. PLS-SEM, özellikle karmaşık yapısal ilişkilerin incelenmesi ve yüksek model karmaşıklığına sahip çalışmalarda tercih edilen bir teknik olup, geleneksel Kovaryans Tabanlı SEM (CB-SEM) yöntemine göre daha esnek varsayımlara dayanır. Bu yöntem, örneklem büyüklüğü veya veri dağılımı gibi kısıtlayıcı varsayımlar içermemesi nedeniyle, veri analizlerinde daha geniş bir uygulama alanı sağlar ve tahmin doğruluğunu artırır (Hair et al., 2021).

PLS-SEM, özellikle keşifsel araştırmalarda ve yordayıcı modellerde kullanılan bir yaklaşımdır. Modelde yer alan değişkenlerin birbirleriyle doğrudan ve dolaylı etkilerinin test edilmesini sağlar, bu da cinsiyetin yaşam beklentisi üzerindeki etkisini ve bu etkileşimin dolaylı yollarla nasıl açıklandığını incelemek için uygundur. PLS-SEM, yol analizi birlikte kullanıldığında, araştırmacılar karmaşık model yapıları oluşturabilir ve bu modeller üzerinden değişkenler arası ilişkileri daha detaylı inceleyebilirler. Yol analizi, gözlemlenen değişkenler arasındaki doğrudan ve dolaylı ilişkileri incelemek için kullanılan bir regresyon tabanlı tekniktir. Yol analizinde her bir yol katsayısı, bağımlı değişkene etki eden bir bağımsız değişkenin etkisini yansıtır. Temel olarak, doğrudan etkiler (bir değişkenin başka bir değişken üzerindeki doğrudan etkisi) ve dolaylı etkiler (bir değişkenin diğer değişken üzerindeki araçlar üzerinden etkisi) olarak ayrılır. Bu araştırmada yol analizi kullanılarak cinsiyetin yaşam beklentisi üzerindeki dolaylı etkilerin çeşitli faktörlerle nasıl şekillendiği detaylı olarak analiz edilebilmiştir.

PLS-SEM ve Yapısal Eşitlik Modelleri (SEM) literatürde, gözlemlenmeyen (latent) yapılar arasındaki ilişkileri incelemek ve değişkenler arası yapısal ilişkileri test etmek için güçlü bir araç olarak kabul edilmektedir. Bu bağlamda, yol analizi de değişkenler arasındaki belirli yol (direkt ve indirekt etkiler) ilişkilerini incelemek için etkin bir çerçeve sunar. Bu çalışmada kullanılan PLS-SEM, yaşam beklentisi gibi çok boyutlu bir çıktıyı, karmaşık değişkenler arası ilişkileri modellemek için ideal bir çözüm sağlamıştır.

Araştırma süreci; veri toplama, veri temizleme, model kurma, istatistiksel analizleri gerçekleştirme ve sonuç çıkarma aşamalarını kapsamıştır. OECD'nin güvenilir ve kapsamlı "Daha İyi Yaşam" verilerine dayalı olarak gerçekleştirilen veri toplama süreci, çalışmanın içsel ve dışsal geçerliliğini güçlendirmiştir. PLS-SEM ile gerçekleştirilen analizler, modeldeki yapısal yolların doğruluğunu ve istatistiksel anlamlılığını sağlamış, elde edilen sonuçların güvenilirliğini ve bilimsel geçerliliğini desteklemiştir.

Bootstrapping yöntemi, tahmin edilen model parametrelerinin güvenilirliğini değerlendirmek amacıyla parametrik olmayan bir yeniden örnekleme tekniği olarak kullanılmıştır. Hair ve arkadaşları, bu yöntemin tahminlerin hata varyansını azaltmak ve istatistiksel gücü artırmak için güçlü bir araç olduğunu belirtmiştir (Hair et al., 2021). SEM modellerinin en az 5000 alt örnekle çalıştırılması önerilir. Çalışmada, güvenilir sonuçlar elde etmek için modelin 10.000 alt örnekleme ve maksimum 300 yeniden örnekleme ile değerlendirilmesi yoluna gidilmiştir. İstatistiksel anlamlılık düzeyi ise 0,05 olarak kabul edilmiştir.

Bu çerçevede, PLS-SEM yöntemi ve kullanılan yol analizi, çalışmanın araştırma sorularına yanıt bulmada ve cinsiyet temelinde yaşam beklentisinin değişimini anlamada güçlü bir temel sunmuş; dolaylı etkilerin belirlenmesine olanak tanımıştır. Bu teknik seçimi, araştırmanın bilimsel geçerliliğini ve güvenilirliğini sağlamak için dikkatle yapılmıştır.

### 3.2.1. Model Varsayımları

Kurulan yapısal eşitlik modelinde, ayrışma geçerliliği HTMT değerlerine göre, Çoklu Doğrusal Bağlantı durumu ise VIF değerlerine göre kontrol edilmiştir. Ayrışma geçerliliğinin sağlanabilmesi için elde edilen HTMT değerlerinin 0,90'ın altında olması gerekir (Henseler, Ringle, & Sarstedt, 2015). 10'un üzerindeki VIF değerleriyse, modelde kabul edilemez seviyede çoklu doğrusal bağlantı sorununun bir göstergesidir (Kim, 2019).

Bu araştırmanın modeline ait HTMT değerlerinin 0 ile 0,749 arasında, VIF değerlerinin ise 1 ile 5,294 aralığında olduğu ve buna göre ciddi bir varsayım ihlali olmadığı görülmüştür.

PLS-SEM metodu yeniden örnekleme ile çalışan ve parametrik olmayan bir yöntem olduğu için, çok değişkenli normallik varsayımı, değişen varyans ve uyum iyiliği ölçütleri test edilmemiştir.

## 4. Bulgular

### 4.1. Yaşam Beklentisine Ait Tanımlayıcı İstatistikler

İlk olarak cinsiyete göre yaşam beklentisi ortalamaları incelendiğinde kadınların yaşam beklentilerinin erkeklerden daha yüksek olduğu görüldü. OECD ülkeleri için yaşam beklentisi ortalamaları kadınlar için 83,6 yıl iken erkeklerde 78,3 yıldır. OECD ülkelerine Brezilya, Rusya ve Güney Afrika eklendiğinde bu ortalamalar kadınlarda 82,99 yıl, erkeklerde 77,7 yıl olarak değişti. Türkiye'de ise kadınların yaşam beklentisi ortalaması 81,3 yıl, erkeklerde 75,8 yıldır.

## 4.2. Yaşam Beklentisi Üzerinde Erkek Cinsiyeti Lehine Olan Dolaylı Etkiler

Tasarlanan aracı değişken modeline uygulanan yol analizi sonucunda, cinsiyetin yaşam beklentisi üzerindeki dolaylı etkilerinde istihdam oranı önemli bir rol oynamaktadır. İstihdam oranı üzerinden gerçekleşen tüm dolaylı etkilerin katsayısının negatif olması (Bkz. Tablo 1), iş dünyasında erkeklere kıyasla daha düşük istihdam oranına sahip kadınların, bu durumun yaşam beklentisi üzerinde dezavantaj yaratabileceğini göstermektedir.

Modelde referans grup olarak erkekler (0) ve kadınlar (1) olarak kodlanmıştır. Bu kodlama yapısı altında, kadınların işgücüne katılım oranının erkeklere kıyasla düşük olması, onların yaşam memnuniyeti, destek ağı kalitesi ve gece yürürken güvenli hissetme gibi faktörlerde dezavantaj yaşamalarına yol açmaktadır. Bu faktörlerdeki dezavantaj, kadınların yaşam beklentisini dolaylı yoldan olumsuz etkileyebilmektedir.

*Tablo 1. Kadınların Yaşam Beklentisine Olumsuz Etki Eden Dolaylı Etkiler*

Yol (Dolaylı Etki)	Dolaylı Etki Açıklaması	B katsayısı	t değeri	p değeri
Cinsiyet → İstihdam Oranı → Yaşam Memnuniyeti → Yaşam Beklentisi	İstihdam oranı düşük olan kadınların yaşam memnuniyeti azalmaktadır, bu da yaşam beklentisini olumsuz etkileyebilir.	-0,093	2,600	0,009
Cinsiyet → İstihdam Oranı → Destek Ağı Kalitesi → Yaşam Memnuniyeti → Yaşam Beklentisi	Destek ağı kalitesinin yetersizliği yaşam memnuniyetini düşürerek kadınların yaşam beklentisine olumsuz yansiyabilir.	-0,053	2,081	0,037
Cinsiyet → İstihdam Oranı → Gece Yürürken Güvende Hissetme → Yaşam Beklentisi	Gece güvenli hissetmenin azalması, kadınların yaşam beklentisini olumsuz etkileyebilir.	-0,090	1,981	0,048

- **Yaşam Memnuniyeti Üzerinden Etki:** Kadınların istihdam oranının düşük olması, yaşam memnuniyetinde bir azalmaya yol açmakta ve bu durum yaşam beklentisini olumsuz etkilemektedir (B = -0,093, p = 0,009).
- **Destek Ağı Kalitesi Üzerinden Etki:** Yetersiz sosyal destek ağı, kadınların yaşam memnuniyetini ve dolayısıyla yaşam beklentilerini azaltabilmektedir (B = -0,053, p = 0,037).



- **Güvenlik Algısı Üzerinden Etki:** Kadınların gece yürürken güvenli hissetme oranlarının düşük olması, yaşam beklentisi üzerinde olumsuz bir etkiye yol açmaktadır ( $B = -0,090$ ,  $p = 0,048$ ).

Bu bulgular, kadınların iş gücüne katılımının artmasının, onların yaşam beklentisi üzerindeki dolaylı etkilerini iyileştirebileceğine işaret etmektedir. Modelin sonuçları, cinsiyet farklılıklarının yaşam beklentisi üzerindeki etkilerini analiz etmek için kullanılan yol analizinin kapsamlı bir görünümünü sunmaktadır.

OECD Ülkeleri için istihdam oranları kadınlarda %59 erkeklerde %74; OECD ülkelerine ek olarak Brezilya, Rusya ve Güney Afrika için istihdam oranları kadınlarda %62 erkeklerde %73; Türkiye için ise istihdam oranları kadınlarda %30 erkeklerde %65 idi.

Yapılan analizler sonucunda, bireylerin istihdam oranı yükseldikçe, destek ağı kalitelerinin de yani kişilerin ihtiyaç anında güvенеbilecekleri birini tanıma olasılıklarının da artma eğilimi olduğu görüldü. Çalışmaya katılan 41 ülke içerisinde istihdam oranı en yüksek 5 ülkede (İsviçre, İzlanda, Hollanda, Almanya, Japonya) katılımcıların ortalama %93'ü destek alabilecekleri bir sosyal çevre ağına sahip olduklarını belirtirken, en düşük 5 ülkede (Güney Afrika, Türkiye, Kosta Rika, Şili, Yunanistan) bu oran 84,4'e düştü.

Destek Ağı kalitesinin ise "Yaşam Memnuniyeti" değişkeni üzerinde pozitif etkisi görüldü. Yüksek kalitede sosyal ağın görüldüğü 5 ülkede (İzlanda, İrlanda, Finlandiya, Çek Cumhuriyeti ve Norveç) katılımcıların yaşam memnuniyeti skoru 10 üzerinden ortalama 7,34 iken, sosyal çevre bağlarının en zayıf olduğu 5 ülkede (Kosta Rika, Meksika, Kolombiya, Kore, Yunanistan) bu ortalama 5,92'ye düştü.

Bu dolaylı etkinin zincirinin son halkası olarak ise yaşam memnuniyetinin yaşam beklentisi üzerindeki pozitif etki görüldü. Yaşam memnuniyetinin en yüksek olduğu 5 ülkede (İzlanda, Hollanda, Finlandiya, Danimarka ve İsviçre), yaşam beklentisi ortalama 82,6 sene iken, yaşam memnuniyetsizliğinin en yüksek olduğu 5 ülkenin (Güney Afrika, Türkiye, Rusya, Kolombiya ve Yunanistan) vatandaşlarında ortalama ömür 74,9 olarak hesaplandı.

Diğer yandan, istihdam oranı yüksek olan toplumlarda, vatandaşların geceleri yalnız yürürken daha yüksek oranlardan kendilerini güvende hissettikleri görüldü. İsviçre, İzlanda, Hollanda, Almanya, Japonya vatandaşlarının ortalama %81,4'ü geceleri tek başına güvenle yürüyebiliyorken Güney Afrika, Türkiye, Kosta Rika, Şili, Yunanistan için bu oran %51,2 idi.

Vatandaşların geceleri en yüksek güvenle tek başına yürüyebildikleri 5 ülkede (Norveç, Slovenya, Finlandiya, Lüksemburg, İsviçre) katılımcıların



yaşam memnuniyeti skoru 10 üzerinden ortalama 7,32 iken, geceleri yalnız yürümeye en çok çekinilen ülkelerde (Kosta Rika, Meksika, Kolombiya, Şili, Brezilya) bu ortalama 5,9'a düştü.

İstihdam oranı yüksek olan toplumlarda, vatandaşların geceleri yalnız yürürken daha yüksek oranlardan kendilerini güvende hissettikleri görüldü. İsviçre, İzlanda, Hollanda, Almanya, Japonya vatandaşlarının ortalama %81,4'ü geceleri tek başına güvenle yürüyebiliyorken Güney Afrika, Türkiye, Kosta Rika, Şili, Yunanistan için bu oran %51,2 idi. Ayrıca 41 ülkedeki erkeklerin %80,51'i geceleri güvenli bir şekilde tek başlarına yürüyebildiklerini belirtirken, gece dışarı çıkan kadınların sadece %64,16'sı kendilerini güvende hissediyor.

Vatandaşların geceleri en yüksek güvenle tek başına yürüyebildikleri 5 ülkede (Norveç, Slovenya, Finlandiya, Lüksemburg, İsviçre) katılımcıların yaşam beklentisi 82,68 yıl iken, geceleri yalnız yürümeye en çok çekinilen ülkelerde (Kosta Rika, Meksika, Kolombiya, Şili, Brezilya) bu ortalama 75,26 yıla düştü.

Bu iki dolaylı etkiyi birlikte değerlendirmek gerekirse, kadınların işgücüne katılımının erkeklere kıyasla daha düşük olması, onların güçlü bir sosyal çevre edinme ve geceleri güvenli bir şekilde dışarıya çıkabilme oranlarını düşürmekte, bu olumsuz durumlar da kadınların yaşam memnuniyetini düşürmektedir. Yaşam memnuniyetindeki düşüş, kadınların yaşam beklentisi süresini dolaylı yoldan olumsuz etkileyebilir.

### **4.3. Yaşam Beklentisi Üzerinde Kadın Cinsiyeti Lehine Olan Dolaylı Etkiler**

Yapısal eşitlik modeli ile gerçekleştirilen yol analizinde, cinsiyetin yaşam beklentisi üzerindeki dolaylı etkilerinde yaşam memnuniyeti ve destek ağı kalitesi öne çıkan faktörler olarak belirlenmiştir. Her iki dolaylı etkinin de katsayısı pozitif olup, bu durum kadınların yaşam memnuniyeti ve sosyal destek ağlarının yaşam beklentisi üzerinde olumlu bir etki yarattığını göstermektedir (Bkz. Tablo 2).

Modelde, referans grup olarak erkekler (0) ve kadınlar (1) olarak kodlanmıştır. Bu kodlama yapısı altında, kadınların erkeklere kıyasla daha yüksek yaşam memnuniyeti ve destek ağı kalitesine sahip olmaları, onların yaşam beklentisini olumlu yönde etkilemektedir.

Tablo 2. Kadınların Yaşam Beklentisine Olumlu Etki Eden Dolaylı Etkiler

Yol (Dolaylı Etki)	Dolaylı Etki Açıklaması	B katsayısı	t değeri	p değeri
Cinsiyet → Yaşam Memnuniyeti → Yaşam Beklentisi	Yaşam memnuniyeti erkeklere kıyasla daha yüksek olan kadınların, yaşam beklentisi artış eğilimi gösterebilir.	0,11	2,195	0,028
Cinsiyet → Destek Ağı Kalitesi → Yaşam Memnuniyeti → Yaşam Beklentisi	Güçlü destek ağına sahip olan kadınların yaşam memnuniyetinin artmasıyla, yaşam beklentilerinde olumlu bir dolaylı etki görülebilir.	0,076	2,263	0,024

- **Yaşam Memnuniyeti Üzerinden Etki:** Kadınların yaşam memnuniyetinin erkeklere kıyasla daha yüksek olması, onların yaşam beklentisini arttırıcı bir etki yapmaktadır (B = 0,11, p = 0,028).
- **Destek Ağı Kalitesi Üzerinden Etki:** Kadınların çevrelerinde güvenebilecekleri kişilerin varlığı, onların yaşam memnuniyetini ve dolayısıyla yaşam beklentisini olumlu yönde etkilemektedir (B = 0,076, p = 0,024).

41 ülkenin tamamı dahil olmak üzere OECD verileri kullanılarak yapılan analizde, kadınların yaşam memnuniyeti skoru ortalaması 10 üzerinden 6,63 olarak hesaplanırken, erkeklerde bu oran 6,58 olarak bulunmuştur (B=0,24, t=2,49, p=0,013). Bu anlamlı fark, kadınların yaşam beklentisinin erkeklere kıyasla daha yüksek olmasında yaşam memnuniyetinin aracı bir rol oynayabileceğini göstermektedir.

Araştırmada elde edilen verilere göre, kadınların %91,41'i hayatlarında güvenebilecekleri en az bir kişinin olduğunu belirtmiştir; bu oran erkeklerde %90,41 olup, iki değişken arasındaki yola ait katsayı istatistiksel olarak anlamlıdır (B=0,397, t=4,408, p<0,001). Buna göre çevreleriyle güçlü sosyal bağlar kurmaları, kadınların yaşam memnuniyetini artırarak onların yaşam beklentisi üzerinde olumlu bir etki gösterebilmektedir.

Bu bulgular, kadınların yaşam beklentisini etkileyen faktörlerin belirlenmesi açısından değerli bilgiler sunmakta olup, yaşam memnuniyeti ve destek ağı kalitesinin cinsiyet farklılıklarına dayalı yaşam beklentisi üzerindeki etkilerini daha net bir şekilde ortaya koymaktadır.

Yapılan yol analizi sonucunda cinsiyet ve yaşam beklentisi arasında; uzun saatler boyu çalışma oranı, eğitime katılım, öğrenci becerileri, eğitimde geçirilen yıllar, bildirilen sağlık durumu, cinayet oranı, boş ve kişisel bakıma

ayrılan zaman değişkenleri üzerinden gerçekleşen istatistiksel olarak anlamlı bir endirekt etki tespit edilmemiştir.

## 5. Sonuç ve Tartışma

Bu çalışma, cinsiyet temelinde yaşam beklentisi üzerindeki faktörleri inceleyerek önemli içgörüler sunmaktadır ve bu bulgular literatürle ilişkilendirildiğinde daha da anlamlı hale gelmektedir.

Öncelikle, cinsiyet temelinde yaşam beklentisi farklılıkları konusu literatürde geniş bir şekilde ele alınmıştır. Araştırmalar, kadınların genellikle erkeklere kıyasla daha uzun yaşam beklentilerine sahip olduğunu ortaya koymaktadır (Oksuzuyan et al., 2008). Bu bulgu, bu çalışmanın sonuçlarıyla uyumludur; nitekim elde edilen veriler kadınların yaşam beklentisindeki bu avantajı doğrulamaktadır.

Cinsiyet temelinde iş gücüne katılım ve iş bulma oranları da toplumsal cinsiyet eşitsizliği alanındaki literatürle yakından ilişkilidir. Kadınların işgücünde daha az temsil edilmeleri, sosyal çevre kalitelerini ve güven duygusunu zayıflatarak yaşam memnuniyetini azaltabilmektedir (York & Bell, 2014). Bu sonuç, cinsiyet temelindeki işgücü eşitsizliğinin kişilerin yaşam deneyimlerini nasıl etkileyebileceğini ortaya koymaktadır.

Aynı şekilde, istihdam oranları ile sosyal ağlar arasındaki ilişki de belirgin bir şekilde görülmüştür. İstihdam oranının yüksek olduğu ülkelerde, kişilerin sosyal ağlarının daha güçlü olduğu ve bu ağların yaşam memnuniyetini artırdığı gözlenmektedir (Montgomery, 2022). Bu durum, iş piyasasının ve kişiler arası ilişkilerin yaşam memnuniyeti üzerindeki etkisinin altını çizmektedir.

Güvenlik algısı da cinsiyet eşitsizliği alanında önemli bir konu olarak ele alınmaktadır. Araştırmalar, kadınların özellikle gece saatlerinde kendilerini güvensiz hissetme olasılıklarının erkeklere göre daha yüksek olduğunu göstermiştir (Pain, 1997). Bu çalışma sonucunda istihdam oranı yüksek toplumlarda, bireylerin gece dışarı çıkarken kendilerini daha güvende hissettikleri tespit edilmiştir. Bu durum istihdamda dezavantajlı durumda bulunan kadınların gece dışarı tek başlarına çıkarken daha fazla endişe duymaya meyilli oldukları gibi bir neticeyi yansıtmaktadır. Nitekim cinsiyet temelindeki bu güvenlik farklılıkları, yaşam memnuniyetini etkileyen önemli bir faktör olarak da göze çarpmaktadır (Fisher et al., 2003). Bu çalışma da kadınların gece sokakta daha güvensiz hissettiklerini ve bu hissin yaşam beklentisini olumsuz etkileyebileceğini doğrulamıştır.

Son olarak, çalışmada yaşam memnuniyeti ve cinsiyet arasındaki ilişki de incelenmiştir. Elde edilen sonuçlar, kadınların erkeklere kıyasla daha yüksek yaşam memnuniyetine sahip olduklarını göstermektedir (*Research Demystifies the Gender Gap in Life Satisfaction*, 2022). Bu fark, cinsiyet temelindeki yaşam beklentisi farklılıklarını daha iyi anlayabilmemiz için önemli bir bulgudur (Lawrence, Rogers, & Wadsworth, 2015; Steptoe, Deaton, & Stone, 2015).

Bu bulguların önemi büyüktür, nitekim cinsiyet eşitsizliği toplumsal bir sorundur ve bu çalışma, cinsiyet temelinde yaşam beklentisinin nasıl etkilendiğine dair önemli veriler sunmaktadır. Bu tür bilgiler, politikacıların ve toplum liderlerinin cinsiyet eşitsizliği ile mücadele eden stratejiler geliştirmelerine yardımcı olabilir. Ayrıca, bu çalışma, cinsiyet temelinde yaşam beklentisini etkileyen faktörlerin karmaşıklığını anlamamıza yardımcı olacak daha fazla araştırmaya ihtiyaç duyulduğunu da vurgulamaktadır.

## Kaynakça

- Arulampalam, W. (2001). Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal*, 111(475), 585-606.
- Autor, D. H. (2019). Work of the Past, Work of the Future. AEA Papers and Proceedings,
- Ball, L. M. (2009). *Hysteresis in unemployment: old and new evidence*.
- Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, 106(2), 407-443.
- Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, 104, 184-198.
- Berkman, L. F., & Glass, T. (2000). Social integration, social networks, social support, and health. *Social epidemiology*, 1(6), 137-173.
- Blanchard, O., & Katz, L. F. (1992). YRegional Evolutions, \$ Brooking Papers on Economic Activity. *Economic Studies Program, The Brookings Institution*, 23(1), 76.
- Blanchard, O., & Wolfers, J. (2000). The role of shocks and institutions in the rise of European unemployment: the aggregate evidence. *The Economic Journal*, 110(462), 1-33.
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & company.
- Canning, D. (2012). Progress in Health around the World. *The Journal of Development Studies*, 48(12), 1784-1798.
- Card, D., Kluve, J., & Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, 120(548), F452-F477.
- Case, A., & Deaton, A. (2015). Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proceedings of the national academy of sciences*, 112(49), 15078-15083.
- Clark, A., Georgellis, Y., & Sanfey, P. (2001). Scarring: The psychological impact of past unemployment. *Economica*, 68(270), 221-241.
- Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological bulletin*, 98(2), 310.
- Cutler, D., & Miller, G. (2005). The role of public health improvements in health advances: the twentieth-century United States. *Demography*, 42(1), 1-22.
- Cutler, D. M., & Lleras-Muney, A. (2006). Education and health: evaluating theories and evidence. In: National bureau of economic research Cambridge, Mass., USA.
- De Bloom, J., Kompier, M., Geurts, S., De Weerth, C., Taris, T., & Sonnentag, S. (2009). Do we recover from vacation? Meta-analysis of vacation effects on health and well-being. *Journal of occupational health*, 51(1), 13-25.

- Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American psychologist*, 55(1), 34.
- Esping-Andersen, G. (1990). The three worlds of welfare capitalism. *Polity*.
- Fajnzylber, P., Lederman, D., & Loayza, N. (2002). Inequality and violent crime. *The journal of Law and Economics*, 45(1), 1-39.
- Ferraro, K. F. (1995). Fear of crime: Interpreting victimization risk. In: State University of New York Press.
- Fisher, B. S., Daigle, L. E., Cullen, F. T., & Turner, M. G. (2003). Reporting sexual victimization to the police and others: Results from a national-level study of college women. *Criminal justice and behavior*, 30(1), 6-38.
- Glei, D. A., & Horiuchi, S. (2007). The narrowing sex differential in life expectancy in high-income populations: effects of differences in the age pattern of mortality. *Population studies*, 61(2), 141-159.
- Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family.
- Gomis, R., Kapsos, S., & Kuhn, S. (2020). World employment and social outlook: trends 2020. *ILO: Geneva, Switzerland*, 127.
- Gruber, J., & Wise, D. A. (2002). Social security programs and retirement around the world: Micro estimation. In: National Bureau of Economic Research Cambridge, Mass., USA.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. SAGE Publications.
- Hale, C. (1996). Fear of crime: A review of the literature. *International review of Victimology*, 4(2), 79-150.
- Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of economic literature*, 46(3), 607-668.
- Hanushek, E. A., & Woessmann, L. (2011). The economics of international differences in educational achievement. *Handbook of the Economics of Education*, 3, 89-200.
- Hanushek, E. A., & Woessmann, L. (2015). *The Knowledge Capital of Nations Education and the Economics of Growth*. The MIT Press. <http://www.jstor.org/stable/j.ctt17kk9kq>
- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. *Handbook of the Economics of Education*, 1, 307-458.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, 24(3), 411-482.

- Helliwell, J. F., Layard, R., & Sachs, J. D. (2019). Happiness and community: An overview. *World Happiness Report 2019*.
- Henderson, K. A., & Bialeschki, M. D. (2005). Leisure and active lifestyles: Research reflections. *Leisure sciences*, 27(5), 355-365.
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the academy of marketing science*, 43, 115-135.
- Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS medicine*, 7(7), e1000316.
- Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, 21-37.
- Jylhä, M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social science & medicine*, 69(3), 307-316.
- Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the national academy of sciences*, 107(38), 16489-16493.
- Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean J Anesthesiol*, 72(6), 558-569. <https://doi.org/10.4097/kja.19087>
- Kivimäki, M., Jokela, M., Nyberg, S. T., Singh-Manoux, A., Fransson, E. I., Alfredsson, L., Bjorner, J. B., Borritz, M., Burr, H., & Casini, A. (2015). Long working hours and risk of coronary heart disease and stroke: a systematic review and meta-analysis of published and unpublished data for 603 838 individuals. *The lancet*, 386(10005), 1739-1746.
- Krueger, A. B., & Meyer, B. D. (2002). Labor supply effects of social insurance. *Handbook of public economics*, 4, 2327-2392.
- LaFree, G. (1999). A summary and review of cross-national comparative studies of homicide. *Homicide: A Source book of Social research/Sage*.
- Lawrence, E. M., Rogers, R. G., & Wadsworth, T. (2015). Happiness and longevity in the United States. *Social science & medicine*, 145, 115-119.
- Layard, R., Nickell, S., & Jackman, R. (2005). *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199279166.001.0001>
- Life expectancy at birth*. (2020). World Bank. <https://data.worldbank.org/indicator/SP.DYN.LE00.IN>
- Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *The Review of economic studies*, 72(1), 189-221.
- Madsen, P. K. (2005). *How can it possibly fly?: The paradox of a dynamic labour market in a Scandinavian welfare state*. Citeseer.



- Marmot, M., & Wilkinson, R. (2006). *Social Determinants of Health* Oxford University Press: New York. In: USA.
- Mattingly, M. J., & Blanchi, S. M. (2003). Gender differences in the quantity and quality of free time: The US experience. *Social forces*, 81(3), 999-1030.
- Montgomery, M. (2022). Reversing the gender gap in happiness. *Journal of Economic Behavior & Organization*, 196, 65-78.
- Nickell, S. (1997). Unemployment and labor market rigidities: Europe versus North America. *Journal of Economic perspectives*, 11(3), 55-74.
- OECD. (2019). *PISA 2018 Results (Volume I)*. <https://doi.org/doi:https://doi.org/10.1787/5f07c754-en>
- OECD. (2020). *How's Life? 2020 Measuring Well-being*. <https://doi.org/doi:https://doi.org/10.1787/9870c393-en>
- OECD. (2023). *OECD Data Explorer - Archive*. OECD. Retrieved 5 Ağustos 2023 from [https://data-explorer.oecd.org/vis?tenant=archive&df\[ds\]=DisseminateArchiveDMZ&df\[id\]=DF\\_BLI&df\[ag\]=OECD&dq=...&to\[TIME\]=false](https://data-explorer.oecd.org/vis?tenant=archive&df[ds]=DisseminateArchiveDMZ&df[id]=DF_BLI&df[ag]=OECD&dq=...&to[TIME]=false)
- OECD. (2021). *OECD Employment Outlook 2021 Navigating the COVID-19 Crisis and Recovery*. OECD publishing.
- Oksuzyan, A., Juel, K., Vaupel, J. W., & Christensen, K. (2008). Men: good health and high mortality. Sex differences in health and aging. *Aging clinical and experimental research*, 20, 91-102.
- Pain, R. (2001). Gender, race, age and fear in the city. *Urban studies*, 38(5-6), 899-913.
- Pain, R. H. (1997). Social geographies of women's fear of crime. *Transactions of the Institute of British geographers*, 231-244.
- Pissarides, C. A. (1992). Loss of skill during unemployment and the persistence of employment shocks. *The quarterly journal of economics*, 107(4), 1371-1391.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon Schuster.
- Research Demystifies the Gender Gap in Life Satisfaction*. (2022). Gallup. Retrieved 10 Ekim from <https://news.gallup.com/opinion/gallup/402839/research-demystifies-gender-gap-life-satisfaction.aspx>
- Riley, J. C. (2001). *Rising life expectancy: a global history*. Cambridge University Press.
- Sonnentag, S. (2001). Work, recovery activities, and individual well-being: A diary study. *Journal of occupational health psychology*, 6(3), 196.



- Sparks, K., Faragher, B., & Cooper, C. L. (2001). Well-being and occupational health in the 21st century workplace. *Journal of occupational and organizational psychology*, 74(4), 489-509.
- Stebbins, R. A. (2016). *Leisure and positive psychology: Linking activities with positiveness*. Springer.
- Stephoe, A., Deaton, A., & Stone, A. A. (2015). Subjective wellbeing, health, and ageing. *The lancet*, 385(9968), 640-648.
- Veblen, T. (1934). The theory of the leisure class: an economic study of institutions. 1899. *New York: The Modern Library*.
- Wilkinson, R., & Pickett, K. (2010). The spirit level. *Why equality is better for everyone*.
- York, R., & Bell, S. E. (2014). Life satisfaction across nations: The effects of women's political status and public priorities. *Social Science Research*, 48, 48-61.

## Avrupa Birliği Ülkelerinde Temel Gelir Sorgulamasında Çoklu Uyum İyiliği Analizi Kullanımı

Zehra Zeynep Şahinbaşoğlu<sup>1</sup>

### Özet

Çok değişkenli istatistiksel analizler birden fazla değişkenin incelendiği bir araştırma yöntemidir. Bu yöntem ile değişkenlerin birbirleriyle olan ilişkileri ve etkileşim içerisinde olup olmadıkları incelenir. Bu yöntemde tek değişkenli istatistiksel analizlerden farklı olarak, veri olarak kabul edilen birçok faktör birer değişken olarak sisteme dahil edilir. Bu sayede değişkenler arasındaki ilişkiler daha ayrıntılı ve detaylı bir şekilde edilir.

Özellikle kategorik verilerin analizinde çoklu uyum analizi kullanılması uygun ve açıklayıcı sonuçlar elde edilmesini sağlar. Elde edilen sayısal sonuçların yanı sıra, çizilen grafikler sayesinde zengin bir bilgi elde edilmiş olur ve elde edilen bu bilgiler karar verme sürecinde oldukça etkilidir. Grafikler sayesinde değişkenler arasındaki yakınlıklar (ilişkiler) bir harita üzerinde, araştırmacıya görsel olarak tablolarda ifade edilemeyen bilgiler sunar.

Uygulaması gerçekleştirilen veri setine Kaggle’s veri tabanından ve “Basic Income” linkinden ulaşılabilir. Veriler Dalia Research tarafından 2016 senesinde Avrupa Birliğine üye ülkelere anket çalışması yapılarak elde edilmiştir. Bu ankette Avrupa genelinde halka “Temel Gelir” fikrine nasıl baktıkları, eğer sandığa gidilirse bu fikre oy verip vermeyecekleri sorulmuştur. Temel Gelir, kişilerin diğer gelir kaynakları ya da çalışıp çalışmadıklarına bakılmaksızın, hükümet tarafından her bir bireye koşulsuz olarak ödenen bir gelirdir. Diğer sosyal güvenlik ödemelerinin yerini alır ve tüm ihtiyaçları karşılayacak kadar yüksek olduğu düşünülmüştür. Bu çalışmadaki amaç, bu sistemi destekleyen katılımcıların tüm değişkenlere göre hangi kategorilerde yer aldığının araştırılmasıdır.

1 Öğr. Gör. Dr., Özyeğin Üniversitesi, zeynepsahinbasoglu@gmail.com,  
ORCID ID:0000-0001-6848-6447

Bu arařtırmaya 28 Avrupa Birliđine üye devlet katılmıřtır. Katılımcılar 14 – 65 yař aralıđında, kadın ve erkeklerden oluřmaktadır ve 9649 kiřiye uygulanmıřtır.

Elde edilen sonuřlara göre katılımcıların üçte biri bu sistemi desteklemektedir. Bunun yanı sıra bu fikre oy vermeyecek kiřilerin konu hakkında yeterli bilgisi olmayan ve kırsal kesimde yařayan kiřiler ile yüksek gelire ve eđitime sahip kiřiler olduđu görülmüřtür.

## 1. Giriř

Avrupa Birliđi (AB) ülkelerinde, Temel Gelir uygulaması son yıllarda önemli bir tartıřma ve arařtırma konusu haline gelmiřtir. Bu tür bir politikayı deđerlendirmek amacıyla kullanılan çeřitli yöntemler arasında, Çoklu Uyum İyiliđi Analizi özellikle sosyal politikaların çok boyutlu etkilerini anlamak için giderek daha fazla ilgi görmekte ve kullanılmaktadır.

Temel Gelir, belirli bir gelir seviyesinin, tüm vatandařlara, belirli kořullara bađlı olmaksızın, eřit olarak sađlanması fikridir. AB ülkeleri arasında, bu öneri bazen sosyal eřitlik, iř gücü piyasası, gelir dađılımı, yoksullukla mücadele ve bireysel özgürlüklerin artırılması gibi çok farklı hedeflere yönelik bir araç olarak tartıřılmaktadır.

Temel Gelir uygulamasının başarıyla hayata geçirilip geçirilemeyeceđi ve bu tür politikaların potansiyel etkilerinin anlaşılmasında, karmařık sosyal, ekonomik ve psikolojik dinamiklerin dikkate alınması gereklidir.

Bu bilgiler dođrultusunda Avrupa Birliđi ülkelerinde temel gelir sorgulaması ekonomik boyutta, sosyal boyutta çoklu uyum analizi kullanılarak incelenmiřtir. Avrupa Birliđi ülkelerinde Temel Gelir tartıřmaları, her bir ülkenin kendi ekonomik, sosyal ve kültürel yapısına bađlı olarak farklı dinamikler gösterir. Bununla birlikte, temel gelir önerisinin uygulanabilirliđini ve etkilerini anlamada Çoklu Uyum İyiliđi Analizinin kullanılması, farklı paydařların perspektiflerinden faydalanmayı sađlar. Farklı deđerkenlerin analize dahil edilmesi, karmařık sosyal politikaların daha dođru ve kapsamlı bir řekilde deđerlendirilmesini sađlar. Bu yaklařım, örneđin Temel Gelir gibi toplumsal ve ekonomik etkileri geniş kapsamlı olabilecek politikaların etkilerini çok daha derinlemesine anlaşılmasına yardımcı olur.

Ekonomik faktörler, temel gelir gibi sosyal politikaların temel etkilerini anlamada kritik bir rol oynar. Bu deđerkenler, politika uygulamalarının bireylerin gelir seviyeleri, iř gücü piyasası, vergi sistemleri ve sosyal güvenlik üzerine etkilerini deđerlendirirken önemlidir.

## 2. Çoklu Uyum Analizi Tanımı ve Varsayımları

İki ya da daha fazla kategorik değişken arasındaki ilişkinin grafiksel olarak ifade edilmesi Uyum analizi (Correspondence Analysis-CA) ile elde edilebilir. Çoklu Uyum Analizi (Multiple Correspondence Analysis –MCA) CA'nın genelleştirilmiş halidir. MCA 0 ve 1'lerden oluşan indikatör(gösterge) matrisine basit uyum analizi uygulanması ile elde edilir. Yapılması gereken açıklanan varyans yüzdelерinin iyileştirilmesi ve uyum analizi sonucu tahmin edilen değişken uzaklıklarının yorumlanmasıdır(Abdi & Béra, 2014).

Çoklu Uyum Analizi nominal değişkenlerden oluşan gözlemleri açıklamak için kullanılır. Her nominal değişken birkaç seviye içerir ve bu seviyelerin her biri ikili değişken olarak kodlanır. Örneğin cinsiyet değişkeni Kadın ve Erkek olarak iki seviye içerir. Yani cinsiyet değişkeni 2 seviyeli bir nominal değişkendir (Nominal değişkenler kalitatif sınıflandırmalarda kullanılırlar. Örneğin; ırk, cinsiyet, renk, vb.).

Süreç iki adımdan oluşur:

- Değişkenlerin birinin kategorileri dikkate alınarak çözümleme yapılırken; ikinci aşamada bu durum diğer değişkenler içinde uygulanır.
- Satır ve sütun içinde yapılan işlemler kendi aralarında da üç adımdan oluşur (Hoffman & de Leeuw, 2011):
  - a. Kategorik profiller
  - b. Ağırlıklar (marjinal olasılıklar)
  - c. Kategorik profiller arasındaki uzaklıklar

Bu adımlar sonrasında en iyi uyumu gösterecek  $\min [(satur\ sayısı-1), [(sütun\ sayısı-1)]]$  boyutlu uzay elde edilir ve bu uzayın genellikle ilk iki boyutundaki koordinatlar kullanılarak; satır ve sütun arasındaki ilişkiyi daha basit bir şekilde açıklayan grafikler elde edilir. Uyum analizinin boyut bulma sürecinde Temel Bileşenler Analizinden yararlanır (Benammou & Saporta, 2003). Dolayısıyla Uyum Analizi aynı zamanda “negatif değerli veri içermeyen bir veri matrisini çözümlmek için tekil değer ayrıştırmasını kullanan bir yöntem” olarak da ifade edilebilir.

İlk olarak satır ve sütun profilleri hesaplanmakta, ikinci aşamada satır ve sütun profilleri iki-boyutlu uzayda ayrı ayrı resmedilmekte ve son aşamada ise satır ve sütun profilleri iki boyutlu ortak bir harita üzerinde gösterilmektedir. Uyum analizinde harita adı verilen bu grafiklerin gözlemlenen frekanslara göre değil, frekansların satır içindeki nispi önemini gösteren satır profillerine

ve benzer şekilde frekansların sütun içindeki nispi önemini gösteren sütun profillerine göre çizildiđi görölmektedir. Uyum analizi hesaplanırken dört temel hesaplama kullanılır: profil, mass, ki-kare uzaklıđı ve toplam inertia (toplam deđişkenlik) (Hoffman & de Leeuw, 2011). Amaç, modeldeki varyansı yani inertia'nın büyük bir bölümünü minimum sayıda boyutta açıklamaktır.

Kontenjans tablosunu meydana getiren satır ve sütunları oluşturan deđerleri, grafiksel gösterimde “noktalar(points)” olarak ifade edilir. Yani deđerşkenlerden birinin bir kategorisi olarak tanımlanır. Birbirine benzer kategoriler uzayda birbirine daha yakın noktalar olarak belirirken, birbirlerinden farklı kategoriler bu uzayda birbirlerine uzak noktalar olarak belirirler. Bu şekilde deđerşkenler arasındaki ilişkiler bu haritalar yardımıyla kolayca açıklanabilir.

Aşağıdaki varsayımların gerçekleşip gerçekleşmediđi kontrol edilmeli ve gerçekleşmemesi halinde, deđerşkenler arasındaki ilişkiler ile biplot verilerinin analizinin geçersiz hale gelebileceđi unutulmamalıdır. Bu varsayımlar:

- Varyansın Homojenliđi: Satır ve sütun deđerşkenleri arasında varyansın homojenliđinin sağlanması gerekir. Uyum analizi, istatistiksel özelliklerin, satırlar ve sütunlar arasında benzer olduđunu varsayar (örneğin hiç boş deđerşken olmamalıdır; tüm verilerin sıfırlardan oluştuduđu deđerşkenler gibi)
- Ayrık Veri Kuralı: Uyum analizi, analiz edilen verilerin ayrık olduđunu varsayar.
- Kategori Sayısı: Veri setini oluşturan deđerşkenler birkaç kategoriden oluşmalıdır (genellikle 3 ve 3'ten fazla). Eğer uyum analizi 2 ya da 3 kategoriyi analiz ederse, bu analiz orijinal verilerin oluşturduđu tablodan daha fazla bilgi verici olmayabilir).
- Frekans tablosundaki tüm deđerler negatif olmamalıdır, böylece biplot'ta noktalar arasındaki uzaklıklar her zaman pozitif olur.
- Uyum analizinin dağılım ile ilgili herhangi bir varsayımı yoktur (normallik varsayımları gibi). Dolayısıyla kategorik verilerin analizinde herhangi bir varsayım gerektirmeden deđerşkenler arasındaki ilişkinin açıklanmasında ve yorumlanmasında kullanılır.

### 3. Uygulama

#### 3.1. Veriler ile İlgili Genel Bilgi

Veriler Temel Gelir ile ilgili olarak Avrupa’da yapılan bir anket sonucu elde edilmiştir. Katılımcılara “Temel Gelir Sistemini destekleyip desteklemedikleri sorulmuştur. Elde edilen sonuçlara göre katılımcıların üçte biri bu sistemi desteklemektedir. Bu çalışmadaki amaç, bu sistemi destekleyen katılımcıların hangi kategorilerde yer aldığıın araştırılmasıdır. Verilere Kaggle’s veri tabanından ve “Basic Income” linkinden ulaşılabilir. Araştırma da kullanılan değişkenlerin tanımları ve soru formu detaylarına da buradan ulaşılabilir.

Temel Gelir, kişilerin diğer gelir kaynakları ya da çalışıp çalışmadıklarına bakılmaksızın, hükümet tarafından her bir bireye koşulsuz olarak ödenen bir gelirdir. Diğer sosyal güvenlik ödemelerinin yerini alır ve tüm ihtiyaçları karşılayacak kadar yüksek olduğu düşünülmüştür (ALPAR, 2017).

Temel Gelir Verisi 15 değişken ve 9649 gözlemden oluşmaktadır. Uygulanacak olan MCA analizi için bu değişkenlerden çeşitli ve farklı seviyelere sahip olan 8 değişken seçilmiştir. Bunlar; “Country”, “Gender”, “Education”, “Job”, “Children”, “Awareness”, “Vote” ve “Age-Group” değişkenleridir.

Anket çalışması Dalia Research tarafından Nisan 2016’da Avrupa Birliği Devletlerinde gerçekleştirilmiştir. 28 Avrupa Birliğine üye devlet açısından 14 – 65 yaş aralığında, Cinsiyet, Bölge ve Ülke değişkenleri kullanılarak 9649 kişiye uygulanmıştır. Analiz R programı kullanılarak gerçekleştirilmiştir.

#### 3.2. Verilere Çoklu Uyum Analizi Uygulaması

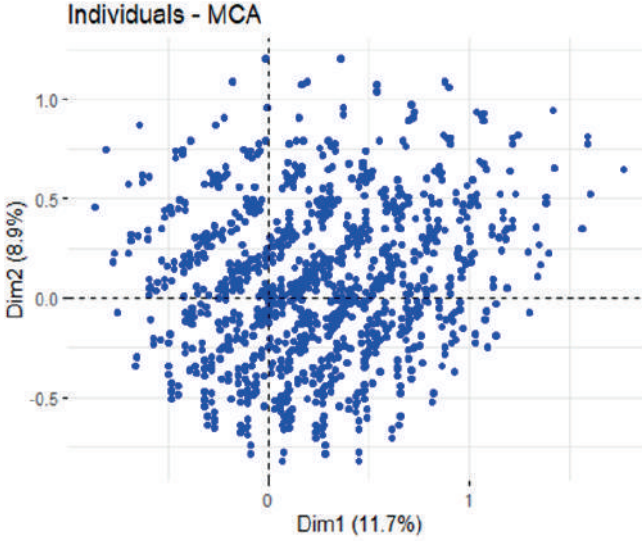
Çoklu Uyum Analizi ile inertia’nın her bir boyut tarafından ne kadar açıklandığı yüzde olarak görülebilir. Buna göre tabloya bakıldığında inertianın açıklanan yüzdeleri görülebilir:

- 1. Boyutta inertia’nın %11,74 lük kısmı,
- 2. Boyutta inertia’nın %8,90 lük kısmı,
- 3. Boyutta inertia’nın %8,45 luk kısmı açıklanmıştır.

*Tablo 1. Boyutlar Tarafından Açıklanan Varyans Yüzdeleri*

##	eigenvalue	percentage of variance	cumulative percentage of variance
## dim 1	0.23489407	11.744704	11.74470
## dim 2	0.17803731	8.901865	20.64657
## dim 3	0.16904043	8.452021	29.09859

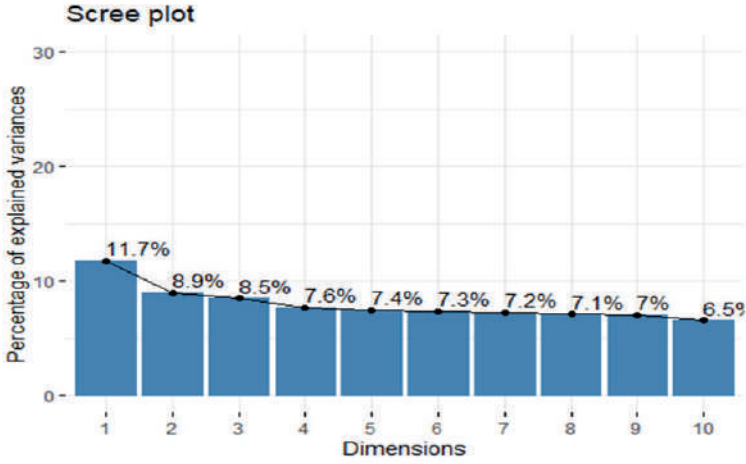
Analize katılan 9649 kiřinin görsel haritası 2 boyutta Grafik 1'de gösterilmiřtir:



*Grafik 1. Analize katılan gözlemlerin görsel haritası*

Grafik 2'de özdeđerler ve varyanslar tablosuna bakıldıđında toplamda 14 boyut olduđu görölmektedir. Genellikle MCA analizlerinde ilk 3 boyut deđiřkenliđin(varyansın) açıklanması için yeterli görölmürken, dataya bakıldıđında varyansın %36,71'i, 6. boyut ile %51,43'ünün açıklandığı görölmüřtür.

Özdeđerler her eksen tarafından açıklanan miktara karřılık gelir. Boyutlar azalan řekilde sıralanır ve analizde açıklanan varyans miktarına göre listelenir. 1. Boyut analizde varyans oranını en çok açıklayan boyuttur. Daha sonra onu 2. ve 3. Boyut ve diđerleri sırasıyla takip eder. Kümülatif varyans yüzdeleri ise; açıklanan varyans oranları ardışık olarak eklenerek bulunur. Buna göre varyansın %51'lik kısmı ilk 6 boyutla açıklanmıřtır.



Grafik 2. Özdeğer ve varyansların büyükten küçüğe doğru gösterimi

Grafik 3'te en büyüğünden en küçüğüne göre sıralanan özdeğerlerin ve varyansın grafiğidir. Özdeğerler oldukça küçük ve karşılaştırılabilir olduğu noktada boyut sayısı elde edilir. Grafiğe bakıldığında en büyük değişim 1. Boyutta sonra 2. Boyut ve en son 3. Boyutta görülmektedir.



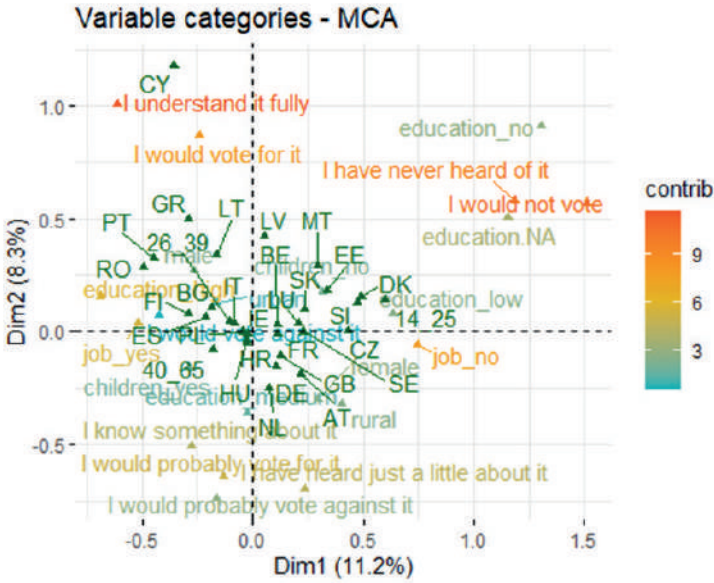
Grafik 3. İşsizlik verisinin boyutlar içerisindeki dağılımı

Grafikte işsizlik verisinin Boyutlar içerisindeki dağılımı görülmektedir. Bu grafiğe göre işsiz kesimin Boyut 1'e katkısı iş sahibi kişilerden daha fazla olduğu görülmektedir. Temel Gelir fikrine sıcak baktıkları söylenebilir.



Çoklu uyum analizini açıklamanın bir diğer yolu boyutlar baz alınarak aşağıdaki simetrik grafiğin çizilmesidir (Sourial et al., 2010). Bu grafikte satır ve sütunlar temel koordinatlar üzerinde yer alır. Noktalar arası Ki-kare uzaklıkları satır ve sütunlar arasında yaklaşık olarak ifade edilir. Grafik dağılımı bir desen halinde verir. Satırlar gözlemleri, Sütunlar değişken kategorilerini temsil eder. Herhangi satır ya da sütunlar arasındaki uzaklıklar benzerlikleri ortaya çıkarır. Aynı profile sahip satır noktaları faktör haritasında birbirine yakın konumlanır. Aynı durum sütun noktaları içinde geçerlidir. Birbirine benzer profile sahip sütun noktaları da faktör haritasında yakın konumlanır (Josse et al., 2008).

Bu grafiğe göre inertia'nın 1. Boyuttaki açıklanma yüzdesi %11,23'tür. 2. Boyutta ise inertia %8,31 oranında açıklanmıştır.



Grafik 4. Gözlemlerin ve değişkenlerin faktör haritası

- Kategorilerle belirlenmiş Grafik 4'e bakıldığında eğitim seviyesi düşük ve işsiz kişilerin Temel Gelir fikrini hiç duymadıkları ve oy kullanmayacakları görülmektedir.
- "CY" ile ifade edilen Kıbrıs'ta Temel Gelir uygulaması hakkında fikir sahibi oldukları ve oy kullanacakları anlaşılmaktadır. Grafiğe bakıldığında "I understand it fully (Tamamen bilgi sahibiyim)" ve "I would vote against it (Ona karşı oy kullanacağım)" ile Kıbrıs'ın aynı

boyutta olduğu ve oy kullanma konusunda pozitif düşünceye sahip oldukları görülmektedir.

- “AT” ile ifade edilen Avusturya’ya bakıldığında “Orta Eğitim Düzeyi”ne Sahip “Bayanlar” ın Temel Gelir konusu hakkında bilgi sahibi oldukları ve oy kullanacakları görülmektedir.
- 40-65 yaş arasında “Hırvatistan (HU) ve Polonya(PL) da yaşayan katılımcıların konu hakkında bilgisi olduğu ve oy kullanma fikrine sıcak baktıkları söylenebilir.
- “Yunanistan (GR)”, “Portekiz (PT)” ve “Romanya (RO)” Eğitim seviyesi yüksek olan Erkek Katılımcıların Temel Gelir fikrine karşı oy kullanacakları belirlenmiştir.
- “I have never heard of it (hiç duymadım)”, “education\_no(egitimsiz)”, “I would not vote (Oy kullanmayacağım)” kategorilerinin 1. Boyuta pozitif yönlü katkısı bulunurken, “I understand it fully(herşeyi anlıyorum)”, “I would vote for it(Onun için oy veririm)” 1. boyuta negatif yönlü katkısı bulunmaktadır.
- 14-25 yaş arası, eğitim seviyesi az on ve işsiz kişilerin “Temel Gelir” fikrine karşı oy kullanacakları,
- Bulgaristan (BG), Finlandiya(FI), Yunanistan(GR), İtalya(IT), Litvanya(LT), Portekiz(PT), Romanya(RO) ve İspanya(ES)’dan 26-39 yaş aralığında “Eğitim seviyesi yüksek” “Erkek” katılımcıların Temel Gelir fikrine oy verecekleri görülmektedir.

Temel gelir, insanların daha iyi eğitim alma fırsatlarını artırabilir. Özellikle düşük gelirli aileler, çocuklarının eğitimine daha fazla yatırım yapabilecek durumdayken, yetişkinler de iş gücü piyasasındaki değişen ihtiyaçlara uygun beceriler kazanmak için yeniden eğitim alabilirler. Örneğin, bir kişi, ek gelire bir kursa kaydolabilir veya online eğitimler alabilir. Eğitim için harcanan zamanın ve kaynağın artması, uzun vadede kişisel gelişim ve ekonomik fırsatlar yaratabilir.

Temel gelir, ülkelerin eğitim sistemlerine daha fazla yatırım yapmalarını teşvik edebilir, ancak uygulamaya koyulması ciddi maliyetler gerektirebilir. Bazı eleştirmenler, temel gelirin eğitim sisteminin yeterince finansman almasına engel olabileceğini savunurlar, çünkü temel gelir için gereken bütçe büyük olabilir.

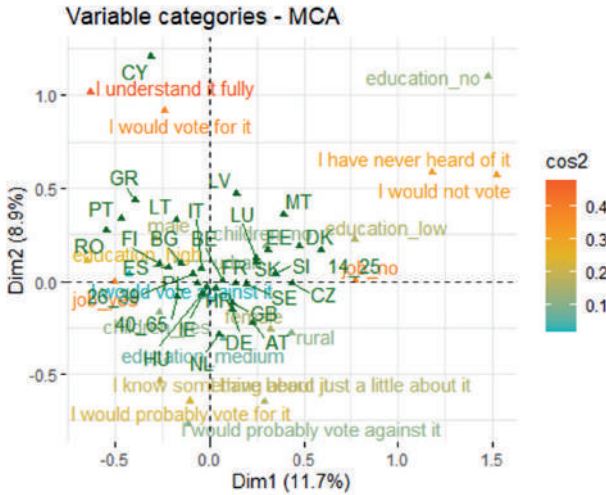
### 3.3. Değişken Kategorilerinin Temsil Kalitesi (cos2 değerleri)

Kategori noktaları ile başlangıç noktası arasındaki mesafe, faktör haritasında değişken kategorisinin kalitesini ölçer (Markos et al., 2009). Başlangıç noktasından uzakta olan kategori noktaları, faktör haritasında iyi temsil edilir. Grafik 5'te görüldüğü gibi 1.ve 2. Boyut toplam inertia'nın %20,64 ünü açıklamaktadır.

Tablo 2. Boyutlara Göre Değişkenlerin Temsil Kalitesi

##	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
## female	0.09238982	0.06018623	0.061756184	0.35597005	0.04237855
## male	0.09238982	0.06018623	0.061756184	0.35597005	0.04237855
## rural	0.07364483	0.03032410	0.008589266	0.03694573	0.23306012
## urban	0.07364483	0.03032410	0.008589266	0.03694573	0.23306012

Temsil Kalitesi *square cosine* (*cos2*) değeri ile ölçülür ve değişken kategorileri ile belirli bir eksen arasındaki ilişki derecesini ölçer.

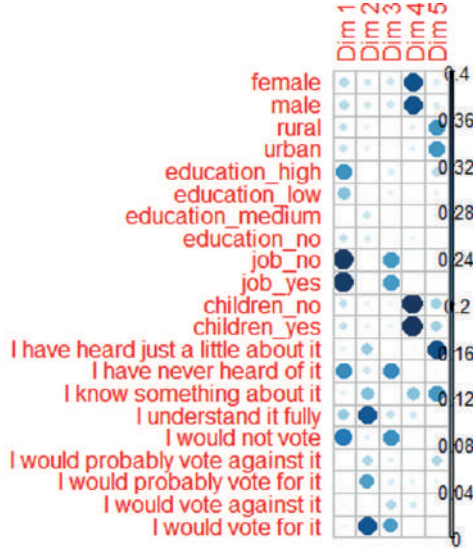


Grafik 5. Değişkenlerin temsil kalitesi görseli

Eğer bir değişkenin kategorisi iki boyut tarafından iyi açıklanmışsa *cos2* değerleri toplamı "1" e eşittir. Bazı satır değişkenleri için veriyi mükemmel biçimde temsil etmek için 2'den fazla boyuta ihtiyaç duyulur. Grafik 5'te "cos2" değerleri her değişken kategorisi için ayrı renklerde belirtilmiştir. Bu sayede her değişken kategorisi ayırt edilebilir. Az *cos2* değerine sahip değişken kategorileri "beyaz" ile, orta *cos2* değerine sahip değişken kategorileri

“mavi” ile, yüksek  $\cos^2$  değerine sahip değişken kategorileri “kırmızı” ile renklendirilmiştir.

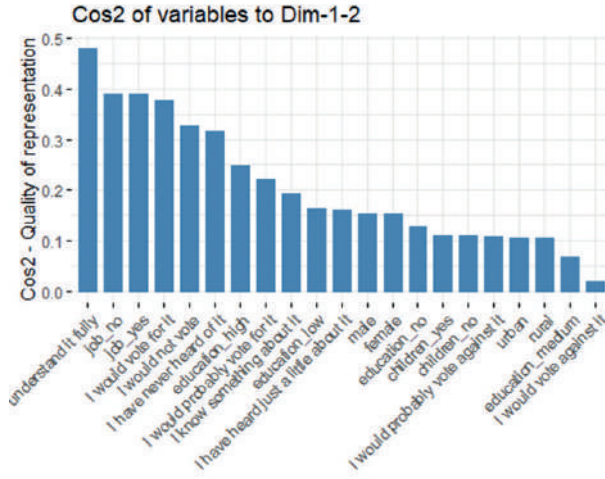
Grafik 6’da pozitif korelasyonlar mavi ile negatif korelasyonlar ise kırmızı ile gösterilmiştir. Araştırmada korelasyonlar yuvarlak ile ifade edilmiştir. Kategoriler ve korelasyon renkleri aşağıdaki grafikte ayrıntılı olarak görülmektedir.



Grafik 6. Boyutlar ve değişkenler arası korelasyonlar

Grafik 7’ye bakıldığında “ I would vote against it (karşı tarafında oy kullanırdım)” kategorisi ilk 2 boyut tarafından iyi temsil edilmemiştir. Aynı şekilde “education-medium” yani orta eğitime sahip kişilerde iyi temsil edilememiştir. Aslında son 10 kategoriye bakıldığında temsil değerlerinin 0,1 etrafında olduğu görülmektedir. Dolayısıyla bu kategorilerin bu boyutlar kullanıldığında çok da iyi açıklanmadığı görülmüştür. Daha yüksek boyutta bir çözüm yöntemi denenebilir. Diğer kategoriler incelendiğinde “I understand it fully (kesinlikle anladım)”, her iki seviyesi ile “iş durumu” değişkeni iyi bir şekilde boyutlar tarafından açıklanmıştır.

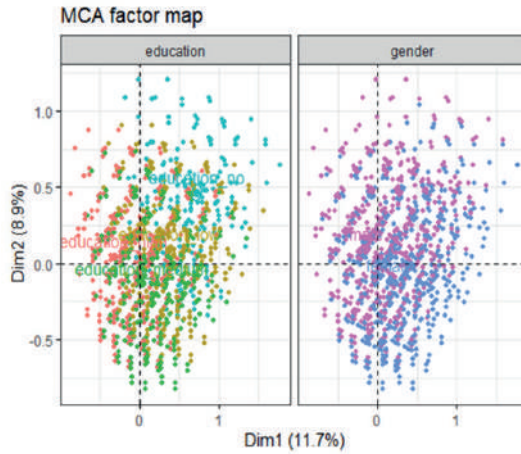
Eğitim seviyesi düşük olan bireyler, genellikle daha düşük ücretli ve daha güvencesiz işlerde çalışmaktadır. Temel gelir, eğitim düzeyi düşük olan bireylerin maddi olarak daha az bağımlı olmalarını sağlayarak, meslek değiştirme veya daha iyi eğitim fırsatlarına yönelmelerini teşvik edebilir.



Grafik 7. Cos 2 değerlerinin boyutlar tarafından temsili

### 3.4. Cinsiyetin ve Eğitim Durumunun Boyutlara Göre Dağılımı

Grafik 8'e göre eğitim seviyesi yüksek kişilerin 2. Boyuta katkısı pozitif yönlüdür. Eğitim seviyesi düşük ve olmayan kişilerin ise 1. Boyuta katkısı pozitifdir. Eğitim seviyesi yüksek ve orta olan kadın katılımcıların temel gelir için pozitif oy kullanacakları anlaşılmaktadır. Çoğunlukla Eğitim seviyesi düşük ve eğitimi olmayan erkek katılımcıların temel gelir fikrine negatife oy verdiği anlaşılmaktadır. Grafik 8'e bakıldığında kadın katılımcıların erkek katılımcılara göre 1. Boyuta katkısının daha fazla olduğu görülmektedir.



Grafik 8. Cinsiyet ve eğitim durumunun boyutlara göre dağılımı

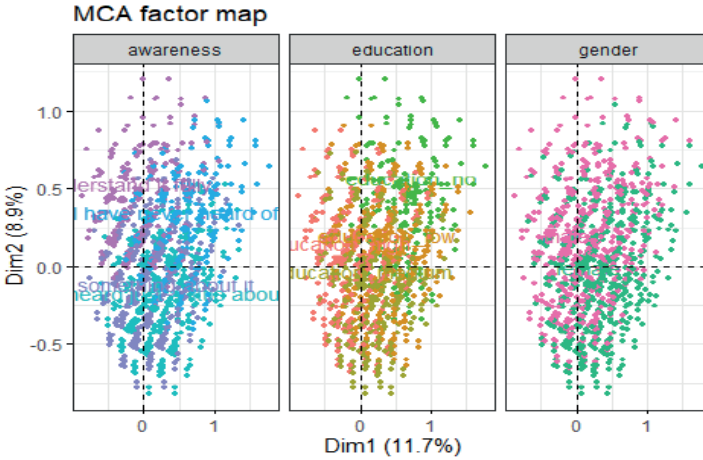
Kategori noktaları ile başlangıç noktası arasındaki mesafe, faktör haritasında değişken kategorinin kalitesini ölçer. Başlangıç noktasından uzakta olan kategori noktaları, faktör haritasında iyi temsil edilir. Bu bilgiler ele alındığında, Grafik 8'de 1.ve 2. Boyut toplam inertia'nın %20,64 ünü açıklamaktadır. Tüm noktalar iki boyutta iyi bir şekilde açıklanamadığı görülmüştür.

Temel gelir, sadece kadınlar için değil, erkekler için de eşitsizlikleri giderici bir etki yaratabilir. Ancak, temel gelir uygulamalarının özellikle kadınlar üzerinde büyük bir dönüştürücü etkisi olması beklenmektedir çünkü kadınlar genellikle ekonomik olarak daha kırılgan gruptur.

### 3.5. Cinsiyet, Eğitim ve Farkındalık

Cinsiyet, eğitim ve farkındalık değişkenleri açısından Grafik 9'a bakıldığında, yüksek eğitime sahip ve konu hakkında fikri olan kişilerin 2. Boyuta pozitif yönlü etkisi vardır.

Grafik 9'a bakıldığında temel gelir konusunda fikir sahibi olan farkındalığı yüksek, eğitim seviyesi orta ve yüksek seviyede bulunan kadın katılımcıların olumlu oy kullandıkları anlaşılmaktadır.

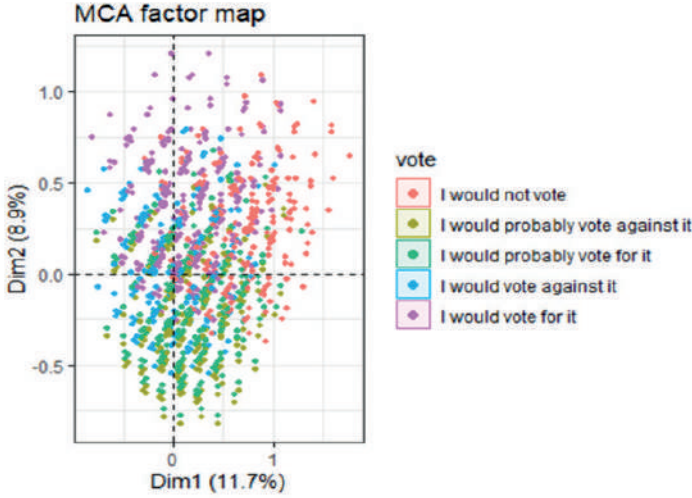


Grafik 9. Cinsiyet, eğitim ve farkındalık değişkenlerinin boyutlara göre dağılımı

### 3.6. Uyum iyiliği Analizi Faktör Haritasında Oy Durumunun Dağılımı

Grafik 10'a göre faktör haritasına bakıldığında oy verme konusunda karşıt fikirlerin birlikteliği görülmektedir. Oy vermeyi düşünen katılımcılar ile

kesinlikle oy vermeyi düşünmeyen katılımcıların yakın gruplarda yer aldığı görülmektedir. Temel Gelir konusunun kesinlikle karşısında oy kullanacak kişiler ile muhtemelen bu fikre karşı oy kullanacak olan kişilerin birbirlerine yakın gruplandığı görülmektedir.



Grafik 10. Faktör haritasında oy durumunun dağılımı

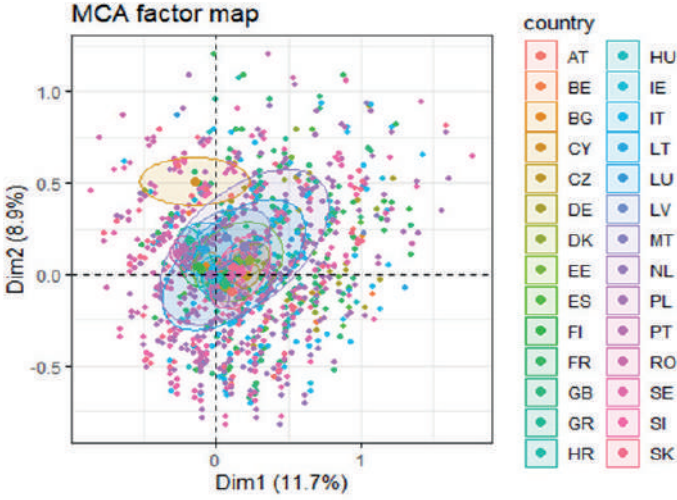
### 3.7. Uyum iyiliđi Analizi Faktör Haritasının Ülkelere Göre Dağılımı

Grafik 11’de ülkeler bazında katılımcı dağılımı görülmektedir. Buna göre; Portekiz (PT), Romanya (RO), Slovakya (SK), Slovenya (SI), İsveç (SE) aynı boyutta yer almışlardır. Temel Gelir fikrinin karşısında oy kullanacakları görülmektedir.

Temel gelir, ekonomik eşitsizliği azaltmayı hedeflerken, eğitim politikaları, toplumun her bireyine kaliteli eğitim fırsatları sunmaya çalışır. Her iki politika da özellikle dezavantajlı gruplar için fırsat eşitliğini teşvik eder.

Finlandiya, İspanya, Hollanda ve Almanya gibi ülkeler, çeşitli pilot projeler aracılığıyla temel gelir üzerine önemli deneyler yapmaktadır. Ancak, temel gelir henüz AB genelinde yaygın olarak kabul edilen bir politika değildir. Her ülke, kendi ekonomik durumu, sosyal politikaları ve toplumsal ihtiyaçları doğrultusunda bu konuda farklı yaklaşımlar sergilemektedir.





Grafik 11. Faktör baritasının ülkeler bazında dağılımı

#### 4. Sonuç

Çoklu Uyum Analizi kategorik değişken arasında grafiksel olarak yorum yapabilmeyi kolaylaştıran çok değişkenli ve çok boyutlu bir analiz şeklidir. Günümüzde çeşitli araştırmalarda (sosyo-ekonomik, tıp, vb) bu yöntem sıklıkla başvurulmaktadır. Bu yöntemde analizi yapılacak değişkenlerin kategorik olması istendiğinden sosyal araştırmalarda bu yöntem sıklıkla başvurulur. Özellikle Evet /Hayır ya da ölçek aralığı olan anket çalışmalarında da bu yöntem kullanılır.

Çoklu Uyum İyiliği Analizi, Avrupa Birliği ülkelerinde Temel Gelir tartışmalarına derinlik kazandırmak ve uygulamaların çok boyutlu etkilerini daha kapsamlı bir şekilde incelemek için önemli bir araçtır. AB ülkelerinin farklı sosyal yapıları, ekonomik seviyeleri ve kültürel özellikleri göz önünde bulundurularak, temel gelir gibi büyük sosyal reformların uygulanabilirliğini ve etkilerini anlamada bu tür analizlerin kullanılması gereklidir.

Bu analizler, sadece ekonomik etkileri değil, aynı zamanda sosyal, psikolojik ve çevresel etkileri de değerlendirerek, politikaların daha dengeli ve sürdürülebilir olmasına yardımcı olabilir. AB'nin çok farklı sosyal yapıları, Temel Gelir gibi yenilikçi politikaların her ülkede aynı sonuçları vermediğini, dolayısıyla bu tür analizlerin her ülke için özelleştirilmesi gerektiğini gösterir.

Uygulaması gerçekleştirilen veri "<https://www.kaggle.com/daliaresearch/basic-income-survey-european-dataset>" linkinden elde edilmiş olup, Dalia Research tarafından 2016 senesinde Avrupa Birliğine Üye ülkelere anket



çalışması yapılarak elde edilmiştir. Bu ankette Avrupa genelinde “Temel Gelir” fikrine nasıl baktıkları, eđer sandıđa gidilirse bu fikre oy verip vermeyecekleri sorulmuştur.

Bu araştırmaya 28 Avrupa Birliđine üye devlet katılmıştır. Katılımcılar 14 – 65 yaş aralığında 9649 kişiye uygulanmıştır.

Elde edilen veri setine çoklu uyum analizi uygulanmıştır. Analiz için uygun deđişkenler seçilmiş, eksik veri barındıran satırlar datadan çıkarılmıştır ve analizler bu şekilde yapılmıştır.

Yapılan analiz sonucu;

İnertia (varyans) 14 boyutta açıklanmıştır. Genelde ilk iki boyutta varyans oranı büyük ölçüde açıklanırken, uygulamada 14 boyutta inerta açıklanmıştır. Birçok makalede ve araştırmalarda görüldüğü üzere sosyal araştırmalarda boyut sayısı arttıkça inertia açıklanma oranının da arttığı belirtilmiştir. 1. Boyutta inertia'nın %11,2 lik kısmı, 2. Boyutta ise %8,90'lık kısmı açıklanmıştır.

Yapılan analizlere ve özdeđer sonuçları ile grafiđine bakıldığında kırılma miktarına göre 3 boyut kullanılması uygun görülmüştür.

Deđişken kategorileri sonuçlarına göre;

- Eđitim seviyesi az olan katılımcılar ile oy vermeyecek ve Temel Gelir konusunda fikir sahibi olamayan katılımcıların aynı boyutta gruplandığı görülmüştür.
- 14-25 yaş arası, eđitim seviyesi düşük ve işsiz katılımcılarında 1. Boyutta yer aldığı ve Temel Gelir fikrine karşı oldukları görülmüştür.
- Avrupa Birliđi üyesi olan Kıbrıs'ın açık bir şekilde Temel Gelir konusu hakkında fikir sahibi olan ve bu konuda kesinlikle pozitif yönlü oy kullanacak katılımcılardan oluştuđu görülmektedir.
- 26-39 ve 40-65 yaş arası Hırvatistan, Polonyave İrlanda'daki iş ve çocuk sahibi olan katılımcıların bu fikre sıcak bakmadıkları anlaşılmaktadır.

Bazı boyutlar için bazı deđişkenlerin iyi temsil edilmediđi görülmüştür. Bu durumda çok boyutlu bir analiz yöntemi ya da deđişkenleri kümeleme yöntemi ile ayrıştırma ve detaylandırma yöntemine gidilebileceđi gibi analize farklı deđişkenler eklenebileceđi görülmüştür.

Bu fikre Oy vermeyecek kişilerin çoğunlukla;

- a. Konu hakkında yeterli bilgisi olmayan, kırsal kesimdeki kişiler olduğu gibi enteresan bir biçimde şehirli olan ve eğitim seviyesi yüksek kişiler olduğu da görülmüştür.
- b. Eğitimsiz katılımcıların oy kullanmak istemedikleri, yüksek eğitime sahip kişilerin oy kullanacakları ama oy durumunun olumlu ya da olumsuz olabileceği görülmüştür.
- c. Konu hakkında bilgi sahibi olan katılımcılar ile konuya karşı olumlu ya da olumsuz oy kullanacak kişiler de aynı boyutta yer almışlardır.
- d. Yüksek eğitilmiş, iş sahibi kişilerin bu fikre karşı oldukları görülmektedir.

AB ülkelerinde, Temel Gelir uygulamaları farklı ulusal ve bölgesel bağlamlara göre değişiklik gösterebilir. Her ülkenin sosyal yapısı, ekonomik durumu ve politika öncelikleri farklıdır. Bu nedenle, çoklu uyum analizi gibi çok boyutlu analiz yöntemleri, Temel Gelir'in potansiyel etkilerini ülke bazında karşılaştırmalı bir şekilde değerlendirebilmek için kullanışlıdır. Yapılan uygulamada farklı ekonomik durumlara sahip farklı ülke vatandaşlarının konu hakkında farklı fikirlere sahip olduğu görülmektedir.

Uygulamada eğitim değişkeninin farklı kırılımlara sahip olduğu görülmektedir. Farklı ülke vatandaşlarının farklı eğitim seviyelerine sahip oldukları halde temel gelir konusunda ortak görüşlere sahip olduğu görülmüştür. Avrupa'da eğitimdeki eşitsizlikler, bireylerin iş gücü piyasasında karşılaştıkları fırsatları doğrudan etkilemektedir. Temel gelir, düşük eğitilmiş kişilerin eğitim alabilmesi için finansal bir güvence sağlayarak, sosyal mobilitiyeyi artırabilir. Bu, özellikle düşük gelirli ailelerden gelen ve eğitim fırsatları kısıtlı olan bireyler için önemli bir fırsat olabilir.

Sonuç olarak Avrupa Birliği ülkelerinde Temel Gelir uygulamasının çoklu uyum iyiliği analizi, bu tür politikaların çok boyutlu etkilerini daha doğru ve kapsamlı bir şekilde anlaşılmasına olanak tanıdığı görülmektedir. Bu analizler, temel gelirin yalnızca ekonomik değil, aynı zamanda toplumsal, psikolojik, çevresel ve kültürel sonuçlarını değerlendirerek, daha dengeli ve sürdürülebilir bir politika geliştirilmesine yardımcı olabilir.

Cinsiyet ve eğitim durumu, temel gelir tartışmalarında önemli faktörlerdir. Avrupa Birliği ülkelerinde temel gelir uygulamaları, kadınların ve düşük eğitilmiş bireylerin ekonomik bağımsızlıklarını artırabilir, toplumsal eşitsizlikleri azaltabilir ve daha geniş bir eğitim fırsatları yelpazesi sunabilir. Bu etkileşimlerin politika yapımcılar tarafından dikkate alınarak, temel gelir uygulamalarının daha adil ve etkili bir şekilde hayata geçirilmesi sağlanabilir.

Temel Gelir uygulamasının AB ülkelerinde potansiyel etkilerini çok boyutlu bir şekilde deđerlendirmek için Çoklu Uyum İyiliđi Analizi, önemli bir yöntem olarak öne çıkmaktadır. Çoklu Uyum Analizi sadece gelir dağılımı üzerinde deđil, sosyal, psikolojik, çevresel ve toplumsal düzeydeki etkileri de dikkate alarak, Temel Gelir'in geniş kapsamlı sonuçlarını daha iyi anlaşılmasına olanak tanır. Bu sayede, politika yapıcılar, Temel Gelir uygulamalarının potansiyel faydalarını ve zorluklarını daha derinlemesine analiz edebilir ve daha etkili sosyal politikalar geliştirebilirler.

## Kaynakça

- Abdi, H., & Béra, M. (2014). Correspondence Analysis. In ALPAR, C. (2017). Uygulamalı çok değişkenli istatistiksel yöntemler.
- Benammou, S., & Saporta, G. (2003). On the connection between the distribution of eigenvalues in multiple correspondence analysis and log-linear models. *REVSTAT-Statistical Journal*, 1, 42-79.
- Beken, Hikmet Gülçin. (2020). 21.yüzyılda Temel Gelir Tartışmaları
- Bohun, Lukasz (2020). The Feasibility of Universal Basic Income in the European Union: Challenges and Opportunities., *Journal of European Social Policy*, 30(1), 70-84.
- Fiedler, Michael & Gimpel, Peter (2019)., Basic Income in Europe: A Comparative Analysis of Implementation Prospects., *European Journal of Social Security*, 21(2), 119-141.
- Fitzgerald, R.; Goldblatt, D. (2018). Basic Income and the Welfare State: A Multi-Criteria Decision Approach, *Journal of Social Policy*
- Harvey, D.; Plummer, M. (2015). Universal Basic Income: A Solution to the European Crisis?, *Global Social Policy*
- Hoffman, D. L., & de Leeuw, J. (2011). Geometrical aspects of multiple correspondence analysis: implications for the coordinate scaling debate.
- Josse, J., Husson, F., & Lê, S. (2008). Multiple Correspondence Analysis. Applied Mathematics Department: Dortmund, 7.
- Keenan, D.; Lio, M.; et al. (2019). Exploring the Implications of a Universal Basic Income in Europe. *Journal of European Social Policy*
- Lennart J. S. H. Martein, Sebastian H. P. E. Müller, (2013), Multi-Criteria Decision Analysis in the Management of Sustainable Development Projects, *Sustainable Development Journal*
- Lepage, F (2016). Basic Income in Europe: Prospects and Problems. *International Social Security Review*
- Markos, A., Menexes, G., & Papadimitriou, T. (2009). Multiple correspondence analysis for “tall” data sets. *Intelligent Data Analysis*, 13(6), 873-885.
- Matthias Ehrgott, Xinhua Li, Roman S. Salo, (2010), Multi-Criteria Decision Analysis: Methods and Software, Springer
- Sáez, E., & Zucman, G. (2020). *Progressive Wealth Taxation*. Brookings Papers on Economic Activity.
- Schuett, J.; Kallis, G.; Otero, I.; et al. (2020). The Feasibility of Universal Basic Income in Europe
- Sourial, N., Wolfson, C., Zhu, B., Quail, J., Fletcher, J., Karunanathan, S., Banteen-Roche, K., Béland, F., & Bergman, H. (2010). Corresponden-

ce analysis is a useful tool to uncover the relationships among categorical variables. *Journal of clinical epidemiology*, 63(6), 638-646.

Standing, G. (2017). *Basic Income: And How We Can Make It Happen*. Pelican.

Van Parijs, Philippe; Vanderborght, Yannick(2017). *Universal Basic Income in the European Union: A Comparative Study*, *Oxford Handbook of Economic Inequality*

# Gerçek Yaşam Temelli Matematik Etkinliklerine Katılan Dezavantajlı Ortaokul Öğrencilerin Matematiğe Yönelik Bakış Açılarının İstatistiksel Olarak İncelenmesi

Elif Bahadır<sup>1</sup>

Şevval Gökçen<sup>2</sup>

Elif Tuna<sup>3</sup>

Mehmet Şamil Güneş<sup>4</sup>

## Özet

Bu çalışma, sosyal hizmetler himayesinde bulunan dezavantajlı öğrencilerin gerçek yaşam durumlarına yönelik, çevrelerinde olan matematik hakkındaki farkındalıklarını ve bilgilerini artırmaya odaklanmıştır. TÜBİTAK 4004- Doğa Eğitimi ve Bilim Okulları Destekleme Programı kapsamında gerçekleştirilen bu çalışmada, öğrencilerin matematiksel anlayışlarının gelişmesine katkı sağlamak, matematiğe karşı olumlu tutum geliştirmek, matematik yapma yönünde etkin katılım isteği kazanmalarını sağlamak ve işbirlikli öğrenmenin gelişmesini sağlamak amaçlanmıştır. Çalışmanın örneklemini İstanbul ili Avrupa yakasında 4 farklı Çocuk Evleri Koordinasyon Merkezi'nden (ÇEKOM) katılan 30 öğrenci (5, 6. ve 7. Sınıflar) oluşturmaktadır. Dezavantajlı öğrencilerle yapılan bu çalışma, öğrencilerin akademik becerilerini geliştirmenin yanında sosyalleşme bağlamında da önemli görülmüştür. Gerçek yaşam temelli matematik etkinliklerinin uygulandığı projede, verilerin toplanması amacıyla anket tekniğinden faydalanılmıştır. Bu bağlamda farklı ölçekler kullanılmıştır. Matematik ve Ben, Matematiksel Motivasyon Ölçeği ve Ortaokul Matematik Okuryazarlık-Öz yeterlilik

1 Doç. Dr., Yıldız Teknik Üniversitesi, ebahadir@yildiz.edu.tr, 0000-0002-1154-5853

2 Arş. Gör., Yıldız Teknik Üniversitesi, sgokcen@yildiz.edu.tr, 0000-0002-3552-0298

3 Doç. Dr., Yıldız Teknik Üniversitesi, cozturk@yildiz.edu.tr, 0000-0001-8572-3109

4 Dr. Öğr. Üyesi, Yıldız Teknik Üniversitesi, msgunes@yildiz.edu.tr, 0000-0001-5842-5181

Ölçeği kullanılmıştır. Ölçeklerden elde edilen veriler, SPSS 18 istatistiksel veri analizi paket programı kullanılarak analiz edilmiştir. Bu ölçeklerle, öğrencilerin matematiksel etkinliklere katılımları sonucu düzeylerinin belirlenmesi amaçlanmıştır. Ayrıca cinsiyet ve sınıf düzeyi gibi demografik özelliklere göre gruplar arasında anlamlı farklar olup olmadığı incelenmiştir. Öne çıkan sonuçlarda, 6. ve 7. Sınıf öğrencilerinin matematiğe karşı olumlu tutumlar sergilediği görülür. Cinsiyet özelliğinin uygulanan ölçekler ile aralarında istatistiksel olarak anlamlı bir fark olmadığı ortaya çıkmıştır. Ancak sınıf düzeyi ile ortaokul matematik okuryazarlık-öz yeterlilik ölçeği arasında istatistiksel olarak anlamlı bir fark olduğu görülmüştür.

## 1. Giriş

Matematik dersinin amacı, öğrencilerin eleştirel düşünme ve problem çözüme becerilerini geliştirmek için gerekli olan beceri, tutum, anlayış, bilgi ve değerleri kazandırmaktır (Duran vd., 2018). Bununla birlikte matematik, öğrencileri kariyerlerine hazırlamak için okullarda öğretilmesi gereken bir ders olmasının yanı sıra öğrencilerin yaşadıkları dünyayı anlamalarına yardımcı olan eğlenceli bir alandır (Švecová ve & Pavlovičová, 2014; Bahadır, 2017). Matematiğin önemi yalnızca bilimsel ve teknolojik gelişime yaptığı katkılardan değil, aynı zamanda toplumun hem okur-yazar hem de okur-yazar olmayan üyeleri tarafından pazar yerlerindeki, ulaşımdaki, her türlü işteki günlük etkileşimlerdeki faydasından da kaynaklanmaktadır (Golji & Dangpe, 2016). Matematiğin gerçek hayat ile ilişkilendirilmesinde kullanılacak problemler, örnekler, etkinlikler ve görevlerin öğrenciyi keşfetmeye, modelleme ve araştırma yapmaya, kanıt sunmaya ve teknoloji kullanmaya yöneltmesi ve en önemlisi de öğrencinin zevk aldığı ve kendine güvenerek elde edeceği deneyimleri kapsamalı gerekmektedir (Romberg & Kaput, 1999). Sürekli değişen ve gelişen dünyamızda matematiği anlayan ve günlük hayat ile ilişkiler kurabilen kişiler, kendi geleceklerini şekillendirebilecek fırsatlar yaratabilecek ve bu fırsatları değerlendirmede daha ön planda olabileceklerdir (National Council of Teachers of Mathematics, 2000). Yani aslında matematiğin, bireylerin yaşamı için olmazsa olmaz bir yerde olduğu görülebilir (Ergül ve Doğan, 2022).

Her ne kadar matematik yaşamımızın tüm alanlarında var olsa da (Van De Walle vd., 2012), gerçekte, öğrenciler matematik öğrenmekten korkmakta ve çoğu zaman bundan kaçınmaktadır (Laurens vd., 2018). Matematik, eskiden öğrenilmesi gereken soyut kavramların ve becerilerin bir toplamı olarak görülmüştür, ancak günümüzde problem çözüme ve anlamlandırma süreci sonucunda oluşan bilgi ve gelişen becerileri işaret eden gerçekliğin modellenmesini temel alan bir alandır (Üzel, 2007). Model ve materyaller, matematikteki soyut ifadeleri görselleştirir, somut bir hale getirir ve böylece

öğrencilerin yaratıcı düşünme becerilerine katkı sağlar (Moyer, 2001). Ayrıca matematik, öğrencilerin neden sonuç ilişkilerini görebilecekleri, meraklarını giderecekleri, araştırma yapabilecekleri, tahminde bulunabilecekleri fırsatlar verir ve onların gelişimlerini destekleyicidir (Duran vd., 2018). Bu açıklamalar göz önünde bulundurulduğunda, problem çözme becerisinin matematiksel düşünme becerisi ile bağı görülebilir ve öğrencilerde matematiksel düşünme becerilerinin gelişmesi amacıyla bilişsel süreçlerin desteklenmesi büyük önem taşır. (Bahadır, 2021). Güncel matematik öğretim programlarını incelediğimizde de öğrencilere hazır bilgiler vermekten çok öğrencileri sorgulamaya yöneltmek ve 21. Yüzyıl becerilerini kazandırmak amaçlanmıştır. Bu amacı gerçekleştirmek öğrencilerde üst düzey düşünme becerilerinin gelişmesini sağlamak ile mümkün olup (Gökçe vd., 2022), matematik öğrenmenin önemini ve gereğini ortaya çıkarttığı söylenebilir (Laurens vd., 2017). Bahsedilenler göz önüne alındığında, matematiği öğrenmek için, matematiksel kavramlar öğrencilere günlük yaşamda karşılaşılabilecek problemler şeklinde sunulmalıdır (Gravemeijer, 1994; Fauzan vd., 2002; Evcan vd., 2020). Olumlu matematiksel öğrenme deneyimleri sağlayan yenilikçi öğretim yöntemleri, öğrencilerin matematikteki başarılarını artırmaya yardımcı olabilir (Riley vd., 2017). Gerçekçi Matematik Eğitimi (GME), öğrencileri durumlarla bağlantılı olarak problemlerle tanıştırır; öğrenciler için matematik öğrenimini daha eğlenceli ve anlamlı hale getirmeyi amaçlar (Laurens vd., 2017; Noreen, 2019).

GME, günlük aktivitelerin matematikselleştirilmesi sürecine yönelik matematik öğrenimi ile ilgili bir teoridir (Yuhariati vd., 2022). Bu kuramının kurucusu Freudenthal, öğrencilerin matematiği anlamlandırması için öğrenme süreçlerine başlarken gerçek yaşam durumlarından elde edilen informal bilgilerle başlanmasının çok önemli ve gerekli olduğunu açıklamıştır (Van den Heuvel-Panhuizen, 1996). GME'nin temel amacı, gerçek problemler sunarak matematik öğrenimini öğrenciler için daha eğlenceli ve değerli hale getirmektir (Üzel, 2007; Laurens vd., 2018).

Freudenthal'a göre matematik öğrencilerin sosyal yaşamları ve yaşam alanları ile ilgili olmalıdır; matematiğin onların deneyimlerine yakın, yaşadıkları toplum ile ilişkili olması gerektiğini savunur (Sezer, 2019). Bu sebeple Freudenthal matematiği bir "insan etkinliği" olarak yorumlamıştır (Freudenthal, 1973; Gravemeijer, 1994). Buradaki temel nokta ise matematiğin kapalı bir sistem olarak nitelendirilmemesi, matematikleştirme ile matematik eğitiminin her zaman gelişen bir yapı olduğunun açıklanmasıdır (Freudenthal, 1968). Yani matematikleştirmenin aşamalı olarak gerçekleştiği söylenebilir. Treffers, aşamalı matematikleştirmeyi iki farklı matematiksel etkinlik dizisi şeklinde tanımlar; yatay matematikselleştirme ve dikey matematikselleştirme



(Rasmussen vd., 2005). Treffers'a (1987) göre yatay matematikleştirme, öğrencilerin kendi informal stratejilerini kullanarak bağlamsal bir problemi anlaması ve çözmesi; dikey matematikleştirme ise öğrencilerin kendilerine ait informal stratejilerinin onları problemi çözmeye (matematikselsel bir dil kullanmaya veya algoritma oluşturmaya) yönlendirmesidir. GME'yi açıklayan diğer araştırmacılarından Olkun ve Toluk Uçar (2003) ise benzer şekilde GME'nin üç temel prensibini ortaya koymuştur. Bu ilkelerden birincisi öğretime başlarken çocuğun gerçek yaşamıyla ilişkili olarak anlamlı bir matematikselsel etkinliğe katılmasını sağlamak olmalıdır. İkincisi, öğretimi planlama sürecinde öğrencilerin önceki bilgileri göz önünde bulundurulmalı ve girişte yapılan etkinlik hedeflenen matematikselsel kavram ve becerilerle uyumlu olmalıdır. Üçüncü önemli prensibi ise matematikselsel öğrenme etkinlikleri öğrencilerin kendi modellerini oluşturmaya ve geliştirmesine fırsat vermelidir.

Ortaya çıktığından bu yana GME ile ilgili yürütülen birçok çalışma mevcuttur (Akkaya, 2010; Demirdöğen ve Kaçar, 2010; Ünal ve İpek, 2010; Ersoy, 2013; Palinussa, 2013; Çilingir ve Dinç Artut, 2016; Demir, 2017; Korkmaz ve Korkmaz, 2017; Korkmaz ve Tutak, 2017; Uça ve Saracoğlu, 2017; Erdoğan ve Ayvaz Tuncel, 2018; Saleh, vd., 2018; Berkant ve Yaren, 2020; Chairil Hikayat vd., 2020; Paroqi vd., 2020; Yılmaz, 2020; Kösece ve Doğanay, 2021; Yonucuoğlu ve Bindak, 2021; Güner ve Karataş, 2022; Filiz, 2023;). Bu çalışmaların büyük çoğunluğunun ortaokul seviyesindeki gruplar ile gerçekleştirildiği görülebilir (Demirdöğen ve Kaçar, 2010; Ersoy, 2013; Korkmaz ve Korkmaz, 2017; Korkmaz ve Tutak, 2017; Erdoğan ve Ayvaz Tuncel, 2018; Berkant ve Yaren, 2020; Chairil Hikayat vd., 2020; Özkan, Aksoy ve Çınar, 2023). İlkokul seviyesinde (Akkaya, 2010; Çilingir ve Dinç Artut, 2016; Uça ve Saracoğlu, 2017; Saleh vd., 2018; Filiz, 2023;), lise seviyesinde (Palinussa, 2013; Çelik, 2016) ve üniversite seviyesinde (Paroqi vd., 2020) de gerçekleştirilen çalışmalar mevcuttur. Ayrıca öğretmen adayları ile gerçekleştirilmiş çalışmalara da rastlanmıştır (Çilingir, Artut ve Tarım, 2015; Meletiou-Mavrotheris ve Prodromou, 2016; Sevinc ve Lesh, 2018; Yılmaz, 2020). Yapılan çalışmalar incelendiğinde, araştırmacıların erişebildiği kadarıyla dezavantajlı öğrencilerle yürütülmüş çalışmalara rastlanmamıştır. Bu alanda yürütülen çalışmalara ihtiyaç olduğu açıktır.

Eğitimsel dezavantaj, anne karnında başlar, anne ve çocuk sağlığı hizmetleri bir eğitim zorunluluğudur. (UNESCO, 2010). Araştırmalar, ebeveyn uygulamalarının çocukların matematik başarısı üzerindeki etkisinin son derece cesaret verici olduğunu göstermektedir (Banerjee, 2016). Ebeveynlerinin eğitim beklentileri yüksek olan ve anne eğitim yılı yüksek olan çocuklar, anaokulu ve sonrasında öğrenmeye hazır olma durumlarını

iyileştiren temel matematik becerilerini öğrenmede tipik öğretmen destekli fırsatlardan daha iyi yararlanabilmektedir (Wang, Shen ve Byrnes, 2013). Ancak aileden uzak, korunmaya muhtaç çocukları düşündüğümüzde, bu çocukların matematik becerilerinin gelişme fırsatları diğer tüm öğrenciler ile eş değer olmayabilir. Türkiye’de çok sayıda korunmaya muhtaç çocuk, Aile ve Sosyal Hizmetler Bakanlığı’na bağlı Çocuk Hizmetleri Genel Müdürlüğü yuva ve yurtlarında barınmaktadır. Bu öğrencilerin küçük yaşlarda ilgili kurum bakımı altında olması, korunmaya muhtaç olması, onların bireysel ilgi ve sevgi ihtiyaçlarının karşılanması ve nitelikli bir yaşam sürdürebilmesi açısından bu çocukların dezavantajlı olduğu söylenebilir. (Topalsan, Türk ve Güler, 2019). Pateman ve Lim (2013), matematik öğreniminin adil ve erişilebilir olması için; sosyal ve kültürel geçmiş, cinsiyet, dini inançlar, etnik köken, coğrafi konum ve ailenin mali durumu ne olursa olsun, tüm öğrencilerin aynı ‘öğrenme fırsatına’ sahip olması gerektiğini açıklamışlardır. Bu da çocukların seviyelerine göre düzenlenmiş, mümkün olduğunca çok duyunun öğrenme sürecinde yer alacağı “matematik etkinlikleri” ile mümkündür (Görgün, 2017; Çelik, 2018). Çocukların matematik alanında ilerlemelerini desteklemek için onlara araştırma yapabilecekleri, meraklarını giderebilecekleri, neden-sonuç ilişkilerini görebilecekleri ve tahminlerde bulunabilecekleri fırsatlar sunulmalıdır. Yaparak yaşayarak öğrenme stratejileri ile tasarlanmış etkinlikler, öğrencilerin sosyoekonomik durumları-öğrenme stilleri veya kültürleri farklı olsa bile grup ile beraber olarak iyi anlaşabilmeyi destekler (Görgün, 2017). Bu bağlamlarda, TÜBİTAK 4004 Doğa Eğitimi ve Bilim Okulları kapsamında gerçekleştirilen projelerin topluma ve bireylere sağladığı kazanımlar, projelerin önemi ve yeni projelerin gerçekleştirilmesinin gerekliliği ortaya çıkmaktadır (Evcan vd., 2020). Bu tür projelerin dezavantajlı çocuklara ulaştırılması, çalışmaların yararlarını önemli ölçüde arttırmaktadır (Topalsan vd., 2019). Bu bağlamlarda bu çalışmada da anne-baba mahrumiyeti olan ve sosyal hizmetlerin himayesindeki dezavantajlı öğrencilerle çalışılması uygun görülmüştür. Çalışmanın amacı, TÜBİTAK 4004 Doğa Eğitimi ve Bilim Okulları “Yaşamdaki Matematikğin Farkındayım-II” projesi kapsamında dezavantajlı öğrenci grubuyla bir hafta boyunca gerçek yaşam temelli matematiksel etkinliklerin uygulandığı atölyeler gerçekleştirilmesidir. Öğrencilerin katıldığı bu etkinliklerin öğrencilerin matematiğe yönelik bakış açılarını, motivasyonlarını ve matematik okuryazarlık-öz yeterlilik düzeylerini olumlu yönde etkilemesi beklenmiştir. Bunun için öğrencilere farklı tipte ölçekler uygulanmıştır. Bu amaçlara yönelik olarak aşağıdaki araştırma sorularına yanıt aranmıştır:

1. Çalışmaya katılan öğrencilerin ölçeklere verdiği cevaplara ait genel düzeyler (bakış açısı, motivasyon, okuryazarlık-özyeterlilik) nelerdir?

2. Öğrencilerin matematiğe bakış açıları cinsiyet ve sınıf düzeyi değişkenlerine göre anlamlı bir farklılık gösterir mi?

3. Öğrencilerin okuryazarlık-özyeterlikleri cinsiyet ve sınıf düzeyi değişkenlerine göre anlamlı bir farklılık gösterir mi?

## 2. Yöntem

Bu bölümde bu çalışmada kullanılan araştırma deseni, örneklem, verilerin toplanması ve verilerin analizi başlıklarına yer verilecektir.

### 2.1. Araştırma Deseni ve Örneklem

Bu çalışmada parametrik olmayan yöntem tasarımı kullanılmıştır. Çalışma örneklem olarak seçilen anne baba mahrumiyeti olan ve sosyal hizmetlerin himayesindeki öğrencilerle gerçekleştirilmiştir. İstanbul ili Avrupa yakası 4 farklı ÇEKOM'dan 30 öğrenciye (5, 6 ve 7. sınıflar olmak üzere) uygulanmıştır. Anket uygulamalarında yargısal örnekleme yöntemi kullanılmıştır.

### 2.2. Verilerin Toplanması

Araştırmada veri toplama yöntemi olarak anket tekniğinden yararlanılmıştır. Çalışmada çeşitli Matematik ve Ben ölçeği (Aktaş Cansız ve Tabak, 2018), Matematik Motivasyon Ölçeği (Yavuz, Özyıldırım ve Doğan, 2012) ve Ortaokul Matematik Okuryazarlık-Öz yeterlilik Ölçeği (Baypınar ve Tarım, 2019) kullanılmıştır. Bu ölçeklerin ilk bölümünde katılımcıların demografik özelliklerini (cinsiyet, sınıf, okuduğu okul tipi) belirlemeyi amaçlayan 3 kapalı uçlu soru yer almaktadır. Anketin ikinci bölümünde katılımcıların ilgili ölçek düzeylerini ölçmek amacıyla çeşitli sayılarda kapalı uçlu sorulara yer verilmiştir. Uygulanan ankette demografik ifadeler hariç diğer bütün ifadeler 5'li Likert ölçekleri ve 2 seçenekli anket soruları (Matematisel motivasyon ölçeği) şeklinde düzenlenmiştir. Araştırmanın amacı açısından en olumlu ifade en yüksek puanı, en olumsuz ifade ise en düşük puanı simgelemektedir. Bazı sorular, araştırmanın güvenilirliğini sağlamak amacıyla olumsuz olarak sorularak ters kodlanmıştır.

### 2.3. Verilerin Analizi

Soru formu yardımı ile toplanan veriler SPSS 18 istatistiksel veri analizi paket programı ile analiz edilmiştir. Elde edilen verilere göre demografik özelliklerin frekans ve yüzde dağılımları bulunmuştur. Ankette sorulan her soru için frekans, yüzde, ortalama ve standart sapma hesaplanmıştır. Ayrıca demografik ölçekler arasında Mann-Whitney U ve Kruskal Wallis testi uygulanmıştır.

### 3. Bulgular

Bu bölümde, ölçeklerden elde edilen bulgular alt başlıklar halinde verilecektir.

#### 3.1. Matematik ve Ben Ölçeği Bulguları

Tablo 1'den de görüldüğü gibi örnekleme oluşturan katılımcıların %51,7'si erkek, %48,3'ü kızdır. Katılımcıların sınıf grupları arasında en yüksek oranla %57,7 'si 6. sınıf, en düşük oranla %19,2'si 7. sınıflardır. Bireylerin okudukları okul tiplerine göre ise %92,3'ü devlet okullarından %7,7'si ise özel okullardan çalışmaya katılmıştır.

*Tablo 1. Katılımcıların demografik özelliklerine ilişkin bulgular*

Değişkenler	Gruplar	Frekans	%
Cinsiyet	Erkek	15	51,72
	Kız	14	48,28
Sınıf	5. Sınıf	6	23,08
	6. Sınıf	15	57,69
	7. Sınıf	5	19,23
Okul Tipi	Devlet	24	92,31
	Özel	2	7,69

#### 3.1.1. Ölçeğe Ait Güvenirlilik Analizi

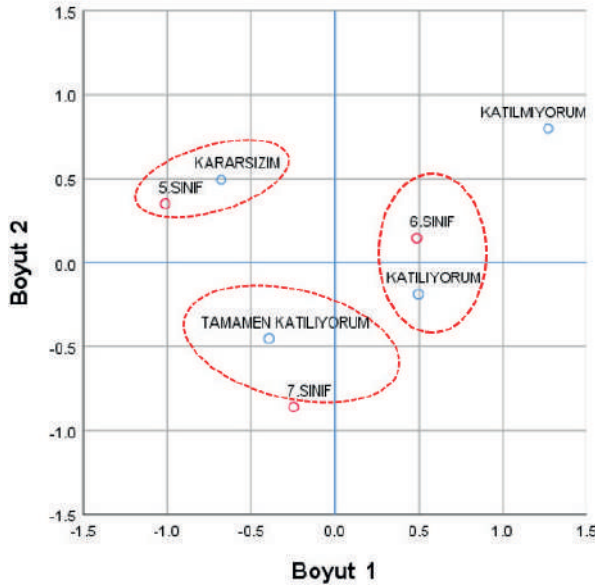
Çalışmaya ait Cronbach's Alpha güvenilirlik katsayısı %90 (0,904) olarak hesaplanmıştır ve Matematik ve Ben ölçeği için güvenilirliğinin yüksek seviyede olduğu söylenebilir.

*Tablo 2. Matematik ve Ben Ölçeği Güvenirlilik Analizi*

Cronbach's Alpha	Soru Sayısı
0,904	20

#### 3.1.2. Sınıf Seviyesi ile Ölçeğin Genel Tutumunun Harita ile İncelenmesi

Bu ilişkinin harita üzerinde görüntülenmesi Correspondence (Uyum) Analizi ile yapılmıştır.



Şekil 1. Matematik ve Ben Ölçeği ile Sınıf Seviyesi Uyum Haritası

Bu ilişki haritası incelendiğinde Matematik ve Ben ölçeğine karşı 5. sınıflar kararsız statüye daha yakın olarak görülmektedir. Diğer yandan 6. ve 7. sınıflar ise ölçeğe karşı daha olumlu (katılıyorum ve Tamamen katılıyorum kümesine yakın olarak) tutum sergilemişlerdir.

### 3.1.3. Ölçek İçi İstatistikler

İlgili araştırma hipotezleri aşağıdaki gibidir. Hipotez 1:

$H_0$ : Cinsiyet ile matematik ve ben ölçeği arasındaki ilişki istatistiksel olarak anlamsızdır.

$H_1$ : Cinsiyet ile matematik ve ben ölçeği arasındaki ilişki istatistiksel olarak anlamlıdır.

Bu ilişkinin istatistiksel anlamlılığı Mann-Whitney U testi ile ölçülmüştür.

Tablo 3. Cinsiyet ve Matematik ve Ben Ölçeği Mann Whitney U Testi Sonuçları

İstatistiksel Analiz	Değer
Mann-Whitney U	96
Wilcoxon W	201
Z	-0,393
Asymp. Sig. (2-tailed)	0,694

Tablodaki verilere göre ( $p=0,694>0,05$ ) olduğundan cinsiyet ile matematik ve ben ölçeği arasında istatistiksel anlamlı bir ilişki bulunamamıştır ( $H_0$  reddedilemez).

Hipotez 2 ise:

$H_0$  : Sınıf seviyesi ile matematik ve ben ölçeği arasındaki ilişki istatistiksel olarak anlamsızdır.

$H_1$  : Sınıf seviyesi ile matematik ve ben ölçeği arasındaki ilişki istatistiksel olarak anlamlıdır.

Bu ilişkinin istatistiksel anlamlılığını Kruskal Wallis testi ile ölçülmüştür.

**Tablo 4. Sınıf Seviyesi ve Matematik ve Ben Ölçeği Kruskal Wallis Testi Sonuçları**

İstatistiksel Analiz	Değer
Kruskal-Wallis H	0,572
df	2
Asymp. Sig.	0,751

Tablodan verilere göre ( $p=0,751>0,05$ ) olduğundan sınıf seviyesi ile matematik ve ben ölçeği arasında istatistiksel anlamlı bir ilişki bulunamamıştır ( $H_0$  reddedilemez).

### 3.2. Matematiksel Motivasyon Ölçeği Bulgular

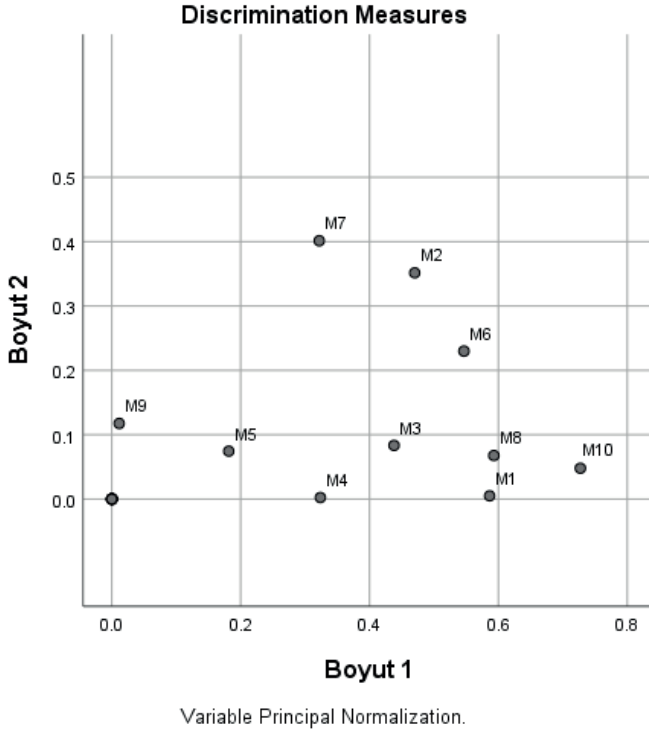
Tablo-1'den de görüldüğü gibi örnekleme oluşturan katılımcıların %51,72'si erkek, %48,28'i kızdır. Katılımcıların sınıf grupları arasında en yüksek oranla %53,57 'si 6. sınıf, en düşük oranla %21,43'ü 7. sınıflardır. Bireylerin okudukları okul tiplerine göre ise %92,86'sı devlet okullarından %7,14'ü ise özel okullardan çalışmaya katılmıştır.

**Tablo 5. Katılımcıların Demografik Özelliklerine İlişkin Bulgular**

Değişkenler	Gruplar	Frekans	%
Cinsiyet	Erkek	15	51,72
	Kız	14	48,28
Sınıf	5. Sınıf	7	25,0
	6. Sınıf	15	53,57
	7. Sınıf	6	21,43
Okul Tipi	Devlet	26	92,86
	Özel	2	7,14

### 3.2.1. Sınıf Seviyesi ile Ölçeğin Genel Tutumunun Harita ile İncelenmesi

Bu ilişkinin harita üzerinde görüntülenmesi Uyum (Correspondence) Analizi ile yapılmıştır.



*Şekil 2. Sınıf Seviyesi ve Matematiksel Motivasyon Ölçeği Uyum Haritası*

Toplam 10 sorudan oluşan bu iki seçenekli yapıda oluşan harita, merkez noktası (0,0) dikkate alınarak yorumlanır. Ölçekteki en etkili sorular merkeze en yakın sorulardır. Bunlar M9, M5 ve M4 olarak sayabiliriz.

### 3.3. Ortaokul Matematik Okuryazarlık-Öz yeterlilik Ölçeği Bulgular

Tablo 6'dan da görüldüğü gibi örnekleme oluşturan katılımcıların %55,2'si erkek, %44,8'i kızdır. Katılımcıların sınıf grupları arasında en yüksek oranla %57,7 'si 6. sınıf, en düşük oranla %19,2'si 7. sınıflardır. Bireylerin okudukları okul tiplerine göre ise %92,3'ü devlet okullarından %7,7'si ise özel okullardan çalışmaya katılmıştır.

**Tablo 6. Katılımcuların Demografik Özelliklerine İlişkin Bulgular**

Değişkenler	Gruplar	Frekans	%
Cinsiyet	Erkek	16	55,17
	Kız	13	44,83
Sınıf	5. Sınıf	6	23,08
	6. Sınıf	16	57,69
	7. Sınıf	5	19,23
Okul Tipi	Devlet	24	92,31
	Özel	2	7,69

### 3.3.1. Ölçeğe Ait Genel ve Alt Ölçek Bazında Güvenirlik Analizi

Çalışmaya ait Cronbach's Alpha güvenilirlik katsayısı %86 (0,864) olarak hesaplanmıştır ve bu ölçek için güvenilirliğinin yüksek seviyede olduğu söylenebilir.

**Tablo 7. Ölçeğe Ait Genel Cronbach's Alpha**

Cronbach's Alpha	Soru Sayısı
0,864	30

Tablo 8'de görüldüğü gibi, alt ölçeklerin Cronbach's Alpha katsayıları 0,70'den büyüktür. Bu sebeple alt ölçekler “güvenilirdir” denir.

**Tablo 8. Alt Ölçek Bazında Güvenirlik Analizi**

Boyutlar	Cronbach's Alpha	Güvenilirlik Düzeyi
Matematiksel Beceri Boyutu	0,90 (%90)	Çok Yüksek
Kişisel Deneyim Boyutu	0,83 (%83)	Yüksek
Bilimsel Modelleme ve Sosyal Bağlam Boyutu	0,78 (%78)	Yüksek

### 3.3.2. Alt Ölçek Bazında Korelasyon Analizi

Alt ölçek bazında yapılan korelasyon analizi sonuçları Tablo 9'da verilmiştir.



Tablo 9. Alt Ölçekler Arası Korelasyon Tablosu

Korelasyon Katsayısı $r$	Matematiksel Beceri	Kişisel Deneyim	Bilimsel Modelleme ve Sosyal Bağlam
Matematiksel Beceri	1	-0,37	0,70
Kişisel Deneyim	-0,37	1	0,05
Bilimsel Modelleme ve Sosyal Bağlam	0,70	0,05	1

Tablo 9'daki verilere göre:

- Matematiksel Beceri ile Bilimsel Modelleme ve Sosyal Bağlam arasında pozitif yönlü güçlü bir ilişki vardır (0,70- %70). Bu ölçekler ortak olarak artmakta ve azalmaktadır.
- Matematiksel Beceri ile Kişisel Deneyim arasında ise düşük seviyeli negatif yönde ilişki vardır (-0,37- %37). Matematiksel Beceri arttıkça Kişisel Deneyim azalmaktadır.
- Kişisel Deneyim ile Bilimsel Modelleme ve Sosyal Bağlam arasında ise ilişki yok denecek azdır.

### 3.3.3. Ölçek İçi İstatistiksel Testler

İlgili araştırma hipotezleri aşağıdaki gibidir. Hipotez 1:

$H_0$ : Cinsiyet ile ortaokul matematik okuryazarlık-öz yeterlilik ölçeği arasındaki ilişki istatistiksel olarak anlamsızdır.

$H_1$ : Cinsiyet ile ortaokul matematik okuryazarlık-öz yeterlilik ölçeği arasındaki ilişki istatistiksel olarak anlamlıdır.

Bu ilişkinin istatistiksel anlamlılığı Mann-Whitney U testi ile ölçülmüştür.

Tablo 10. Matematik Okuryazarlık-Öz Yeterlilik Ölçeği ve Cinsiyet Mann-Whitney U Testi Sonuçları

İstatistiksel Analiz	Değer
Mann-Whitney U	97
Wilcoxon W	188
Z	-0,309
Asymp. Sig. (2-tailed)	0,758

Tablodan verilere göre ( $p=0,758>0,05$ ) olduğundan cinsiyet ile ortaokul matematik okuryazarlık-özyeterlilik ölçeği arasında istatistiksel anlamlı bir ilişki bulunamamıştır ( $H_0$  reddedilemez).

Hipotez 2 ise:

$H_0$  : Sınıf seviyesi ile ortaokul matematik okuryazarlık-özyeterlilik ölçeği arasındaki ilişki istatistiksel olarak anlamsızdır.

$H_1$  : Sınıf seviyesi ile ortaokul matematik okuryazarlık-özyeterlilik ölçeği arasındaki ilişki istatistiksel olarak anlamlıdır.

Bu ilişkinin istatistiksel anlamlılığı Kruskal Wallis testi ile ölçülmüştür.

*Tablo 11. Sınıf seviyesi ile ortaokul matematik okuryazarlık-özyeterlilik ölçeği Kruskal Wallis testi sonuçları*

İstatistiksel Analiz	Değer
Kruskal-Wallis H	6,371
df	2
Asymp. Sig.	0,041

Tablodan verilere göre ( $p=0,041<0,05$ ) olduğundan sınıf seviyesi ile ölçek arasında istatistiksel anlamlı bir ilişki vardır.

#### 4. Tartışma ve Sonuç

Bu çalışmanın amacı, sosyal hizmetler himayesinde bulunan korunmaya muhtaç dezavantajlı bir öğrenci grubuyla gerçek yaşam temelli matematiksel etkinliklerin uygulandığı atölyeler gerçekleştirilmesidir. Öğrencilerin katıldığı bu etkinliklerin öğrencilerin matematiğe yönelik bakış açılarını, motivasyonlarını ve matematik okuryazarlık-öz yeterlilik düzeylerini olumlu yönde etkilemesi de çalışmanın amacını oluşturmaktadır.

Çalışmanın sonuçları incelendiğinde, öğrencilerin "Matematik ve Ben" ölçeğine verdiği cevaplarda 6 ve 7. sınıf seviyelerinde daha olumlu tutum sergilediği görülmüştür. Bu durum, 5. Sınıf öğrencilerine göre daha üst sınıflardaki öğrencilerin matematikle daha uzun bir geçmişlerinin olması ile açıklanabilir. Bu bağlamda yürütülen etkinliklere olumlu tutum sergilemeleri beklenen bir durumdur. Benzer şekilde literatürde Duran vd., (2018) ilkokul 4. Sınıf öğrencileri ile gerçekleştirdiği çalışmada, öğrencilerin matematik etkinliklerine yönelik tutumlarını ölçmeyi amaçlamıştır. Öğrencilerin matematik etkinliklerine yönelik tutumlarının yüksek seviyede olduğu belirlenmiştir. Yang vd., (2016) küçük yaş gruplarıyla gerçekleştirdiği

çalışmada, matematik etkinliklerinin matematiğe yönelik tutumu olumlu yönde etkilediğini belirtmiştir. Kurt ve Doğan (2019) öğrencilerle GME etkinlikleri içeren öğretimin yürütülmesi sonucunda genel olarak öğrencilerin matematik dersini daha çok sevdiğini ve dersten daha çok keyif aldıklarını ortaya koymuşlardır. Çelik (2018) çalışmasında deney ve kontrol grupları kullanarak, geleneksel ve etkinlik temelli eğitim alan öğrencilerin tutumlarını karşılaştırmıştır. Mevcut çalışmada elde edilen sonuçlardan farklı olarak, etkinlik temelli eğitim alan öğrencilerin tutum puanlarının düşük ve geleneksel eğitim alan kontrol grubunun puanları yükselmiştir. Korkmaz ve Korkmaz (2017) da gerçekçi matematik eğitimi yaklaşımı ile işlenen derslerin diğer bir yaklaşıma göre öğrencilerin tutumlarına etkisinin olmadığını belirtmiştir. Cinsiyet ile Matematik ve Ben ölçeğine verilen cevaplar arasında istatistiksel anlamlı bir ilişki bulunamamıştır. Çalışmaya katılan gruptaki öğrencilerin matematiğe karşı bakış açılarının cinsiyete göre farklılaşmamaktadır. Benzer şekilde, Duran vd. (2018) de, cinsiyet ve öğrencilerin matematiğe yönelik tutumları arasında anlamlı bir ilişki olmadığını ifade etmiştir. Elde edilen diğer bir sonuç ise, sınıf seviyesi ile matematik ve ben ölçeği arasında istatistiksel olarak anlamlı bir ilişki olmadığıdır. Ancak öğrencilerin genel durumları uyum analizi ile incelendiğinde, elde edilen genel haritada, daha üst sınıftaki öğrencilerin matematiğe bakış açılarının daha olumlu olduğu sonucuna ulaşılmıştır.

Öğrencilerin matematik okuryazarlığı ve öz yeterliliği için yapılan ölçümlerde, matematiksel beceri ile bilimsel modelleme ve sosyal bağlam arasında pozitif yönlü güçlü bir ilişki vardır. Bu ölçekler ortak olarak artmakta ve azalmaktadır. Bu sonuç öğrencilerin matematiksel becerileri arttıkça bağlı olarak modelleme becerileri ve sosyal becerilerinin de arttığını gösterir. Benzer şekilde, Çilingir ve Dinç Artut (2016) tarafından gerçekleştirilen çalışmada öğrencilerin görsel matematik okuryazarlığı-özyeterlik algılarında gelişim gösterdikleri belirtilmiştir. Matematiksel Beceri ile Kişisel Deneyim arasında ise düşük seviyeli negatif yönde ilişki ortaya çıkmıştır. Yani matematiksel beceri arttıkça kişisel deneyim azalmaktadır. Bu durum, öğrencilerin katıldıkları etkinlik süreçlerinde genellikle grup etkinliklerinde bulunmaları ile açıklanabilir. İşbirlikli çalışmalar daha öne çıkmaktadır. Paroqi vd. (2020) de GME öğreniminin öğrencilerin matematiksel iletişim becerilerini geliştirebilir olduğunu belirtmiştir. Kişisel deneyim ile bilimsel modelleme ve sosyal bağlam arasında ise ilişki yok denecek azdır. Ancak Yorulmaz ve Doğan (2019) tarafından yapılan çalışmada matematiği gerçek yaşamda kullanmanın eğlenceli bir öğrenme ortamı sağlayacağı belirtilmiş ve gerçek hayattaki örnekleri matematikte kullanmak amacıyla gerçekçi matematik eğitimi etkinliklerinin kullanılmasının önemi üzerinde durulmuştur.

Sonuçta, bu çalışmaya katılan öğrencilerin uygulanan etkinliklerde akademik kazanımlardan çok sosyalleşme ve üniversite ortamında bulunmanın heyecanını yaşadıkları söylenebilir. Çalışma boyunca, öğrencilerin sosyal becerilerinde gelişmelerin olması için etkinliklerin pek çoğu bireysel değil grup olarak yapılacak şekilde tasarlanmıştır. İşbirlikli olarak faaliyetleri gerçekleştirdikleri için akranları ile çalışmanın hedeflerine daha etkin bir şekilde ulaştıkları düşünülmektedir. Bu çalışmalar sayesinde öğrencilerde, akran öğrenmesi ve işbirlikli öğrenmenin gelişmesini sağladığı düşünülmektedir. Katılımcı öğrencilerimiz açısından etkinliklerin akademik boyutundan ziyade sosyal boyutunun daha fazla ilgilerini çektiği görülmüştür. Bu sonuçların da dezavantajlı öğrenciler göz önünde bulundurulduğunda önemli olduğu düşünülmektedir. Bu gruplarla yürütülen çalışmaların arttırılması gerektiğine inanıyoruz. Farklı yaş gruplarında yürütülen çalışmalar ile farklı sonuçlar elde edilebilir. Ayrıca disiplinlerarası etkinliklerin arttırılması, öğrencilerin ilişkiler kurup matematiğe yönelik bakış açılarının geliştirilmesinde faydalı olacaktır.

### **Notlar**

Bu çalışma, TÜBİTAK-4004 Doğa ve Bilim Okulları kapsamında gerçekleştirilen “Yaşamdaki Matematiğin Farkındayım-II” başlıklı projeden üretilmiştir.

## Kaynakça

- Akkaya, R. (2010). Olasılık ve istatistik öğrenme alanındaki kavramların gerçekçi matematik eğitimi ve yapılandırıcılık kuramına göre bilgi oluşturma sürecinin incelenmesi. Doktora Tezi. Uludağ Üniversitesi
- Aktaş Cansız, M. & Tabak, S. (2018). Turkish adaptation of Math and Me Survey: A validity and reliability study. *European Journal of Educational Research*, 7(3), 707-714.
- Bahadır, E. (2017). Associating Mathematical Stories That Are Written By the 8th Grade Students Who Are Studying At Advantageous And Disadvantageous Regions'schools With Their Mathematical Perceptions: Istanbul Case. *European Journal of Education Studies*. 3(4).
- Bahadır, E. (2021). Göçmen öğrencilerin matematik eğitiminde etnomatematik yaklaşımının kullanılması. *Milli Eğitim Dergisi*, 50(1), 577-59.
- Banerjee, P. A. (2016). A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools. *Cogent Education*, 3(1), 1178441.
- Baypınar, K., & Tarım, K. (2019). The development of mathematical literacy self-efficacy scale for middle school: a reliability and validity study. *Cukurova University Faculty of Education Journal*, 48(1), 878-909.
- Berkant, H. G., & Yaren, R. (2020). Altıncı sınıf tam sayılar konusunda uygulanan gerçekçi matematik eğitiminin öğrencilerin matematik motivasyonlarına etkisi. *Kahramanmaraş Sütçü İmam Üniversitesi Sosyal Bilimler Dergisi*, 17(2), 543-571.
- Chairil Hikayat, S., Hairun, Y., & Suharna, H. (2020). Design of realistic mathematics education approach to improve critical thinking skills. *Universal Journal of Educational Research*, 8(6), 2232-2244.
- Çelik, A. (2016). Koniklerin gerçekçi matematik eğitimi yaklaşımı ile öğretimi üzerine bir araştırma. Yüksek Lisans Tezi. Bilecik Şeyh Edebali Üniversitesi.
- Çelik, H. C. (2018). The effects of activity based learning on sixth grade students' achievement and attitudes towards mathematics activities. *EURASIA Journal of Mathematics, Science and Technology Education*, 14(5), 1963-1977.
- Çilingir, E., & Dinç Artut, P. (2016). Gerçekçi matematik eğitimi yaklaşımının ilkökul öğrencilerinin başarılarına, görsel matematik okuryazarlığı öz yeterlik algılarına ve problem çözme tutumlarına etkisi. *Turkish Journal of Computer and Mathematics Education*, 7(3), 578-600.
- Çilingir, E., Dinç Artut P. & Tarım, K. (2015). Sınıf öğretmeni adayları üzerinde gerçekçi matematik eğitimine ilişkin bir uygulama örneği. *Eğitim Kuram ve Uygulama Araştırmaları Dergisi*, 1(1), 1-12.

- Demir, G. (2017). Gerçekçi matematik eğitimi yaklaşımının meslek lisesi öğrencilerinin matematik kaygısına, matematik özyeterlik algısına ve başarısına etkisi. Yüksek Lisans Tezi. Adnan Menderes Üniversitesi.
- Demirdöğen, N., & Kaçar, A. (2010). İlköğretim 6. Sınıfta Kesir Kavramının Öğretiminde Gerçekçi Matematik Eğitimi Yaklaşımının Öğrenci Başarısına Etkisi. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 12(1), 56-74.
- Duran, C., Sidekli, S., & Yorulmaz, A. (2018). İlkokul Dördüncü Sınıf Öğrencilerinin Matematik Etkinliklerine Yönelik Tutumlarının İncelenmesi. *International Primary Education Research Journal*, 2(1), 17-26.
- Erdoğan, H., & Ayvaz-Tüncel, Z. (2018). Gerçekçi Matematik Eğitimine Davalı Matematik Öğretiminin Akademik Başarı, Kalıcılık Ve Yansıtıcı Düşünme Becerisi Üzerine Etkisi. *Electronic Turkish Studies*, 13(19).
- Ergül, E., & Doğan, M. (2022). İlkokul Matematik Öğretiminde Oyun Temelli Yaklaşımın Öğrenci Başarısına Etkisi. *Millî Eğitim Dergisi*, 51(235), 1935-1960.
- Ersoy, E. (2013). Gerçekçi matematik eğitimi destekli öğretim yönteminin 7. sınıf olasılık ve istatistik kazanımlarının öğretiminde öğrenci başarısına etkisi. Doktora Tezi. Sakarya Üniversitesi.
- Evcan, S. S., Adilov, G., Eken, Z., Barut, S., Kemali, S., & Tınaztepe, G. (2020). TÜBİTAK 4004-Doğa eğitimi ve bilim okulları kapsamında 7. sınıf öğrencilerine yönelik gerçekleştirilen "hayatımızdaki matematik: tarım" projesinin değerlendirilmesi. *International Journal of Scholars in Education*, 3(1), 28-41.
- Fauzan, A., Slettenhaar, D., & Plomp, T. (2002). Traditional mathematics education vs. realistic mathematics education: Hoping for changes. In *Proceedings of the 3rd International Mathematics Education and Society Conference* (pp. 1-4).
- Filiz, T. (2023). Gerçekçi Matematik Eğitimi Yaklaşımının İlkokul Öğrencilerinin Matematik Performansına Etkisi: Meta-Analiz. *Korkut Ata Türkiyat Araştırmaları Dergisi*, (Özel Sayı 1 (Cumhuriyetin 100. Yılına)), 1062-1081.
- Freudenthal, H. (1968). Why to teach mathematics so as to be useful? *Educational Studies in Mathematics* 1(1), 3-8.
- Freudenthal, H. (1973). *Mathematics as an educational task*. Reidel.
- Golji, G. G., & Dangpe, A. K. D. (2016). Activity-based learning strategies (ABLS) as best practice for secondary mathematics teaching and learning. *International Advanced Journal of Teaching and Learning*, 2(9), 106-116.
- Gökçe, Z., Gökçe, H., Büyükata, M. & Güneri, E., (2022). Dezavantajlı Okullarda TÜBİTAK Tarafından Düzenlenen Bilim ve Teknoloji Haftası Etkinliklerine İlişkin Öğretmen Görüşleri. *ODÜSOBİAD* 12 (2), 1295-1318.

- Görgün, G. (2017). Hands-on aktivitelerin ilköğretim 7. sınıf matematik öğretiminde cebir alanına uygulanmasının öğrenci başarısına etkisi. Yüksek Lisans Tezi. Akdeniz Üniversitesi
- Gravemeijer, K.P.E. (1994). Developing Realistic Mathematics Education. Freudenthal Institute.
- Güner, D., & Karataş, İ. (2022). Matematik Öğretmenlerinin Gerçekçi Matematik Eğitime Yönelik Algılarının İncelenmesi: Bir Ölçek Geliştirme Çalışması. *Gazi Eğitim Bilimleri Dergisi*, 8(2), 155-173.
- Korkmaz, E., & Korkmaz, C. (2017). Ebob-Ekok Konusunun Gerçekçi Matematik Eğitimi Etkinlikleriyle Öğretiminin Başarı Ve Tutuma Etkisi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 14(39), 504-523.
- Korkmaz, E., & Tutak, T. (2017). Dönüşüm geometrisi konularının gerçekçi matematik eğitimi etkinlikleriyle işlenmesinin öğrenci başarısına ve matematik tutumuna etkisi. *Disiplinlerarası Eğitim Araştırmaları Dergisi*, 1(2), 30-42.
- Kurt, E. S., & Doğan, M. (2019). Gerçekçi matematik öğretimi ile gerçekleştirilen uzunlukları ölçme konusunda öğrenci görüşleri. *Temel Eğitim*, 1(3), 33-38.
- Laurens, T., Batlolona, F. A., Batlolona, J. R., & Leasa, M. (2018). How does realistic mathematics education (RME) improve students' mathematics cognitive achievement? *Eurasia Journal of Mathematics, Science and Technology Education*, 14(2), 569-578. <https://doi.org/10.12973/ejmste/76959>
- Meletiou-Mavrotheris, M., & Prodromou, T. (2016). Pre-service teacher training on game-enhanced mathematics teaching and learning. *Technology, Knowledge and Learning*, 21, 379-399.
- Moyer, P. S. (2001). Are we having fun yet? How teachers use manipulatives to teach mathematics. *Educational Studies in Mathematics*, 47, 175-197.
- National Council of Teachers of Mathematics (NCTM). (2000). Principles and Standards for School Mathematics.
- Olkun, S., & Toluk-Uçar, Z. (2003). İlköğretimde etkinlik temelli matematik öğretimi. Anı Yayıncılık.
- Özkan, M., Aksoy, N. C., & Çınar, C. (2023). Cebir öğretiminde gerçekçi matematik eğitimi yaklaşımının öğrenci başarısına ve kalıcılığına etkisi [Effect of realistic mathematics education approach in algebra teaching on student success and permanence]. *Kalem Eğitim ve İnsan Bilimleri Dergisi*, 13(2), 201-221.
- Palinussa, A. L. (2013). Students' Critical Mathematical Thinking Skills and Character: Experiments for Junior High School Students through Re-

- alistic Mathematics Education Culture-Based. Indonesian Mathematical Society Journal on Mathematics Education, 4(1), 75-94.
- Paroqi, L. L., Mursalin, M., & Marhami, M. (2020). The Implementation of Realistic Mathematics Education Approach to Improve Students' Mathematical Communication Ability in Statistics Course. *International Journal for Educational and Vocational Studies*, 2(10).
- Pateman, N. A. & Lim, C. S. (2013). The Politics of Equity and Access in Teaching and Learning Mathematics. In M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Third International Handbook of Mathematics Education* (pp. 243-263). Springer.
- Rasmussen, C., Zandieh, M., King, K., & Teppo, A. (2005). Advancing mathematical activity: A practice-oriented view of advanced mathematical thinking. *Mathematical thinking and learning*, 7(1), 51-73.
- Riley, N., Luban, D., Holmes, K., Gore, J., & Morgan, P. (2017). Movement-based mathematics: Enjoyment and engagement without compromising learning through the easy minds program. *EURASIA Journal of Mathematics Science and Technology Education*, 13(6):1653-1673.
- Romberg, T. A., & Kaput, J. J. (1999). Mathematics worth teaching, mathematics worth understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 3-17). Lawrence Erlbaum.
- Saleh, M., Prahmana, R. C. I., & Isa, M. (2018). Improving the Reasoning Ability of Elementary School Student through the Indonesian Realistic Mathematics Education. *Journal on Mathematics Education*, 9(1), 41-54.
- Sevinc, S., & Lesh, R. (2018). Training mathematics teachers for realistic math problems: a case of modeling-based teacher education courses. *ZDM*, 50, 301-314.
- Sezer, E. N. (2019). Gerçekçi matematik eğitimi çerçevesinde tasarlanan etkinliklerin uygulama sürecinin klasik ve elektronik portfolyo ile değerlendirilmesi. Yüksek Lisans Tezi. Yıldız Teknik Üniversitesi.
- Švecová, V., Rumanová, L., & Pavlovičová, G. (2014). Support of Pupil's Creative Thinking in Mathematical Education. *Procedia Social and Behavioral Sciences*, 116, 1715-1719.
- Topalsan, A. K., Türk, Z. & Güler, G. (2019). Korunmaya muhtaç çocuklara yönelik gerçekleştirilen "Doğada Bilim Yapıyorum!" TÜBİTAK 4004 projesinin değerlendirilmesi. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 8(1), 581-607.
- Treffers, A. (1987). *Three Dimensions- A Model of Goal and Theory Description in Mathematics Instruction*. Kluwer Academic.



- Treffers, A. (1991). Realistic mathematics education in the Netherlands 1980-1990. L. Streefland içinde, Realistic Mathematics Education in Primary School. CD-β Press.
- Uça, S., & Saracoğlu, A. S. (2017). Öğrencilerin ondalık kesirleri anlamlandırmasında gerçekçi matematik eğitimi kullanımı: Bir tasarı araştırması. *İlköğretim Online*, 16(2), 469-496.
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (2010). Education for all global monitoring report 2010: Reaching the marginalized.
- Ünal, Z. A., & İpek, A. S. (2010). Gerçekçi matematik eğitiminin ilköğretim 7. sınıf öğrencilerinin tam sayılarla çarpma konusundaki başarılarına etkisi. *Eğitim ve Bilim*, 34(152).
- Üzel, D. (2007). Gerçekçi matematik eğitimi (RME) destekli eğitimin ilköğretim 7. sınıf matematik öğretiminde öğrenci başarısına etkisi. Doktora tezi. Balıkesir Üniversitesi
- Van De Walle, J. A., Karp, K. S. & Bay-Williams, J. M. (2012). İlkokul ve ortaokul matematiği gelişimsel yaklaşımla öğretim (Soner Durmuş, Çev.). Nobel Yayıncılık.
- Van Den Heuvel-Panhuizen, M. (1996). Assessment and realistic mathematics education (Vol. 19). Doctoral Dissertation. Utrecht University.
- Wang, A. H., Shen, F., & Byrnes, J. P. (2013). Does the opportunity- propensity framework predict the early mathematics skills of low-income pre-kindergarten children? *Contemporary Educational Psychology*, 38, 259-270.
- Yang, H. S., Park, Y. S., & Cho, K. H. (2018). The effects of mathematical activities using 4D-Frame on young children's mathematical ability and attitude towards mathematics. *J. Korea Acad. Industr. Coop. Soc.*, 19, 146-159.
- Yavuz, G., Ozyıldırım, E., & Doğan, N. (2012). Mathematics motivation scale: a validity and reliability. *Procedia-Social and Behavioral Sciences*, 46, 1633-1638.
- Yılmaz, R. (2020). Prospective Mathematics Teachers' Cognitive Competencies on Realistic Mathematics Education. *Journal on Mathematics Education*, 11(1), 17-44.
- Yorulmaz, A., & Doğan, M. C. (2019). İlkokul Dördüncü Sınıf Öğrencilerinin Gerçekçi Matematik Eğitimine İlişkin Görüşlerinin İncelenmesi. *Eğitim Kuram ve Uygulama Araştırmaları Dergisi*, 5(2), 153-162.
- Yuhariati, Y., Johar, R., Khairunnisak, C., Rohaizati, U., Jupri, A., & Zubaidah, T. (2022). Students mathematical representation ability in learning algebraic expression using realistic mathematics education. *Jurnal Didaktik Matematika*, 9(1), 151-169.

## Faktör Analizi Çözümlemesine Ortak Entropi Yaklaşımı

Nida Oruç<sup>1</sup>

Doğan Yıldız<sup>2</sup>

### Özet

Faktör analizi, belirli bir olguyla ilgili ve birbirleriyle ilişkili çok sayıda değişkenin birbirleriyle olan ilişkisini inceleyerek, değişkenlerin sayısını azaltmak ve ilişkisiz yeni değişkenler yaratmak için kullanılan çok değişkenli istatistiksel bir analizdir. Bu çalışmada farklı bir yaklaşım olarak, faktör çıkarmada birleşik entropiden yararlanılmıştır. Buradan elde edilen “Faktörleşme” ile “Açıklayıcı Faktör Analizi” sonuçları karşılaştırılarak benzerlik ve farklılıkları tartışılmıştır.

İstatistik ile entropiyi birleştiren bir model kullanılmıştır. Uygulama kısmında ilk olarak istatistikte yaygın olarak kullanılan faktör analizi sonuçları elde edilmiştir. Ardından, kovaryans matrisi yerine birleşik entropi matrisi kullanılan entropi tabanlı yaklaşımla elde edilen matrisle bir karşılaştırma yapılmıştır. Bu bağlamda birleşik entropi matrisi ile elde ettiğimiz sonuçlar ile faktör analizi ile elde ettiğimiz sonuçlar karşılaştırılmıştır. Son durumda ise birleşik entropi matrisi ile elde edilen sonuçlar gösterilmiştir. Ek olarak, bu çalışma ile entropiye teorik destek ile aslında varyans-kovaryans ölçülerinin kullanım alanlarına uygulanabileceği ve desteklenebileceğini göstermek hedeflenmiş ve çözülmesi gereken bazı sorunlara rağmen, alternatif olarak bilinen yöntemlerin yerine önerilmektedir.

### 1. Giriş

Bilim insanları, istatistiksel gözlemler veya deneylerde bilgi arayışından yararlanır. Claude Shannon, bu bağlamda öne çıkan bilim adamlarından biridir. Bu çerçevede bakıldığında, 1948 yılında Claude Shannon’ın “A

1 Araştırma Görevlisi, İstanbul Topkapı Üniversitesi, nidaoruc@topkapi.edu.tr, 0009-0005-0230-2234

2 Dr. Öğr. Üyesi, Yıldız Teknik Üniversitesi, dyildiz@yildiz.edu.tr, 0000-0001-7430-2368

Mathematical Theory of Communication” adlı kitabında istatistikte entropi kavramı ilk kez ortaya çıktı. Bilgi teorisinde bu kavram Shannon entropisi olarak da bilinir. Shannon’nun makalesi yazılmadan önce, bilgi genellikle belirsiz bir sisli akış olarak görülüyordu. Ancak Shannon’ın makalesi ile, bilginin iyi tanımlanmış ve en önemlisi ölçülebilir bir nicelik olduğunu gösterdi (Altaylıgil, 2008).

Boltzmann tarafından ele alınan entropi terimi, 1948 yılına kadar fizik alanında kullanılan bir kavram olmuştur. Bell Laboratuvarlarında bu tarihlerde “Enformasyon Teorisi” üzerinde çalışan Claude Shannon, yüzyılın en ünlü matematikçilerinden biri olan John von Neumann’ın da yardımıyla bu terimi bilgedeki belirsizliğin ölçüsünü göstermek için kullanmıştır (Shannon, 1948).

Entropinin tanımı için ortaya atılan her fikir her alanda düzensizliğin ölçülmesi amacıyla kullanılması için tanımlanmıştır. Bilgi teorisi üzerindeki tanımında ise bir veri kaynağındaki belirsizlik (rastgelelik) miktarını temsil eden bit sayısıdır (Koşan vd., 2019).

İstatistikte entropi kavramını, bilgi teorisinden türetilmiştir. Entropi kavramı başlangıçta termodinamikte düzensizliği veya belirsizliği ölçmek için ortaya çıkmıştır. Bilgi teorisi, ele alınacak şekliyle, matematiksel olasılık ve istatistik teorisinin bir alt dalıdır. Bu bağlamda, soyut matematiksel formülasyonları, herhangi bir olasılıksal veya istatistiksel gözlem sistemi için genel geçerli olan bir nitelik taşımaktadır. Dolayısıyla, bilgi teorisi, olasılık ve istatistik gibi çeşitli disiplinlerde uygulanabilmektedir. Modern iletişim teorisinde önemli bir rol oynar, iletişim sistemini stokastik veya rastgele bir süreç olarak formülize eder. Tuller, iletişim teorisinin genellikle bilgi teorisi olarak adlandırıldığını belirtir. Rothstein, bilgi teorisini ölçülebilir kümelerle ilgilenen, belirli bir doğası belirtilmemiş alternatifler arasındaki seçimlerle ilgilenen soyut matematik şeklinde tanımlamıştır (Kullback, 1960).

Bir sistemin düzensizliğinin ölçülmesinde bilgi teorisindeki entropi kullanılmaktadır. Bilgi teorisinde entropi kavramı  $H$  ile ifade edilmektedir. Entropi, bir bilgi ortamındaki düzensizlik seviyesini belirlemek amacıyla da kullanılmaktadır. Bilginin monoton bir düzen içerisinde akması durumunda düzensizliği azaltabilir. Entropi formülündeki  $n$  genellikle veri miktarını temsil etmektedir (Koşan vd., 2019). Ek olarak  $P_i$  ise ilgili verideki olasılığı temsil etmektedir.

$$H = - \sum_{i=1}^n p_i \log p_i \quad (1)$$

Yukarıda (1) ile ifade edilen formül esas alındığında,  $p$  ve  $1-p$  olasılık değerlerine sahip olan iki durum varsayımı altında, bu iki olasılığın eşit

olduğu durumda entropi değerinin maksimum olduğu noktadır (Koşan vd., 2019).

Bu çalışmada asıl ilgilenilen konu bilgi teorisi, olasılık, istatistik ve matematik teorilerinin bir dalıdır. Bu nedenle bilgi teorisine bu alanlarda oldukça sık rastlanmaktadır. Ek olarak modern iletişim teorisinde önemli bir rol oynamakla birlikte iletişim sistemini stokastik veya rastgele bir yöntem olarak formüle eder.

Bir sistemde girilen bilgi miktarındaki artış, sistemdeki belirsizlik seviyesinin azalması sebebiyle entropi değerinin düşmesine sebep olacaktır. Bunun tersi olarak diğer taraftan ise bilgi miktarındaki azalış, belirsizlik seviyesinin artması sebebiyle entropi değerinin yükselmesine neden olacaktır (Koşan vd., 2019).

Bir sistemin bilinmeyen yapısındaki artan bilgi miktarı, sistemin belirsizliğini azaltacaktır. Bu bağlamda istatistiksel bir bakış açısından ele alınan kutulardan top çekme örneğinde ve zar atma deneyinde bilgi miktarları ile entropi düzeyleri arasındaki ilişki, örneklerde detaylı bir şekilde incelenmiştir (Öncül, 2023).

Entropi terimi bilgi teorisinde tanımlanmasında ise belli başlı özellikler altında şekillendirilmiştir (Koşan vd., 2019). Bu bağlamda özelliklerin ilkinde eğer bir sistemin entropisi sıfırsa, istisna bir durum haricinde tüm durumların olasılıklarının sıfır olduğu belirtilebilir. Bu istisna durum için ise tek bir durum kalmaktadır. Bu durum ise, tüm olasılıkların tamamına sahiptir. Diğer bir özellik ise her bir  $p_i$  değeri belirli  $n$  için eşit olduğu zaman entropi en yüksek değere ulaşır ve  $\log(n)$ 'e eşittir. Bu özellik sezgisel açıdan en belirsiz olan durum olarak ifade edilmektedir. Üçüncü özellikte ise ele alınacak bir diğer nokta ise  $x$  ve  $y$  olacak şekilde iki olay ele alındığında ve bu olayların sırasıyla olasılık değerleri  $m$  ve  $n$  olarak belirtildiğinde;

Ortak olasılığı  $p(i, j)$ 'dir. Bu ortak oluşumun entropisi;

$$H(x, y) = -\sum_{i,j} P(i, j) \log_p(i, j) \quad (2)$$

$$H(x) = -\sum_{i,j} P(i, j) \log_p(i, j) \quad (3)$$

$$H(y) = -\sum_{i,j} P(i, j) \log_p(i, j) \quad (4)$$

Bağımsız olaylar söz konusu olduğunda, bireysel entropilerin toplamının ortak olayın belirsizliğine eşit ya da daha büyüktür şeklinde ifade edilebilir.

$$H(x, y) \leq H(x) + H(y) \quad (5)$$

$$p(i, j) = p(i) p(j) \quad (6)$$

Öne çıkan dördüncü özellik olarak  $\{1,0,0,\dots,0\},\{0,1,0,\dots,0\},\dots\{0,0,0,\dots,1\}$  şeklindeki dağılımların entropi değeri  $0$  dir.

Bir başka özelliğe değinmek gerekirse bir olasılık dağılımındaki tüm  $p_i$  değerlerinin birbirine eşit olmasına yönelik bir değişim entropi değerini arttırır.

Shannon entropisi “A Mathematical Theory of Communication” başlıklı araştırma makalesi sayesinde öne sürmüştür. Shannon’nun tanımladığı iletişim sisteminde bilgi kaynağı, alıcı birime iletilmesi amacıyla ileti ya da ileti dizisi üreten sistemdir. Verici ise, bilgi kaynağı ile üretilmiş olan mesajı, kanal boyunca iletim için uygun sinyal oluşturan bir sistemdir. Mesajı kanalda iletmek amacıyla kullanılan biçime sinyal adı verilmektedir. Vericiden alıcıya iletilen sinyalin ortamı kanal olarak adlandırılmaktadır. Alıcı ise bu sinyali yeniden mesaj formatına dönüştürür ve genellikle vericinin yaptığı işlemin tersini yapmaktadır. Alıcı terminal ile iletişim kurmak için bir ileti ya da ileti dizisi üreten sistem, bilgi kaynağı olarak isimlendirilmektedir. Sinyalin vericiden alıcıya iletilmesinde kullanılan ortam kanal olarak adlandırılmaktadır. Bilginin varış yeri yani bir diğer adı hedef, iletinin ulaştığı son noktadır şeklinde tanımlanabilmektedir (Değirmenci, 2011).

Shannon’nun entropisi iletişim kuramı açısından bir ölçü olarak tanımlanmış olmasına rağmen, bu ölçünün geniş çaplı bir kullanım alanı mevcuttur. Bu alanlardan bazıları matematik, mühendislik, fizik, biyoloji, sosyal bilimlerdir. Entropinin bu önemli başarısı, özellikler olasılıksal sistemde belirsizliği ölçebilmesindeki olarak söylenebilmektedir.

Shannon, bir olay hakkındaki bilgi edinme sürecini, olayın belirsizlik içermesi gerektiği durumlarda gerçekleşeceğini belirtmiştir. Bu durumlarda, yüksek olasılıklı olayların ortaya çıkması genellikle çok az bilgi verir, ancak düşük olasılıklı olayların ortaya çıkması genellikle daha fazla bilgi verir. Bu bağlamda, belirli bir alternatfin gerçekleşme olasılığı, ilgili çözümü gösteren işaret, sembol ya da sayının belirsizliğini göstermektedir. Sonuç olarak, edinilen bilgi, giderilen belirsizliğin miktarını dolaylı olarak ölçer (Bahadır vd., 2021).

Shannon, entropi terimini bu yöntemle, bir olayın farklı seçeneklerinin beklenen değeri olarak matematiksel bir ilişkiyle tanımladı. Entropi bit veya desibel gibi birimlerle ölçülebilen bir nicel büyüklüktür. Bu sebeple bir rasgele sürecin olasılıklı yapısının belirlenmesinin ardından bu sürecin entropisi hesaplanır ve belirli birimler ile ifade edilebilmektedir. Shannon geliştirmiş olduğu entropi formülündeki logaritmayı 2 tabanında kullanmıştır. Logaritma kullanılmasının amacı ise bir bağımsız sistem içindeki entropinin toplanabilir

nicelik formatında getirilebilmesi için kullanılmaktadır. Bu kapsamda, logaritma işleminin bilimsel çalışmalarda da sayısal kodlama sistemlerinde de uygun olduğu kanısına varılmıştır ve bu sebeple iki tabanında logaritma kullanılmıştır. İki tabanında logaritmanın tercih edilmesinde en büyük etken ise sayısal kodlama sistemlerine uygun olmasından kaynaklanmaktadır. Ek olarak entropi değerlerinin hesaplanmalarında elde edilen sonuçlarının ikili sayı sisteminde kısaltılarak ifade edilmesine olanak sağladı. Bu şekildeki kısaltma birimine bit denmektedir (Cover, 1999).

Olası durumları  $p_1, p_2, p_3, \dots, p_n$  kümesi olarak tanımlayan Shannon, bu kümedeki hangi olayın gerçekleşeceği bilinmediği için bu oluşan belirsizliği entropi yardımıyla ölçülebileceğini ifade etmiştir. Entropinin genel formülü (7) eşitliğinde gösterildiği gibidir. Denklemden H entropiyi ifade ederken, k sabiti Boltzmann sabiti olarak tanımlanmaktadır (Kullback, 1960).

$$H = -K \sum p_i \log_2 p_i \quad (7)$$

Shannon, k sabitini 1 kabul ederek ortaya Shannon entropisi ya da bir diğer ifadeyle Shannon belirsizliğini ortaya koymuştur. Shannon entropisi (8) denklemi ile ifade edilmektedir.

$$H = -\sum p_i \log_2 p_i \quad (8)$$

Bu bağlamda bu ifadeler ışığında varılan sonuç, rasgele bir değişkenin değerini tahmin ederken ortaya çıkmış olan belirsizliğin sayısallaştırılan entropi değeridir.

Bilgi teorisinin temellerini ortaya atmış olan Claude Shannon, ortak entropi ile Shannon entropisinin arasındaki ilişkiyi açıklamıştır. Bir önceki bölümlerde yazıldığı gibi bir olayın meydana gelme olasılığına göre bulunan bilgi ölçüsü Shannon entropisi ile elde edilmektedir. Bununla birlikte, eğer değişken sayısı iki ya da daha fazla olması halinde rasgele değişkenlerin kendi aralarındaki belirsizliğin ölçümüne ortak entropi denir. Shannon bu kavramlar üzerinde çalışmaları sayesinde gelişen bu alan, günümüzde istatistik, matematik, bilgisayar bilimleri ve birçok başka alanda uygulamalarda kullanılabilir.

Ortak entropinin genel anlamda tanımı  $p(x, y)$  olasılık dağılımına sahip ve  $(X, Y)$  kesikli rasgele değişkeninin entropisi (9) denklemi ile tanımlanır (Cover, 1999).

$$H(X, Y) = -\sum_{x \in X} \sum_{y \in Y} P(x, y) \log P(x, y) \quad (9)$$

Bu çalışmada temel hedefi, belirli bir veri seti ile ilk olarak faktör analizi gerçekleştirilmesidir ve bu analizin hemen ardından ise aynı veri seti kullanılarak faktör analizinin başındaki gözlemlenen değişkenlerin arasındaki bağımlılığın ölçülmesinde kullanılan korelasyon matrisi yerine entropik tabanlı yöntem kullanılacaktır. Değişkenlerin entropi değerlerinden faydalanılarak birleşik entropi matrisi oluşturulacak ve elde edilen bu matrisi faktör analizindeki varyans-kovaryans matrisi yerine kullanarak bu doğrultuda bu matris yardımı ile boyut indirgeme yönteminin etkinliği incelenecektir. Son aşamada, elde edilen sonuçların karşılaştırılması yapılacak olup bu doğrultuda farklı yaklaşımların avantajları ve dezavantajları incelenecektir.

Bu çalışma kapsamında, kullanılan verinin faktör analizi ve entropi temelli yaklaşımın analizinin sonuçlarının karşılaştırılmasını değerlendirmeyi hedeflenmiştir. İlk aşamada faktör analizi yapılacak ve faktör analizindeki aynı amaç için entropi tabanlı yaklaşım kullanılarak birleşik entropi matrisi elde edilecek ve faktörler bulunacaktır. Bu iki sonuç karşılaştırmalı olarak ölçülecektir. İkinci olarak ise kovaryans matrisi yerine birleşik entropi matrisi kullandığımız entropi tabanlı yaklaşımla elde edilen matris yardımı ile elde edilen sonuçların avantajları ve dezavantajları üzerinde durulacak ve değerlendirilecektir. Bu çalışmasının çıktıları ise entropinin istatistiksel anlamda daha sık kullanılmasına veya yapılacak olan araştırmalarda farklı bir yöntemin de kullanılabilir olması şeklinde yardımcı olması beklenebilir.

## 2. Metodoloji

Çok değişkenli istatistiksel analiz teknikleri, çok sayıda değişkenin bulunduğu durumlarda genellikle tercih edilmektedir. Araştırmacılar bu teknikleri kullanarak, değişkenler arasındaki karmaşık ilişkileri anlamaya çalışmaktadırlar. Analize başlamadan önce ilk yapılması gereken şey değişkenler arasındaki ilişkileri iyi bir şekilde kavrayabilmektir. Sınıflandırılmış verilerin olduğu durumlar için farklı yöntemler geliştirilmiştir. Bu tür veriler üzerinde çalışılırken, doğrusal olmayan yaklaşımların veya sınıflandırılmış değişkenlerin seviyelerinin kukla değişken olarak kullanılması oldukça yaygın kullanılan yöntemlerdendir. Entropiye bağlı hesaplamalarda ise, değişkenler hakkındaki yapılan hesaplamaların kesin sonuçları yerine sınıflandırılmış verilerin olasılıkları dikkate alınarak işlem yapılmaktadır. Bu gibi durumlarda analizin belirsizlik düzeyi veya analizin düzensizliği hesaplanarak, değişkenler arasındaki ilişki çözümlenmeye çalışılmaktadır. Entropiye bağlı hesaplamaları kullanarak değişkenlerin yapısını anlayabilmek ve verileri daha sade bir biçimde yazabilmek için temel bileşenler analizi kullanılmaktadır.

Analiz yaparken çok sayıda değişkenin kullanılması, araştırmacının işini çok daha karmaşık ve güç bir duruma getirebilmektedir. Bu gibi zorlukların önüne geçmek ve analizi daha kolay yönetilebilir hale getirmek, araştırmacının işini çok daha kolay hale getirecektir. Bu sebeple, çok sayıda değişkenin daha az sayıda değişkene indirgenmeye çalışılması oldukça önemlidir. Daha az değişken olmasıyla birlikte, analizin uygulanabilirliği yüksek ölçüde arttırılabilir ve daha net, daha kolay yorumlanabilir sonuçlar elde edilebilmektedir. Değişken sayısının indirgenmesine yardımcı olan yöntemlerden biri temel bileşenler analizidir. Bu yöntemle değişkenler daha az sayıdaki temel bileşenlerine kadar indirgenmekte ve veri seti özet haline getirilmektedir (Büyüköztürk, 2002).

Faktör analizi ile birlikte,  $p$  tane değişkeninin bulunduğu durumlarda birbirleriyle ilişkili olan değişkenlerin birbirleriyle yanyana getirilmesi ve daha az sayıda elde edilecek olan yeni (ortak) ilişkisi olmayan değişkenlerin ortaya çıkması amaçlanmaktadır. Bu sebeple faktör analizi, boyut indirgeme konusunda, araştırmacılar arasında oldukça popüler olan bir seçenektir (Shannon, 1948).

Faktör analizleri yapılar iken, tüm değişkenler bir arada ele alınmaktadır. Sonrasında ortaya çıkabilecek herhangi bir faktör ise, gözlemlenmiş değişkenlerin bir fonksiyonu olarak yani bağımsız bir değişken olarak kabul edilmektedir (Tatlıdil, 1992).

Entropi ve varyans benzer yapıya sahiptir. Benzer özelliklere sahip olan entropi ve varyansı incelediğimizde ilk olarak ancak ve ancak  $p_i$ 'den biri 1 ise, o zaman  $H = 0$ . Bu özellik varyans içinde geçerlidir. Bir diğer ortak özellik ise  $x$  ve  $y$  değişkenleri bağımsız ise  $H(x, y) = H(x) + H(y)$  şeklindedir. Bu durum da varyans içinde geçerlidir.  $x$  ve  $y$  bağımsız ise  $var(xy) = var(x) + var(y)$  şeklinde ifade edilebilir (Wei, 1987).

Belirsizlik ölçüsü için varyans ve entropi arasındaki ilişki ve farklılıklar mevcuttur. Bilgi teorisi çalışmaları ve uygulamaları sırasında, farklı verilere göre varyans ve entropi arasında uygun bir istatistik seçilmelidir. Bu aşamada Tablo 1'den faydalanılabilmektedir (Wei, 1987).

*Tablo 1. Veri tipi ve İstatistik*

Veri Tipi	İstatistik
Değerler (frekans dağılımında dönüştürülmesi güç değerler)	Varyans
Değerler ve olasılıklar	Varyans ya da entropi
Olasılıklar (kukla değişken uygulamak kolay)	Varyans ya da entropi
Olasılıklar (kukla değişken uygulamak zor)	Entropi



Ortak entropi hesaplanmasında kullanılan formül (10) ile gösterilmiştir.

$$H(X, Y) = -\sum_{i,j} p(i, j) \log p(i, j) \quad (10)$$

Buradaki sonuç ise  $X_\alpha$  rassal değişkeninin marjinal entropisine eşittir. Bu durumda ulaşılabacak sonuç ise (11) eşitliğidir. (Çilingir, 2019)

$$H(X, X) = H(X) \quad (11)$$

Birleşik entropi ile her bir olayın ayrı entropileri arasındaki ilişki incelendiğinde, bu ilişki (12) ile ifade edilir.

$$H(X, Y) \leq H(X) + H(Y) \quad (12)$$

Bu eşitsizlik Schwarz-Cauchy eşitsizliğine denktir. Bu eşitsizliğe ulaşabilmemizdeki sebep olayın bağımsız olması sebebiyle  $p(i, j) = p(i)p(j)$  denkleminin geçerli olmasıdır. Sonuç olarak birleşik entropi olayının belirsizliği, X ve Y'nin belirsizliklerinin ayrı ayrı toplamına eşit veya daha küçüktür (Çilingir, 2019).

Bir veri setinde bağımlılık ölçüsünün hesaplanmasında ise (12) eşitliğinden faydalanarak;

$$\beta_{x_\alpha, x_\beta} \geq \frac{H(x_\alpha) + H(x_\beta)}{H(x_\alpha, x_\beta)} \quad (13)$$

denklem (13) ile bağımlılık ölçüsünün hesaplanması mümkündür (Cover, 1999).

Değişkenlerin entropi değerleri ile birleşik entropi değerleri elde edildikten sonra,  $p \times p$  boyutlarında simetrik bir birleşik entropi matrisi oluşturmak mümkündür (Erbaş, 2010).

*Tablo 2. Birleşik entropi matrisinin görselleştirilmiş hali*

	$X_1$	...	$X_p$
$X_1$	$H(X_1)$	...	$H(X_1, X_p)$
...	...	$\ddots$	...
$X_p$	$H(X_1, X_p)$	...	$H(X_p)$

Bu çalışma doğrultusunda ülkelerin sosyo-ekonomik ve sağlık alanındaki faktörleşmesi üzerinde durulması amaçlanmıştır. Literatürde faktör analizi ve entropi alanında çeşitli araştırmalar mevcuttur. Ancak istatistiksel anlamda yapılan entropi tabanlı analizi yöntemlerinin uygulamaları açısından literatür maalesef oldukça zengin değildir. Bu sebeple bu çalışmasıyla birlikte istatistiksel anlamda entropi tabanlı bir uygulama ile hem dikkat çekilmesi hem de literatüre katkıda bulunması amaçlanmaktadır.

Bu çalışmada hedef, veri setini analiz etmek ve veri setindeki temel faktörleri belirlemek için faktör analizi ve entropi tabanlı bir yaklaşım olan birleşik entropi yönteminin kullanılmasıdır. Bu uygulama kapsamında, her iki yöntem de kullanılarak elde edilen faktörlerin belirlenmesi hedeflenmiştir. Bu bağlamda, her iki analiz yöntemiyle elde edilen faktörler karşılaştırılması amaçlanmıştır. Bu karşılaştırma, farklı analiz yöntemlerinin veri setindeki yapıyı anlama ve yorumlama yeteneklerini değerlendirmeyi amaçlamıştır.

Özetle, bu çalışma ile hem faktör analizi hem de entropi tabanlı bir yaklaşım olan birleşik entropi kullanarak veri setindeki yapıyı incelemiş ve bu iki farklı analiz yönteminin sonuçlarını karşılaştırarak araştırmanın genel amacını oluşturmuştur.

Bu çalışma Kaggle üzerinden “Unsupervised Learning on Country Data” başlıklı veri seti alınarak yapılmıştır. Veri setinin ülkeler bazında genel gelişimini belirleyen sosyo-ekonomik ve sağlık faktörlerini kullanarak kategorize edilmesi amaçlanmıştır. Bu bağlamda 167 ülke üzerinden bir değerlendirme söz konusudur. Veri setinde bulunan değişkenler sırasıyla çocuk ölüm oranı, ihracat, sağlık, ithalat, gelir, enflasyon, ortalama yaşam süresi, doğurganlık oranı ve kişi başına gayri safi yurtiçi hasıla (GDPP) şeklinde 9 değişkendir.

Tablo 3. Veri setinde kullanılan değişkenler

Değişkenler	Değişken Adı	Açıklama
Çocuk Ölümü	$X_1$	Her 1000 doğumda, 5 yaş altı ölen çocuk ölümleri.
İhracat	$X_2$	Kişi başına düşen mal ve hizmet ihracatı. Kişi başına düşen GSYİH'nin yüzdesi olarak verilmiştir.
Sağlık	$X_3$	Kişi başına düşen toplam sağlık harcamaları. GSYİH'nin yüzdesi olarak verilmiştir.
İthalat	$X_4$	Kişi başına düşen mal ve hizmet ithalatı. Kişi başına düşen GSYİH'nin yüzdesi olarak verilmiştir.
Gelir	$X_5$	Kişi başına net gelir.
Enflasyon	$X_6$	Toplam GSYİH'nin yıllık büyüme oranı
Beklenen Yaşam Süresi	$X_7$	Mevcut ölüm oranları aynı kalsa, yeni doğan bir çocuğun ortalama yaşayacağı yıl sayısı
Toplam Doğurganlık	$X_8$	Mevcut yaş-doğurganlık oranları aynı kalsa, herkadına doğacak çocuk sayısı
GDPP	$X_9$	Kişi başına düşen GSYİH. Toplam GSYİH'nin toplam nüfusa bölünmesiyle hesaplanır.

## 2.1. Faktör Analizi Uygulaması

Bu başlık altında ilk olarak veri seti SPSS ile faktör analizine sokulup kaç faktöre indirgendiği hesaplanmıştır ve isimlendirilmiştir. İlk aşamada SPSS ile faktör analizi incelemesi yapılmıştır. Bu bağlamda ilk olarak KMO değeri hesaplanmış olup, bu değer 0,68 ile orta seviyede çıkmıştır. Tablo 4'te görüldüğü gibi KMO değerinin faktör analizi için uygun olduğunu göstermektedir.

Tablo 4. Kaiser-Meyer-Olkin ve Bartlett Testi

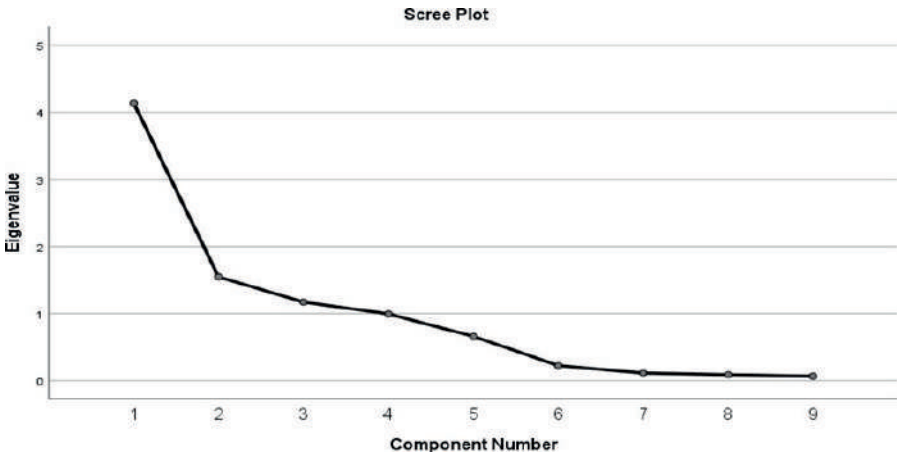
Kaiser-Meyer-Olkin ve Bartlett Testi		
Kaiser-Meyer-Olkin Örnekleme Yeterliliğinin Ölçüsü.	0,678	
Bartlett'in Küresellik Testi	Tahmini Ki-Kare	1169,737
	Serbestlik Derecesi	36
	p-değeri	0,000

Sonrasında ise değişkenler incelenmiş olup bu kapsamda veri setindeki 9 değişkenin de analiz sonucunda kalması kararına varılmıştır. Buna Tablo 5’de görüldüğü gibi 0,5 altında değer olmaması sebebiyle karar verilmiştir.

*Tablo 5. Toplam Varyans*

Değişkenler	Initial	Extraction
cocuk_olumu	1,000	0,786
ihracat	1,000	0,939
saglik	1,000	0,602
ithalat	1,000	0,911
gelir	1,000	0,764
enflasyon	1,000	0,638
yasam_suresi	1,000	0,842
toplam_dogurganlik	1,000	0,712
gdpp	1,000	0,659

Faktör analizine geçiş yapılmasının ardından ise Şekil 6’da gösterilmiş olan Scree Plot ile 3 faktöre indirgenmesine gerektiği gözlemlenmektedir. Scree Plot’a bakıldığında faktör sayısının 3 olarak belirlenmesinin sebebi özdeğerin 1’den büyük olan değerler kadar faktöre indirgeneyeceğidir. Bu kapsamda özdeğerin 1’den büyük olduğunu gösteren 3 nokta mevcuttur. Bu bilgilere ek olarak indirgenen faktörlerin açıklanan toplam varyansı ise Tablo 7 ile gösterilmektedir.



*Şekil 6. Scree plot*

Tablo 7. Açıklanan toplam varyans

Başlangıç Özdeğerleri				Özütleme Karekök Yüklemeleri Toplamı		
Bileşen	Toplam	Varyans Yüzdəsi	Kümülatif %	Toplam	Varyans Yüzdəsi	Kümülatif %
1	4,136	45,952	45,952	3,666	40,734	40,734
2	1,546	17,182	63,133	1,794	19,938	60,672
3	1,170	13,004	76,138	1,392	15,466	76,138
4	,995	11,053	87,191			
5	,661	7,340	94,531			
6	,224	2,484	97,015			
7	,113	1,260	98,276			
8	,088	,981	99,257			
9	,067	,743	10,000			

Şekil 6'da ve Tablo 7'de görüldüğü gibi faktör analizi ile 9 değişkenden 3 faktöre indirgenmiştir. Bu bağlamda ilk faktörün açıklanan toplam varyansı %40,7 iken ikincinin %19,9 olduğu gözükmektedir. Bu bağlamda ilk iki faktörün açıklanabilir toplam varyansının %60,6 olduğu gözlemlenmektedir. Üçüncü faktörün ise %15,5 iken üç faktörün toplam açıklanabilir varyans oranının %76,1 olduğu gözlemlenmektedir. Faktör analizinde Varimax dönüşümü kullanılmıştır.

Tablo 8. SPSS ile faktör analizinden elde edilen faktörler

	Sosyal	Ekonomi	Sağlık
yasam_suresi	0,902	-0,005	-0,169
cocuk_olumu	0,852	-0,037	0,244
gelir	0,834	0,241	0,100
toplam_dogurganlik	0,793	-0,078	0,276
gdpp	0,785	0,180	-0,097
ithalat	-0,022	0,929	-0,219
ihracat	0,375	0,880	0,154
enflasyon	-0,138	-0,216	0,756
saglik	0,158	-0,111	0,752

Mevcut 9 adet değişkenin 3 faktör ile temsil edilebileceğine karar verilir. Bu 3 faktör Tablo 8’de gösterilmiştir. Bu faktörlerin isimlendirilmesi ise değişkenlere bağlı olarak yapılmıştır. Bu bağlamda faktör 1’i temsilen sosyal faktörü olarak isimlendirilmiştir. Sosyal faktöründeki değişkenler ise sırasıyla yaşam süresi, çocuk ölümü, gelir, toplam doğurganlık ve gdp (kişi başına düşen gayri safi yurtiçi hasılat) süresi şeklindedir. Faktör 2’yi temsilen ekonomi faktörü ele alınmıştır. Bahsi geçen faktördeki değişkenler ise ithalat ve ihracat değişkenleridir. Son olarak faktör 3’ü temsilen ise sağlık faktöründe sadece sağlık değişkeni bulunmaktadır.

## 2.2. Ortak Entropi Uygulaması

Bu başlık altında ilk olarak veri seti Python ve Excel yardımları ile kaç faktöre indirgendiği hesaplanmıştır ve isimlendirilmiştir. Bu bağlamda ilk olarak veri setine ait olan 9 değişkenin tek tek entropisi hesaplanmaktadır. Bu değerler sırasıyla Tablo 9’da gösterilmiştir.

*Tablo 9. Değişkenlere ait entropi değerleri*

Değişken	Entropi Değeri
$X_1$	7,0183
$X_2$	7,1351
$X_3$	7,1351
$X_4$	7,1921
$X_5$	7,2520
$X_6$	7,2474
$X_7$	6,8775
$X_8$	6,9944
$X_9$	7,2520

Tablo 10’da ise değişken çiftleri için hesaplanan birleşik entropi değerleri ve değişken hakkında kararlar gösterilmektedir.

Tablo 10. SPSS ile faktör analizinden elde edilen faktörler

	$H(X_\alpha, X_\beta)$	$\frac{H(X_\alpha) + H(X_\beta)}{H(X_\alpha, X_\beta)}$	Değişkenler Hakkında Karar
$X_1, X_2$	7,952450	1,7798	Bağımlı.
$X_1, X_3$	7,998094	1,7696	Bağımlı.
$X_1, X_4$	7,954710	1,7864	Bağımlı.
$X_1, X_5$	8,135143	1,7542	Bağımlı.
$X_1, X_6$	8,029995	1,7766	Bağımlı.
$X_1, X_7$	7,890984	1,7610	Bağımlı
$X_1, X_8$	7,968946	1,7584	Bağımlı
$X_1, X_9$	8,135143	1,7542	Bağımlı
$X_2, X_3$	8,117179	1,7580	Bağımlı.
$X_2, X_4$	7,964127	1,7990	Bağımlı.
$X_2, X_5$	8,193555	1,7559	Bağımlı.
$X_2, X_6$	8,117179	1,7719	Bağımlı.
$X_2, X_7$	7,949890	1,7626	Bağımlı
$X_2, X_8$	8,056506	1,7538	Bağımlı
$X_2, X_9$	8,193555	1,7559	Bağımlı
$X_3, X_4$	8,157627	1,7563	Bağımlı.
$X_3, X_5$	8,193555	1,7559	Bağımlı.
$X_3, X_6$	8,092434	1,7773	Bağımlı.
$X_3, X_7$	8,006342	1,7502	Bağımlı

$X_3, X_8$	7,978662	1,7709	Bağımlı
$X_3, X_9$	8,193555	1,7559	Bağımlı
$X_4, X_5$	8,222028	1,7568	Bağımlı.
$X_4, X_6$	8,169603	1,7675	Bağımlı.
$X_4, X_7$	7,889517	1,7833	Bağımlı
$X_4, X_8$	8,093227	1,7529	Bağımlı
$X_4, X_9$	8,222028	1,7568	Bağımlı
$X_5, X_6$	8,249708	1,7576	Bağımlı.
$X_5, X_7$	8,064754	1,7520	Bağımlı
$X_5, X_8$	8,123167	1,7538	Bağımlı
$X_5, X_9$	8,151639	1,7793	Bağımlı
$X_6, X_7$	8,062494	1,7519	Bağımlı
$X_6, X_8$	7,995534	1,7812	Bağımlı
$X_6, X_9$	8,249708	1,7576	Bağımlı
$X_7, X_8$	7,935953	1,7480	Bağımlı
$X_7, X_9$	8,064754	1,7520	Bağımlı
$X_8, X_9$	8,123167	1,7538	Bağımlı

Tablo 9'da görüldüğü gibi ortak entropi değerleri hesaplanmıştır. Bu bilgiler ışığında bağımlı olup olmadığı da kontrol edilmiştir. Birleşik entropi değerlerinin 1 ile 2 arasında olması beklenmektedir (Çilingir, 2019). Bir sonraki adım ise birleşik entropi matrisinin elde edilmesidir. Birleşik entropi matrisi Tablo 11'de gösterilmektedir.



**Tablo 11. Birleşik Entropi Matrisi**

	1	2	3	4	5	6	7	8	9
1	7,01832	7,95245	7,99809	7,95471	8,13514	8,03000	7,89098	7,96895	8,13514
2	7,95245	7,13514	8,11718	7,96413	8,19356	8,11718	7,94989	8,05651	8,19356
3	7,99809	8,11718	7,13514	8,15763	8,19356	8,09243	8,00634	7,97866	8,19356
4	7,95471	7,96413	8,15763	7,19209	8,22203	8,16960	7,88952	8,09323	8,22203
5	8,13514	8,19356	8,19356	8,22203	7,25197	8,24971	8,06475	8,12317	8,15164
6	8,03000	8,11718	8,09243	8,16960	8,24971	7,24745	8,06249	7,99553	8,24971
7	7,89098	7,94989	8,00634	7,88952	8,06475	8,06249	6,87754	7,93595	8,06475
8	7,96895	8,05651	7,97866	8,09323	8,12317	7,99553	7,93595	6,99437	8,12317
9	8,13514	8,19356	8,19356	8,22203	8,15164	8,24971	8,06475	8,12317	7,25197

Birleşik entropi matrisi yardımıyla entropi-korelasyon matrisine geçiş yapılacaktır. Bu bağlamda entropi-korelasyon matrisi Tablo 12'de gösterilmektedir.

**Tablo 12. Entropi- Korelasyon Matrisi**

	1	2	3	4	5	6	7	8	9
1	1,00000	0,13110	0,03670	0,16343	-0,20953	0,02072	0,12739	0,02144	-0,20953
2	0,13110	1,00000	-0,14976	0,22976	-0,27336	-0,10754	0,09559	-0,09307	-0,27336
3	0,03670	-0,14976	1,00000	-0,20145	-0,24017	-0,01110	-0,02150	0,10495	-0,24017
4	0,16343	0,22976	-0,20145	1,00000	-0,29959	-0,19006	0,24392	-0,13526	-0,29959
5	-0,20953	-0,27336	-0,24017	-0,29959	1,00000	-0,35093	-0,07466	-0,15029	-0,05301
6	0,02072	-0,10754	-0,01110	-0,19006	-0,35093	1,00000	-0,08959	0,11875	-0,35093
7	0,12739	0,09559	-0,02150	0,24392	-0,07466	-0,08959	1,00000	0,03768	-0,07466
8	0,02144	-0,09307	0,10495	-0,13526	-0,15029	0,11875	0,03768	1,00000	-0,15029
9	-0,20953	-0,27336	-0,24017	-0,29959	-0,05301	-0,35093	-0,07466	-0,15029	1,00000

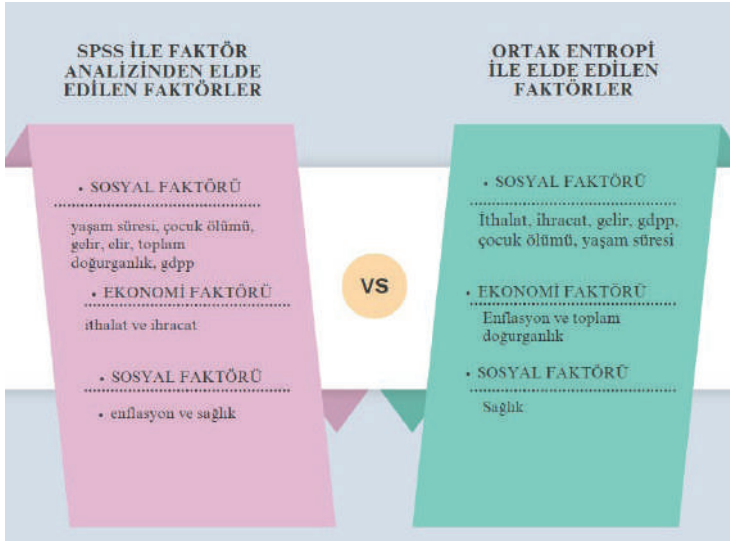
Entropi-korelasyon matrisinin elde edilmesinin ardından ise 9 değişkenden kaç tane faktöre indirgeneceği bulunacaktır. Faktör analizinde de olduğu gibi buradaki işlemlerde de Varimax dönüşümü kullanılmıştır. Mevcut 9 adet değişkenin 3 faktör ile temsil edilebileceğine karar verilir. Bu 3 faktör Tablo 13'de gösterilmiştir. Bu faktörlerin isimlendirilmesi ise değişkenlere bağlı olarak yapılmıştır.

Tablo 13. Ortak entropi ile elde edilen faktörler

	Sosyal	Ekonomi	Sağlık
ithalat	0,765	-0,494	0,070
ihracat	0,613	-0,287	-0,018
gelir	0,603	-0,255	-0,005
gdpp	0,603	-0,255	-0,005
cocuk_olumu	0,539	0,005	0,072
yasam_suresi	0,340	-0,270	0,115
enflasyon	0,332	0,767	-0,566
toplam_dogurganlik	0,034	0,443	-0,018
saglik	0,090	0,693	0,720

Faktör 1'i temsilen toplum faktörü olarak isimlendirilmiştir. Sosyal faktöründeki değişkenler ise sırasıyla ithalat, ihracat, gelir, ithalat, gdpp (kişi başına düşen gayri safi yurtiçi hasılat), çocuk ölümü ve beklenen yaşam süresi şeklindedir. Faktör 2'yi temsilen ekonomi faktörü ele alınmıştır. Bahsi geçen faktördeki değişkenler ise enflasyon ve toplam doğurganlık değişkenleridir. Son olarak faktör 3'ü temsilen ise sağlık faktöründe sadece sağlık değişkeni bulunmaktadır.

Sonuç olarak yapılan bu iki uygulama sonrasında elde edilmiş olan faktörlerin karşılaştırılması Şekil 14'te gösterilmiştir.



Şekil 14. Faktör Analizi ve Ortak Entropi yöntemlerinden elde edilen faktörlerin karşılaştırılması

### 3. Sonuç ve Tartışma

Çalışma kapsamında öncelikli hedef, istatistiksel anlamda entropi bazı yaklaşımların da uygulamalarda bir alternatif olarak kullanılabilmesinin gösterilmesi hedeflenmiştir. Bu bağlamda ilk olarak entropi kavramı hakkında bilgiler verilmiş ve ortak entropi konusu ele alınmıştır. Bu bağlamda değişkenler arası ilişki entropi ölçüsü ile incelenmiştir ve bu yöntem ile birleşik entropi matrisi kullanılarak boyut indirgemenin sağlanabileceği incelenmiştir. Bu bilgilere ek olarak bibliyometrik analiz üzerinde durulmuş ve bu amaç doğrultusunda yapılmış olan çalışmalara ait yazarlar, anahtar kelimeler ve literatür bilgileri kazanılmıştır.

Uygulama bölümünde ülkeler bazında genel gelişimini belirleyen sosyo-ekonomik ve sağlık faktörlerini kullanarak kategorize edilmiş değişkenlerin bulunduğu veri seti ile farklı analiz yöntemleri kullanılarak karşılaştırma yapılmıştır. İlk olarak veri seti faktör analizinin sonuçlarıyla entropi tabanlı yaklaşımın sonuçlarının karşılaştırması amacıyla faktör analizinin sonuçları elde edilmiştir. Ardında ise değişkenlere ait entropi ölçüleri ile elde edilen birleşik entropi matrisi yardımıyla boyut indirgeme sağlanıp sağlanmadığı teyit edilmiştir. Bu bilgiler doğrultusunda elde edilen bilgiler ile faktör analizinden elde edilen bilgilerin karşılaştırılması yapılmıştır.

Kullanılan veri setiyle entropi yöntemi sayesinde elde edilen birleşik entropi matrisi ile faktör indirgeme yöntemi gerçekleştirilmesi hedeflenmiştir. Bu sayede birleşik entropi yöntemiyle faktör indirgemenin sağlanıp sağlanmadığı araştırılmıştır.

Sonuç olarak bu çalışma, entropinin teorik bir temelini ve varyans-kovaryans ölçümlerini kullanmanın yararlı olduğu alanlara ek olarak katkıda bulunmuştur. Bu çalışmanın, entropinin önemini ve entropiye dayalı araştırma yöntemlerini vurguladığı düşünüldüğünde, istatistik alanında entropi konusuna daha fazla önem verilmesi ve bu alanda daha fazla çalışma yapılması önerilmektedir. Bununla birlikte, çalışmanın ortaya çıkan zorluklar, akademik alanda yeni çalışmaların gerekli olduğunu göstermektedir. Çözülmesi gereken bazı sorunlara rağmen, alternatif olarak bilinen yöntemlerin yerine önerilmektedir.

## Kaynakça

- Altaylıgil, Y. B. (2008). Entropi ölçüsü ve bazı ekonometri uygulamaları.
- Bahadır, O., & Türkmencehalıkoğlu, H. (2021). Bilgi Kuramında Shannon Entropisi ve Uygulamaları. *Avrupa Bilim ve Teknoloji Dergisi*, 32, 497-497.
- Büyükoztürk, Ş. (2002). Faktör Analizi: Temel Kavramlar ve Ölçek Geliştirme de Kullanımı. *Kuram ve uygulamada eğitim yönetimi*, 32(32), 470-483.
- Cover, T. M. (1999). *Elements Of Information Theory*. John Wiley & Sons.
- Çilingir, N. B. (2019). Faktör Analizi: Teori ve Uygulama [Master's thesis]. Sosyal Bilimler Enstitüsü.
- Değirmenci, İ. (2011). Entropi Ölçümleri ve Maksimum Entropi İlkesi (Doctoral dissertation, Yüksek Lisans Tezi, Hacettepe Üniversitesi). [Yüksek Lisans Tezi]. Hacettepe Üniversitesi.
- Erbaş, Ü. (2010). Entropi ilkelerinin boyut indirgeme uygulamaları. <https://search.proquest.com/openview/c98c77beacb361f03877297736fc38be/1?pq-origsite=gscholar&cbl=2026366&cdiss=y>
- Koşan, M. A., Coşkun, A., & Karacan, H. (2019). Yapay Zeka Yöntemlerinde Entropi. *Journal of Information Systems and Management Research*, 1(1), 15-22.
- Kullback, S. (1960). *Information Theory and Statistics*.
- Oruç, N. (2024). Faktör analizi çözümlemesinde birleşik entropi yaklaşımı (Yüksek lisans tezi), Yıldız Teknik Üniversitesi
- Öncül, E. (2023). Çok Değişkenli Yöntemlerde Entropi Kullanımı: Mutluluk İndeksi Üzerine Bir Uygulama. Yıldız Teknik Üniversitesi.
- Shannon, C. E. (1948). A Mathematical Theory Of Communication. *The Bell system technical journal*, 27(3), 379-423.
- Tatlıdil, H. (1992). *Uygulamalı Çok Değişkenli İstatistiksel Analiz*. Engin Yayınları.
- Wei, Y. (1987). *Variance, Entropy, and Uncertainty Measure*. Dept. of Statistics, People's Univ. of China.



## Banka Sektöründe Otomatik Makine Öğrenme Yöntemi (DataBlender) ile Vadeli Mevduat Talep Tahmini

Gizem Aydın<sup>1</sup>

Mehmet Yalçın<sup>2</sup>

### Özet

Günümüzde şirketlerin, müşteri havuzlarından doğru kişilere ulaşarak ürünleri ve hizmetleri pazarlaması oldukça önemli hale gelmiştir. Bu durum, bankacılık ve finans kuruluşları için de geçerlidir, çünkü büyük müşteri kitlesi içinden doğru kişilere ulaşmak maliyetli ve zaman alıcı bir süreç olabilmektedir. Bu doğrultuda kurumlar, ürün ve hizmetleri satın alan müşterileri içerisinden demografik ve finansal verilerini kullanarak makine öğrenmesi tabanlı modellerini arttırmışlardır. Bu modellerin hem doğru kişileri tespit ederek başarılı sonuçlar çıkarması hem de zaman ve para maliyetini azaltması bu teknolojiye olan ilgiyi arttırmıştır. Makine öğrenmesi tabanlı modeller bu doğrultuda sıklıkla kullanım imkânı bulmaktadır.

Bu çalışmada, Portekiz'deki finans kuruluşunu tercih eden müşterilerin vadeli mevduatı alıp almayacağı üzerine tahmin modelleri kurulmuştur. Böylece kuruluş, vadeli mevduata abone olma şansı daha yüksek olan mevcut müşterileri belirleyecek ve pazarlama çabaları ile bu müşterilere odaklanacaktır. Çalışmada; Portekiz'deki finans kuruluşunun, müşterilerini arayarak vadeli mevduat hesabı isteyip istemediği bilgisinin tutulduğu veri seti kullanılmıştır. Veri ön işleme adımları ile modelleme işlemleri otomatik bir şekilde yapılmış ve bunun için "DataBlender" otomatik makine öğrenmesi uygulaması kullanılmıştır. Çalışma sonucunda modeller içerisinden "LightGBM" modeli en başarılı model çıkmıştır. Sonuçlar neticesinde ağaç temelli modellerin başarı ölçütlerinin birbirine yakın olduğu gözlenmiştir. Vadeli mevduat hesabı isteyip istemediği hedef değişkeni üzerindeki etkili değişkenler, model sonucunda bulunarak müşteriye ait özellikler sıralanmıştır.

1 Veri Bilimci, ERETEAM, gaydin@ereteam.com, 0000-0001-6353-0648

2 Yıldız Teknik Üniversitesi, myalcinmh@gmail.com, 0000-0002-8162-6085

## 1. Giriş

Günümüzde, hızla değişen ve gelişen teknolojik ortam işletmelerin pazarlama stratejilerini ve müşteri odaklı yaklaşımlarını temelden değiştirmektedir. Özellikle bankacılık ve finans sektörü, dijitalleşme ve teknolojik ilerlemelerin etkisiyle müşterilere daha özelleştirilmiş ve hedef odaklı hizmetler sunmaya çalışmaktadır. Bu sektörlerdeki kurumlar için, müşterilerle daha derin ve anlamlı bağlantılar kurmak, doğru zamanda doğru mesajları iletmek ve uygun ürün ve hizmetleri sunmak giderek önem kazanmaktadır. Müşteri havuzundan doğru kişilere ulaşmak ve onların gereksinimlerini anlamak, finansal kurumlar için kritik bir görev haline gelmiştir. Ancak, geniş müşteri tabanları arasından doğru hedef kitlenin belirlenmesi ve bu kişilere uygun ürün veya hizmetleri sunmak, zaman alıcı ve maliyetli bir süreç olabilir. Bu nedenle, finans kuruluşları ve benzeri sektörler, veri odaklı yaklaşımlarla müşteri davranışlarını analiz etmeye ve makine öğrenmesi gibi teknolojileri kullanarak daha etkin bir hedefleme stratejisi geliştirmeye yönelik çabalarını artırmaktadır (Moro, Laureano, & Cortez, 2011; Moro vd., 2014).

Makine öğrenmesi, finans sektöründe önemli bir araç haline gelmiştir. Müşteri profilleri, finansal veriler ve dijital izler üzerinde yapılan analizler sayesinde, vadeli mevduat hesaplarından kredi riskine kadar geniş bir yelpazede modeller geliştirilebilmektedir. Özellikle, vadeli mevduat ürünlerine olan ilgiyi öngörmek ve bu alanda müşteri hedefleme stratejilerini iyileştirmek için makine öğrenmesi tabanlı modellerin kullanımı artmaktadır (A.Elsalamony, 2014).

Literatürdeki önemli çalışmalarla birlikte değerlendirildiğinde, finans sektöründe makine öğrenmesi tabanlı modellerin kullanımının çeşitliliği ve başarı potansiyeli hakkında kapsamlı bir anlayış sunmayı amaçlamaktadır. Ghatasheh ve arkadaşlarının yaptığı çalışma, dengesiz veri setlerinin etkisini hafifletmek ve müşterilerin vadeli mevduat hesabı açıp açmayacağını öngörmek amacıyla çok katmanlı bir algılayıcı (MLP) sınıflandırıcı kullanarak derin öğrenme uygulamıştır (Ghatasheh vd., 2020). Okur ve Çetin'in çalışması da Microsoft Azure makine öğrenmesi platformunda regresyon ve ikili sınıflandırma türünde makine öğrenmesi algoritmaları kullanarak kredi risk değerlendirmesi yapmıştır. Bu çalışma, büyük veri setlerinin etkin bir şekilde analiz edilebileceğini ve kredi riski tahmininde makine öğrenmesinin etkili bir araç olduğunu vurgulamıştır (Okur & Cetin, 2019). Coşkun ve Turanlı'nın yaptığı çalışma da ise kullanıcıların kredi riski değerlendirmesi için CatBoost, XGBoost ve LightGBM gibi topluluk öğrenme yöntemlerini karşılaştırarak, kredi skorlamasında makine öğrenmesi tabanlı modellerin etkinliğini ortaya koymaktadır (Coşkun & Turanlı, 2023).

Bu çalışmada, Portekiz'deki bir finans kuruluşunun müşterilerinin vadeli mevduat ürünlerini tercih etme olasılığını öngörmek amacıyla makine öğrenmesi tabanlı tahmin modelleri oluşturulmuştur. Çalışma, kurumun mevcut müşteri tabanında vadeli mevduata abone olma olasılığı yüksek olan müşterileri tespit ederek pazarlama çabalarını optimize etmeyi hedeflemektedir. Veri seti üzerinde yapılan veri ön işleme ve modelleme işlemleri, DataBlender otomatik makine öğrenimi aracıyla gerçekleştirilmiştir.

Çalışmanın sonucunda, en başarılı modelin LightGBM modeli olduğu belirlenmiş ve ağaç tabanlı modellerin benzer başarı kriterlerine sahip olduğu gözlemlenmiştir. Bu modeller, vadeli mevduat hesabı isteme olasılığı üzerinde etkili olan değişkenleri belirlemiş ve müşterilerin özelliklerini sıralamıştır. Bu sonuçlar, finans kuruluşlarının hedefleme stratejilerini iyileştirmek ve pazarlama çabalarını daha etkili bir şekilde yönlendirmek için makine öğrenmesi modellerini kullanmalarını teşvik edebilir. Ayrıca bu çalışmanın DataBlender ile gerçekleştirilen ilk çalışma olması sebebiyle literatüre bu anlamda da katkı sunacaktır.

## 2. Metodoloji

Firmaların hizmetlerini pazarlamak için doğru kişilere ulaşması önemli bir durumdur. Müşteri tabanından doğru kişilere ulaşmak maliyetli ve zaman alıcı bir süreç olabilmektedir. Bu sebeple makine öğrenimi algoritmalarına olan ilgi ve kaynak artmıştır. Çalışma kapsamında da makine öğrenmesi algoritmaları kullanılarak kullanıcı odaklı olarak kod yazmadan otomatik olarak veri ön işleme ve modelleme yapılmasını sağlayan DataBlender ile problemin çözümü gerçekleştirilmiştir. Metodoloji kapsamında makine öğrenimi yöntemlerinden ve DataBlender uygulamasından faydalanılmıştır.

### 2.1. Makine Öğrenimi

Makine öğrenimi, bilgisayar sistemlerinin insanlar gibi öğrenme yeteneğini kazanmaları için geliştirilen algoritmalar, modeller ve tekniklerin incelendiği bir bilim dalıdır. Bu alanda, "makine" olarak adlandırılan yazılım ve istatistiksel algoritmalarından oluşan sistemler, verileri analiz ederek öğrenme sürecini gerçekleştirirler (Alpaydın, 2010). Sade bir açıklama ile, geçmiş verilerden yola çıkarak hem gelecekteki olayları tahmin etmek hem de mevcut veri setindeki örüntüleri ve farklılıkları keşfetmek için kullanılırlar. İstatistiksel ve matematiksel modellere dayalı bu süreçler, çeşitli programlama dilleri kullanılarak yürütülür. Makine öğrenimi, esas olarak üç ana kategori altında incelenir: Denetimli öğrenme (Supervised Learning), Denetimsiz öğrenme (Unsupervised Learning) ve Pekiştirmeli öğrenme (Reinforcement Learning). Çalışmada denetimli öğrenme modelleri kullanılmıştır. Denetimli



öğrenme; hem girdi (input) hem de çıktı (output) değerlerinin yer aldığı veri setlerinde girdi ve çıktılar arasındaki ilişkiyi açıklamak için fonksiyonların elde edilmesidir. Sınıflandırma ve Regresyon modelleri denetimli öğrenme modelleri altında yer almaktadır.

## 2.2. Veri Seti

Çalışma kapsamında kullanılan veriler, UC Irvine Machine Learning Repository'den alınmıştır. Portekizli bir bankacılık kurumunun doğrudan pazarlama kampanyalarına (telefon görüşmeleri) ilişkindir. Veri içerisinde, 41188 gözlem ve 21 değişken içererek kaggle.com üzerinde yetkili kuruluşun izni ile yayınlanmıştır. Veri seti; 11 kategorik ve 10 numerik değişkenden oluşmaktadır. Vadeli Mevduatı İsteme Durumu hedef değişken olarak değişkeni altında bu bilgi mevcuttur.

*Tablo 1. Veri İçerisindeki Değişkenler ve Açıklamaları*

Değişken	Açıklama
Age	Kullanıcı yaşı
Job	İşin türü
Marital	Medeni durumu
Education	Eğitimi
Default	Kredinin temerrüde düşmesi var mı?
Housing	Konut kredisi var mı?
Loan	Bireysel krediniz var mı?
Contact	İletişim türü
Month	Yılın son iletişim ayı
Day_of_week	Haftanın son iletişim günü
Duration	Son iletişim süresi
Campaign	Bu kampanya sırasında ve bu müşteri için gerçekleştirilen iletişim sayısı
Pdays	Önceki bir kampanyadan müşteriyle son iletişime geçildikten sonra geçen gün sayısı
Previous	Bu kampanyadan önce ve bu müşteri için gerçekleştirilen iletişim sayısı
Poutcome	Önceki pazarlama kampanyasının sonucu
Emp.var.rate	İstihdam değişim oranı
Cons.price.idx	Tüketici fiyat endeksi- aylık gösterge
Cons.conf.idx	Tüketici güven endeksi- aylık gösterge
Euribor3m	Euribor 3 aylık oran- günlük gösterge
Nr.employed	Çalışan sayısı- üç aylık gösterge
y	Müşteri vadeli mevduata abone oldu mu?

Tablo 2’de numerik değişkenlere ait istatistiksel tanımlayıcı bilgiler yer almaktadır. Değişkenlerin değer aralıkları oldukça birbirinden farklılık göstermektedir.

*Tablo 2. Numerik Değişkenlerin İstatistiksel Tanımlayıcı Bilgileri*

	gözlem sayısı	ortalama	standart sapma	minimum	25%	50%	75%	maximum
age	41188.0	40.024.060	10.421.250	17.000	32.000	38.000	47.000	98.000
duration	41188.0	258.285.010	259.279.249	0	102.000	180.000	319.000	4.918.000
campaign	41188.0	2.567.593	2.770.014	1.000	1.000	2.000	3.000	56.000
pdays	41188.0	962.475.454	186.910.907	0	999.000	999.000	999.000	999.000
previous	41188.0	172.963	494.901	0	0	0	0	7.000
emp.var.rate	41188.0	81.886	1.570.960	-3.400	-1.800	1.100	1.400	1.400
cons.price.idx	41188.0	93.575.664	578.840	92.201	93.075	93.749	93.994	94.767
cons.conf.idx	41188.0	-40.502.600	4.628.198	-50.800	-42.700	-41.800	-36.400	-26.900
euribor3m	41188.0	3.621.291	1.734.447	634	1.344	4.857	4.961	5.045
nr.employed	41188.0	5.167.035.911	72.251.528	4.963.600	5.099.100	5.191.000	5.228.100	5.228.100

Tablo 3’de kategorik değişkenlere ait istatistiksel tanımlayıcı bilgiler yer almaktadır. En çok benzersiz sınıf sayısına sahip job değişkeni olurken en az sınıf sayısına sahip değişken y olmuştur.

*Tablo 3. Kategorik Değişkenlerin İstatistiksel Tanımlayıcı Bilgileri*

	gözlem sayısı	benzersiz sınıf sayısı	en çok gözlenen sınıf
job	41188	12	admin
marital	41188	4	married
education	41188	8	university.degree
default	41188	3	no
housing	41188	3	yes
loan	41188	3	no
contact	41188	2	cellular
month	41188	10	may
day_of_week	41188	5	thu
poutcome	41188	3	nonexistent
y	41188	2	no

### 2.3. Metot

Çalışma kapsamında “DataBlender” uygulaması altında makine öğrenmesi modellerinden denetimli öğrenme modelleri uygulanmıştır. Makine öğrenmesi çalışma döngüsü altında yer alan veri ön işleme adımları ve modelleme adımları “DataBlender” üzerinde gerçekleştirilmiştir. Modelleme için “Python Sklearn” kütüphanesi, veri ön işleme için de yine

“Python” altında yer alan kütüphaneler uygulama tarafından otomatik olarak kullanılmıştır. Modelleme için Lojistik Regresyon, Destek Vektör Makineleri (Support Vector Machines, SVM), Torbalama (Bagging) ve Boosting (Hızlandırma) modelleri kullanılmıştır.

### 2.3.1. Lojistik Regresyon

Lojistik Regresyon, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi inceleyen ve yaygın olarak kullanılan bir istatistiksel analiz yöntemidir. Lojistik Regresyon, bağımlı değişkenin kategorik olması ve sınıflandırma işlemlerinde kullanılması bakımından Doğrusal Regresyondan ayrılır. Bu analiz türü, bağımlı değişkene uygulanan logit dönüşümünden adını alır ve amacı, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi açıklayacak bir model oluşturmaktır. Lojistik Regresyon, Doğrusal Regresyonda gereken varsayımları gerektirmediği için daha esnek bir şekilde kullanılabilir (Mertler & Reinhart, 2005).

Lojistik Regresyon Modeli:

$$L = \left[ \frac{P_i}{1 - P_i} \right] = Z_i = b_0 + b_1 X_i + e_i \quad (1)$$

Burada olma ihtimalini  $P_i$  ve olmama ihtimali  $1 - P_i$ ,  $\frac{1}{1 + e^{-z}}$  denkliği ile hesaplanır.

$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n$  şeklinde ifade edilir.  $\beta$ 'lar regresyon katsayısını göstermektedir (Özer, 2004).

### 2.3.2. Destek Vektör Makineleri (Support Vector Machines, SVM)

Destek Vektör Makineleri (SVM), hem regresyon hem de sınıflandırma işlemleri için kullanılan popüler bir makine öğrenmesi yöntemidir. Ancak, özellikle sınıflandırma amacıyla daha yaygın olarak tercih edilmektedir. Bu yöntem hem doğrusal hem de doğrusal olmayan veri setlerini etkili bir şekilde sınıflandırabilir. Genellikle, verileri doğrusal olarak sınıflandırmaya odaklanır ve az hesaplama gücüyle yüksek doğruluk oranları sağlaması nedeniyle tercih edilir. Destek Vektör Makineleri'nin temel işlevi, iki sınıfa ait verileri en uygun şekilde ayırmaktır. Bu, karar sınırları veya hiper düzlemler belirleyerek gerçekleştirilir ve sınıfları en iyi şekilde ayıran çizgiyi bulmaya çalışırlar (James vd., 2013).

Destek Vektör Makineleri, iki sınıfa ait verileri ayırmak için etkin bir sınıflandırma sağlamayı hedefler. Bu yöntem, iki sınıf arasındaki en uygun

karar fonksiyonunu, yani hiper düzlemi belirlemeye odaklanır. İki sınıflı verilerin sınıflandırılmasında, mevcut sonsuz sayıdaki hiper düzlemler arasından en etkili ayrımı yapacak olan hiper düzlemi seçmek önemlidir. Bu seçim, sınıflandırma işleminin doğruluğunu ve verimliliğini artırır (Vapnik, 1995).

### 2.3.3. Rassal Ormanlar (Random Forest)

Rassal orman algoritması makine öğrenmesinde karar ağaçlarını ana model olarak kullanan ve yaygın bir torbalama tekniğidir. Bu yöntemde, rastgele orman, daha önceden belirtilen torbalama prosedürü kullanılarak elde edilen bootstrap örneklerine dayalı bir dizi karar ağacından oluşturulur. Torbalama metodunun etkililiği, temel alınan modellerin çeşitliliğine büyük oranda bağlıdır. Eğer temel modeller arasında yüksek korelasyon varsa, birleşik model muhtemelen benzer sonuçlar üretecektir (Coşkun & Turanlı, 2023; Breiman, 2001).

Rassal Ormanlar algoritmasında, Bagging yöntemi gibi ağaçlar için gözlemler, bootstrap ile rastgele örnek seçimi kullanılarak belirlenir. Ancak, Rassal Ormanlar yönteminde farklı olarak, değişkenler arasından Random Subspace (Rastgele Alt Uzay) metodu ile rastgele bir alt grup seçilir. Bu seçilen alt grup, ilgili düğümün dallanmasında kullanılır. Bu rastgele seçim, orman içindeki ağaçların birbirine benzeme olasılığını azaltır. Bu nedenle, Rassal Ormanlar, Bagging metodunun geliştirilmiş bir versiyonu olarak kabul edilir (Breiman, 2001).

Rassal Ormanlar algoritmasının oluşturulması ve uygulanması kolay olması, bu yöntemin temel avantajlarından biridir. Son zamanlarda yapılan araştırmalar, birçok veri yapısı ve problemde karşılaşılan regresyon ve sınıflandırma hatalarının Rassal Ormanlar yöntemi kullanılarak azaltılabileceğini göstermiştir. Ayrıca, daha az parametre içerdiği için bu algoritmanın eğitimi daha hızlıdır. Minimum veri hazırlığı gereksinimiyle büyük veri setlerinde bile etkili ve verimli çalışabilirler. Ayrıca, veri setinde bulunabilecek aykırı gözlemlere ve aşırı öğrenmeye karşı dirençlidirler (Yalçın & Kalkan, 2022; Lewis, 2017).

### 2.3.4. Gradyan Artırma Makineleri (Gradient Boosting Machines)

Gradyan Artırma, tarafından önerilen ve daha güçlü bir model elde etmek amacıyla birden fazla karar ağacını birleştiren popüler bir ensemble yöntemidir. Bu yöntemde, her bir tahminci, önceki tahmincinin hatalarını etiket olarak kullanarak eğitilir. Gradyan destekli ağaçlar makine öğrenmesinde yaygın olmasına karşın, bu ağaçlar oldukça sığ yapıdadır. Örneğin, ağaçların

derinliği genellikle bir ila beş arasında değişir, bu da modelin bellek açısından daha az yer kaplamasını sağlar ve tahmin süreçlerini hızlandırır. Bu sığ ağaçlar, zayıf öğreniciler olarak görev yapar ve tahmin ediciye daha fazla sığ ağaç eklenerek performans sürekli olarak iyileştirilir(Coşkun & Turanlı, 2023;Friedman, 2001).

### **2.3.5. Ekstrem Gradyan Artırma (Extreme Gradient Boost, XGBoost)**

Ekstrem Gradyan Artırma (XGBoost) algoritması, performans ve işlem hızı açısından Gradyan Artırma algoritmasının ileri bir versiyonudur. Büyük veri setlerini içeren makine öğrenmesi problemlerinde kullanılmak üzere tasarlanmıştır. Tianqi Chen tarafından 2014 yılında geliştirilen bu algoritma, açık kaynak kodlu makine öğrenmesi algoritmalarının yer aldığı DMLC (Dağıtık Makine Öğrenimi Topluluğu) kütüphanesine eklenmiştir(Chen & Guestrin, 2016).

XGBoost, milyonlarca veri setini kısa sürede, diğer algoritmalara kıyasla daha az işlemci gücü ve geçici bellek (RAM) kullanarak analiz edebilir. Günümüzde veri miktarının artması ve veri analizi için gerekli işlem gücünün maliyetlerini azaltma ihtiyacı nedeniyle bu modele olan ilgi artmıştır. XGBoost, veriyi tamamen incelemek yerine parçalara ayırarak işler ve bu yüzden Gradyan Artırma algoritmasına kıyasla daha hızlı bir performans sunar.

### **2.3.6. Kategorik Artırma (Category Boosting, CatBoost)**

CatBoost (Kategorik Artırma), Yandex mühendisleri tarafından geliştirilen ve karar ağaçlarına dayalı bir gradyan artırma algoritmasıdır.

CatBoost, kategorik verileri işleyebilen bir gradyan artırma kütüphanesidir. Kategorik değerlerin ikili yerine konulması yerine, bu yöntem veri setinin rastgele bir permutasyonunu gerçekleştirir ve aynı kategori değerine sahip örneğin verilen sıradan önce yer alan ortalama etiket değerini hesaplar(Jhaveri vd., 2019). Bu, CatBoost'ta tanıtılan ve yeni bir artırma şeması olan sıralı artırma olarak adlandırılan önemli bir ilerlemedir. Bu yöntem, gradyan yanlılığı nedeniyle oluşan tahmin kaymasını aşabilir ve modelin genelleme yeteneğini daha da artırabilir(Coşkun & Turanlı, 2023;Zhang vd., 2020).

### **2.3.7. Hafif Gradyan Arttırma Makinesi (LightGBM)**

LightGBM, Microsoft tarafından geliştirilmiş ve XGBoost algoritmasının eğitim süresi performansını artırmak için tasarlanmış hafif bir gradyan artırıcıdır. LightGBM, dört açıdan geliştirilmiş bir XGBoost sürümüdür.

İlk olarak, LightGBM algoritması gradyan tabanlı tek taraflı örnekleme algoritmasını içerir. İkinci olarak, optimal segmentasyon noktasını belirlemek için bir histogram kullanır ve özel özellik paketlemesi yoluyla özelliği belirli bir ölçüde azaltır. Son olarak ise, geleneksel seviye bazında yerine derinlik sınırlaması ile yaprak bazında bir algoritma kullanır, bu da hem doğruluk artışı sağlar hem de aşırı öğrenmeyi önler(Ke vd., 2017).

## 2.4. Uygulama

Uygulama kapsamında kullanılan veri seti içerisinde, 41188 gözlem ve 21 değişken içererek kaggle.com üzerinde yetkili kuruluşun izni ile yayınlanmıştır. Veri seti; 11 kategorik ve 10 numerik değişkenden oluşmaktadır. Vadeli Mevduatı İsteme Durumu hedef değişken olarak y değişkeni altında bu bilgi mevcuttur.

Bağımlı (hedef) değişken olan Vadeli Mevduatı Kabul Etme Durumu (y) incelendiğinde imbalance (dengesiz) olduğu görülmektedir. Vadeli Mevduat isteme durumuna 36548 Hayır (%89) ve 4640 Evet (% 11) diyen müşteri mevcuttur.

DataBlender uygulaması altında makine öğrenmesi modellerinden denetimli öğrenme modelleri uygulanmıştır. Makine öğrenmesi çalışma döngüsü altında yer alan veri ön işleme adımları ve modelleme adımları DataBlender üzerinde gerçekleştirilmiştir. Modelleme için “Python Sklearn” kütüphanesi, veri ön işleme için de yine “Python” altında yer alan kütüphaneler uygulama tarafından otomatik olarak kullanılmıştır.

Modelleme kapsamında, kategorik değişkenler otomatik olarak “Label Encoder” ve “One Hot Encoder” ile sayısallaştırılmışlardır. Veri ön işleme adımları tamamlandıktan sonra model kurma işlemi gerçekleştirilmiştir.

Veri seti, bağımlı değişkeni, Vadeli Mevduatı İsteme Durumu olmak üzere %80 eğitim seti, %20 test seti şeklinde ayrılmıştır. 5’li “cross validation” uygulanmıştır. Çalışma kapsamında kullanılan modeller “DataBlender” altında yer alan “Sklearn” üzerinden alınan sınıflandırma modellerdir. Modeller AUC metriğine göre eğitilmiştir.

DataBlender uygulaması Resim 1’deki gibi veri ön işleme adımlarında kullanılan istatistiksel yöntemlerini sunmaktadır. Kullanıcının kararına göre bu işlemler seçilerek ön işleme adımları hızlı bir şekilde geçilmektedir. Veri ön işleme ve veri hazırlığı adımlarının makine öğrenmesi modellemelerinde oldukça zaman alması göz önüne alındığında, DataBlender uygulamasının bu yardımı oldukça verimli olmaktadır.



*Resim 1. "DataBlender" Uygulaması İçerisindeki "Preprocessing" Ara Yüzü*

Outlier yöntemlerinden (Tukey, Z-Score, Hampel Filter, Isolation Forest, DBScan, Grubb, Standard Deviation, Median Absolute Deviation, Double Median Absolute Deviation olmak üzere toplam 9 yöntem bulunmaktadır) Tukey seçilmiştir. Tukey yöntemi, veri setindeki sürekli değişkenlerin dağılımı hakkında varsayımlar yapmaksızın medyan, alt ve üst çeyrekler ile alt ve üst uç gibi temel istatistiksel bilgileri görselleştirmek için kullanılan etkili bir araçtır (Seo & Gary M. Marsh, 2006). Bu kural, aykırı değerlerin, veri setinin çeyrekler arası aralığının 1,5 katından fazla olan değerleri tanımlar.

Veri setinde eksik veri bulunmadığından eksik veri yöntemleri içerisinde yer alan Silme, KNN, Simple Imputer, Iterative Imputer, 0 ile doldurma yöntemleri seçilmeden devam edilmiştir.

Normalization sekmesi altında veri setindeki değişkenlerin Shapiro-Wilk ve Kolmogrov Smirnov testine göre normallik testleri yapılmaktadır. Çarpıklık ve Basıklık değerleri gösterilmektedir.

Correlation altında Spearman korelasyon testine göre pozitif ve negatif en yüksek değişken çiftleri gösterilmektedir. Korelasyon katsayısı, iki değişken arasındaki ilişkinin yönünü ve şiddetini ölçen bir istatistiksel değerdir. Bu katsayı, -1 ile +1 arasında bir değer alır. Pozitif değerler, aralarında pozitif yönlü bir doğrusal ilişki bulunan değişkenleri gösterirken, negatif değerler negatif yönlü bir doğrusal ilişkiyi işaret eder. İlişkinin şiddeti, genellikle şu şekilde yorumlanır: 0 ile 0.29 arasındaki korelasyon katsayıları düşük düzeyde ilişkiyi, 0.30 ile 0.70 arasındakiler orta düzeyde ilişkiyi ve 0.71 ile 1 arasındakiler ise yüksek düzeyde ilişkiyi temsil eder (James vd., 2013).



Tablo 4’de numerik değişkenler arasındaki ilişki ve ilişki gücü gösterilmektedir. Emp.var.rate olarak adlandırılan istihdam değişim oranı ve günlük bir gösterge olan Euribor 3 aylık oranı (euribor3m) arasında belirgin bir şekilde yüksek pozitif bir ilişki bulunmaktadır. Bu ilişki, günlük Euribor oranı aracılığıyla temsil edilen kısa vadeli ekonomik değişimlerin, üç aylık istihdam değişim oranı üzerinde önemli bir etkisi olduğunu göstermektedir. Pdays’ değişkeni, bir önceki kampanyadan beri müşterinin son temas edilme tarihini ifade eder. ‘Previous’ ise bu kampanyadan önce ve bu müşteri için yapılan temasların sayısını gösterir. Bu iki değişken arasında belirgin bir şekilde negatif bir ilişki bulunmaktadır. Bu güçlü negatif ilişki, önceki temas sayısı arttıkça son temasın üzerinden geçen sürenin azaldığını gösterir, bu da kampanyalar arasında daha sık ve kısa süreli etkileşim olasılığını işaret eder.

*Tablo 4. Değişkenler Arasındaki Pozitif ve Negatif Yönlü Korelasyon Tabloları*

Pozitif Yönlü En Yüksek Korelasyon Listesi			Negatif Yönlü En Yüksek Korelasyon Listesi		
degisken1	degisken2	korelasyon	degisken1	degisken2	korelasyon
emp.var.rate	euribor3m	0.972245	pdays	previous	-0.587514
euribor3m	nr.employed	0.945154	previous	nr.employed	-0.501333
emp.var.rate	nr.employed	0.906970	previous	euribor3m	-0.454494
emp.var.rate	cons.price.idx	0.775334	previous	emp.var.rate	-0.420489
cons.price.idx	euribor3m	0.688230	previous	cons.price.idx	-0.203130
cons.price.idx	nr.employed	0.522034	pdays	cons.conf.idx	-0.091342
pdays	nr.employed	0.372605	campaign	previous	-0.079141
pdays	euribor3m	0.296899	duration	campaign	-0.071699
cons.price.idx	euribor3m	0.277686	previous	cons.conf.idx	-0.050936
pdays	emp.var.rate	0.271004	duration	pdays	-0.047577

### 3. Bulgular

Sınıflandırma modellerinin değerlendirme kriterleri arasında genellikle Karmaşıklık Matrisi (Confusion Matrix) ve ROC Eğrisi değerleri kullanılmaktadır (Han vd., 2012).

*Tablo 5. İki Sınıflı Karmaşıklık Matrisi*

Karmaşıklık Matrisi (Confusion Matrix)		Tahmin Edilen Sınıf	
		0	1
Gerçek Sınıf	0	Doğru Negatif (DN)	Yanlış Negatif (YN)
	1	Yanlış Pozitif (YP)	Doğru Pozitif (DP)



- Doğru Pozitif (DP) durumu, bir sınıflandırıcının pozitif olarak sınıflandırılması gereken verilerden kaç tanesini doğru şekilde pozitif olarak sınıflandırdığını gösterir.
- Doğru Negatif (DN) ise, negatif sınıfa ait verilerin kaç tanesinin sınıflandırıcı tarafından doğru şekilde negatif olarak sınıflandırıldığını ifade eder.
- Yanlış Negatif (YN) durumu, gerçekte pozitif sınıfa ait olan ancak sınıflandırıcı tarafından yanlışlıkla negatif sınıf olarak işaretlenen verileri tanımlar.

Yanlış Pozitif (YP), gerçekte negatif sınıfa ait bir verinin, sınıflandırıcı tarafından yanlışlıkla pozitif sınıf olarak etiketlenmesi durumudur (Özdemir, 2021; ALAN & KARABATAK, 2020).

Modellerin karşılaştırılması için Karmaşıklık Matrisi'nden elde edilen değerler ile Doğruluk, Kesinlik, Hassasiyet, F1 skoru, AUC, Kappa, MCC metrikleri göz önünde bulundurulmuştur.

- Doğruluk (Accuracy): Tahminlerin genel doğruluk oranıdır.

$$\text{Doğruluk} = \frac{DP + DN}{DP + DN + YP + YN} \quad (2)$$

- Kesinlik (Precision): Pozitif olarak tahmin edilen örneklerin gerçekte pozitif olma oranıdır.

$$\text{Kesinlik} = \frac{DP}{DP + YP} \quad (3)$$

- Hassasiyet (Recall): Gerçekte pozitif olan örneklerin pozitif olarak tahmin edilme oranıdır.

$$\text{Hassasiyet} = \frac{DP}{DP + YN} \quad (4)$$

- F1 Ölçütü: Kesinlik ve Hassasiyet değerlerinin harmonik ortalamasıdır. Hedef değişkenin dengesiz olduğu durumlarda özellikle incelenen bir ölçüttür.

$$F1 \text{ Ölçütü} = \frac{2 * K * G}{K + G} \quad (5)$$

- Eğri Altında Kalan Alan (Area Under the Curve, AUC): Pozitif örnekleri negatif örneklerden ayırma yeteneğini ölçer. ROC eğrisinin altında kalan alan değeridir.
- KAPPA Ölçütü: Accuracy'nin beklenen doğruluk oranından sapmasını ölçer.
- Matthews Korelasyon Katsayısı (Matthews Correlation Coefficient, MCC): Gerçek sınıflandırmanın tahmin edilen sınıflandırmaya yakınlığını ölçer (Burkov, 2019)

“DataBlender” ile birden fazla model kullanıcı kararıyla oluşturulabilmektedir. Çalışma kapsamında ağaç temelli modeller ve klasik modeller seçilerek performansları kıyaslanmıştır. Oluşturulan modellerin hata metrikleri aşağıdaki tabloda gösterilmektedir.

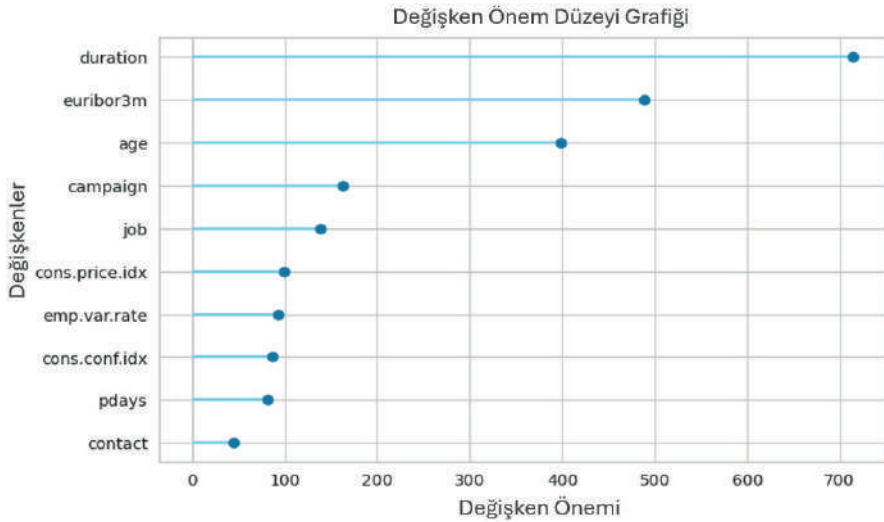
*Tablo 6. Model Performansları*

Model	Doğruluk	Eğri Altında Kalan Alan	Hassasiyet	Kesinlik	F1 Ölçütü	Kappa Ölçütü	MCC Ölçütü
Light Gradient Boosting Machine	0.9179	0.9502	0.5493	0.6653	0.6010	0.5558	0.5594
Catboost	0.9161	0.9492	0.5576	0.6495	0.5995	0.5530	0.5553
Gradient Boosting Machine	0.9165	0.9464	0.5339	0.6604	0.5900	0.5441	0.5482
Extreme Gradient Boosting(XgBoost)	0.9129	0.9460	0.5425	0.6332	0.5837	0.5355	0.5378
Rassal Ormanlar (Random Forest)	0.9135	0.9410	0.4701	0.6644	0.5502	0.5039	0.5134
Lojistik Regresyon	0.9105	0.9328	0.4129	0.6671	0.5093	0.4631	0.4797
Destek Vektör Makineleri	0.8207	0.0000	0.4758	0.5421	0.3987	0.3372	0.3710

Topluluk öğrenme modelleri (Bagging ve Boosting), genel olarak başarılı sonuçlar vermiştir. Tablodaki performans ölçütleri incelendiğinde, modellerin sınıflandırma performanslarını değerlendirebilmek adına çeşitli metriklerin dikkate alındığı görülmektedir. En yüksek AUC skoru (%95) ile LightGBM modeli, verilen veri seti üzerinde en iyi sınıflandırma yeteneğini

sergilemiştir. Veri dengesiz olduğundan F1 skoru, veri desenine hiç müdahale edilmemesine rağmen %60,1'dir. Bu da modellemenin iyi bir başarı oranı yakaladığını göstermektedir.

Diğer modeller arasında, Catboost Classifier ve Gradient Boosting Classifier modelleri de oldukça yüksek başarı oranlarına sahiptir. Bu modeller, Accuracy (Doğruluk) ve AUC değerlerinde diğer modellerle yakın performans gösterirken, Recall (Geri Çağırma) ve Precision (Hassasiyet) ölçütlerinde de dengeli sonuçlar elde etmişlerdir. Ayrıca, model performansı yanında işlem süreleri de dikkate alındığında, LightGBM gibi yüksek performanslı modellerin yanı sıra Gradient Boosting ve Random Forest gibi modellerin de oldukça kabul edilebilir işlem sürelerinin olduğu görülmektedir. Bu sonuçlar, topluluk öğrenme modellerinin genel olarak iyi performans gösterdiğini ve bu tür modellerin bu belirli veri seti üzerinde daha etkili sınıflandırma yetenekleri sergilediğini vurgulamaktadır.



**Resim 2. Değişken Önem Düzeyi Grafiği**

En başarılı çıkan model üzerinden hedef değişken üzerinde etkili olan değişkenleri önem sırasına göre sıralamak mümkündür. “LightGBM” modeli en başarılı model çıktığı için bu model üzerinden, rastgele seçilen değişkenler arasında kurulan modeller sonucunda bağımlı değişken üzerinde en fazla etkisi olan faktörler belirlenmiştir. Bu analizler neticesinde ‘Duration’ (Son iletişim süresi) en etkili değişken olarak belirlenmiştir. Yani, müşterilerin vadeli mevduat alma davranışı üzerinde en önemli değişken olarak öne

çıkmiştir. Ardından sırasıyla euribor3m (Euribor 3 aylık oran- günlük gösterge) ve age (yaş) değişkenleri gelmektedir.

#### 4. Sonuç ve Tartışma

Bu çalışma, finansal hizmetlerde müşteri tercihlerini anlama ve vadeli mevduat talebini öngörme amacıyla yapılmıştır. “DataBlender AutoML” ile veri seti detaylıca incelenmiş ve uygun ön işleme adımları ve modelleme çalışması kod yazmadan yapılmıştır. İncelenen veri seti, dengesizlik içeren bir dağılım göstermiştir bu da doğru sınıflandırmayı zorlaştırmıştır. Ancak, yapılan modelleme çalışmaları, topluluk öğrenme modellerinin, özellikle “LightGBM” gibi yüksek performanslı modellerin, bu dengesizlikle başa çıkabilme yeteneğini ortaya koymuştur. Özellikle AUC skoru üzerinden değerlendirildiğinde, “LightGBM” modeli en yüksek sınıflandırma yeteneği sergilemiştir. Çalışma, veri setindeki dengesizliği göz önünde bulundurarak F1 skoru üzerinden modele daha dengeli bir değerlendirme yapma imkânı sunmuştur. Topluluk öğrenme modelleri genel olarak yüksek başarı oranları göstermiş ve özellikle “Catboost Classifier” ve “Gradient Boosting Classifier” modelleri, dengeli sonuçlar elde etmiştir. Ayrıca, işlem süreleri de dikkate alındığında, yüksek performanslı modellerin yanı sıra “Gradient Boosting Classifier” ve “Random Forest Classifier” gibi modellerin de kabul edilebilir işlem süreleriyle etkileyici sonuçlar verdiği gözlemlenmiştir.

Çalışmanın sonucunda, başarılı çıkan “LightGBM” modeli üzerinden, vadeli mevduat hesabı isteme olasılığı üzerinde etkili olan değişkenleri belirlemiş ve müşterilerin özelliklerini sıralamıştır.

Bu çalışma ile müşterilerin model sonuçlarına göre atanacağı tahmin olasılık skorlarına göre sıralanıp en yüksek vadeli mevduatı alacak müşterilere odaklanması sağlanacaktır. Böylece hedef kitle daraltılarak zaman ve maliyet yönünden tasarruf sağlanıp doğru strateji ile müşteriye ulaşılması sağlanacaktır. Bu sonuçlar, finans kuruluşlarının hedefleme stratejilerini iyileştirmek ve pazarlama çabalarını daha etkili bir şekilde yönlendirmek için makine öğrenmesi modellerini kullanmalarını teşvik edebilir.

## Kaynakça

- A.Elsalamony, H. (2014). Bank Direct Marketing Analysis of Data Mining Techniques. *International Journal of Computer Applications*, 85(7), 12–22. <https://doi.org/10.5120/14852-3218>
- ALAN, A., & KARABATAK, M. (2020). Veri Seti - Sınıflandırma İlişkisinde Performansa Etki Eden Faktörlerin Değerlendirilmesi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 32(2), 531–540. <https://doi.org/10.35234/fumbd.738007>
- Alpaydın, E. (2010). *Introduction to Machine Learning*. USA: Massachusetts Institute of Technology Press.
- Breiman, L. (2001). Random Forests. In *Machine Learning*. Springer.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Coşkun, S. B., & Turanlı, M. (2023). Credit risk analysis using boosting methods. *Journal of Applied Mathematics, Statistics and Informatics*, 19(1), 5–18. <https://doi.org/10.2478/jamsi-2023-0001>
- Friedman, J. H. (2001). *Greedy function approximation: A gradient boosting machine*. *The Annals of Statistics*.
- Ghatasheh, N., Faris, H., AlTaharwa, I., Harb, Y., & Harb, A. (2020). Business Analytics in Telemarketing: Cost-Sensitive Analysis of Bank Campaigns Using Artificial Neural Networks. *Applied Sciences*, 10(7), 2581. <https://doi.org/10.3390/app10072581>
- Han, Jiawei Kamber, Micheline Pei, J. (2012). *Data mining: Concepts and techniques*. Morgan Kaufmann Publishers.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer.
- Jhaveri, S., Khedkar, I., Kantharia, Y., & Jaswal, S. (2019). Success Prediction using Random Forest, CatBoost, XGBoost and AdaBoost for Kickstarter Campaigns. *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, 1170–1173. IEEE. <https://doi.org/10.1109/ICCMC.2019.8819828>
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *NIPS*.
- Lewis, D. (2017). *Machine Learning Made Easy with R: An Intuitive Step by Step Blueprint for Beginners*. CreateSpace Independent Publishing Platform.
- Mertler, C. A., & Reinhart, R. V. (2005). *Advanced and multivariate statistical methods: practical application and interpretation*.

- Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62, 22–31. <https://doi.org/10.1016/j.dss.2014.03.001>
- Moro, S., Laureano, R. M. S., & Cortez, P. (2011). Using data mining for bank direct marketing: An application of the CRISP-DM methodology. *ESM 2011 - 2011 European Simulation and Modelling Conference: Modelling and Simulation 2011*, (Figure 1), 117–121.
- Okur, H., & Cetin, A. (2019). Credit Risk Estimation With Machine Learning. *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 1–6. IEEE. <https://doi.org/10.1109/ISMSIT.2019.8932917>
- Özdemir, R. (2021). *Makine öğrenmesindeki sınıflandırma yöntemlerinin karşılaştırılması ve e-ticaret üzerinde bir uygulama*. İstanbul Ticaret Üniversitesi.
- Özer, H. (2004). *Nitel Değişkenli Ekonometrik Modeller*. Ankara: Nobel Yayın Dağıtım.
- Seo, S., & Gary M. Marsh, P. D. (2006). A review and comparison of methods for detecting outliers in univariate data sets. *Department of Biostatistics, Graduate School of Public Health*, 1–53. Retrieved from <http://d-scholarship.pitt.edu/7948/>
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York, USA: Springer-Verlag.
- Yalçın, M., & Kalkan, S. B. (2022). DETERMINING THE BEST ESTIMATION MODEL WITH TREE-BASED MACHINE LEARNING METHODS: IMPLEMENTATION ON CUSTOMER SPENDING FOR E-COMMERCE WEBSITES. *Advances and Applications in Statistics*, 75, 91–109. <https://doi.org/10.17654/0972361722029>
- Zhang, Y., Zhao, Z., & Zheng, J. (2020). CatBoost: A new approach for estimating daily reference crop evapotranspiration in arid and semi-arid regions of Northern China. *Journal of Hydrology*, 588(May), 125087. <https://doi.org/10.1016/j.jhydrol.2020.125087>



# Türkiye’de Teknolojik Ürünlerin Hedef Pazarlarının İstatistiksel Yöntemlerle Değerlendirilmesi<sup>1</sup>

Aylin Uçan<sup>2</sup>

Nuran Çakır Yıldız<sup>3</sup>

## Özet

Günümüzde teknoloji alanında yaşanan büyük ve hızlı değişimler teknolojik ürünlerinin çeşitliğini artırıp kullanım alanlarını çeşitlendirmektedir. Teknoloji perakende sektöründe beyaz eşyadan klimaya, elektrikli ev aletlerinden telefona, bilgisayardan tablete kadar geniş bir ürün yelpazesi sunulmaktadır. Tüketiciler bu ürünleri satın almada fiziksel mağazaları sorularına ayrıntılı ve net yanıtlar alabilmek için daha çok tercih etmektedir. Türkiye’de teknolojik ürün satan çok sayıda firma olması dolayısı ile artan rekabet sonucu firmalar satış rakamlarını ve stratejilerini gizli tutmaktadır. Ulusal literatür incelendiğinde, teknolojik ürün gruplarının satış performansına göre iller düzeyinde farklı olup olmadığının tespiti ve bu farklılaşmayı etkileyen faktörleri inceleyen çalışmaların sayısının yok denecek kadar sınırlı sayıda kaldığı görülmektedir. Teknolojik ürünlerin iller düzeyinde ürün bazlı satış performanslarını etkileyen önemli makro değişkenleri (demografik, ekonomik, eğitim vb.) saptamak ve ürünlerin kendi içerisinde benzer ve farklı yönlerini belirleyerek pazar analizinin yapılması bu çalışmanın özgünlüğünü oluşturmaktadır. Saha çalışması, Türkiye’de teknolojik ürünler pazarında sektörü temsil edebilme özelliğine sahip ve önemli bir pazar payı olan büyük ölçekli bir işletme örneklemini üzerinden yapılmıştır. Çalışmada teknolojik ürünlerin iller düzeyinde satış performanslarını etkileyen önemli makro değişkenleri saptamak için diskriminant analizi kullanılmıştır. Bunun yanı sıra iller ve ürünlerin satış performanslarındaki benzerlik ve farklılıklarını görsel olarak daha kolay anlaşılabilir hale getirmek amacıyla çok boyutlu ölçekleme analizi kullanılmıştır.

- 1 Bu çalışma, “Türkiye’deki Teknolojik Ürünlerin Çok Değişkenli İstatistik Yöntemlerle Pazar Analizi” başlıklı yüksek lisans tezinden üretilmiştir.
- 2 Öğrenci (Doktora), Yıldız Teknik Üniversitesi, aynucn@gmail.com, 0000-0003-3004-1076
- 3 Dr. Öğr. Üyesi, İstanbul Üniversitesi-Cerrahpaşa, nurcak@iuc.edu.tr, 0000-0003-1770-9561



## 1. Giriş

Teknoloji alanında yaşanan hızlı ve köklü değişimler teknolojik ürünlerin hem sayısını hem de farklı kullanım alanları yaratarak çeşitliliğini artırmıştır. İş yaşantısından eğitime, eğitimden sosyal ilişkilere sağlıktan eğlenceye, hayatın her alanında teknolojinin etkisi derinden hissedilmektedir.

Aynı şekilde farklı kullanım amaçlarına sahip ev içerisinde kullanılan elektronik ev ürünleri de günlük insan yaşamının vazgeçilmezi olmuştur.

Teknolojide meydana gelen hızlı değişimler ulusal ve uluslararası piyasada rekabet avantajı sağlamak isteyen işletmeleri rakiplerine göre daha kaliteli ve yenilikçi ürün üretebilmeyi ve satabilmeyi zorunlu hale getirmiştir. Artan rekabet ortamı, ürün çeşitliliğini artırırken, müşteri memnuniyetini sağlamak adına sunulan ürün ve hizmetler arasındaki farklılıkları da neredeyse yok denecek kadar azaltmıştır. Bu durum özellikle teknolojik ev ürünlerinin pazarlama faaliyetinin satın alım sürecinde yüksek düzeyde bilgi sağlandığı ve fiziksel incelemenin mümkün olduğu gerekçesiyle fiziksel alışverişin sağlandığı perakende sektörünün gün geçtikçe öneminin artmasına sebep olmuştur. Bu noktada tüketiciler teknolojik ürün alışverişlerinde fiyat ve kalite dışında ihtiyaçlarının tamamını tek bir noktadan karşılayabilecekleri ve daha çok seçeneğin sunulduğu perakendecileri tercih etmektedirler (Polat & Külter, 2007).

Marketing Türkiye adına Pulside Research tarafından gerçekleştirilen “Teknoloji Araştırması” konulu çalışmada Türkiye’de tüketicilerin yüzde 47,9’u teknoloji alışverişi için fiziksel mağazaları tercih ederken yüzde 32,6’sı online pazaryerlerine ve yüzde 18,5’le markaların kendi online mağazaları tercih ettikleri tespit edilmiştir<sup>4</sup>.

Teknolojik ürünler içinde yer alan elektronik ev aletleri beğenmeli ürünler grubunda değerlendirilmektedir. Tüketici satın alma davranışları açısından markada ısrar etme derecesi, alışveriş için dolaşma çabası, marka karşılaştırma çabası, fiyat değerlendirmesi, satış çabası ve mağaza hizmetleri, satış temsilcilerinden bilgi alma gibi yöntemlere başvurması beğenmeli ürünlerin ayırt edici özellikleridir (Mayer vd.1971).

Bu sektörde, dağıtım kanalı halkası olarak perakendeci mağaza üretici ve tüketiciler arasında sağladığı bilgi akışı sonucu, ürün veya hizmeti doğru hedef kitlesine ulaştırmada kilit rol oynar. Demografik çevre faktörü olarak bilinen nüfusun yaş ve cinsiyet dağılımı, eğitim düzeyi, nüfusun kentlere ve kırsal alanlara dağılımı vb. unsurlar perakendecilik sektörünün yapısını

4 <https://www.marketingturkiye.com.tr/haberler/teknoloji-alisverisi-hayaller-ve-gercekler/>

belirlenmektedir (Arıkbay, 1996). Bu yönüyle bakıldığında tüketici pazarlarının hacmi ve büyüklüğü hakkında genel ancak önemli bir bilgi vermesi nedeniyle pazarlamacılar tarafından sürekli izlenmektedir. Tüketicilerden ürünler hakkında ne kadar geri bildirim alabilirse, üretici işletmelerde bir o kadar tüketici tercih ve taleplerindeki değişime uyum sağlama kabiliyetine sahip olacaktır (Cengiz & Özden, 2002).

Ürünün, ürün fiyatının ve satış sonrası hizmetlerin benzer yapıda olduğu teknolojik ev ürünlerinde firmalar farklılaşmayı dağıtım kanalları üzerinden de sağlamaya çalışmaktadırlar. Bu açıdan yeni müşteriler kazanma, mevcut müşterileri elde tutma ve mağaza bağlılığı oluşturma amaçlı çeşitli satış artırıcı faaliyetlerde de bulunabilmektedirler. Dolayısıyla dağıtım kanalındaki perakendecilerin güçlenmesi dağıtım maliyetlerini düşürme açısından önemli olmaktadır (Rosenbloom , 2004).

Teknolojik ev ürünlerinin dağıtım kanalındaki perakendeci mağaza, bir yandan ürünleri nihai tüketicilere ulaştırırken, diğer yandan üstlendikleri görevlerle üretici ve tüketicilere fayda sağlayarak üreticilerin temsilcisi, tüketicilerin de bir çeşit garantörü pozisyonundadır. Bu yönüyle ekonomik yapı içerisinde büyük önem taşımaktadır (Gököz, 2018).

Literatürde, teknolojik ürünlerin pazarlanması üzerine yapılan çalışmalar çoğunlukla tüketicinin satın alma karar süreci üzerinden teknolojik ürünlerin satış hacmini artıran faktörlerin belirlenmesine yöneliktir.

Ulusal literatür incelendiğinde, beğenmeli ürün sınıfında yer alan teknolojik ev ürünlerinin dağıtım kanalı üzerinden ürün gruplarının iller düzeyinde farklılaşma olup olmadığının tespiti ve bu farklılaşmayı etkileyen faktörleri inceleyen çalışmaların sayısının yok denecek kadar sınırlı sayıda kaldığı görülmektedir.

Bu çalışmanın amacı teknolojik ev ürünlerinin iller düzeyinde ürün bazlı satış performanslarını etkileyen önemli makro çevresel değişkenlerin (demografik, ekonomik, eğitim vb.) etkisini anlayabilmek ve illerin ve ürünlerin kendi içerisinde benzer ve farklı yönlerini tespit etmektir.

Bu araştırma, teknolojik ev ürünleri perakende sektörü için pratik katkılar sunmaktadır. Elde edilen bulgular söz konusu perakendeci mağazaların satışlarının ürünler ve iller bazında nasıl farklılaştığını ortaya koyarak pazarlama karması stratejilerine rehberlik edebileceği düşünülmektedir.

Çalışma, Türkiye’de teknolojik ürün pazarındaki yaklaşık %30’luk payıyla ilk üç arasında sayılan bir teknolojik ürün mağazası örneklemeden gerçekleştirilmiştir. Verilerin elde edildiği teknolojik ürün mağazasının

seçiminde mağazanın yaygınlığı (mağaza sayısının çokluğu), ürün çeşitliliği, hizmet kalitesi ve güvenilirlik gibi ölçütler dikkate alınmıştır.

## 2. Kavramsal Çerçeve

### 2.1. Ürün ve Tüketicilerin Alışkanlıklarına Göre Ürün Sınıflandırılması

Tüketicilerin istek ve ihtiyaçların karşılanmasına yönelik olarak işletmelerin pazara sunmuş oldukları ve tüketiciye fayda sağlayan fiziksel ve fiziksel olmayan varlıklara ürün denilmektedir (Kotler ve Armstrong, 2008; Kotler, 2003; Avlonitis, 2000). Üretici ürünü, kar sağlama amaçlı olarak bir fikir, düşünce ya da fiziksel bir varlık olarak ele alırken, tüketici aynı ürünü/hizmeti, ihtiyacı karşılama özelliğinden dolayı kendisine sağlayacağı fayda boyutu ile düşünmektedir. Aracı taraf ise, bu ürünü tekrar başkalarına satmak amacı ile alarak kendisine bir kazanç getiren bir araç olarak görülmektedir (Değer, 2010). İlk olarak 1923 yılında Copeland tarafından tüketicilerin alışkanlıklarına yönelik olarak yapmış olduğu ürün sınıflandırması “kolayda ürünler”, “beğenmeli ürünler” ve “özellikli ürünler” olmak üzere üç temel başlık altında ele alınmış ve ilgili literatürde birçok araştırmacı tarafından kabul görmüştür ve önerilmiştir (Kavak & Sığındı, 2012). Temellerini Copeland’ın atmış olduğu, tüketici alışkanlıklarına göre yapılan ürün sınıflarını aşağıdaki gibi temel çerçevede genel hatlarıyla tanımlamak mümkündür.

**Kolayda Mallar:** Tüketicilerin satın alma sürecinde çok fazla zaman ve çaba harcama niyetinde olmadığı, düzenli olarak satın alınan ürün grubudur. Ek olarak diğer ürün grubundaki ürünlere göre görece olarak daha düşük fiyatlıdır. Tüketiciler kolayda malları satın almadan önce genelde herhangi bir plan yapmamaktadır (Kavak & Sığındı, 2012). Kolayda malların satın alınma sürecinde tüketicilerin algıladığı risk düşüktür (Avlonitis, Hart, & Tzokas, 2000).

**Beğenmeli Ürünler:** Tüketiciler bu ürünleri satın almadan önce kendilerine en fazla tatmini sağlayacak olan ürünü bulmak için daha önce belirlemiş oldukları alternatifler arasında hem fiyat hem de kalite karşılaştırması yapmaktadırlar. Benzer beğenmeli malları satan birçok mağaza tüketiciye kolaylık sağlamak ve satış hacmini artırma amacıyla birbirlerine yakın konumlarda faaliyet göstermektedir (Kavak & Sığındı, 2012). Bu ürünler daha az sıklıkta satın alınan daha ve uzun süreli kullanılan ürünlerdir. Alışverişe referans gruplarından diğer insanlar ile gidip, satış temsilcilerinden bilgi alma gibi yöntemlere başvururlar (Harrell, 2002). Ek olarak, başka eşyalarla da birlikte kullanılabilme gerekçesiyle, alacakları ürünlerin diğer eşyalar uyumunu isterler. Benzer beğenmeli malları satan

birçok mağazanın birbirlerine yakın konumlarda faaliyet göstermelerinin sebebi, tüketicilere bu noktada kolaylık sağlamak ve satış imkanını artırmaktır (Kotler & Armstrong, 1994).

Beğenmeli mallarda kendi içinde türdeş (homojen) ve türdeş olmayan (heterojen) mallar olarak iki alt sınıfa ayrılmaktadır. Türdeş olmayan ürünler (mobilya, ev eşyaları vb.) fiyat, kalite ve diğer özellikleri önemli farklılıklar gösterebilmektedir. Beyaz eşyalar gibi türdeş ürün grubunda yer alan beğenmeli ürünlerde ise, tüketiciler aradıkları özelliklere sahip seçenekler içerisinde en düşük fiyata sahip olan ürünü tercih etme eğilimindedirler (Kırcova, 2018). Bu açıklamalar dikkate alındığında bu çalışmanın konusu olan beyaz eşya, elektronik ev aletleri, telefon bu kategoride bulunan mallardandır.

**Özellikli Mallar:** tüketicilerin nereden satın alabilecekleri konusunda fikir sahibi oldukları ve bu ürünleri satın alma sürecinde genelde karşılaştırma yapmadığı ürünlerdir. Özellikle mallar 'lüks ürünler' olarak da bilinmektedir (Kotler, 2003). Ünlü saat markaları, mücevher ve spor arabalar bu sınıf içinde yer almaktadır.

Pazarlama yazınında uzun süredir geçerliliğini koruyan bu sınıflandırmayı tartışan çalışmaların odak noktası ise, belirli bir ürün tipinin çeşitli durumlarda farklı ürün sınıflarına dâhil olup olamayacağının tartışılmasıdır. Buna göre aynı ürünün ürün sınıflandırması, tüketicinin niteliğine göre ve aynı tüketici için farklı zamanlarda farklılık gösterebilecektir. Avlonitis'e (Avlonitis, Hart, & Tzokas, 2000) göre, ürün sınıflandırması dinamik bir özellik taşımaktadır.

Ürün sınıflandırılması üzerinden yürütülen pazarlama stratejilerinde birçok ürün türünün, tanımlanan ürün sınıflarından herhangi birine tam olarak dahil olmadığı görülür. Dolayısıyla bu özelliğin göz önünde bulundurulması gerekmektedir. Sebebine gelince, bir ürünün ait olduğu ürün grubu tüketiciden tüketiciye değişen bireysel alışkanlıklara, farklı coğrafik bölgelere göre ve ürün hayat seyir aşamalarına bağlı olarak bir ürün sınıfından diğer ürün sınıflarına geçişkenlik gösterebilmektedir. Teknolojisi sürekli ve hızlı gelişen akıllı cep telefonu, LCD veya LED televizyon gibi ürünler bu duruma örnek verilebilir. Söz konusu bu ürünler pazara ilk sunuldukları dönemde "özellikli" olarak tanımlanırken, ürün hayat eğrisinin olgunluk safhasında, "beğenmeli" ürün olabilirler (Avlonitis, Hart, & Tzokas, 2000). Bu ürünler ürün-yaşam eğrisinin ilk safhalarında erişiminin zor ve nispeten pahalı olmasından ötürü "özellikli" olarak tanımlanırken ürünün ilerleyen aşamalarında ise görece geniş dağıtım ağına sahip olduğu ve daha ucuz olduğu ya da tüketici tarafından artık özel olarak algılanmadığı için "beğenmeli" olabilmektedir.

## 2.2. Literatür Özeti

Teknolojik ürün sektöründeki küresel rekabetin artması sonucu işletmeler pazar paylarını artırmak ya da korumak amaçlı olarak ürünlerini sürekli geliştirmeye zorlanırken bir yandan da fiyat rekabeti içindedirler. İleri teknoloji işletmelere ürün geliştirmeyi daha kısa sürede yapılabilme olanağını yaratırken teknolojik ürün perakendeciliğinde her yeni ürün modeli bir öncekinin demode olmasına neden olmaktadır. Dolayısıyla teknolojik ürün perakendeciliğinde, benzer ürünler farklı kanal ve satış noktalarında benzer fiyatlarla satışa sunulduğu için rekabet açısından zor bir sektördür.

Teknolojik ürün sektörünün temsilci kuruluşu TÜBİSAD’ın “Bilgi ve İletişim Teknolojileri Sektörü 2019 Yılı Pazar Verileri” ne göre sektör, TL bazında bir önceki yıla oranla %14’lük büyüme ile 152,7 milyar TL’lik hacme ulaşmıştır. İstihdamını 143 bin kişiye çıkaran sektörün ihracatı da 6,5 milyar TL olarak gerçekleşmiştir. Türkiye Sanayi ve Teknoloji Bakanlığı’nın açıkladığı 2020 yılının en çok satılan teknolojik ürün grubu akıllı telefonlardır. Yıl toplamında 10,8 milyon akıllı telefon, 7,6 milyon beyaz eşya, 2,6 milyon televizyon, 1,3 milyon tablet ürünü satılmıştır (Kalelioğlu, 2021).

Türkiye’de teknolojik ürünler satan çok sayıda küçük ve büyük işletme olduğu için bu pazarın analizini yapabilmek oldukça zor bir süreçtir. Bu alanda rekabet eden çok sayıda firma olması nedeniyle firmalar satış rakamlarını ve stratejilerini gizli tutmaktadır. Bu konuda yapılan araştırmaların amacı, kişi ve kurumların veri ve bilgilerinin gizliliğine saygı ilkesi gereğince sahadan elde edilen verilerden yararlanarak işletmelere sadece danışmanlık hizmeti vermektir.

Ulusal literatür incelendiğinde iller düzeyinde teknolojik ürün gruplarının satış adet verileri üzerinden istatistiksel analiz yöntemleri kullanılarak hem iller bazında hem de ürün grupları arasında farklılaşma olup olmadığının tespiti ve bu farklılaşmayı etkileyen faktörleri inceleyen çalışmaların sayısı yok denecek kadar sınırlı sayıdadır.

Ulusal literatürde konuyla ilgili çalışmaların büyük bir kısmı tüketici satın alma davranışları ile teknolojik ürünlerin satış hacmini artıran faktörlerin belirleyicileri üzerinedir.

Acar ve Çizmeçi (2015) tarafından kolayda örnekleme yöntemiyle seçilen 450 denek üzerinden yapılan bir çalışmada demografik özelliklerin teknolojik ürün perakendecisi seçiminde farklılık gösterip göstermediği araştırılmıştır. Yapılan faktör analizi sonucunda yaş, eğitim düzeyi ve gelir durumuna göre farklılık olduğu tespit edilmiştir.

Özbek ve Koç (2009), kırsal ve kentsel bölge ayırımı üzerinden dayanıklı tüketim mallarına yönelik satın alma davranışlarını incelemişlerdir. Veri seti üzerinden Ki-kare testi analizi sonucunda kentsel yerleşim yerlerinde dayanıklı tüketim eğiliminin kırsal yerleşim bölgelerine göre daha fazla olduğu sonucuna varmışlardır.

Erdal'ın (2000), yüksek teknoloji ürünlerin pazarlanması konulu çalışmasında ise yüksek teknoloji ürünlerin üretiminde ve pazarlanmasında makro çevre faktörlerinin önemi ortaya konulmuştur.

Şahin (2007), tüketicilerin beyaz eşya talebini etkileyen faktörlerin tespitine yönelik bir çalışma yapmıştır. Yaptığı saha çalışması sonucunda beyaz eşya talebini etkileyen en önemli faktörlerin “ilk satın alma” ve “yenileme-değiştirme talebi” olduğunu gözlemlemiştir. Bunun yanında pazarın büyümesinde etkin ve tüketiciye uygun imkanlar sunan kampanyaların önemli olduğunu vurgulamıştır.

Sarangı (2014), 300 hane halkının katıldığı anket yöntemiyle elde ettiği veri seti üzerinden yaptığı çalışmasında dayanıklı tüketim mallarını büyük ve küçük dayanıklı tüketim malı olarak sınıflamış ve bu mallara yönelik talebi incelemiştir. Çalışmasında analiz yöntemi olarak logit model uygulamış ve sonuç olarak birkaç ürün dışında (düdüklü tencere, saat vb.) gelirin ve eğitim düzeyinin sahip olunan dayanıklı tüketim mallarının talebini pozitif etkilediği sonucuna ulaşmıştır.

Türkiye’de perakende mağaza seçimini etkileyen faktörler üzerine yapılan çalışmaların önemli bir kısmı, daha çok uluslararası literatürde bahsedilen mağaza özelliklerini içermektedir.

Araştırma firması Twentify (2021), 1040 kişinin anket yoluyla görüşünün alındığı ve tüketici eğilimlerinin tespitine yönelik olarak “elektronik aletlerin tercih edilmesindeki kriterlerin tespiti” isimli bir çalışma gerçekleştirmiştir.

Bu çalışmanın sonuçlarına göre, yenilikçi ürün arayışında ilk tercihin evin vazgeçilmezleri olarak sıralanan küçük ev aletleri (%17 kahve makinesi ve %15 blender) olmakla birlikte %83 gibi bir oranla televizyon açık ara en çok talep gören üründür. En az sahip olunan elektronik ürünler ise sırasıyla %28 akıllı saat ve %21 akıllı bilekliktir. Bu çalışmada dikkat çeken en önemli bulgulardan birisi de satın alma kanalları üzerinedir. Buna göre, elektrikli ürün alışverişlerinde tüketicilerin %48’i fiziksel mağazaları, %51’i de elektronik ortamları tercih etmiştir. Elektronik ortamda alışveriş yapanların birinci tercihi ise %29’la farklı ve geniş ürün yelpazesine sahip olan platformlardır. Ürünlerin (markaların) web sitelerinde yapılan alışveriş ise sadece %19’dur. Çalışmanın önemli bir bulgusu olarak teknolojiye yönelik ilgi ve bilgi düzeyi

düşük kişilerin ürün özellikleri konusunda sorularına ayrıntılı ve net cevap alabilmelerini gerekçe göstererek fiziksel mağazaları tercih ettikleri ortaya konulmuştur.

Kaya (2022), “Elektrikli Küçük Ev Aletleri Sektöründe Markanın Tüketici Tercihlerindeki Rolü ve Tüketici Karar Sürecinde Marka Bağlılığı” isimli çalışmasında, elektrikli ev aletlerinin satış yerleri ile algılanan kaliteleri arasındaki ilişkiyi araştırmıştır. Çalışmada elektrikli küçük ev aletleri sektöründe üretici ve perakendeci firmaların faaliyetleri için marka bilinirliğinin satın alma davranışında en önemli bir kriter olduğu ortaya konmuştur.

Altunışık ve Mert (2001), perakende mağaza seçimini etkileyen temel değişkenlerin belirlenmesi üzerine yaptıkları çalışmada, mağaza içi ürün promosyonlarının ve müşteriye göre uygun fiyatlandırmanın etkili olduğu saptanmıştır.

Uluslararası literatüre bakıldığında ulusal literatüre benzer şekilde çoğunlukla tüketicilerin perakende mağaza seçimini etkileyen faktörler üzerinedir. Bu çalışmalara göre, mağaza atmosferi, mağaza imajını, müşterilerin ruh halini, çalışanın performansını, çalışanın psikolojisini, alışveriş süresini, alışveriş süresini, alışveriş tarzını, mağazada geçen süreyi, satın alma davranışını, satın alma sonrası değerlendirmeleri ve müşteri sadakatini etkilemektedir (Polat & Külter, 2007). Yine benzer şekilde uluslararası yazın çalışmalarının odak noktası çoğunlukla tüketicilerin üretici markalı ürünler arasından yapılan tercihlerinde perakendeci markalı ürünlere yönelik eğilimlerini etkileyen faktörler üzerinedir.

Cho vd. (2015), temel bileşenler analizi ve faktör analizini kullanarak, üretici firma markalarının, satıldığı perakendeci mağazalarda diğer markaların bulunmasının, müşteri davranışlarında üretici firmaya yönelik farkındalığın artmasında ve mağaza bağlılığında etkili olduğu sonucuna ulaşmıştır.

Zboja ve Voorhees (2006), çalışmalarında tüketicilerin perakendecinin mağazasında sattığı markalara yönelik olumsuz düşünceye sahip olması durumunda iyi bir hizmet alsalar bile perakendeciye yönelik memnuniyetleri ve güvenlerinin zarar görebileceğini vurgulamışlardır.

Sharma ve Gautam (2016), çoğu tüketicinin hem ürün çeşidi hem de zengin müşteri portföyü gerekçesiyle satın alma kararlarında perakende işletmelerine dikkat ettiklerini vurgulayarak perakende sektörünün öneminden bahsetmektedirler.

Beggs ve Cardell (1981), hane halkına yönelik uyguladıkları anket tekniğiyle elde ettikleri veri seti üzerinden sıralı logit analiziyle kullanarak



elektrikli araçlar için potansiyel tüketici talebini incelemişlerdir. Bu çalışmanın sonucunda elektrikli araçların daha düşük işletme maliyetli olmasına rağmen sınırlı yol kat etme ve uzun depo doldurma süresi nedeniyle tüketicilerin istekli olmadığı tespit edilmiştir.

P. Fernandez (2000), olasılık yoğunluk modeli kullanarak bireylerin dayanıklı tüketim mallarını yenileme ve değiştirmelerine ilişkin kararlarında bireylerin yaşının önemli bir faktör olduğu, istatistiksel olarak anlamlı bulunmuştur.

### 3. Metodoloji

#### 3.1. Araştırmanın Amacı

Çalışmanın amacı, Türkiye’de teknolojik ürün pazarındaki %30’luk payıyla ilk sırada olması, yaygınlığı (mağaza sayısının çokluğu), ürün çeşitliliği, hizmet kalitesi ve güvenilirlik gibi ölçütlerin dikkate alınmasıyla seçilen bir teknolojik ürün mağazası örneklemeden hareketle, iller düzeyinde ürün bazlı satış performanslarını etkileyen önemli makro çevresel değişkenlerin (demografik, ekonomik, eğitim vb.) etkisini anlayabilmek ve illerin, ürünlerin kendi içerisinde benzer ve farklı yönlerini tespit etmektir.

#### 3.2. Araştırmanın Yöntemi

Çalışmanın bu kısmında uygulamada kullanılan analizler hakkında bilgi verilecektir.

##### 3.2.1. Diskriminant Analizi

Diskriminant analizi pazarlama alanında çok sık kullanılan ve başarılı sonuçlarda elde edilebilen analiz türlerindedir. Analiz sayesinde örneğin, “farklı pazar bölümünü oluşturan müşterilerin reklama bakış açıları arasında farklılık var mıdır ya da farklı satış seviyelerindeki il grupları arasında en ayırıcı olan bağımsız değişken hangisidir?” gibi sorulara cevap bulunabilmektedir.

Diskriminant analizi, iki veya daha fazla grup arasındaki farklılıkların birden çok değişken ile ilişkisinin araştırılmasına yardımcı olan çok değişkenli istatistiksel analiz türüdür. Diskriminant analizinde gruplar (gözlemler) ve ayırt edici değişkenler önemli unsurlardır. Gruplar arası ayırma en çok etki eden değişkenleri belirlemek, hangi gruptan geldiği bilinmeyen bir birimin tahmini grubunu minimum hata ile tahmin etmek, bağımlı değişkenin varyansının ne kadarının bağımsız değişkenler tarafından açıklanabildiğini belirlemek analizin temel amaçlarından bazılarıdır. Araştırmacı, diskriminant analizine başlamadan önce, veri setinin; örneklem büyüklüğü, normal



dağılımı, varyans-kovaryans matrislerinin homojenliği, uç değer analizi, çoklu doğrusal bağlantı problemi olmaması gibi varsayımların sağladığını test etmelidir.

Diskriminant analizinde p tane özelliği bilinen gözlemler, bu özellikleri doğrultusunda gruplara ayrılmak istenirken bir diskriminant fonksiyonu kullanılır.

$Z$  : bağımlı yapıdaki kategorik değişken,

$x_1, x_2, \dots, x_p$  : bağımsız metrik yapıdaki değişkenler,

$w_1, w_2, \dots, w_p$  ; bu değişkenlere ait katsayılar,

$\alpha$  ; sabit terim olmak üzere diskriminant fonksiyonu denklem (1) ’deki gibi gösterilir:

$$Z = \alpha + w_1 x_1 + w_2 x_2 + \dots + w_k x_k \quad (1)$$

Diskriminant analizinde toplam değişim grup içi ve gruplar arası değişimlerin toplamından oluşur. Amaç grup içi kareler toplamının minimizasyon kısıtı altında parametrelerin tahmin edilmesidir.

### 3.2.2. Çok Boyutlu Ölçekleme Analizi

Çok boyutlu ölçekleme analizi (ÇBÖ), sosyal bilimler, tıp, pazarlama araştırmaları gibi birçok alanda kullanılır. Pazarlama araştırmalarında daha çok tüketici tercihi, tutumu, beklentisi gibi durumları analiz etmek amacı ile ya da pazar bölümlendirme farklı ve benzer yapıdaki kümeleri (örneğin; il, ürün, marka düzeyinde) belirlemek amacı ile kullanılır.

Çok boyutlu ölçekleme analizinin genel amacı veriyi görsel olarak keşfedilebilir yapmak, farklı nesnelere arasındaki farkı ortaya koymak, benzemezlik yargılarının altında yatan boyutları keşfetmektir. Çok boyutlu ölçekleme analizi, veriler ile ilgili dağılım varsayımı gerektirmeyen bir yöntemdir. Fakat analizde kullanılan değişkenlerin ölçümleri farklı ise (örneğin; TL, birim, kilogram gibi) veriler standartlaştırılmalıdır (Borg & Groenen, 2010). Analizde temel ayrım, metrik ve metrik olmayan modeller arasındaki farktır. Metrik ve metrik olmayan modeller, çok boyutlu ölçekleme modeli tarafından tahmin edilebilen koordinatlarla hesaplanan, veri ve uzaklıklar arasındaki ilişkiler hakkında farklı varsayımlara sahiptir. Metrik modeller ilişkinin en az aralık ölçeği ile ölçülmüş niteliklere sahip olduğunu varsayarken, metrik olmayan modeller sadece ordinal ölçekle ölçülmüş nitelikleri gerektirmektedir (Mackay & Zinnes, 1986).

ÇBÖ’de amaç olabildiğince az boyutta uzaysal haritayı bulmaktır. Boyut sayısına karar verme aşamasında araştırmacı önceki bilgilerini kullanarak, farklı

boyut durumlarında nesnelere yorumlanabilme durumlarını karşılaştırabilir ya da aşağıda belirtilen Stress değerlerine karşı olarak boyut sayısını bir grafik üzerinde analiz ederek de boyut sayısına karar verebilir. Genel olarak iki boyutlu analizler daha kolay ve anlaşılır görsel sonuçlar verebilmektedir. Analizin güvenilirliğinin ve geçerliliğinin test edilmesinde dikkat edilmesi gereken iki durum söz konusudur. Uyumun veya uyum iyiliğinin bir ölçüsü olan Stress ölçüsü analiz sonucu elde edilen koordinat düzleminin boyut sayısının uygunluğunun belirlenmesinde önemli bir ölçüttür.

Stress ölçüsü denklem (2)'ki gibidir.

$$\frac{\sum (d_{ij} - \hat{d}_{ij})^2}{\sum \hat{d}_{ij}}^{1/2} \quad (2)$$

$\hat{d}_{ij}$  terimi i. ve j. bireyler arasındaki veri uzaklığı,  $d_{ij}$  de i. ve j. bireyler arasındaki konfigürasyon uzaklığı olarak yansıtılmaktadır (Doğan, 2003).

Kruskal'ın 1964 yılında gerginlik değerinin yorumlanması için düzenlemiş olduğu tabloya göre 0,05'e eşit olan bir Stress değerinin iyi uyum, 0'a yakın bir Stress değerinin tam uyum gösterdiği söylenebilir (Wickelmaier, 2003).  $R^2$  (korelasyon endeksinin karesi), değeri, çok boyutlu ölçekleme analizi sonucunda elde edilen modelin, veri setini ne derece iyi temsil edip etmediği gösterir.  $R^2$ 'nin değerinin rakamsal olarak 0,60 ve üstü değerleri kabul edilebilir olarak yorumlanabilmektedir (Nakip, 2003).

### 3.3. Veri Toplama ve Değişken Analizi

İşletmeler, müşterilerine yönelik elde ettiği bilgileri, kişisel verilerin korunma kanunu gereği, gizlemek zorundadır. Ayrıca işletmeler kendi karlılıklarını ve sektörel paylarını gizli tutma gerekçesiyle de bu tip verileri paylaşma konusunda mesafeli yaklaşmaktadır. Çalışmanın bu kısıtına (güçlüğüne) rağmen ikincil veriler üzerinden analizler yapılmıştır. İkincil veride kullanılan kaynaklardan ilki teknolojik ürünlerin satışını yapan bir firmanın iller ve ürün düzeyinde ürün birim satış adetleridir. Kişisel verilerin korunması kanunu gereği firmanın adı çalışma boyunca belirtilemeyecektir ve çalışmada sadece analiz çıktılarına yer verilecektir. Diğer kaynak ise Türkiye İstatistik Kurumu (TÜİK)'dur. Firmanın birim satış miktarını teorik olarak etkileyeceği düşünülen demografik, ekonomik, eğitim gibi makro değişkenler TÜİK'in internet sitesinden 2019 yılı için il düzeyinde elde edilmiştir. Tüm istatistiksel analizler SPSS 25.0 kullanılarak, tablolar ise M.S. Excel programları yardımıyla düzenlenmiştir.

Firmanın satışa sunduğu sayıca çok fazla ürün çeşidi olduğu için cesametli bir birim satış adet verisi elde edilmiştir. Ayrıca bazı ürünlerin illere göre farklı satış performans göstermelerinden dolayı analiz için hangi ürün çeşidinin satış adet verilerinin kullanılacağı konusunda bir kısıt getirilmiştir. Bu kısıt belirlenirken “temel ihtiyaç ürün grubu olup, lüks ürün özelliği taşımayan” ürünler olmasına dikkat edilmiştir. Ayrıca seçilen bu ürün grupları araştırmanın yapıldığı zamana kadar devam eden ve yıllara göre benzer satış eğilime sahip ürün gruplarıdır.

Bu ürünler:

- i. Elektrikli ev aletlerinden; elektrik süpürgesi, çamaşır makinesi ve televizyon,
- ii. Bilgisayar sektöründen; tabletler,
- iii. Telekom sektöründen; akıllı telefonlardır.

Analize dahil edilen illerin her biri için, TÜİK’ in internet sitesinden, toplam nüfus sayısı, ilin okuma, yazma oranı, aldığı ve verdiği toplam göç sayısı, ildeki toplam konut satışları, imalat sanayisinin gayri safi yurt içi hasıla içindeki payı, (cari fiyatlarla, 2009 yılı düzeyinde), ildeki kişi başına gayri safi yurtiçi hasıla (cari fiyatlarla, 2009 bazlı) değerleri dikkate alınmıştır.

Veri setinde kullanılan bağımlı ve bağımsız değişkenler aşağıda belirtilmiştir. Bağımlı değişken gruplarındaki gözlem sayılarının uygunluğu ve diskriminant analizi varsayımlarından da birisi olan normallik varsayımını sağlamak amacı ile veri setindeki değişkenler normalize edilmiştir. Ayrıca diskriminant analizine başlamadan önce analiz için gerekli olan aykırı değer tespiti, kayıp değer kontrolü, örneklem yeterliliği varsayımları da test edilmiştir ve aşağıda belirtilmiştir.

- i. **Bağımlı Değişkenlerin Belirlenmesi:** Bağımlı (açıklanan) değişkenler belirlenirken, analize dahil edilen beş ürün grubu için ayrı gruplamalar yapılmıştır. Satış adeti seviyelerine göre hiyerarşik olmayan kümeleme analizi yöntemlerinden “k-ortalama” yöntemi kullanılmıştır. Veri setinde gruplar ve gözlem sayıları Tablo 1’deki gibidir.

Tablo 1. Bağımlı Değişken Grupları ve Gözlem Sayıları

Grup belirlemede kullanılan bağımlı değişken adı	Grup Adı		
	A (Yüksek Satış Seviyesi)	B (Orta Satış Seviyesi)	C (Düşük Satış Seviyesi)
Elektrik süpürgesi satış adetlerine göre	17	18	33
Çamaşır makinesi satış adetlerine göre	12	38	18
Televizyon satış adetlerine göre	20	22	26
Tablet satış adet seviyelerine göre	20	20	28
Akıllı telefon satış adet seviyelerine göre	14	19	35

Kümeler sırası ile A, B ve C olarak adlandırılmışlardır ve gruplar küme merkez noktalarına göre yüksek, orta ve düşük satış performansına sahip illeri temsil etmektedirler.

- ii. **Bağımsız Değişkenlerin Belirlenmesi:** İller düzeyinde Tablo 2’de belirtilen makro değişkenler bağımsız (açıklayıcı) değişkenler olarak analize dahil edilmişlerdir ve analiz boyunca yorumlanan çıktı tablolarında temsili değişken kodları ile ifade edilmişlerdir.

Tablo 2. Bağımsız Değişken Kodları ve Değişken Adları

Değişken kodu	Değişken adı
$X_1$	İlin toplam nüfus sayısı (yıl)
$X_2$	İlin aldığı toplam göç sayısı
$X_3$	İlin verdiği toplam göç sayısı
$X_4$	İldeki toplam konut satışları
$X_5$	Sektörlerin Gayri Safi Yurt İçi Hasıla İçindeki Payı, İmalat Sanayi, Cari fiyatlarla, 2009 Bazlı
$X_6$	İldeki kişi başına gayri safi yurtiçi hasıla (Cari fiyatlarla, 2009 Bazlı)
$X_7$	İlin okuma, yazma oranı (nüfusa oranlı)

- iii. **Aykırı değer tespiti:** Aykırı değerlerin tespitinde çok değişkenli veri setlerinde sıklıkla kullanılan Mahalanobis Uzaklık (MU) ölçüsü kullanılmıştır. Analiz sonucu elde edilen MU değerleri incelendiğinde 0,001’den küçük olan ve aykırı değere sahip iller Tablo 3’te belirtilmiştir.

Tablo 3. Aykırı Değerlerin Mahalanobis Uzaklık Değerleri

İl Adı	Mahalanobis Uzaklık Değerleri
ADANA	0,00051
ANKARA	0,00000
İSTANBUL	0,00000
KOCAELİ	0,00094

Bu durum illerin gelişmişlik düzeyinden kaynaklanan bir durumdur. Bu illeri analiz dışında tutmanın çok doğru bir tutum olmayacağı düşünülerek değerlerin analize olan etkilerini kısıtlamak amacıyla literatürde kırpmaya yöntemi olarak da bilinen Winsorization yöntemi ile veri setindeki değerler %90’lık bir kazanç sağlanacak şekilde kırılmıştır ve bu şekilde veri setindeki aykırı değer sorununa çözüm üretilmeye çalışılmıştır.

- iv. **Kayıp değer kontrolü:** Firmadan elde edilen ilgili iller ve ürünlerin satış adet verilerinde ve TÜİK’ten alınan il düzeyindeki makro verilerde herhangi bir kayıp değer bulunmamaktadır.
- v. **Örneklem Yeterliliği:** Veri setinin örneklem yeterliliğinde kullanılan Kaiser Meyer Olkin (KMO) test oranının 0,80’in üstünde çıkması iyi bir örneklem uygunluğuna işaret eder. Örneklem yeterliliği için test ettiğimiz veri setinin Tablo 4’e göre KMO test istatistik değeri 0,87’dir ve veri setinin örneklem yeterliliğinin uygun olduğunu söylemek mümkündür.

Tablo 4. Örneklem Yeterliliği Testi

KMO ve Bartlett Testi		
Kaiser-Meyer-Olkin Örneklem Yeterliliği Ölçümü		0,878
Bartlett’in Küresellik Testi	Yaklaşık Ki-kare	1471,458
	df	91
	Anlamlılık	0,000

## 4. Bulgular

### 4.1. Diskriminant Analizine İlişkin Bulgular

Çalışma kapsamında seçilen beş ürün için ayrı ayrı diskriminant analizi yapılmıştır. Ancak, çalışmanın boyutunu hacimsel olarak aşmamak adına, bu ürün gruplarından sadece bir ürünün (elektrik süpürgesi) diskriminant analizi çıktıları detaylı yorumlanmıştır. Diğer ürün gruplarındaki çıktıların

detaylı bir yorumu verilmeyip, analiz sonucu elde edilen “Wilks’ Lambda” ve “Özdeğer” büyüklüklerine ilişkin sonuçlar tek bir tabloda sunulmuştur. Her bir ürün için sınıflandırma sonuçlarına yönelik tablolar ayrı ayrı belirtilmiştir.

Elektrik süpürgesi için yapılan diskriminant analiz sonucunda elde edilen ilk çıktılardan birisi olan Wilks’ Lambda’nın F testi tablosu, bağımsız değişkenler açısından gruplar arasında bir farkın olup olmadığını anlaşılmasında bilgi verecektir. Bu teste ilişkin hipotez aşağıdaki gibi belirtilmiştir

$$H_0 : \mu_{uX_1} = \mu_{uX_2} = \dots = \mu_{uX_7}$$

$$H_1 : \mu_{uX_1} \neq \mu_{uX_2} \neq \dots \neq \mu_{uX_7}$$

$H_0$  gruplar arası değişkenlerin ortalamalarının eşit olduğunu,  $H_1$  ise gruplar arası değişkenlerin ortalamalarının farklı olduğunu göstermektedir. Tablo 5’e göre Wilks’ Lambda’nın F Testi anlamlılık değeri  $x_7$  değişkeni için 0,05’ten büyüktür ve  $H_0$  hipotezi  $x_7$  değişkeni için reddedilememiştir. Diğer değişkenler için ise anlamlılık değerleri  $p = 0,000 < 0,05$  olduğu için  $H_0$  hipotezi reddedilebilir ve gruplar arasında ortalamaların farklı olduğu şeklinde yorumlanabilir.

*Tablo 5. Elektrik Süpürgesi için Wilks’ Lambda’nın F Testi*

	Wilks’ Lambda	F	df <sub>1</sub>	df <sub>2</sub>	Anlamlılık
$X_1$	0,335	64,56	2	65	0,000
$X_2$	0,297	76,9	2	65	0,000
$X_3$	0,405	47811	2	65	0,000
$X_4$	0,289	80,08	2	65	0,000
$X_5$	0,914	3044	2	65	0,055
$X_6$	0,777	9304	2	65	0,000
$X_7$	0,94	2077	2	65	0,134

Analiz sonucunda elde edilen diğer bir çıktı olan Box’s M İstatistiği tablosu gruplar arası varyans-kovaryans matrislerinin eşitliğinin anlaşılmasında yardımcı olacaktır.

$H_0$  : Grupların varyans-kovaryans matrisleri homojendir.

$H_1$  : Grupların varyans-kovaryans matrisleri homojen değildir;

şeklinde kurulabilen hipotez Tablo 6’ya göre anlamlılık değeri  $p = 0,000 < \alpha = 0,05$  olduğuna göre grupların varyans-kovaryans matrislerinin homojen olduğunu söyleyen sıfır hipotezi tüm anlamlılık düzeylerinde reddedilmektedir.

**Tablo 6. Elektrik Süpürgesi için Box’s M İstatistiği**

<b>Box’s M</b>		137,326
<b>F</b>	Yaklaşık	2,02
	df <sub>1</sub>	56
	df <sub>2</sub>	7211,163
	Anlamlılık	0,000

Grupların varyans-kovaryans matrislerinin homojen olması durumu diskriminant analizine devam edebilmek için başlıca sağlanması gereken varsayımlarından birisidir ve varsayım sağlanamamıştır. Ancak bu noktada Lachenbruch’un ileri sürdüğü eşit kovaryans varsayımlarının ihlalinin diskriminant analizi sonuçlarını önemli ölçüde etkilemediği yaklaşımından hareketle analize devam edilmiştir (Berberoğlu, 2010).

Diskriminant analizinde varsayılması gereken bir diğer koşul da çoklu doğrusal bağlantı sorunu olmaması durumudur. “Çoklu doğrusal bağlantı değişkenler arasında 0,90 ve üzeri korelasyon olduğunda, tekillik ise, değişkenlerden bir veya birkaçı gereksiz olduğunda ortaya çıkar” (Güzeller, 2016). Tablo 7’ye göre gruplar içi korelasyon matrisinde bakıldığında değişkenler arasında 0,90 sınırına göre kabul edilebilir seviyelerde bir korelasyon ilişkisi bulunmaktadır şeklinde yorum yapılabilir.

**Tablo 7. Elektrik Süpürgesi için Gruplar İçi Değişkenler Arası Korelasyon Matrisi**

<b>Değişkenler</b>	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
$X_1$	1,00	0,884	0,872	0,799	0,083	-0,167	-0,276
$X_2$	0,884	1,00	0,875	0,798	0,073	0,054	-0,121
$X_3$	0,872	0,875	1,00	0,652	-0,092	-0,281	-0,235
$X_4$	0,799	0,798	0,652	1,00	0,409	0,251	-0,003
$X_5$	0,083	0,073	-0,092	0,409	1,00	0,496	0,368
$X_6$	-0,167	0,054	-0,281	0,251	0,496	1,00	0,512
$X_7$	-0,276	-0,121	-0,235	-0,003	0,368	0,512	1,00

Elektrik süpürgesi için Tablo 8’de görüldüğü üzere satış adet seviyelerine göre üç grup söz konusu olduğu için iki fonksiyon oluşturulmuştur. Tabloda her iki fonksiyonun değişkenliği açıklama yüzdesi ve kanonik korelasyon değerleri verilmiştir. Birinci fonksiyonun kanonik korelasyon değeri 0,86’dır ve bağımlı değişkendeki varyansın %73’ü açıklanmaktadır. İkinci fonksiyon için varyansın açıklama oranı ise %7’dir. Birinci fonksiyonun daha yüksek değerlerde açıklayıcı olması analizin devamında tercih edileceğinin bir göstergesidir.

Tablo 8. Elektrik Süpürgesi için Kanonik Ayırma Fonksiyonları (Özdeğerler)

Fonksiyon	Özdeğer	Varyans %	Kümülatif %	Kanonik Korelasyon
1	3,027	97,2	97,2	0,867
2	0,088	2,8	100,0	0,284

Tercih edilen fonksiyonun gruplara göre önem katsayıları Tablo 9'daki gibidir.

Tablo 9. Elektrik Süpürgesi için Fisher'in doğrusal ayırt edici fonksiyonları

	A	B	C
$X_1$	837,059	825,775	828,097
$X_2$	-807,438	-809,294	-815,567
$X_3$	591,706	596,719	597,489
$X_4$	-413,608	-416,103	-423,612
$X_5$	-96,576	-92,54	-92,102
$X_6$	1031,577	1020,613	1024,364
$X_7$	8173,174	8058,955	7981,905
<b>Sabit</b>	-4632,385	-4490,79	-4447,566

Tablo 9'daki katsayılar kullanılarak elektrik süpürgesi ürünü için grupların doğrusal ayırt edici tahmin Diskriminant Fonksiyonları (DF) denklem (3) – (4) – (5)'teki gibi yazılır.

$$DF_{(GRUP A)} = -4632,385 + 837,059x_1 - 807,438x_2 + 591,706x_3 - 413,608x_4 - 96,576x_5 + 1031,577x_6 + 8173,174x_7 \quad (3)$$

$$DF_{(GRUP B)} = -4490,790 + 825,775x_1 - 809,294x_2 + 596,719x_3 - 416,103x_4 - 92,540x_5 + 1020,613x_6 + 8058,955x_7 \quad (4)$$

$$DF_{(GRUP C)} = -4447,566 + 828,097x_1 - 815,567x_2 + 597,489x_3 - 423,612x_4 - 92,102x_5 + 1024,364x_6 + 7981,905x_7 \quad (5)$$

Elektrik süpürgesi için yapılan diskriminant analizinde son çıktı olan sınıflandırma sonuçları Tablo 10'daki gibidir. Tablo 10'a göre analizde kullanılan veride 17 tane A, 18 tane B ve 33 tane C grubu olmak üzere toplamda 68 il vardır. Oluşturulan diskriminant fonksiyonunun doğru sınıflandırılma oranları A, B, C grupları için sırası ile %82, %91 ve %56'dır. Modelin bağımsız değişkenler ile toplam sınıflandırma öngörü gücü ise %79'dur.



Tablo 10. Elektrik Süpürgesi için Tahmin Edilmiş Grupların Sınıflandırma Sonuçları

		Tahmini grup üyeliği			Toplam
		A	B	C	
Miktar	A	14	3	0	17
	B	2	10	6	18
	C	0	3	30	33
Orijinal	A	82,4	17,6	0	100
	B	11,1	55,6	33,3	100
	C	0	9,10	90,9	100

Toplam Sınıflandırma Başarısı (54/68) %79

Tablo 10’a göre A yüksek, B orta ve C düşük satış seviye gruplarını göstermek üzere; tahmin edilmiş grupların sınıflandırma sonuçları tablosundaki sayısal verilerinin Türkiye haritasındaki görünümü Şekil 1’deki gibidir.



Şekil 1. Elektrik Süpürgesi için Tahminlenmiş Grupların Türkiye Haritası Üzerinde Görünümü

Seçili beş ürün grubu için diskriminant analizi ayrı ayrı tekrarlanmıştır ve satış adetlerine göre oluşturulan diskriminant fonksiyonlarının geçerlilik değerlerinde kullanılan Wilks’ Lambda ve özdeğer büyüklükleri özete Tablo 11’deki gibidir. Fonksiyonların öz değerlerinin 0,40 ve yukarısında olması, Wilks’ Lambda değerinin küçük, ki-kare değerinin yüksek, anlamlılık değerlerinin 0,005 anlamlılık düzeyinde küçük olması fonksiyonların etkili bir ayrıştırma gücüne sahip olduğunu gösterir.

*Tablo 11. Seçili Beş Ürün Grubu için Wilks' Lambda ve Özdeğer Büyüklükleri*

	Wilks' Lambda	Özdeğer	Ki-kare	Anlamlılık
Elektrik süpürgesi	0,228	3,027	91,584	0,000
Çamaşır makinesi	0,343	1,583	66,382	0,000
Televizyon	0,262	2,622	83,13	0,000
Tablet	0,167	4,261	110,949	0,000
Akıllı telefon	0,237	2,775	89,328	0,000

Çalışmanın bu aşamasından sonra diğer dört ürün için etkili ayrıştırma gücüne sahip olan fonksiyonlar yardımı ile tahmin edilmiş grupların sınıflandırma sonuç tabloları ve Türkiye haritası üzerindeki görünümleri paylaşılmıştır.

Çamaşır makinesi için diskriminant analizinin sınıflandırma sonuçları Tablo 12'de yer almakta olup, sınıflandırma sonuçları sayısal verilerinin Türkiye haritasındaki görünümü de Şekil 2'deki gibidir.

*Tablo 12. Çamaşır Makinesi için Tahmin Edilmiş Grupların Sınıflandırma Sonuçları*

		Tahmini Grup Üyeliği			Toplam	
		A	B	C		
Orijinal	Miktar	A	10	2	0	12
		B	1	34	3	38
		C	0	11	7	18
	Yüzde	A	83,3	16,7	0	100
		B	2,60	89,5	7,90	100
		C	0	61,1	38,9	100
Toplam Sınıflandırma Başarısı (51/68) %75						



Şekil 2. Çamaşır Makinesi için Tahmin Edilmiş Grupların Türkiye Haritası Üzerinde Görünümü

Televizyon için diskriminant analizinin sınıflandırma sonuçları Tablo 13’te yer almakta olup, sınıflandırma sonuçları sayısal verilerinin Türkiye haritasındaki görünümü de Şekil 3’teki gibidir.

Tablo 13. Televizyon için Tahmin Edilmiş Grupların Sınıflandırma Sonuçları

		Tahmini Grup Üyeliği			Toplam	
		A	B	C		
Orijinal	Miktar	A	18	2	0	20
		B	4	13	5	22
		C	0	6	20	26
	Yüzde	A	90	10	0	100
		B	18,2	59,1	22,7	100
		C	0	23,1	76,9	100
Toplam Sınıflandırma Başarısı (51/68) %75						



Şekil 3. Televizyon için Tahminlenmiş Grupların Türkiye Haritası Üzerinde Görünümü

Tablet için diskriminant analizinin sınıflandırma sonuçları Tablo 14'te yer almakta olup, sınıflandırma sonuçları sayısal verilerinin Türkiye haritasındaki görünümü de Şekil 4'teki gibidir.

Tablo 14. Tablet için Tahmin Edilmiş Grupların Sınıflandırma Sonuçları

		Tahmini Grup Üyeliği			Toplam	
		A	B	C		
Orijinal	Miktar	A	19	1	0	20
		B	1	13	6	20
		C	0	4	24	28
	Yüzde	A	95	5	0	100
		B	5	65	30	100
		C	0	14,3	85,7	100
Toplam Sınıflandırma Başarısı (56/68) %82						



Şekil 4. Tablet için Tahminlenmiş Grupların Türkiye Haritası Üzerinde Görünümü

Akıllı telefon için diskriminant analizinin sınıflandırma sonuçları Tablo 15’te yer almakta olup, sınıflandırma sonuçları sayısal verilerinin Türkiye haritasındaki görünümü de Şekil 5’teki gibidir.

Tablo 15. Akıllı Telefon için Tahmin Edilmiş Grupların Sınıflandırma Sonuçları

		Tahmini Grup Üyeliği			Toplam	
		Grup	A	B	C	
Orijinal	Miktar	A	12	2	0	14
		B	0	13	6	19
		C	0	1	34	35
	Yüzde	A	85,7	14,3	0	100
		B	0	68,4	31,6	100
		C	0	2,9	97,1	100
Toplam Sınıflandırma Başarısı (59/68) %86						



Şekil 5. Akıllı Telefon için Tahminlenmiş Grupların Türkiye Haritası Üzerinde Görünümü

## 4.2. Çok Boyutlu Ölçkleme Analizine İlişkin Bulgular

Çok boyutlu ölçkleme analizi, iller ve ürünler düzeyinde olmak üzere iki ayrı şekilde yapılmıştır.

### 4.2.1. İller için Çok Boyutlu Ölçkleme Analizine İlişkin Bulgular

İller için yapılan çok boyutlu ölçkleme analizi ile üründen bağımsız olarak iller seviyesinde toplam satış adetleri dikkate alınarak iller arasındaki benzerlik ve farklılıklar iki boyutlu grafik ile tespit edilmeye çalışılmıştır. Özetle veri setinde 68 gözlem (il) bulunmaktadır ve analizde uzaklık matrislerinden Öklid uzaklık matrisi kullanılmıştır.

ÇBÖ analizinde elde edilen çıktılarının geçerliliği ve güvenilirliği için Stress değerleri ve  $R^2$  için tablolara bakacak olursak;  $k$  boyut sayısını temsil etmek üzere,  $k = 2$  için stress istatistiğinin 0,001'den küçük olduğu değere kadar analiz devam etmiştir. Tablo 16'ya göre 3. iterasyonun sonunda Stress değeri 0,0007 sonucuna ulaştığı için iterasyon durmuştur. 0'a yakın Stress değeri veren boyut çözümü uygun olarak yorumlanmaktadır.

Tablo 16. 68 İlin Çok Boyutlu Ölçkleme Analizi Stress Değerleri

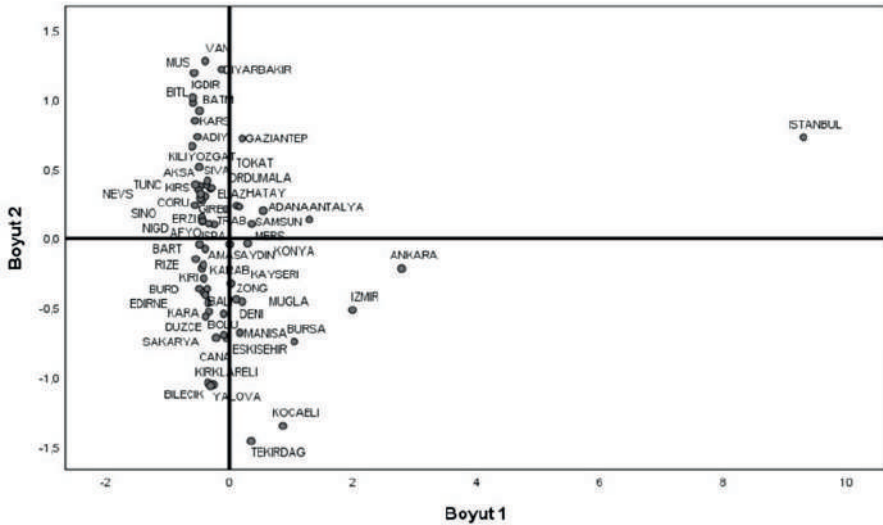
İterasyon	Uyum (Stress)	Gelişme
1	0,03965	
2	0,03111	0,0085
3	0,03041	0,0007

Ayrıca yine Tablo 17’de görüldüğü üzere  $k = 2$  için Stress değeri verileri  $R^2 = 0,98306$  oranında açıklamaktadır.

Tablo 17. 68 İlin Çok Boyutlu Ölçekleme Analizi Kruskal’s Stress İstatistiği

$$R^2 = 0,98306$$

Analiz sonucu illerin iki boyutlu uzayda oluşan grafiksel gösterimi Şekil 6’daki gibidir. Grafikte  $y$  eksenini birinci boyut,  $x$  eksenini ikinci boyut olarak tanımlanmaktadır.



Şekil 6. 68 İl İçin Öklid Mesafesi Modeli

İki boyutlu grafikte merkeze (0,0) yaklaştıkça noktalar arasındaki farklılıklar azalırken, merkezden uzaklaştıkça farklılıklar artmaktadır. Grafike bakıldığında en farklı toplam satış adeti değerine sahip olan ilin İstanbul olduğu söylenebilir. Merkez etrafında ve birbirine yakın konumlanan illerin de satış verilerinin birbirlerine daha benzer olduğu söylenebilir.

İlleri boyut değerleri arasındaki farkın minimum, maksimum ve ekstrem olması durumuna göre üç gruba ayırırsak Şekil 7’deki gibi renklendirilmiş bir Türkiye haritası elde edebiliriz. Aslında bu harita Şekil 6’daki konumlanma

yapısındaki benzerlik ve farklılıkları da görsel olarak daha iyi bir şekilde sunmuştur.



**Şekil 7. 68 İlin Boyut 1 Değerlerine Göre Üç Gruba Ayrılmış Halinin Türkiye Haritasında Görünümü**

İstanbul satış adeti bakımından yüksek performanslı bir il olduğundan dolayı İstanbul dışında kalan iller arasında benzerlik ve farklılıkların daha net anlaşılabilmesi için analiz bir de İstanbul ili çıkarılıp tekrarlanmıştır. Analizin geçerliliği için dikkate alınması gereken ve oldukça uygun çıkan Stress ve güvenilirlik oranları Tablo 18 ve 19'daki gibidir.

**Tablo 18. İstanbul'un Analize Dabil Edilmediği 67 İlin Çok Boyutlu Ölçekleme Analizi Stress Değerleri**

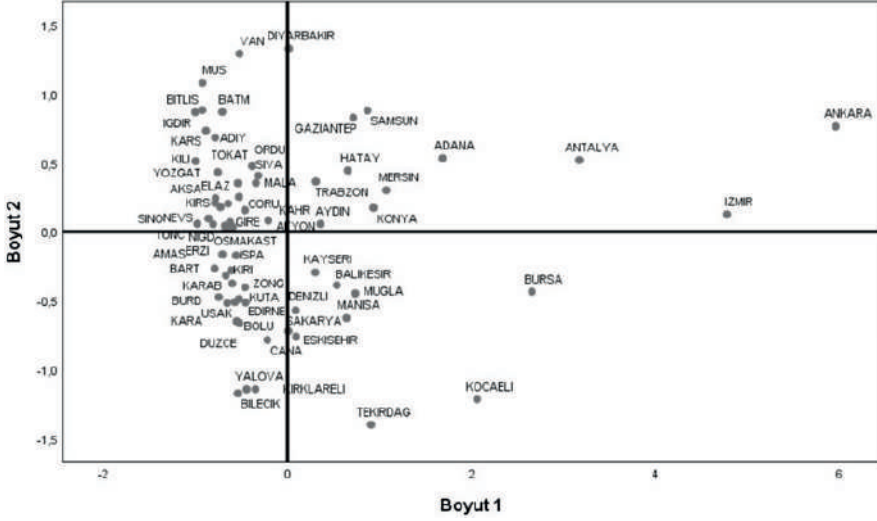
İterasyon	Uyum (Stress)	Gelişme
1	0,0812	
2	0,0655	0,01576
3	0,0641	0,00138
4	0,0640	0,00005

**Tablo 19. İstanbul'un Analize Dabil Edilmediği 67 İlin Çok Boyutlu Ölçekleme Analizi Kruskal's Stress İstatistiği**

$$R^2 = 0,97794$$



İstanbul’un analize dahil edilmediği durum için elde edilen Şekil 8’deki grafiğe bakıldığında Şekil 6’dan farklı bir konumlanma gözlemlenmiştir.



Şekil 8. İstanbul’un analize dahil edilmediği 67 İlin Öklid Mesafesi Modeli

Benzer şekilde yine boyut 1 değerleri kullanılarak illerin üç gruba ayrılmış hali Şekil 9’daki gibidir.



Şekil 9. İstanbul’un Analize Dahil Edilmediği 67 İlin Boyut 1 Değerlerine Göre Üç Gruba Ayrılmış Halinin Türkiye Haritasında Görünümü

#### 4.2.2. Ürünler için Çok Boyutlu Ölçekleme Analizi Sonuçları

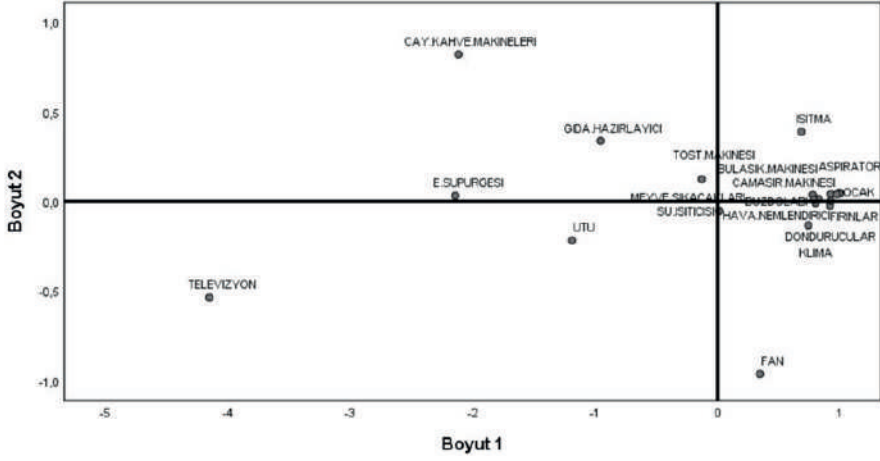
Ürünler için çok boyutlu ölçekleme analizinde firmanın satışa sunduğu tüm teknolojik ürünlerin 2019 yılı toplam satış verileri üzerinden iki boyutlu grafikte konumlanmaları incelenip ürünler arasında işlev benzerliği ve toplam satış adetleri açısından bir bağlantı kurulmaya çalışılmıştır.

Ulaşılabilen veri setinde fazla sayıda farklı sektörlerden ürün grupları olduğu için tek bir gruplama ile analiz çıktısındaki görseller karmaşık bir hal almaktaydı. Veri setindeki ürün grupları sektörel olarak elektrikli ev aletleri, bilgisayar ve telekom ürün grupları olmak üzere üç grupta sınıflanabilirdi. Gelişen teknoloji çağındaki değişimler ile bilgisayarların yaptığı pek çok işlevleri artık telefonlar ile çok daha kısa sürede daha pratik yapılabilen durumlar göz önüne alınarak telefon ve bilgisayar sektörleri altındaki benzer ya da farklı ürünlerdeki konumlanmalar merak edilmiş olup tek grup altında incelenmiştir. Bu noktada analiz öncesi ürünler iki ayrı grup olarak ele alınmıştır ve gruplardaki ürünler aşağıda gibidir.

**Elektrikli ev aletleri (Grup 1):** Klima, hava nemlendiricileri, ocak, buzdolabı, bulaşık makinesi, ısıtma, fan, gıda hazırlayıcıları, dondurucular, aspiratör, çay ve kahve makineleri, ütü, meyve sıkacakları, su ısıtıcıları, mikrodalga fırınlar, televizyon, tost makinesi, kurutucular, elektrikli süpürgeler, çamaşır makineleri.

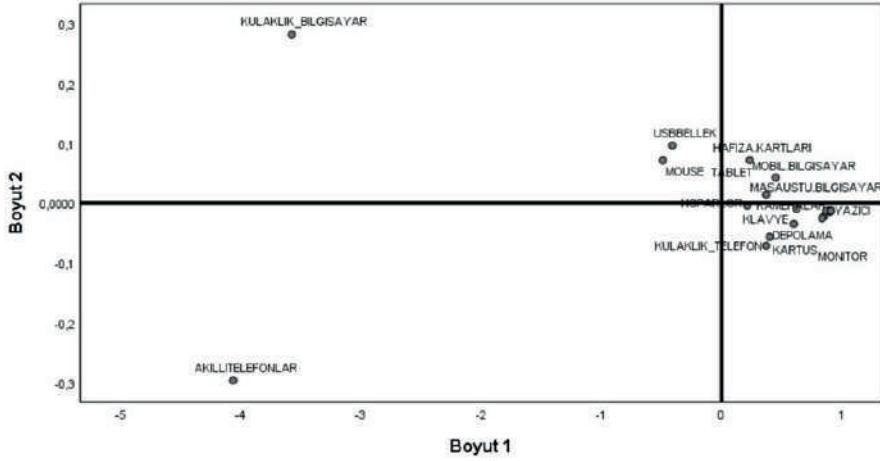
**Bilgisayar ve telekom ürünleri (Grup 2):** Kameralar, masa üstü bilgisayar, telefon kulaklıkları, bilgisayar kulaklıkları, klavyeler, kartuş, tablet, hafıza kartları, hoparlör, mobil bilgisayarlar, monitörler, akıllı telefonlar, mouse, yazıcı, depolama, usb bellek, oyun konsolları, video oynatıcıları.

Şekil 10'da, Grup 1 için Öklid mesafesi modeli gösterilmektedir. Şekle göre televizyon ürün grubu diğer ürünlerden çok farklı bir noktada konumlanmıştır ve en farklı satış adeti seviyesine sahip olan üründür şeklinde yorumlanabilir. Ayrıca ocak ve fırın gibi işlevsel açıdan benzerliği olan ürünlerin yakın konumlanmaları da benzer satış adetlerine sahip olmalarının ve dolayısı ile birlikte satın alınmalarının bir açıklaması olabilir.



Şekil 10. Grup 1 İçin Öklid Mesafesi Modeli

Şekil 11, Grup 2 için Öklid mesafesi modeli gösterilmektedir. Grup 2’de en farklı konumlanan iki ürün akıllı telefonlar ve bilgisayar kulaklıklarıdır. Bilgisayar ve tamamlayıcı ürünleri olarak düşünülebilen masa üstü bilgisayar, klavye, yazıcı, kartuş, gibi genelde birlikte satılan ürünler birbirlerine yakın bir şekilde konumlanmışlardır.



Şekil 11. Grup 2 İçin Öklid Mesafesi Modeli

## 5. Tartışma ve Sonuç

Bu çalışmada ilk olarak, teknolojik ürünlerin perakendeci satışlarında makro çevresel değişkenlerin, iller düzeyinde satış performansı üzerindeki etkilerini saptamak amacı ile diskriminant analizi yapılmıştır.

Diskriminant analizi sonuçlarına göre seçili her bir ürün için makro çevresel (toplam nüfus, aldığı ve verdiği göç sayısı, okuma yazma oranı, imalat sanayi sektörünün gayri safi yurt içi hasıla içindeki payı, ildeki toplam konut satışları ve kişi başına gayr safi yurt içi hasıla) değişkenlerin il düzeyinde satış performansını etkileyebildiği sonucuna ulaşılmıştır. Aynı zamanda diskriminant analizi ile oluşturulan fonksiyonların da anlamlı ve etkili bir ayırıştırma gücüne sahip olduğu gözlenmiştir.

Çalışmada ek olarak hem illerin hem de ürünlerin kendi aralarındaki benzerlik ve farklılıkları görsel olarak daha anlaşılır hale getirmek amacı ile çok boyutlu ölçekleme analizi yapılmıştır. İllerin kendi aralarındaki benzerlik ve farklılıkların tespitinde ise toplam satış adetine göre İstanbul ilinin diğer illerden farklılaştığı gözlemlenmiştir. Diğer iller ise birbirine daha yakın konumlanmıştır. Bu illerin satış adetlerinin birbirlerine daha benzer olduğu söylenebilir. İstanbul'un analize dahil edilmediği durumda ise sonuç dikkat çekicidir. Antalya, Bursa, Kocaeli, Tekirdağ gibi illerin diğer illerden farklılaştığı gözlemlenmiştir. Ürünlerin çok boyutlu ölçekleme analizi sonuçlarında ise işlevsel açıdan birbirine benzeyen, tamamlayıcı ürün olarak da düşünülebilen, genelde beraber satın alınan ürünlerin konumlanmalarındaki yakınlık dikkat çekmiştir. Örneğin; grup 1'in içinde yer alan ürünler açısından açıklamak gerekirse televizyon ürün grubu diğer ürünlerden çok farklı bir noktada konumlanmıştır ve en farklı satış adeti seviyesine sahip olan üründür şeklinde yorumlanabilir. Ayrıca ocak ve fırın gibi işlevsel açıdan benzerliği olan ürünlerin yakın konumlanmaları da benzer satış adetlerine sahip olmalarının ve dolayısı ile birlikte satın alınmalarının bir açıklaması olabilir.

Bu sektörde faaliyet gösteren işletmelerin etkin pazarlama faaliyeti gerçekleştirebilmesi için il ve ürün gruplandırılmasının doğru tespit edilmesine yönelik olarak çalışmada elde edilen bulgulara göre aşağıda sunulan önerilerin rehber olacağı düşünülmektedir.

- i. *Diskriminant analizi ile etkili çıkan fonksiyonlarda satış adet seviyelerini önemli derecede etkileyen katsayısı yüksek olan makro değişkenler takip edilerek hangi dönemlerde pozitif yönde bir eğilim gösterdikleri tespit edilebilir. İlgili dönemler için satış kampanyaları ve reklamlar ile müşteriyi etkileyecek stratejiler geliştirilebilir.*

- ii. *Elde edilen fonksiyonlar ile ilin tahmini satış adet seviyesi belirlenebilir ve bu yönde ilgili gruplarda bulunan illerde satış adetlerinde olumlu yönde fark yaratabilecek ürün grupları üzerine tedarik sağlanabilir.*
- iii. *Firma kendi iç pazarında il bazlı bir bölümlenme yapmak isterse çok boyutlu ölçekleme analizi ile benzer ve farklı illeri belirleyebilir.*
- iv. *Ürünlerin benzer ve farklı yönlerini çok boyutlu ölçekleme analizi ile belirleyip tamamlayıcı olarak gözlemlenen ürün grupları için kampanyalar, reklamlar geliştirerek satış performanslarını arttırabilir veya teknolojik yeni gelişmeler ile ürün özelliklerinde farklılıklar yaratarak daha çok müşterinin dikkatini çekebilir.*

Kullanılan bu analiz yöntemleri, bundan sonraki yapılacak akademik çalışmalarda markalar arasındaki benzerlik ve farklılıkların tespiti üzerine çalışmalarda değerlendirilebilir.

## Kaynakça

- Acar, N., Çizmeçi, B. (2015), Factors Influencing Customer's Choice of Technology Retailers: An Application in Kayseri (Turkey), *Procedia - Social and Behavioral Sciences*, 20(7)
- Altunışık, R., & Mert, K. (2001). Tüketicilerin Alışveriş Merkezlerindeki Satınalma Davranışları Üzerine Bir Saha Çalışması: Tüketiciler Kontrolü Yitiriyor Mu? 6. *Ulusal Pazarlama Kongresi*. Erzurum.
- Arıkbay, C. (1996). *Perakendecilikte Gelişmeler ve Yeni Yaklaşımlar*. Ankara: Milli Prodüktivite Merkezi Yayınları.
- Avlonitis, G. J., Hart, S. J., & Tzokas, N. (2000, January). *An Analysis of Product Deletion Scenarios*. Oxford University Press.
- Berberoğlu, B. (2010). Yaşam Boyu Öğrenme İle Bilgi Ve İletişim Teknolojilerin Açısından Türkiye nin Avrupa Birliği ndeki Konumu. *Bilgi Ekonomisi ve Yönetimi Dergisi*, 5(2), 118. <https://dergipark.org.tr/tr/pub/beyder/issue/3477/47311> adresinden alındı
- Borg, I., & Groenen, P. (2010). *Modern multidimensional scaling : theory and applications*. Springer.
- Catalin, M. C., & ANDREEA, P. (2014, January). Brands as a Mean of Consumer Self-expression and Desired Personal Lifestyle. *Second World Conference on Business, Economics and Management, Procedia* (s. 103-107). Romania: Social and Behavioral Sciences. doi:DOI:10.1016/j.sbspro.2013.12.427
- Cemal, Z., & Özşahin, M. (2006). Stratejik karar verme hızını etkileyen örgütsel, çevresel faktörler ve firma performansı ilişkisi: imalat sektöründe bir saha çalışması. *Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 9(1), 137-157.
- Cengiz, E., & Özden, B. (2002, Mayıs 01). Perakendecilikte büyük alışveriş merkezleri ve tüketicilerin büyük alışveriş merkezleri ile ilgili tutumlarını tespit etmeye yönelik bir araştırma. *Ege Akademik Bakış Dergisi*, 2(1), 1-15. <https://dergipark.org.tr/tr/pub/eab/issue/39832/472236> adresinden alındı
- Değer, A. (2010, Nisan). Yeni Bir Ürün Geliştirmede Ve Ürünün Tanıtılmasında Uygulanan Stratejiler Ve İlaç Sektöründe Bir Uygulama. İstanbul: İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü. <https://tez.yok.gov.tr/UlusalTezMerkezi/tezDetay.jsp?id=9q1G0Wfe-CwvT45cZ-freA&no=z7CXAYkao117gA2cpjDFVw> adresinden alındı
- Ellialtı, Y. (2009). Ürün Özelliklerde, Görece Ürün Kalitesi Ve Tüketici Etnosentrizminin Yerli Ürün Satın Alma Eğilimine Etkisi: Kozmetik Sektöründe Bir Uygulama. Marmara Üniversitesi Sosyal Bilimler Enstitüsü.

- Gundlach, G. T., & Wilkie, W. L. (2009, September). The American Marketing Association’s New Definition of Marketing. *Journal of Public Policy & Marketing*, 28(2), 259-264. doi:https://doi.org/10.1509/jppm.28.2.259
- Güzeller, C. O. (2016). *Herkes İçin Çok Değişkenli İstatistik*. Maya akademi Yayınları.
- Hakan, K., & Bökeoğlu, Ö. (2017). Farklı örneklem büyüklüklerinde uç değerlerle baş etme yöntemlerinin puanların geçerlik ve güvenilirlik kanıtları üzerindeki etkisi. Pegem Atıf İndeksi.
- Harrell, G. (2002). The relationship between consumer characteristics and attitude toward online shopping. *Marketing Intelligence & Planning*, 21(1), 266.
- Harrell, G. (2003). *Marketing: Connecting with customers*. Prentice-Hall.
- Harun, S. (2014). Konjoint analizi tekniğinin pazarlama araştırmalarında kullanım olanakları ve bir uygulama. Anadolu University.
- Işıl, Ş. (2007). Pazarlama araştırmalarında bazı çok değişkenli istatistik yöntemler ve beyaz eşya sektörüne uygulaması. Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü.
- Kalelioğlu, E. (2021). *Webtekno*. Webtekno: <https://www.webtekno.com/turkiye-2020-yili-kategorilere-gore-satilan-teknolojik-urun-sayisi-h107251.html> adresinden alındı
- Kavak, B., & Sığındı, T. (2012). PAZARLAMA'DAKİ ÜRÜN SINIFLANDIRMASINA. *H.Ü. İktisadi ve İdari Bilimler Fakültesi Dergisi*, 30(1).
- Kerin, R. A., Berkowitz, E. N., Hartley, S. W., & Rudelius, W. (2003). *Marketing* (Cilt 7). New York: McGraw-Hill.
- Kırcova, İ. (2018). *Pazarlama Yönetimi*. Beta Yayınevi.
- Kotler, Keller, K. L., Goodman, M., Brady, M., & Hansen, T. (2019). *Marketing management*. Pearson Education.
- Kotler, P. (2003). *Marketing Management*. Prentice-Hall.
- Kotler, P., & Armstrong, G. (1994). *Principles of Marketing*. Pearson Education.
- Kotler, P., & Armstrong, G. (2017). *Principles of Marketing*. Pearson Education.
- Kotler, P., & Keller, K. L. (2009). *Marketing Management*. Pearson Prentice Hall.
- Lamp, C., Boivin, M., Hair, J., Gaudet, D., Mcdaniel, C., & Snow, K. (2020). *MKTG : principles of marketing*. Kendall Hunt.
- Mackay, D., & Zinnes, J. (1986). *Marketing Engineering: Computer*. Marketing Science.
- Murat, E. (2000). Yüksek teknoloji ürünleri pazarlaması. *Öneri Dergisi*, 129-136. doi:https://doi.org/10.14783/maruoneri.734129

- Nakip, M. (2003). *Pazarlama Araştırmaları Teknikler ve (SPSS Destekli) Uygulamalar*: Seçkin Yayıncılık.
- Polat, C., & Külter, B. (2007). TÜKETİCİLERİN PERAKENDE MAĞAZA TERCİHİNİ ETKİLEYEN FAKTÖRLER: NİĞDE İLİ ÖRNEĞİ. *Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 9(3), 110.
- Rosenbloom , B. (2004). *Marketing Channels: A Management View* (Cilt 7). South-Western.
- Schoell, W. F., & Guiltinan, J. P. (1995). *Marketing: Contemporary Concepts and Practices* (Cilt 6). Pearson.
- Sharma, R., & Gautam, A. (2016). Impact of Retail Formats on Consumer Buyer Behavior- A Study of Fast Moving Consumer Goods Market in South Africa. *British Journal of Economics, Management & Trade*, 11(3), 1-8. doi:<https://doi.org/10.9734/bjemt/2016/21388>
- Solomon, M. R. (2011). *Marketing: Real People, Real Choices*. Pearson.
- Süer, İ. (2014). *Pazarlama İlkeleri*. Nobel Akademik Yayıncılık.
- Twentify. (2021). *Twentify*. <https://www.twentify.com/tr/raporlar/teknoloji-ve-elektronik-urunler-arastirmasi-2021> adresinden alındı
- Wedel, M., & Kamakura, W. A. (2003). *Market segmentation: conceptual and methodological foundations*. Boston: Springer Science & Business Media.
- Wickelmaier, F. (2003). *An introduction to MDS*. *Sound Quality Research Unit*. Aalborg Universitetsforlag.
- Wilkie, W. L., & Gundlach, G. T. (2009). The American Marketing Association's New Definition of Marketing: Perspective and Commentary on the 2007 Revision. *Journal of Public Policy & Marketing*, 28(2), 259-264. doi:<https://doi.org/10.1509/jppm.28.2.259>
- Yıldız, D. (1995). Diskriminant Analizine İlişkin Bazı Yöntemler Ve Bir Uygulama. İstanbul: İstanbul Üniversitesi.
- Yükselen, C. (2007). *Pazarlama, İlkeler, Yönetim, Örnek Olaylar*. Detay Yayınları.





## Calculation of Household-Based Socioeconomic Status in Türkiye

**İbrahim Demir<sup>1</sup>**

**Furkan Metin<sup>2</sup>**

**Metin Aytac<sup>3</sup>**

**Turgay Altun<sup>4</sup>**

**Mehmet Şaban Ucari<sup>5</sup>**

**Onur Şentürk<sup>6</sup>**

### Abstract

The determination of the socioeconomic status (SES) of the households is of great importance especially for many studies conducted in the academic field. It is used extensively especially for market and advertising research. Different methods are utilized for calculating SES scores based on education, income and occupation data at household level. The data obtained are record-level real data, not survey data. Variables are rated according to the score scale from 0 to 100.

The data of the education variable is scaled according to the educational attainment of individuals and rankings of university departments, the occupation variable is scaled according to the civil servant coefficient and the number of persons employed, and finally the income variable is scaled according to median income value. All 26 million households in Türkiye were included in this study. According to the findings, the SES scores of the households were grouped into seven groups: A+ (1.44%), A (10.80%),

- 1 Assoc. Prof., Türkiye statistical enstitue (TurkStat), idemir@gmail.com, ORCID –ID 0000-0002-2734-4116
- 2 Assoc. Prof., TurkStat, furkanmetin@gmail.com, ORCID –ID 0000-0001-5782-6581
- 3 Expert, TurkStat, metin.aytac@tuik.gov.tr, 0009-0001-9615-3188
- 4 Expert, TurkStat, turgay.altun@tuik.gov.tr, 0000-0002-6999-9743
- 5 Expert, TurkStat, mehmet.saban.ucari@tuik.gov.tr, 0009-0004-3037-8538
- 6 Expert, TurkStat, onur.senturk@tuik.gov.tr, 0000-0002-6752-4963

B (16.05%), C1 (18.24%), C2 (19.32%), D (17.77%), and E (16.37%). While 16.37 % of the households were in the lowest SES group, the proportion of households in the highest SES group was 1.44%. This article is the first evidence-based study that sheds light on SES scores covering all households in Türkiye. In this way, the socioeconomic levels of people reside in Türkiye were classified by taking advantage of the administrative registers. Besides, in this study, the SES scores of an individual or household without a socioeconomic status score can be calculated by using the multivariate regression model.

## 1. Introduction

Socioeconomic status (SES) is a term that encompasses the social, economic, and educational standing of an individual or group within a society. It is a crucial factor in understanding a wide array of societal issues, from health disparities and educational access to employment opportunities and social mobility. SES indicate the position of an individual or group on the socioeconomic scale, such as income, duration and level of education, occupation according to their residence. In other words, socioeconomic status is a concept that expresses the place of an individual or a group in society based on a combination of factors such as income, education, occupation and well-being. This concept is frequently examined and studied in the social sciences literature. Both the individual components of SES, such as educational attainment, occupational status, and financial resources, and how these components can be combined to create a composite score have been extensively researched and discussed (Baer Et al., 2005). Some researchers choose to look at aspects of SES separately to determine how much each component contributes to any observed effect. For this reason, there is a very extended literature on SES studies. The positions of individuals and/or households in the social hierarchy and what can be used to determine their socioeconomic status is a highly controversial issue (Kalaycıoğlu Et al., 2010).

Globally, countries like the United Kingdom, the United States, and Australia have developed various SES indices to capture these disparities, each tailored to reflect their unique social and economic landscapes. The “Socio-Economic Indexes for Areas” (SEIFA) used in Australia, for example, integrates income, education, and occupation data to analyze socioeconomic trends at regional levels. Similarly, the United States relies on indices derived from census data to assess SES across diverse demographic groups. These international models provide valuable insights and frameworks that can inform SES studies in Türkiye, where socioeconomic disparities present distinct challenges shaped by regional inequalities and the urban-rural divide.

Traditionally, SES has been measured through surveys that capture variables such as income, education, and employment. However, administrative registers, which contain detailed and comprehensive data collected by government agencies for purposes such as taxation, welfare, and public services, offer a promising alternative approach.

Administrative registers provide a wealth of detailed, longitudinal data that can offer a more accurate and comprehensive picture of household SES. Moreover, because these registers often cover entire populations, they offer a unique opportunity to analyze SES in a way that is both representative and free from the biases of traditional survey methods. In this article, we will explore how household socioeconomic status can be calculated using data from administrative registers.

## **2. Methodology**

In the SES literature, socioeconomic status indicators are generally examined within the frame of some basic themes such as health, education, poverty, well-being and ethnic inequality. In the article, some analyzes made by researchers are included. In their study, Sewell and Shah (1967) investigated the relationship between SES, intelligence, and success on higher education. Mueller and Parcel (1981) suggested two occupation-based measures, the Duncan Socioeconomic Index and the Siegel Prestige Scale, as the best measures of the SES of individuals or household respondents.

Bradley and Corwyn (2002) reviewed the history of SES and provided an overview of the relationship between SES and children's well-being for three main developmental domains (cognitive, socioemotional, health). They have drawn attention to models that attempt to explain the link between SES and these aspects of development. Ball and Crawford (2005) examined the relationship between obesity and SES in developed countries in their study. In his study, Vyas (2006) conducted a principal component analysis on the assets owned by the household due to the difficulties in obtaining accurate data, although income and expenditure data are the most appropriate data for socioeconomic status. In the SES factor score calculation in the study, 4 main groups and 28 breakdowns were used in the urban-rural distinction in Brazil and Ethiopia data. The average socioeconomic score was calculated for the quintiles by ranking the calculated factor score from the least amount to the most amount. When looking at the distribution of the results, it was determined that the SES score calculated with the variables used did not have a significant discrimination. At the end of the study, the importance of determining the variables and re-evaluating the variables taken into account as conditions change over time is emphasized.

In their study, Conger and Martin (2010) focus on evidence regarding potential mechanisms proposed to explain these relationships. The study concludes with suggestions for future research on SES, family processes, and individual development in terms of important theoretical and methodological issues that have not yet been addressed. In his study, Patten (2019) aimed to propose and test a conceptual model for socioeconomic status (SES) and to measure variables that are available to researchers in Canada and applicable in other countries.

Powers (2021) provides an overview of the basic concepts of SES in his book, reviews the major approaches to conceptualizing and measuring socioeconomic status in North America over the past thirty-five years and evaluates these measures in light of current theoretical and methodological issues, and examines the origin and development of measures of socioeconomic status and prestige derived from United States (US) Census data. Antonoplis (2023) argues that SES is a construct that cannot be measured under traditional construct concepts and offers an alternative strategy for examining socioeconomic conditions. The study included a literature review consisting of 20 years of psychological research on SES. As a result of the study, SES was reconceptualized as a set of socioeconomic conditions and a measurement strategy was given to evaluate these concepts.

Blau and Duncan presented an analysis regarding to the scope and causes of intergenerational social mobility and argued that occupation has a central position in industrialized societies. Because profession provides a basis from which salaries and wages are earned and gives the professional authority and control over others and resources. As will be discussed, they suggest that different status or prestige is attributed to various occupations. Therefore, Blau and Duncan conclude that occupational hierarchy is the underlying dimension of social stratification. (Peter M. & Duncan, 1967)

### **2.1. Using the Household or Individual as the Unit of Analysis**

Should socioeconomic status measurement be based on the household or the individual as the unit of measurement? The answer to this question is a very important issue for determining the measurement and the indicators to be included in the measurement. Those who examine class within the traditional framework tend to apply class categories to the household and define the man as the person who determines the class structure in the household. Goldthorpe has argued that this approach effectively. According to him, household members share the same class position, and this position can be measured most accurately by looking at the male household head.

Because the male household head is the household member with the most responsibility and continuity in terms of participation in the labor market. Later, following Erikson's ideas, Goldthorpe moved away from the idea that the male determines the class of the family and began to support a method called the dominance approach. Accordingly, the class position of the household is determined by the position of the household member whose participation of household's labor market is at the highest level.

There are also views that argue that the individual should be the object of analysis. According to those who argue that the individual should be the object of analysis, the individual's attitudes and behaviors; According to logic, it should be examined according to the direct experiences of the individual. On the other hand, in studies on consumption behaviors and attitudes where the family consumes together as a whole, the appropriate unit of analysis is the household. If it is decided that the survey is a household-based classification measure rather than an individual-based classification measure, a second stage raises the question of how best to measure the household's class position. In this case, three basic approaches can be listed (Edgell, 1998):

- Goldthorpe's dominance approach: Based on the individual in the household with the highest participation in the labor market (Goldthorpe, 1983).
- Combined classifications: Including men and women in the measurement.
- Separate class schemes for men and women: Applying the schemes developed for men also to women.

There are also some problems and advantages derived from all three approaches. At this point, the scope of the research must be determined in order to decide to unit of analysis correctly. The last step that will enable the concept of class to be operationalized is to decide on the scope of the research. Traditionally, class analyzes include only adults who are economically active full-time. In this case, it becomes clear that groups such as those who do not working full time, unemployed persons, part-time workers, unpaid workers, retirees are not included in the class structure. (Kalaycıoğlu et al., 2010)

## **2.2. Socioeconomic Status Indicators**

Another important issue is also that determining the factors/variables to be included in the index when creating the SES scale. Within this context, it is theoretically accepted that the variables such as education, occupation, income, authority and ownership are the main factors affecting SES.

*a) Education:* It has been one of the most important factors for obtaining certain statuses and vertical mobility. For this reason, education should be one of the basic components of an SES index to be formed for education.

*b) Income:* Income is one of the basic components in SES scales. Income level of a household usually affects the place in the social structure. Income, which is a socioeconomic indicator, closely affects consumption and lifestyle.

*c) Occupation:* Occupational reputation is an almost unchanging part of the SES models. In the modern life, it is often stated that occupation is the most fundamental factor that determines a person's social position. Occupation is also seen as a complementary factor to education and income. (TUSES)

### 2.3. Commonly Used SES Measurements

Although there is disagreement about the conceptual meaning of SES, there appears to be agreement on Duncan's (1972) definition of the tripartite nature of SES, which combines parental income, parental education, and parental occupation as the three primary indicators of SES. (Sirin, 2005)

The most widely used index by those publishing in the field of Child Development and Developmental Psychology is the Hollingshead Index. The index takes into account the fact that social status is a multidimensional concept.

In 1970, Green developed a scoring procedure designed for use in public health research. The scoring guide consists of two steps.

The Siegel Prestige Scale is widely used in sociology. Unlike the Census scale or the Green scale, the development of the Siegel scale was based on the assumption that occupational ranking is socially defined. (Mueller & Parcel, 1981)

According to the Duncan SEI index, developed in the early 1960s, each SEI score is obtained using the regression equation given in Equation 1:

$$SEI_{Score} = 0.59X_1 + 0.55X_2 - 6.0 \quad (1)$$

Here,  $X_1$  is the percentage of men in the occupation who have an income of at least \$3,500, and  $X_2$  is the percentage of men in the occupation who are at least high school graduates. Regression weights were generated from data collected in a national survey conducted by NORC (1947) in North-Hatt (North-Hatt reputation scores). (Duncan, 1961) Featherman and Hauser (1977) updated this scale for three-digit occupational groups of the 1970 US Census. (Featherman & Hauser, 1977)

Another widely used scale, the Kuppuswamy scale, was developed by Kuppuswamy and is still widely used as a measure of socioeconomic status in urban populations. The Kuppuswamy scale is based on a scoring system that takes into account factors such as family income, education level, occupation and social status. (Kuppuswamy, 1981)

## **2.4. Country Case Studies**

When National Statistical Offices and International Organizations are examined, it is seen that statistics on socioeconomic status are not published as indices. In general, statistics on the subjects included in the index are obtained from these sources and used in academic studies.

The “National Statistics Socioeconomic classification” has been published by the United Kingdom Statistics Office on this subject. (UK-1) (UK-2). “Socio-Economic Indexes for Areas (SEIFA)” were calculated by the Australian Office of Statistics. (SEIFA, 2023). Within the scope of the “Programme for International Student Assessment’s (PISA) 2018” conducted by the OECD, it shows how strongly socioeconomic status is related to performance in participating countries and economies. (PISA, 2023). In a study conducted by the US Department of Health and Human Services in 1998, health status was reported according to socioeconomic status. (USA-1, 2023) The reference document for determining socioeconomic status in health research was prepared by the same unit in 2012. (USA-2, 2023)

## **3. Results**

### **3.1. Components of SES Calculation**

Income distribution indicators are needed not only economically but also socially in order to determine the country’s economy and people’s living standards and to make a healthy evaluation of various social systems. Since 2006, within the framework of harmonization with the European Union, the Turkish Statistical Institute (TurkStat) has been carrying out the “Income and Living Conditions Survey (TR-SILC)”, in which the “panel survey” method is used, the aim of which is to produce data on issues such as income distribution, living conditions, social exclusion and relative income poverty. However, the TR-SILC does not allow producing income information at the household level.

The socioeconomic status score, calculated based on administrative registers at the household level, has 3 components. These are education, occupation and income.



In order to calculate the socioeconomic status index at the household level, the 2021 Address-Based Population Registration System (ABPRS), the National Education Statistics Database (NESD) and administrative records containing education, occupation and income information received from institutions were used. Education, occupation and income scores, which are the components of the socioeconomic status index at the household level, were calculated thanks to administrative registers.

### 3.1.1. Scoring the Education for Households

The highest level of education completed for each individual residing in Türkiye is available in the NESD. Within the framework of the education levels in NESD, it was decided to make a scoring for each education level and by using other auxiliary information such as the type of high school graduated, the university, department, and university entrance ranking of the individuals within the education levels. The score of the person with the highest education level in the household represents the household score. The scores for education is shown in Table-1.

*Table 1. Education Scores for SES*

CODE	LEVEL	SCORE
0	Unknown	0.0
1	Illiterate	0.0
2	Literate	10.0
3	Primary school	20.0
4-5	Primary and Secondary School Graduates (Open)	30.0
4-5	Primary and Secondary School Graduate (Formal)	35.0
6	High School (General-Open Education)	40.0
6	High School (Vocational-Open Education)	50.0
6	High School (General-Formal)	55.0
6	High School (Multi-Program)	57.5
6	High School (Vocational-Formal)	60.0
6	Police College	61.0
6	High school (Science)	62.5
7	Associate Degree / Undergraduate (Undifferentiated)	74.0
7	Associate Degree	65.0-73.0
7	Bachelor's Degree	75.0-90.0
8	Master's Degree	91.0
9	Doctorate	100.0

### 3.1.2. Scoring the Income for Households

While calculating the income score; Information regarding premium earnings, pension, social aid income and income declarations in the administrative registers were used. Here, the information in the administrative registers are aggregated annually at the individual level. The distribution of annual total income at the household level is given in Table-2. For households, whose income information was not available in administrative registers, motor vehicle insurance information was used to fill the gap in income.

*Table 2. Number of households according to annual household income classification*

Household Annual Income (TL)	Number of Households
Total	25 329 833
No Household Income in Administrative Records	1 634 999
0-30 000	5 409 708
30 001-100 000	11 150 063
100 001-250 000	4 420 364
250 001-500 000	1 651 154
500 001-1 000 000	843 743
1 000 001 and over	219 802

### 3.1.3. Scoring the Occupation for Households

While calculating the profession score; Social Security Institution paid employee (SSK 4a), self-employed (Bağkur 4b) and civil servant (Emekli Sandığı-4c), Service Tracking Program (HITAP), Directorate General of Public Accounts payroll records, agricultural records, part-time employee records, academic personnel records and private bank chest records were used.

## 3.2. SES Calculation through Administrative Registers in Türkiye

For the three components of the socioeconomic status, education, occupation and income were given scores between 0 and 100. Thus, score values between 0 and 100 were defined for the education, occupation and income components for each household (h).

$$SES_{SCORE_h} = EDUCATION_{SCORE_h} + INCOME_{SCORE_h} + OCCUPATION_{SCORE_h} \quad (2)$$

The household with a SES score of 0 was determined as having the lowest social status, and the household with a SES score of 300 was determined as

the highest social status. Thus, an SES score was assigned to all households in 2021, approximately for 25 million households.

#### 4. Conclusion and Discussion

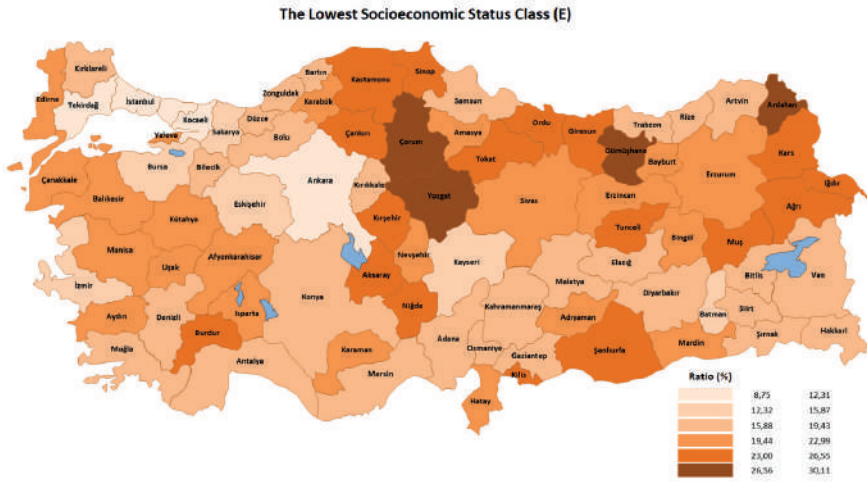
This study aims to determine an SES score for all households in Türkiye based solely on administrative records, using education, occupation and income information. When academic studies in this context are examined, it is seen that status classes are generally evaluated in 5 or 6 classes. In order to make the highest status group more distinct, the 7-level classification is used in this study. The ranges of classes were determined in line with the distribution of SES scores, cluster analysis results and expert opinions.

Lower and upper scores were determined for each class. Thus, SES scores were classified at 7 level. Approximately 360 thousand households in Türkiye were in the top class in terms of socioeconomic status. The ratio of households according to socioeconomic classes is shown in Figure-1.



*Figure 1. Proportion of households by socioeconomic classes*

When socioeconomic status classes are examined at the provincial level; As can be seen in Figure-2, the provinces with the highest proportion of households in the lowest socioeconomic status group (E) are Yozgat, Çorum, Gümüşhane and Ardahan. (Figure-2).



*Figure 2. Distribution of households in the lowest SES class by province*

## References

- Antonoplis, S. (2023). Studying socioeconomic status: Conceptual problems and an alternative path forward. *Perspectives on Psychological Science*, 18(2), 275-292.
- Baer, J., Baldi, S., & Merola, S. S. (2005). The measurement of socioeconomic status (SES) in the social sciences: A review to inform NAEP. *Unpublished paper prepared for the National Center of Education Statistics*.
- Ball, K., & Crawford, D. (2005). *Socioeconomic status and weight change in adults: a review. Social Science & Medicine*, 60(9), 1987-2010.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53(1), 371-399. <https://doi.org/10.1146/annurev.psych.53.100901.135233>
- Catalbas, M. C., & Burken, S. (2022). The mathematical relationship between COVID-19 cases and socio-economic indicators of OECD countries. *Pathogens and Global Health*, 116(6), 376-388.
- Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic status, family processes, and individual development. *Journal of Marriage and Family*, 72(3), 685-704. <https://doi.org/10.1111/j.1741-3737.2010.00725.x>
- Duncan, O. (1961). *A socioeconomic index for all occupations*. tus. New York: Fre.
- Edgell, S. (1998). Sınıf. Dost Kitapevi Yayınları: Ankara
- Edwards, A. (1933,1938). A social-economic grouping of the workers of the U.S with appendix. *Journal of American Statistical Assosiation*.
- Erkenekli, M., Türkiye’de Sosyoekonomik Statü (SES) Gruplarına Göre Temel Değerlerin Farklılaşması, Doktora Tezi, Ankara, 2009.
- Featherman, & Hauser. (1977). Commonalities in social stratification and assumptions about status mobility in the U.S.
- Goldthorpe, J. H. (1983). *Women and Class Analysis: in Defence of the Conventional View*, Sociology
- Hollingshead, A. (2011). Four Factor Index of Social Status. *Yale Journal of Sociology*.
- Kalaycıoğlu, S., Çelik, K., Çelen, Ü., & Türkyılmaz, S. (2010). Temsili Bir Örneklemde Sosyo-Ekonomik Statü (SES). *Journal of Sociological Research*.
- Kuppuswamy, B. (1981). *Manual of socio economic status scale (urban) Delhi*.
- Little, C., Alsen, M., Barlow, J., Naymagon, L., Tremblay, D., Genden, E., ... & van Gerwen, M. (2021). The impact of socioeconomic status on the clinical outcomes of Covid-19; a retrospective cohort study. *Journal of community health*, 1-9.

- Mueller, C. W., & Parcel, T. L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Development*, 52(1), 13–20. <https://doi.org/10.2307/1129211>
- Patten, S. L. (2019). Measures of Socio-Economic Status in Educational Research: The Canadian Context. *International Journal of Education Policy and Leadership*, 14(3), n3.
- Peter M., & Duncan, O. (1967). *The American occupational structure*. Wiley.
- Powers, M. G. (2021). Measures of socioeconomic status: An introduction. In *Measures of socioeconomic status* (pp. 1-28). Routledge.
- PISA. (2023, 07 18). (<https://www.oecd-ilibrary.org/sites/f7986824-en/index.html?itemId=/content/component/f7986824-en>).
- Rossi, P., Sampson, Bose, Jasso, & Passel. (1974). *Measuring household social standing*. Social Science Research.
- SEIFA. (2023, 07 18). (<https://www.abs.gov.au/statistics/detailed-methodology-information/concepts-sources-methods/socio-economic-indexes-areas-seifa-technical-paper/2021/data-underpinning-indexes>)
- Sewell, W. H., & Shah, V. P. (1967). Socioeconomic status, intelligence, and the attainment of higher education. *Sociology of Education*, 40(1), 1–23. <https://doi.org/10.2307/2112184>
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research*, 75(3), 417-453.
- Suna, E. & Özer, M. (2021). The Achievement Gap between Schools and Relationship between Achievement and Socioeconomic Status in Turkey. *Journal of Measurement and Evaluation in Education and Psychology*, 12 (1), 54-70. DOI: 10.21031/epod.860431
- UK-1. (2023, 07 18). (<https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticsocioeconomicclassificationnscrebasedonsoc2010>)
- UK-2. (2023, 07 18). <https://commonslibrary.parliament.uk/find-the-socio-economic-status-of-people-living-in-england-and-wales-by-constituency/>)
- USA-1. (2023, 07 18).(<https://www.cdc.gov/nchs/data/hus/hus98cht.pdf>)
- USA-2.(2023).(<https://ncvhs.hhs.gov/wp-content/uploads/2014/05/120622lt.pdf>)
- Vyas, SE. & Kumaranayake L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health Policy Plan*, 21, 459-468.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3), 461–481. <https://doi.org/10.1037/0033-2909.91.3.461>



# Producing CPI During Covid-19 Period in Türkiye

Furkan Metin<sup>1</sup>

İbrahim Demir<sup>2</sup>

Cem Baş<sup>3</sup>

## Abstract

The first Covid-19 national lockdown was decelerated on 11 March 2020 in Türkiye. The full lockdown process has made it difficult to collect the data necessary for both household budget surveys and consumer price index in Türkiye. Nevertheless, these challenges of Covid-19 have forced national statistical offices to produce reliable, qualified and timely consumer price statistics through using alternative data sources such as scanner data and web scraping data. In this paper, experiences gained by the Turkish Statistical Institute to overcome the difficulties arise by Covid-19 have been shared. It is noteworthy that during the Covid-19 pandemic, Turkish Statistical Institute (TürkStat) successfully raised the proportion of prices essential for the Consumer Price Index (CPI) sourced from alternative data from 0% to approximately 49% within a span of just one to two years.

## 1. Introduction

Since December 2019, the spread of the SARS-Cov-2 virus has reached the level of a global epidemic. In this period, which is called the Covid-19 epidemic, almost all countries around the world have experienced processes called “full lockdown”. Citizens, except for the limited number of sector employees who are essential for human life, especially health, safety and food, are prohibited from leaving their homes, except for compulsory situations.

---

1 Associated Professor, TurkStat Vice President, furkanmetin@gmail.com, ORCID –ID 0000-0001-5782-6581

2 Associated Professor, TurkStat Vice President, ibrahimdemir@tuik.gov.tr, ORCID –ID 0000-0002-2734-4116

3 Dr., TurkStat Expert, cembas@tuik.gov.tr, ORCID –ID 0000-0003-4488-2562



With these restrictions, the spread of the SARS- Covid-2 virus, which caused the Covid-19 epidemic, among humans was tried to be reduced. Nevertheless, the full lockdown process has made it difficult to collect the data necessary for both household budget surveys and consumer price index (CPI) (Andersen, 2020).

During the Covid-19 pandemic, in order to protect public health, face-to-face survey method has been temporarily suspended in all household surveys of Turkish Statistical Institute (TurkStat). Therefore, the Household Budget Survey could not be completed for 2020. CPI weights for 2021 according to the methods given in the EUROSTAT guide; Expenditure shares (HBS data) for the years 2017, 2018, and 2019 have been updated using the 2020 National Accounts Household Final Consumption Expenditure growth rates (EUROSTAT, 2020). In following years EUROSTAT have also advised to use the latest National Accounts Household Final Consumption Expenditure growth rates (EUROSTAT, 2021-2022). This is due to the fact that consumption habits of households have changed during the period of Covid-19 pandemic (Baker et al, 2020).

Furthermore, some outlets were closed or working hours changed as a result of the restrictions of Covid-19; some services were not provided (Dunn et al, 2020). In this period, field studies continued as much as possible in order to produce the CPI in the best way possible (Destatis, 2022). When prices were unavailable due to restrictions, alternative data sources, such as scanner data and web scraping information, were recommended in EUROSTAT's guidance published in April 2020 (EUROSTAT, 2020). As a necessary result, utilizing new alternative digital data sources, such as scanner data and web scraping data, became crucial for statistical offices in calculating the Consumer Price Index (CPI).

Accordingly, TurkStat implemented several measures to address the challenges posed by Covid-19. TurkStat makes great efforts to minimize the impact of Covid-19 measures on CPI quality and takes the necessary measures in a timely manner. Difficulties in data collection, especially during quarantine periods, accelerated the studies on the use of alternative data sources in TurkStat. Within a year, scanner data and web scraping data were incorporated into the production of the CPI by TurkStat.

## **2. Scanner data**

Scanner data, as one of the Big Data sources, has become reasonable popular in national statistical systems to improve the accuracy of CPI. Scanner data is “high in volume and contains information about individual

transactions or summaries, date, quantities and values of products sold, and product descriptions”. Since scanner data is an enhanced data sources for national statistical organizations (NSOs) that can be used to enhance their statistical products and improved, more efficient ways of working by reducing provider burden and also reducing cost of physically collecting data (International Monetary Fund, 2020).

While scanner data sets provide several positive outcomes, there are also several challenges that need to be overcome by NSOs to compile the CPI. Some of these challenges were outlined as follows,

- establishing a sufficient sustainable communication channels among the inside and outside partners,
- developing an IT system to asses and prepare data
  - o such as controlling the scope of items included in the data sets (unit of measurements, date, quantities, price values etc.)
  - o adaptation of the classifying systems of each retailer to classifications of individual consumption (COICOP) quality assurance of the scanner data sets (International Monetary Fund, 2020).

### **3. The Challenges Faced by the Turkstat During Covid-19 Pandemic**

#### **3.1. Obtaining Scanner Data Sets**

In fact, obtaining comprehensive data sets from different retailers is require expertise and skills in many different fields like communication, statistics, information technologies and law. Effective planning and coordination of all processes involving multiple units within the National Statistical Office (NSO) is essential (UnECE, 2021a).

Scanner data can be compiled directly from companies or through intermediary institutions that keep invoice information electronically. In the TurkStat version, data is compiled directly from the companies. The advantages of obtaining data directly from the company are explained in the Consumer Price Index Handbook published by the International Monetary Fund (IMF) as follows:

- Easier negotiations about data (Experiences in the Netherlands, Switzerland and Australia suggest these negotiations will take approximately six months to complete;
- The supply of the data set at no (or minimal) cost

- The scope of items included in the data set
- The level of item aggregation to ensure homogenous information (UnECE,2021b)
- An agreed timetable for the supply of the data set to meet CPI processing requirements
- A contact officer within the retail business who is familiar with the data set to answer NSO data queries (International Monetary Fund, 2020).

A key determinant of share prices and profitability within the sector is the pricing strategy employed by companies. Product features and prices, linked to unique barcode identifiers, are typically available to consumers, and are not regarded as confidential information. Nevertheless, when this pricing data is aggregated from multiple companies, the pricing strategies of these firms become apparent. Additionally, confidential information, or trade secrets, plays a significant role in determining the competitive strength of firms. Thus, safeguarding trade secrets is essential for businesses to maintain their competitive edge (Karataş, 2017).

For these reasons, although barcode scanning data is obtained from firms and utilized by national statistical offices, aggregated data pertaining to individual companies is not disclosed. This aggregated data is used solely for statistical purposes, specifically for the production of the Consumer Price Index (CPI). Four important benefits of using scanner data instead of sampling data have been summarized by the American Bureau of Labor Statistics as follows:

- the ability to account for product and outlet substitutions in some categories of consumer spending,
- the ability to produce indexes with greatly reduced sampling error for some item strata,
- the ability to identify new products and quality enhancements,
- and the ability to ensure a more accurate probability sampling of items with fewer manual steps, a sampling procedure known as “disaggregation” (Bradley et al.,1997).

The TurkStat takes the following steps while compiling the scanner data:

- Examination of legal terms:** Statistics Law of Türkiye Article 7 enacted in 2005, authorized The TurkStat to compile the information serve the needs of Institute. Based on this article, The TurkStat has

the right to request all data to be used in the production of official statistics from all units.

- ii. Clarifying data characteristics:** It is very significant to complete internal processes before requesting data from companies. It should be determined how often and what kind of data is needed in order to decrease response burden. The data pattern should be clarified by determining all units within the Institute that need to use the data. During the scanner data project conducted by TurkStat, it was identified that two departments were the main users of this data source. These departments were both the Price Statistics Department, responsible for calculating the CPI, and the National Accounts Department, which monitors the gross domestic product calculations.
- iii. Selection of companies:** When compiling scanner data, ensuring data flow from all companies operating in the country will be very costly, considering the transfer, processing and storage of the data. Therefore, a small subset of nationwide retailers was selected to begin the scanner data Project. In recent years, the TurkStat was entitled to access to detailed information about the companies through the protocols with various public institutions, in the context of efforts to increase the use of administrative records. This allowed, companies that are predominant in their own sectors for certain variables (turnover, added value, etc.) were determined easily.
- iv. Contacting the companies:** The TurkStat has widespread organizational structure, the expert staff and advanced information technology applications on collecting data. In this way, the Institute communicates with respondents from all over the country quickly and effectively. the TurkStat produces meaningful statistics from compiled data by the authorization of Statistical Law of Türkiye and describes to the respondents how they can take advantage of these statistics that they contributed. In this regard, summary outcomes were planned to be prepared in order to share them merely with the nationwide retailers themselves.
- v. Planning and Performing Data Transfers:** For data transfer, the most appropriate method for the company and the TurkStat infrastructures is determined by negotiating with information technology units of the companies. While determining the appropriate method, attention is paid to protect data privacy and prevent data loss. A platform (FTP, SFTP, etc.) is decided for

transferring of compiled scanner data, and provide data transfer to Institute through this platform. The data is prepared for use by the relevant units through ETL (Extract, Transform, Load) processes conducted following its transfer.

- vi. **Data Editing:** To prepare the transferred data from companies to TurkStat for analysis, potential inconsistencies and data entry errors are identified and reported. These reports include assigning multiple barcodes, definitions or amount to a product, and assigning incorrect amount to the products. As a result of the reports, some imputations such as correcting the number of products, and excluding from relevant population the incorrect barcodes and products are made. After reporting and imputation stages, product codes are assigned to each product so that products can be used in calculations.
- vii. **Data Analyze:** In the analysis phase of code assigned and error free data set, the unit prices that formed on the basis of barcodes and products, and monthly turnover of each company and distribution of this turnover by products are calculated. The unit prices in the scanner data are compared with the unit prices in the Consumer Price Index, and price differences are reported.

### 3.2. Challenges to Compile Scanner Data

As stated above, the difficulties experienced in transferring this data which is of great commercial importance for companies, and the measures taken can be listed as follows:

- **Exhaustive data:** By defining the selection criteria for both sectors and companies, collaboration has begun with the limited number of identified nationwide retailers.
- **Reluctance of companies:** An effective communication strategy is to implement, minimizing the response burden to the lowest possible level, while promptly addressing the needs of the companies. Statistics producers who support and encourage a customer-friendly culture that is mindful of respondent concerns are more likely to be rewarded with improved response rates and careful, accurate responses from respondents (Karataş, 2017).
- **Difficulty in compiling requested data:** Keeping data in different databases, franchises data, etc. are some of the most common problems. Companies should be given sufficient time to complete

these works and technical support should also be provided when it is necessary.

- **Ensuring data privacy:** Privacy provisions that located in The Statistics Law of Türkiye are followed with precision, access authorizations to received data are restricted.
- **Failure to transfer data within specified periods:** In order to accelerate the processes protocols were prepared for each company and submitted to the companies before the ultimate data transfer. Additionally, an e-mail notification system has been developed to ensure scanner data transfers within the specified periods. In the case of exceeding the deadline of data transfer, the e-mail notifications system automatically sends notifications to the people who are responsible for data transfer in both TurkStat and companies.
- **Inadequate product classification of companies:** Assignment studies of the classification code (E-COICOP) are continued. Codes are assigned manually, by programming or by machine learning techniques (ONS,2021).

In the year 2020, the analysis processes of the data obtained from a total number of seven national chain stores (markets), which have a significant share in the retail trade sector, have been completed. Various studies (classification, code matching, etc.) have been carried out to make the data ready for use for CPI calculations. Barcode sales data obtained from chain markets started to be used in 2021 CPI calculations. In this way, prices and quantities covering the whole month were included in the index by using daily sales data on the basis of branches obtained from seven national chain companies instead of prices compiled from the field on certain days of the month. As a result, 21 per cent of the prices for CPI were gathered by scanner data sources and used in CPI calculations in 2021.

- **Local chain market data:** In the 2022 CPI study, sales data obtained from a total of 42 companies are included in the index calculation. In this context, the prices of food, glassware and personal care products from 9 chain markets and 26 local markets in the market sector are followed. Currently, approximately 84.000 item numbers are compiled from 9 chain markets and 8.000 item numbers from 26 local markets.

Nevertheless, the number of item numbers obtained from 26 local markets corresponds to approximately one-tenth of the number of item numbers obtained from 9 national markets. Therefore, covering

26 local markets has significantly increased our communication cost. In fact, national markets are more likely to have their own information technologies (IT) units. Therefore, scanner data transfers from the national markets to TurkStat have been completed usually in a timely manner. On the other hand, most of the time, problems were encountered in terms of timeliness in data transfers made from local markets., which has shortened the CPI analysis periods before the CPI bulletin (ONS,2021).

A further problem regarding to scanner data of local markets has been that there are a few chain stores in local markets and most of them are in just one city.

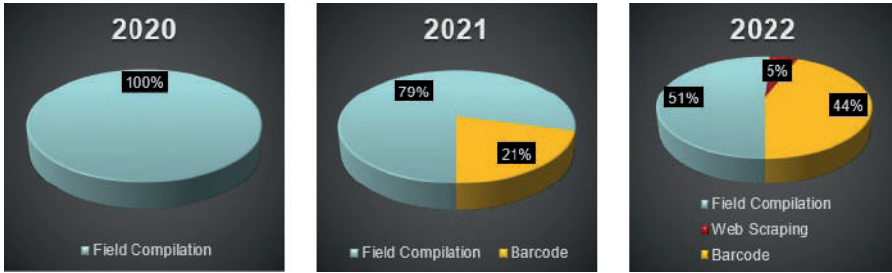
- In order for the barcode numbers in the sales data to be used in the index calculation, first of all, each product is assigned to the appropriate COICOP codes in accordance with the item basket and detailed description. Nevertheless, considering the number of local markets and the number of items tracked from local markets in COICOP assignment processes, workload weight occurs.

The problem of product continuity in local markets increases the need for substitution and COICOP codes must be assigned quickly and accurately. Apart from these, significant errors such as quantity errors are frequently detected in local market sales data. In addition, the fact that the data transfer requested from the local markets covering 1-7 and 1-14 periods of each month is not done regularly and the monthly final sales data requested to be sent on the 26th of the month are delayed then it makes difficult for the CPI calculations to reach the bulletin date. In the year 2022, 44.3% of the prices have been obtained directly by this method (Please see the Figure 1).

### **3.3. Web scrapping data**

The prices compiled over the internet for the prices of household appliances (refrigerator, washing machine, dishwasher, etc.), electronic products, clothing, furniture, brand new cars and bus tickets, which are suitable for pricing using the web scrapping method in terms of item definitions, were used in the CPI calculations in 2022. The number of prices to be obtained by this method constitutes approximately 5.3% of the total number of prices to be compiled (Please see the Figure 1).





*Figure 1. The significant change in the share of CPI data sources during Covid-19*

#### 4. Conclusion

It can be considered a significant achievement that Turkish Statistical Institute (TurkStat) increased the proportion of prices necessary for the Consumer Price Index (CPI) collected from alternative data sources from 0% to nearly 49% in a few years during the Covid-19 pandemic (Please see the Figure 1).

The scope of this study, which has initially commenced with food and clothing retailers, is set to expand in the coming years to encompass telecommunications companies, insurance firms, entities in the healthcare sector, airline companies, and organizations operating within the energy market. Following an evaluation of product information obtained from approximately 90 websites, this data will be utilized for the production of official statistics (TurkStat, 2023).

Using information from all processed products, rather than just from the selected items, will enhance the quality of the generated statistics, mitigate errors stemming from pollster or tagging mistakes, facilitate advanced analyses based on price and quantity, and simplify the selection of substitute products. Although engaging with companies during the establishment of the system, implementing IT and coding systems, and fostering a working culture centred around big data may increase the burden on both TurkStat and the companies involved, it has been observed that, in the medium term, this approach leads to higher-quality statistics with reduced costs and lower response burdens.



## References

- Andersen, A.L., Hansen, E.T., Johannesen, N. and Sheridan, A. (2022), Consumer responses to the COVID-19 crisis: evidence from bank account transaction data\*. *Scand. J. of Economics*, 124: 905-929. <https://doi.org/10.1111/sjoe.12512>
- Scott R Baker, Robert A Farrokhnia, Steffen Meyer, Michaela Pagel, Constantine Yannelis, How Does Household Spending Respond to an Epidemic? Consumption during the 2020 COVID-19 Pandemic, *The Review of Asset Pricing Studies*, Volume 10, Issue 4, December 2020, 834–862. <https://doi.org/10.1093/rapstu/raaa009>
- Bradley, R., Cook, B.A., Leaver, S.G., & Moulton, B.R. (1997). An Overview of Research on Potential Uses of Scanner Data in the US CPI. <https://stats.unece.org/ottawagroup/download/f46.pdf>
- Dunn AC, Hood KK, Driessen A. Measuring the effects of the COVID-19 pandemic on consumer spending using card transaction data. *BEA Working Paper Series*. 2020;5. [https://www.bea.gov/system/files/papers/BEA-WP2020-5\\_0.pdf](https://www.bea.gov/system/files/papers/BEA-WP2020-5_0.pdf).
- Destatis (2022). Impact of the corona crisis on the CPI/HICP price collection. [https://www.destatis.de/EN/Themes/Economy/Prices/Consumer-Price-Index/Methods/corona-cpi-hicp.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/EN/Themes/Economy/Prices/Consumer-Price-Index/Methods/corona-cpi-hicp.pdf?__blob=publicationFile)
- EUROSTAT (2020). Guidance on the Compilation of the HICP in the Context of the Covid-19 Crisis. [https://ec.europa.eu/eurostat/documents/10186/10693286/HICP\\_guidance.pdf](https://ec.europa.eu/eurostat/documents/10186/10693286/HICP_guidance.pdf)
- EUROSTAT (2021). Derivation of HICP Weights for 2022. <https://ec.europa.eu/eurostat/documents/10186/10693286/Derivation-of-HICP-weights-for-2022.pdf/c47503af-e526-0e40-def2-370c501bb595?t=1638266939151>
- EUROSTAT (2022). Derivation of HICP Weights for 2023. <https://ec.europa.eu/eurostat/documents/272892/11336726/Derivation+of+HICP+weights+for+2023.pdf/>
- International Monetary Fund. (2020). *Consumer Price Index Manual-Concepts and Methods*, pp.365
- Karataş D.N. (2017). Tax Confidentiality within the Context of Trade Secret and the Protection of Personal Data, pp.39, 383.
- ONS (2021). Classification of new data in UK consumer price statistics-<https://www.ons.gov.uk/economy/inflationandpriceindices/articles/classificationofnewdatainukconsumerpricestatistics/2021-04-06>
- TürkStat (2023). Consumer Price Index February. <https://data.tuik.gov.tr/Bulten/Index?p=Consumer-Price-Index-February-2023-49656>

- UnECE (2021a). Guide on producing CPI under lockdown. <https://unece.org/sites/default/files/2021-12/Guide%20on%20producing%20CPI%20under%20lockdown%20WEB%20version.pdf>
- UnECE (2021b). Index Compilation Techniques for Scanner Data: An Overview. [https://unece.org/sites/default/files/2021-05/Session\\_1\\_Eurostat\\_Paper.pdf](https://unece.org/sites/default/files/2021-05/Session_1_Eurostat_Paper.pdf)



## Use of Relative Entropy Statistics in Contingency Tables

Atıf Evren<sup>1</sup>

Erhan Ustaoglu<sup>2</sup>

Elif Tuna<sup>3</sup>

Büşra Şahin<sup>4</sup>

### Abstract

There are various information theoretic divergence measures used for determining associations between nominal variables. Among them, Shannon mutual information statistic is especially appealing, since its sampling properties are well-known. Although Shannon mutual information is more frequently used, Rényi and Tsallis mutual informations, as envelopes of various tools, provide much higher flexibility than Shannon mutual information. Indeed, Shannon mutual information is a special case of Kullback-Leibler divergence, Rényi, and Tsallis mutual informations. In this study, large sampling properties of Shannon, Rényi, Tsallis mutual information statistics are considered as well as Pearson, Tschuprow, Sakoda, Cramér, Hellinger, and Bhattacharyya measures. In simulations, the normality of most of the statistics, and the higher positive correlation coefficients between all these tools are observed. Their sampling variabilities are compared. Then by using Rényi and Tsallis mutual information statistics, correlation coefficients are estimated for 8 different scenarios, and 3 bivariate normal distributions.

- 1 Assoc. Prof. Dr., Yildiz Technical University, Faculty of Sciences and Literature, Department of Statistics, Istanbul; aevren@yildiz.edu.tr, ORCID: 0000-0003-4094-7664
- 2 Assoc. Prof. Dr., Marmara University, Faculty of Management, Department of Management of Information Sciences, Istanbul; erhan.ustaoglu@marmara.edu.tr, ORCID: 0000-0002-9077-4370
- 3 Assoc. Prof. Dr., Yildiz Technical University, Faculty of Sciences and Literature, Department of Statistics, Davutpasa, Istanbul; cozturk@yildiz.edu.tr, ORCID: 0000-0001-8572-3109
- 4 Assist. Prof. Dr., Halic University, Faculty of Engineering, Department of Software Engineering, Istanbul; busrasahin@halic.edu.tr, ORCID: 0000-0003-3995-9238

## 1. Introduction

Distance functions in statistics are applied frequently. Some basic statistics like variance, and standard deviation are simply some functions of Euclidean and Minkowski distances. Kolmogorov-Smirnov statistic can be derived from Chebyshev distance, etc. Some tests of independence consider the distance (or divergence) between a joint probability distribution, and the product of two marginal probability functions. Thus, distance or divergence measures are used in hypothesis testing, as well as clustering some multivariate data. In Bayesian methodology, to evaluate sample information, comparisons between conjugate distributions can be made by some divergence measures. Some procedures of the sequential analysis of Wald (Wald, 2004) are based on the divergence between two Bernoulli distributions.

Analysis of dependencies may also be realized by entropy and relative entropy formalism. Mathematical foundations of entropy and relative entropy can be found in Shannon (Shannon and Weaver, 1963) and Khinchin (Khinchin, 1957). Preliminaries can be found also in Rényi (Rényi, 2007), Ash (Robert, 1990), and Cover&Thomas (Cover and Thomas, 2006). Following Shannon's contribution to information theory, studies by Kullback (Kullback, 1997), Gokhale (Gokhale and Kullback, 1978), and Pardo (Pardo, 2006), rather focused on inferential issues of statistics within the framework of entropy.

Kullback considered entropy and mutual information (a special case of relative entropy) as the two key concepts of his system. Pardo discussed various divergence measures (including relative entropy) in goodness of fit testing, loglinear models and contingency tables. Singh (Singh, 1998) and Samilov (Şamilov, 2015) contributed in entropy-based parameter-estimation, and optimizing entropy. From econometrics (Ullah, 1996) to information-theoretic learning (Principe, 2010); the applications of entropy and divergence measures vary on a wide range.

In this study, Rényi and Tsallis mutual informations are considered in the context of two-way contingency tables. In case of nominal bivariate distributions, they can serve as an association measure between variables. In case of continuity of the variables, they may be used in estimating correlations after grouping bivariate continuous observations by the help of a contingency table as well.

## 2. Independence Tests for Contingency Tables

For contingency tables, the test of independence is based on the chi-square statistic

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - n \cdot p_{ij})^2}{n \cdot p_{ij}} \tag{1}$$

where  $p_{ij}$  is the hypothesized probability for the  $i$ th column and  $j$ th row of the contingency table, having  $I$  columns, and  $J$  rows and  $n_{ij}$  is the observed frequency. If the variables are independent,  $p_{ij} = p_{i+} \cdot p_{+j}$  for all  $i, j$

for which  $p_{i+} = \sum_{j=1}^J p_{ij}$ ,  $p_{+j} = \sum_{i=1}^I p_{ij}$ . Based on  $n$  independent observations,

substitution of the estimate  $\hat{p}_{ij} = \hat{p}_{i+} \hat{p}_{+j} = \frac{n_{i+}}{n} \cdot \frac{n_{+j}}{n}$  into (1) yields

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - n_{i+} \cdot n_{+j} / n)^2}{n_{i+} \cdot n_{+j} / n} \tag{2}$$

If the variables are independent, this quantity is chi-square variable with  $(I-1)(J-1)$  degrees of freedom. For an  $I \times J$  contingency table, let  $q = \text{Min}(I, J)$ . It can be shown that  $0 \leq \chi^2 \leq n(q-1)$  (Liebetrau, 1983). (2) is equivalent to

$$\varphi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(p_{ij} - p_{i+} \cdot p_{+j})^2}{p_{i+} \cdot p_{+j}} = \sum_{i=1}^I \sum_{j=1}^J \frac{p_{ij}^2}{p_{i+} \cdot p_{+j}} - 1 \tag{3}$$

Note  $0 \leq \varphi^2 \leq (q-1)$  with  $\varphi^2 = 0$ ; in case of independence; and  $\varphi^2 = q-1$ ; in case of perfect association. Alternatively, Pearson proposed

$$P = \left( \frac{\varphi^2}{1 + \varphi^2} \right)^{1/2} \tag{4}$$

This measure is also called Pearson contingency coefficient. The range of  $P$  depends on the

dimensions of the contingency table. Sakoda (Sakoda, 1976) suggests

$$P^* = \frac{P}{P_{max}} = \left( \frac{q \cdot \varphi^2}{(q-1) \cdot (1 + \varphi^2)} \right)^{1/2} \tag{5}$$

Tschuprow considered the following

$$t = \left( \frac{\phi^2}{\sqrt{(I-1) \cdot (J-1)}} \right)^{1/2} \tag{6}$$

Finally Cramér statistic is

$$v = \left( \frac{\phi^2}{q-1} \right)^{1/2} \tag{7}$$

which is 1; in case of perfect association; and 0; in case of dependence; for all values of J and I.

### 3. Entropy as a Measure of Variability

The entropy of a statistical experiment is a measure of uncertainty. Various entropy formulations of probability distributions and various entropy measures are summarized well by Mihalowicz et al. (Mihalowicz et al.,2014) and Esteban&Morales (Estaban and Morales, 1995).

If the discrete random variable X takes values  $x_1, x_2, \dots, x_K$  with respective probabilities  $p_1, p_2, \dots, p_K$  on sample space S, then Shannon entropy is defined as

$$H_s = -\sum_{i=1}^K p_i \cdot \ln p_i \tag{8}$$

Rényi entropy is a generalization of Shannon entropy. It is defined as

$$H_{Ré}(\alpha) = \frac{\ln \sum_{i=1}^K p_i^\alpha}{1-\alpha} \text{ for } \alpha > 0 \text{ and } \alpha \neq 1 \tag{9}$$

As  $\alpha \rightarrow 1$ , Renyi entropy approaches to Shannon entropy. Another generalized form of

Shannon entropy is due to Tsallis. It is defined as

$$H_{Tsa}(\alpha) = \frac{1 - \sum_{i=1}^K p_i^\alpha}{\alpha - 1}, \text{ for } \alpha > 0 \text{ and } \alpha \neq 1 \tag{10}$$

Tsallis entropy approaches to Shannon entropy as  $\alpha \rightarrow 1$ . Asymptotic properties of several entropy measures are given by Zhang (Zhang, 2013).

#### 4. Bivariate Distributions and Distance Measures

Suppose the pair of discrete random variables X and Y assume  $x_1, x_2, \dots, x_n$  and  $y_1, y_2, \dots, y_m$ . If the joint probability function is denoted by  $P_{X,Y}(x, y)$ , then Shannon entropy is

$$H_S(X, Y) = -\sum_{i=1}^n \sum_{j=1}^m P_{X,Y}(x_i, y_j) \cdot \log P_{X,Y}(x_i, y_j) \tag{11}$$

##### iii.1. Renyi Divergence and Kullback-Leibler Information

Rényi order- $\alpha$  divergence of q from p is defined as

$$D_{Ré}(p \parallel q) = \frac{1}{\alpha - 1} \log \sum \frac{p^\alpha}{q^{\alpha-1}} \tag{12}$$

Here p and q are two discrete probability distribution functions. When  $\alpha \rightarrow 1$

$$\lim_{\alpha \rightarrow 1} \frac{1}{\alpha - 1} \log \sum \frac{p^\alpha}{q^{\alpha-1}} = 0 \quad , \quad \lim_{\alpha \rightarrow 1} D_{Ré}(p \parallel q) = \lim_{\alpha \rightarrow 1} \frac{\frac{d}{d\alpha} \left( \log \sum \frac{p^\alpha}{q^{\alpha-1}} \right)}{\frac{d}{d\alpha} (\alpha - 1)} = \sum p \log \left( \frac{p}{q} \right) \tag{13}$$

The quantity,  $\sum p \log \left( \frac{p}{q} \right)$ , is called relative entropy. It is the divergence between two probability distributions, p and q. Tsallis order- $\alpha$  divergence of q from p is defined as

$$D_{Tsa}(p \parallel q) = \frac{1 - \sum \frac{p^\alpha}{q^{\alpha-1}}}{1 - \alpha} \tag{14}$$

##### iii.2. Bhattacharyya and Hellinger Distances

Alternatively, Bhattacharyya coefficient (B.C.), and Bhattacharyya distance between two distributions, p and q are defined as (Upton and Cook, 2006)

$$B.C. = \sum \sqrt{p \cdot q} \tag{15}$$

$$D_{Bha}(p \parallel q) = \text{Arccos}(B.C.) \tag{16}$$



Hellinger distance can also be calculated directly by Bhattacharyya coefficient;

$$D_{Hel}(p \parallel q) = \sqrt{2(1 - B.C.)} \tag{17}$$

### 4.3. Shannon Mutual Information as a Measure of Independence

Shannon mutual information (or Kullback-Leibler divergence) is defined as

$$I(X, Y) = \sum_{j=1}^m \sum_{i=1}^n P(X = x_i, Y = y_j) \log \frac{P(X = x_i, Y = y_j)}{P(X = x_i) \cdot P(Y = y_j)} \tag{18}$$

If X and Y are independent, it is zero. For bivariate normal distribution, it is

$$I(X; Y) = -\frac{\ln(1 - \rho^2)}{2} \tag{19}$$

Thus, the correlation coefficient can easily be estimated by

$$\hat{\rho} = \sqrt{1 - \exp(-2 \cdot \widehat{I(X; Y)})} \tag{20}$$

### 4.4. Bivariate Version of Rényi and Tsallis Divergences

Rényi order- $\alpha$  divergence of  $P_{X,Y}(x, y)$  from  $P_X(x)P_Y(y)$  is defined analogously,

$$D_{Re}(P_{X,Y}(x, y) \parallel P_X(x) \cdot P_Y(y)) = \frac{1}{\alpha - 1} \log \sum \sum \frac{P_{X,Y}(x, y)^\alpha}{(P_X(x) \cdot P_Y(y))^{\alpha-1}} \tag{21}$$

It is also called as Rényi mutual information. Tsallis order- $\alpha$  divergence is

$$D_{Tsa}(P_{X,Y}(x, y) \parallel P_X(x) \cdot P_Y(y)) = \frac{1 - \sum \sum \frac{P_{X,Y}(x, y)^\alpha}{(P_X(x) \cdot P_Y(y))^{\alpha-1}}}{1 - \alpha} \tag{22}$$

In case of independence, Rényi and Tsallis mutual informations are zero as Shannon mutual information. As  $\alpha \rightarrow 1$ , Rényi and Tsallis mutual informations approach to Shannon mutual information.

### 4.5. Bhattacharyya and Hellinger Measures for Testing independence

Bivariate versions of Bhattacharyya and Hellinger distances are straightforward:

$$D_{Bha}(P_{X,Y}(x,y) || P_X(x).P_Y(y)) = \text{Arccos} \left( \sum \sum \sqrt{\frac{P_X(x).P_Y(y)}{P_{X,Y}(x,y)}} P_{X,Y}(x,y) \right) \quad (23)$$

$$D_{Hel}(P_{X,Y}(x,y) || P_X(x).P_Y(y)) = \sqrt{2 \left( 1 - \sum \sum \sqrt{\frac{P_X(x).P_Y(y)}{P_{X,Y}(x,y)}} P_{X,Y}(x,y) \right)} \quad (24)$$

### 5. Experimental Results

In simulations, 8 discrete bivariate distributions are considered. 5 of them are 2x2, the rest are 3x4 contingency tables. The variables are assumed to be continuous for making comparisons between information-theoretic independence measures and classic correlation measures. The eight bivariate distributions presented are given below:

Table 1.1. 2x2 tables used in simulations having correlations 0.033, 0.88 and -0.88.

Distribution1			Distribution2			Distribution3		
X/Y	1	2	X/Y	1	2	X/Y	1	2
1	0.5	0.4	1	0.5	0.03	1	0.03	0.5
2	0.05	0.05	2	0.03	0.44	2	0.44	0.03
Correlation	0.033		Correlation	0.88		Correlation	-0.88	

Table 1.2 2x2 tables used in simulations having correlations 0.49 and -0.49.

Distribution4			Distribution5		
X/Y	1	2	X/Y	1	2
1	0.45	0.15	1	0.1	0.3
2	0.1	0.3	2	0.45	0.15
Correlation	0.49		Correlation	-0.49	

*Table 1.3 3x4 tables used in simulations having correlations -0.86 and 0.012.*

Distribution6					Distribution7				
X/Y	1	2	3	4	X/Y	1	2	3	4
1	0.01	0.01	0.01	0.27	1	0.1	0.1	0.05	0.05
2	0.01	0.01	0.37	0.01	2	0.1	0.1	0.15	0.05
3	0.27	0.01	0.01	0.01	3	0.12	0.04	0.1	0.04
Correlation	-0.86				Correlation	0.012			

*Table 1.4 3x4 table used in simulations with correlation 0.86.*

Distribution8				
X/Y	1	2	3	4
1	0.27	0.01	0.01	0.01
2	0.01	0.37	0.01	0.01
3	0.01	0.01	0.01	0.27
Correlation	0.86			

1000 independent observations are picked from each distribution. This procedure is repeated 1000 times. Then, the same experiments are repeated with 2000 independent observations, 2000 times. Simulations are realized by the help of Microsoft Excel.

Large sampling properties of 19 association measures studied are given below.

*Table 2. The goodness of fit statistics used in determining associations.*

Measure	Explanation	Eqn. No.	Measure	Explanation	Eqn. No.
I(X;Y)	Mut.Info (Shannon)	18.	R2	Mut. Info (Rényi) $\alpha=2$	21.
Chi-square	Chi-square	3.	T2	Mut. Info (Tsallis) $\alpha=2$	22.
Hellinger	Hellinger Distance	24.	R3	Mut. Info (Rényi) $\alpha=3$	21.
Bhattacharyya	Bhattacharyya Distance	23.	T3	Mut. Info (Tsallis) $\alpha=3$	22.

R0.25	Mut.Info (Rényi) $\alpha=0.25$	21.	Cor. meas.	Correlation measure	20.
T0.25	Mut. Info (Tsallis) $\alpha=0.25$	22.	Pears. cc.	Pearson cont. coef.	4.
R0.5	Mutual Info (Rényi) $\alpha=0.5$	21.	Sakoda	Sakoda coef.	5.
T0.5	Mut. Info (Tsallis) $\alpha=0.5$	23.	Tschupr.	Tschuprow coef.	6.
R1.5	Mut.Info (Rényi) $\alpha=1.5$	21.	Cramér	Cramér coef.	7.
T1.5	Mut. Info (Tsallis) $\alpha=1.5$	22.			

### 5.1. Normality results

The normality of each statistic has been investigated. The seven normality tests used are Shapiro Wilk  $W$ , Anderson Darling, Martinez Iglewicz, Kolmogorov-Smirnov, D’Agostino Skewness, D’Agostino Kurtosis, D’Agostino Omnibus tests. The percentage of normality observed by simulations, for each statistic is given in Table 3. In general, the tendency to distribute normal is obvious with few exceptions, namely, Tsallis and Rényi mutual informations with  $\alpha = 0.25$  and  $\alpha = 0.50$  values.

Table 3. The tendency to normality of each statistics

Measure	% of Normality in Runs	Measure	% of Normality in Runs
I(X;Y)	75	T1.5	75
Chi-square	75	R2	75
Hellinger	87.5	T2	75
Bhattacharyya	87.5	R3	81.25
R0.25	0	T3	75
T0.25	12.5	Cor. meas.	81.25
R0.5	18.75	Pears. cc.	62.5
T0.5	31.75	Sakoda	75
R0.75	56.25	Tschuprow	81.25
T0.75	62.5	Cramér	81.25
R1.5	75		

### 5.2. Sampling variabilities and average correlations

Minimum standard deviations, and maximum average correlations (with the other association measures) are presented in Table 4. Pearson contingency coefficients, and T0.5 have generally the lowest variability. Mutual information, chi-square statistic, and T2 have maximum average correlations with others.

*Table 4. Minimum standard deviations, maximum average correlations*

Distribution	Population Correlation	Min. Std. Deviation	Measure	Maximum Ave. Corr.	Measure
1(1000x1000)	0.033	0.0005	T0.25	0.9134	I(X;Y)
1(2000x2000)	0.033	0.0003	T0.25	0.8914	I(X;Y)
2(1000x1000)	0.88	0.0063	Pears.cc	0.9963	Chi-sq(=T2).
2(2000x2000)	0.88	0.0045	Pears. cc.	0.9972	Chi-sq(=T2).
3(1000x1000)	-0,88	0.0063	Pears.cc	0.9974	Chi-sq(=T2).
3(2000x2000)	-0,88	0.0046	Pears. cc.	0.9974	Chi-sq(=T2).
4(1000x1000)	0.49	0.0045	T0.25	0.9981	R2
4(2000x2000)	0.49	0.0046	Pears. cc.	0.9974	Chi-sq(=T2).
5(1000x1000)	-0.49	0.0044	T0.25	0.998	I(X;Y)
5(2000x2000)	-0.49	0.003	T0.25	0.998	I(X;Y)
6(1000x1000)	-0.86	0.0019	Cramér	0.9826	I(X;Y)
6(2000x2000)	-0.86	0.0013	Pears. cc.	0.9856	I(X;Y)
7(1000x1000)	0.012	0.0025	T0.25	0.9916	R1.5
7(2000x2000)	0.012	0.0016	T0.25	0.9901	I(X;Y)
8(1000x1000)	0.86	0.0019	R0.25	0.9857	Chi-sq(=T2).
8(2000x2000)	0.86	0.0014	Cramér	0.9873	Chi-sq(=T2).

### 5.3. Estimating correlations by mutual information statistics

Correlation coefficients are estimated by plugging Shannon, Rényi and Tsallis mutual information quantities in (20). The best estimates for correlation coefficient, and related statistics are given in Table 5. All mutual information statistics are nonnegative; therefore, they only give information about the magnitudes of correlations. The distributions represented by 8 contingency tables are not bivariate normal, but equation (20) is still used intentionally. As a general tendency, for lower population correlations, Tsallis and Rényi estimates with low  $\alpha$  parameters are observed to be successful. Similarly, for higher population correlations, selecting higher  $\alpha$  parameters may work. For moderate correlations, Shannon mutual information performs well.

*Table 5. Correlation estimates for 8 bivariate tables; normality assumption is violated.*

Distribution	Population Correlation	Estimated Magnitude of Cor.	Best Measure
1(1000x1000)	0.033	0.036	T0.5
1(2000x2000)	0.033	0.031	R0.5
2(1000x1000)	0.88	0.928	R3
2(2000x2000)	0.88	0.928	R3
3(1000x1000)	-0,88	0.887	R2
3(2000x2000)	-0,88	0.887	T2
4(1000X1000)	0.49	0.473	I(X;Y)
4(2000x2000)	0.49	0.471	I(X;Y)
5(1000x1000)	-0,49	0.472	I(X;Y)
5(2000x2000)	-0,49	0.473	I(X;Y)
6(1000x1000)	-0,86	0.864	R0.75
6(2000x2000)	-0,86	0.863	R0.75
7(1000x1000)	0.012	0.144	T0.25
7(2000x2000)	0.012	0.14	T0.25
8(1000x1000)	0.86	0.864	R0.75
8(2000x2000)	0.86	0.864	R0.75

#### 5.4. Correlation estimates for bivariate normal distribution

Correlation estimates by various mutual information statistics, in case of bivariate normality, are made by simulating 1000 pairs. Simulations are realized by Microsoft Excel. Results are in Table 6.

*Table 6. Correlation estimates for bivariate normal*

Population Correlation	Absolute Value of Best Correlation Estimate (9 groups)	Absolute Value of Best Correlation Estimate (16 groups)
0	0.047(R0.25)	0.054(T0.25)
0.5	0.48(T2)	0.499(T2)
-0,5	0.521(R2)	0.501(R1.5)
0,75	0.754(T2)	0.763(R2)
-0,75	0.741(R2)	0.771(R2)

The number of groups used in discretizing normal data seems to be effective on mutual information statistics. In order to see this tendency, 9 and 16 groups are used alternatively. If a mutual information statistic underestimates correlation, two strategies may work: **i)** increasing the

magnitude of  $\alpha$ , or **ii**) increasing the number of categories of each variable in discretizing bivariate normal.

## 6. Discussion

Zografos (Zografos, 1993) gives asymptotic distributions of  $\varphi$  divergences whose special cases are Kullback -Leibler divergence, Rényi order- $\alpha$  divergence, etc. He has shown that the asymptotic distribution of this statistic is either normal, or a linear form in chi-square variables. Agresti (Agresti, 2002) mentions asymptotic normality property of functions of counts of a multinomial distribution. But, selecting  $\alpha$  seems to be decisive in observing normality of mutual information statistics. Smaller  $\alpha$  selections like 0.25 or 0.5 probably prevent normality. But these  $\alpha$  values yielded the measures with least variations, irrespective of the scales of measurement of the variables. Behind this, some mutual information statistics, having  $\alpha$  values near 2, showed the maximum average correlation with other independence statistics (They inform well about the dependencies of variables!). Finally, if the variables are continuous, or bivariate normal, Rényi and Tsallis mutual informations succeeded in estimating correlations with proper selections of  $\alpha$ . Smaller  $\alpha$  will probably yield smaller correlation estimates, whereas higher  $\alpha$  will yield higher correlation estimates.

## References

- A.I. KHINCHIN, Mathematical Foundations of Information Theory, Dover Publications, Inc., New York, USA, 1957.
- Abraham, WALD, Sequential Analysis, Dover Phoenix Editions, Dover Publications, Inc., Mineola, New York, USA, 2004.
- Alaattin ŞAMİLOV, Entropi, İnfomasyon ve Entropi Optimizasyon, Nobel Akademik Yayıncılık Eğitim Danışmanlık Tic. Ltd. Şti., Ankara, Turkey, 2015.
- Alan AGRESTI, Categorical Data Analysis, Second Edition, Wiley-Interscience, New Jersey, USA, 2002.
- Albert M. LIEBETRAU, Measures of Association, Series: Quantitative Applications in the Social Sciences, A Sage University Paper, 32, Sage Publications, The International Professional Publishers, USA, 1983.
- Alfréd RENYI, Probability Theory, Dover Publications, Inc., Mineola, New York, USA, 2007.
- Aman ULLAH, Entropy, Divergence and Distance Measures with Econometric Applications, Journal of Statistical Planning and Inference, 49(1996), 137-162.
- Claude E. SHANNON, Warren WEAVER, The Mathematical Theory of Communication, University of Illinois Press, Urbana and Chicago, USA, 1963.
- D.V. GOKHALE, Solomon KULLBACK, The Information in Contingency Tables, Marcel Dekker, Inc., New York, USA, 1978.
- Graham UPTON, Ian COOK, A Dictionary of Statistics, Second edition, Oxford University Press, New York, USA, 2006.
- José C PRINCIPE, Information Theoretic Learning, Rényi's Entropy and Kernel Perspectives, Springer, USA, 2010.
- Joseph Victor MIHALOWICZ, Jonathan V. NICHOLS, Frank BUCHOLTZ, Handbook of Differential Entropy, CRC Press, Taylor&Francis Group, USA, 2014.
- K ZOGRAFOS, Asymptotic Properties of  $\chi^2$  Divergence Statistic and Its Applications in Contingency Tables, Int. J. Math. & Stat. Sci., Vol. 2, No. 1, June 1993, 5-22.
- Leandro PARDO, Statistical Inference Based on Divergence Measures, Chapman&Hall/CRC, Taylor&Francis Group, New York, USA, 2006.
- M.D. ESTEBAN, D. MORALES, A summary on entropy statistics, Kybernetika, Vol.1(1995), No.4, 337-346.
- Robert B. ASH, Information Theory, Dover Publications, Inc., New York, USA, 1990.



- Sakoda, J. (1976). Measures of association for multivariate contingency tables. In Social Statistics Section, Proceedings of the American Statistical Association (pp. 777-780)
- Solomon KULLBACK, Information Theory and Statistics, Dover Publications, Inc., Mineola, New York, USA, 1997.
- Thomas M. COVER, Joy A. THOMAS, Elements of Information Theory, Second Edition, Wiley-Interscience, USA, 2006.
- Vijay, P. SINGH, Entropy-Based Parameter Estimation in Hydrology, Kluwer Academic Publishers, Dordrecht, Netherlands, 1998.
- Xing ZHANG, Asymptotic Normality of Entropy Estimators, The University of North Carolina at Charlotte, [http://math.uncc.edu/sites/math.uncc.edu/files/2013\\_03.pdf](http://math.uncc.edu/sites/math.uncc.edu/files/2013_03.pdf)

## OpenCV and DeepFace Approach on Image Processing: Applications on Street Interviews

Muzaffer Göztaş<sup>1</sup>

Doğan Yıldız<sup>2</sup>

### Abstract

Using Python-based face recognition technology, this work aims to forecast characteristics including age, gender, and moods of individuals shown in video uploaded on the YouTube platform from 2017 to 2021. This study particularly makes use of this method on individuals engaged in YouTube street interviews, focusing on issues related to the unemployment and economy.

The concordance and distribution of emotional variables alongside demographic variables such as age and gender were investigated using the “correspondence” analysis method, a multivariate technique that visualizes relationships among variables and categorical cross-relationships through graphical mapping.

Data concerning the emotional variance from every year between 2017 and 2021 were gathered. Data were especially separated, and unhappiness rates were computed as percentages from the gathered data. The study looked at the relationship between the unemployment rate for the particular years and the computed degree of discontent. Non-parametric statistical methods were seen more suitable in this phase of the research as the low data volume made normalcy testing impossible.

The value of facial recognition systems in demographic prediction and emotional evaluation is underlined by this work. Furthermore, it emphasizes the possible value of matching emotional forecasts with unemployment data in order to assess how economic changes affect personal emotional responses.

---

1 Researcher Student, Yildiz Technical University, muzaffer.goztas@std.yildiz.edu.tr, 0009-0001-7933-901X

2 Asst. Prof. Dr. Doğan Yıldız, Yildiz technical University, dyildiz@yildiz.edu.tr, 0000-0001-7430-2368

This study emphasizes how improvements in artificial intelligence and data analysis could improve the understanding and answers for society issues. Therefore, it could motivate similar future study with great benefits in marketing, economics, and social sciences as well as highlight developments in data analysis and artificial intelligence.

## 1. Introduction

The study, which aims to determine demographic characteristics such as age, gender, and mood by analyzing the facial expressions of individuals featured in YouTube street interviews from 2017 to 2021, seeks to reveal the relationships between emotional analysis based on facial expressions and social variables such as the unemployment rate during the same period. In the literature review, studies on the use of facial recognition technology in demographic and emotional predictions are examined, and in the originality section, the uniqueness of the study is emphasized within the framework of these findings.

### 1.1. Literature Review:

Facial recognition technology for demographic predictions has garnered interest in artificial intelligence and image processing fields. A notable study by Levi and Hassner (2015) developed a model using convolutional neural networks (CNN) for age and gender prediction, achieving high accuracy rates on a large-scale facial dataset. This work highlights CNN's effectiveness in demographic analysis and classification accuracy for age and gender (Levi & Hassner, 2015).

In emotion analysis, transfer learning and deep learning are frequently used to attain high accuracy on smaller datasets. Sabri and El-Bakry (2021) applied models pre-trained on large datasets to perform facial expression analysis on limited data, demonstrating that transfer learning enables effective emotion predictions even with small datasets. This study emphasizes transfer learning's importance in emotion analysis when large datasets are unavailable (Sabri & El-Bakry, 2021).

### 1.2. Significance & Originality of the Study

Facial recognition systems play an important role in measuring responses to societal emotional rates and examining correlations with various other variables. The DeepFace detection and recognition model used here can be further enhanced by integrating additional parameters, thereby improving the model's accuracy beyond the current 76% prediction success rate, calculated based on the average Female and Male F1 model scores shown in

Figure 13. Additionally, the emotion and age range correspondence analysis shown in Figure 15 can be incorporated with a broader perspective into the correlation analysis in Tables 8 and 9. This integration would allow the development of a predictive model focused on the impact of age-related unemployment and its emotional reflections. When we finally compare the model which we used in this study with studies above, our 76% accuracy rate of DeepFace is considerably higher compared to the transfer learning model by Sabri and El-Bakry, but it falls short of Levi and Hassner’s gender prediction accuracy of 86%. In this context, DeepFace offers broader and more balanced accuracy, whereas Levi and Hassner’s model stands out as more successful specifically in gender prediction.

## 2. Methodology and Results

Face structures from street interviews are identified in this work using Python’s OpenCV module and Deep Face package. Comparative with unemployment statistics, statistical analyses assess model accuracy, prediction, and consistency. The paper also looks at how face traits affect emotional states, gender, and age. Using F1 score, accuracy, and recall criteria, the paper assessed the expected accuracy of the model by contrasting actual gender values with expected ones. To evaluate model performance, a chi-square analysis was done on the correlations among categorical variables—more especially, gender, age, emotion.

By means of correspondence analysis, the study investigated data structures and relationships among variables, therefore clarifying their links. Using a non-parametric correlation approach, it contrasted negative emotions from OpenCV and DeepFace libraries with unemployment rates. Based on interviews posted on social media over years, the study concentrated on sociological and psychological issues pertinent to daily living.

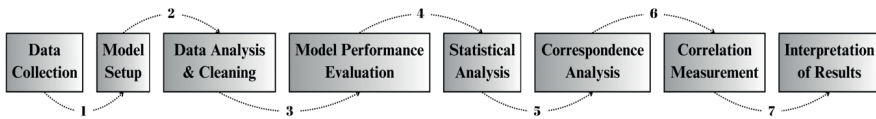


Figure 1: Summarized Workflow for This Study

### 2.1. Recognizing the Core Ideas of the Model Type

In this section, we defined and investigated the fundamentals of our code architecture.

### 2.1.1. Importing the Libraries

With the OpenCV library being the main choice, the Python programming environment was utilized in the data collection phase for its wide range of libraries fit for visual analysis (GitHub, n.d.). Figure 2 shows that at the start of the working process, the necessary libraries were imported into the program environment.

```
import cv2
from deepface import DeepFace
import os
import pandas as pd
```

*Figure 2: Python Code of Essential Libraries for the Software*

OpenCV, or cv2: a Python tool for image processing.

Deep Face is a Python tool for facial detection and analysis.

os: An operating system function execution library.

Pandas: a broad Python tool for data manipulation and analysis.

### 2.1.2. Model Importation

Using a pre-trained facial recognition model (shown in Figure 3) along with Python's facilities, one may ensure exact results from visual data. Using this model was meant to improve data acquisition accuracy.

The next phase will be visual example assessment of the prediction accuracy of the model.

```
face_cascade = cv2.CascadeClassifier('dosya_yolu.xml')
```

*Figure 3: Python Code of Integrating the Model into the System*

### 2.1.3. Output Storage

A Pandas DataFrame was established to store the results, as illustrated in Figure 4.

```

result_df = pd.DataFrame(columns=[
    'YEAR', 'ORDER', 'FEMALE PREDICTION RATE', 'MALE PREDICTION RATE',
    'ANGRY RATE', 'DISGUST RATE', 'FEAR RATE', 'HAPPY RATE',
    'SAD RATE', 'SURPRISE RATE', 'NEUTRAL RATE',
    'DOMINANT GENDER', 'DOMINANT EMOTION', 'AGE PREDICTION'
])

```

*Figure 4: Python Code of Transferred Variables to the Data Frame*

#### 2.1.4. Importing the Image

The import process was carried out with the code shown in Figure 5; the directory of all images for analysis was set aside.

```

image_folder = '/path/to/your/image/folder/'
image_files = os.listdir(image_folder)

```

*Figure 5: Python Code of Importing Visual Content*

#### 2.1.5. Visual Data Manipulation

As Figure 6 shows, a loop was developed to handle every image file (Tekin, M. (n.d.)).

```

for image_file in image_files:
    if image_file.endswith('.png'):

```

*Figure 6: Python Code of Processing Cycle*

The photos in color format were loaded in line with the year using the code in Figure 7.

```

yil, sira = image_file.split('.')[0].split('-')
yil, sira = int(yil), int(sira)

input_image = cv2.imread(os.path.join(image_folder, image_file))

```

*Figure 7: Python Code of Color Loading by Year*

### 2.1.6. Grayscale Technique for Face Detection

I found faces in the pictures using the OpenCV package's `detectMultiScale()` technique. This system uses a multi-scale approach and grayscale photo processing to improve speed and lower computational requirements (Tekin, M. (n.d.)).

```
faces = face_cascade.detectMultiScale(
    gray,
    scaleFactor=1.1,
    minNeighbors=5,
    minSize=(30, 30)
)
```

*Figure 8: Python Code of Grayscale Methodology*

### 2.1.7. Outputs for Variable Extraction Using DeepFace

After detecting the faces, we used the Deep Face library to evaluate the gender, age, and emotions of every found face. These variables are extracted using the following code:

```
result_emotion = DeepFace.analyze(input_image, actions=['emotion'])
result_age = DeepFace.analyze(input_image, actions=['age'])
result_gender = DeepFace.analyze(input_image, actions=['gender'])
```

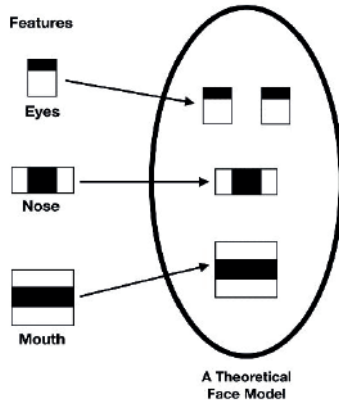
*Figure 9: Python Code of Variable Extraction Utilizing DeepFace*

### 2.1.8. The Fundamental Mathematical Framework

DeepFace uses advanced deep learning techniques to perform facial recognition, evaluating attributes like age, gender, and emotional expression. It classifies emotions based on facial features and accurately estimates age and gender.

### 2.1.9. The Underlying Mathematical Framework

The foundation of Deep Face's approach is Convolutional Neural Network (CNN) application. The model extracts many components at different levels -including edges, textures, and complex patterns- as an input image passes several convolutional layers which are showed in Figure 10 (Kaur, G., & Kaur, P. (2020)). The obtained features are arranged into feature maps that, with each next layer, record even more intricate properties.



*Figure 10: HAAR Face Detection Sample*

## 2.2. Accuracy of Model Predictions on Gender Variable

Understanding accuracy and success rates helps one evaluate machine learning models by means of real-world data comparison of model prediction accuracy versus. These steps show whether a model can provide accurate projections in relevant contexts for making decisions. Development of models depends on improvement and modification; they assist to pinpoint areas that need work and help to avoid false positives and false negatives. Model success rates measure the quality of a dataset, and if a model underperforms, it requires close examination for potential flaws or errors. Several metrics are used to compare and compare the performance of various models, identifying the model with the best results.

## 2.3. Confusion Matrix

Examining the dependent variables in the dataset usually helps one to estimate successful predictions from our model using classification algorithms by means of accuracy rates. In this case, the performance measures of the model were calculated solely depending on the Gender variable without considering any dependent or independent variables (Encord. (n.d.)).



		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	<b>Sensitivity</b> $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	<b>Specificity</b> $\frac{TN}{(TN + FP)}$
		<b>Precision</b> $\frac{TP}{(TP + FP)}$	<b>Negative Predictive Value</b> $\frac{TN}{(TN + FN)}$	<b>Accuracy</b> $\frac{TP + TN}{(TP + TN + FP + FN)}$

Figure 11: Details of the Confusion Matrix

### 2.3.1. Components of the Confusion Matrix

#### 2.3.1.1. True Positive (TP)

This indicates cases when the model correctly recognized the positive class. The model fairly projected a positive class; the actual class was positive. For example; A real positive results from a medical test correctly identifying a patient with an ailment.

#### 2.3.1.2. False Positive (FP)

This happens when the model misfits the positive class. Although the model labeled it as positive, the actual class was negative. Sometimes this is referred to as a “type I error”. For example; A false positive in medicine is the result of a test suggesting, mistakenly, that a healthy person has a disease.

#### 2.3.1.3. True Negative (TN)

This happens when the model correctly spots the negative class. The true class is negative as is the prediction of the model. A real negative would be the outcome of a medical test, for example, if it correctly marks a healthy patient as free of the condition.

#### 2.3.1.4. False Negative (FN)

False negative transpires when the model mistakenly forecasts the negative class. Though the model assigned it as negative, the actual class was

positive. We call this a “type II error.” For example, this happens when a medical test misses an ill patient, so falsely suggesting that they are healthy.

### 2.3.2. Performance Metrics Obtained from the Confusion Matrix

In this section of the study, the focus is on calculating model score metrics.

#### 2.3.2.1. Sensitivity (Recall or True Positive Rate)

This assesses the model’s capacity to accurately detect positive instances. It is computed as:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)$$

Sensitivity equals true positives divided by the sum of true positives and false negatives. High sensitivity indicates that the model accurately detects true positive cases.

#### 2.3.2.2. Specificity

Specificity is defined as TN divided by the sum of TN and FP.

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

High specificity indicates that the model excels at identifying real negatives

#### 2.3.2.3. Precision (Positive Predictive Value)

Precision denotes the proportion of projected positive situations that are genuinely positive. Precision is computed as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

High precision indicates that when the model predicts a positive class, it is typically accurate.

#### 2.3.2.4. Negative Predictive Value (NPV)

This measure reflects the proportion of projected negative situations that are genuinely negative. It is computed as:

$$NPV = \frac{TN}{TN + FN} \quad (4)$$

A high NPV indicates that when the model forecasts a negative class, it is typically precise.

#### 2.3.2.5. Accuracy

Accuracy quantifies the overall correctness of a model's predictions, indicating the ratio of true results (including both true positives and true negatives) to the total number of forecasts. Accuracy is computed as:

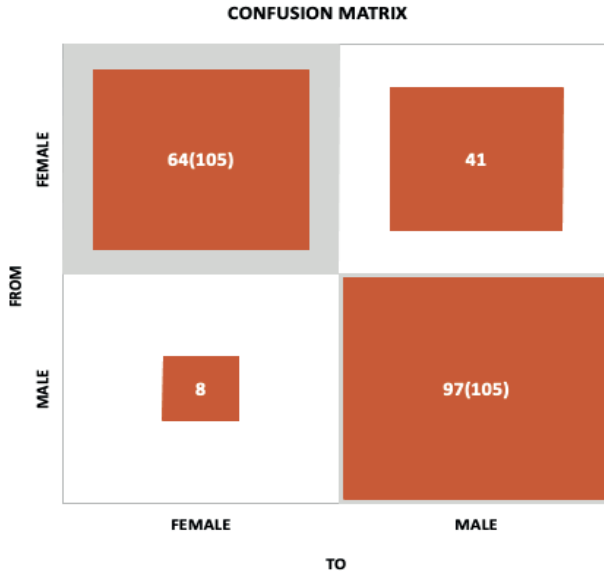
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

The high accuracy indicates that the model effectively identifies both positive and negative cases accurately.

Each statistic offers a distinct viewpoint on the model's performance, facilitating the assessment of its strengths and limitations in prediction accuracy.

### 2.3.3. Illustration of a Confusion Matrix Utilizing Our Data

The objective observability of the gender variable for this evaluation drives much of the choice of it. Unlike emotions or age, which depend on the subjective interpretation of the researcher, gender may be ascertained scientifically, therefore guaranteeing more accurate and significant findings. Figure 12 exemplifies the matrix comprising the parameter values for assessing the performance of any model.

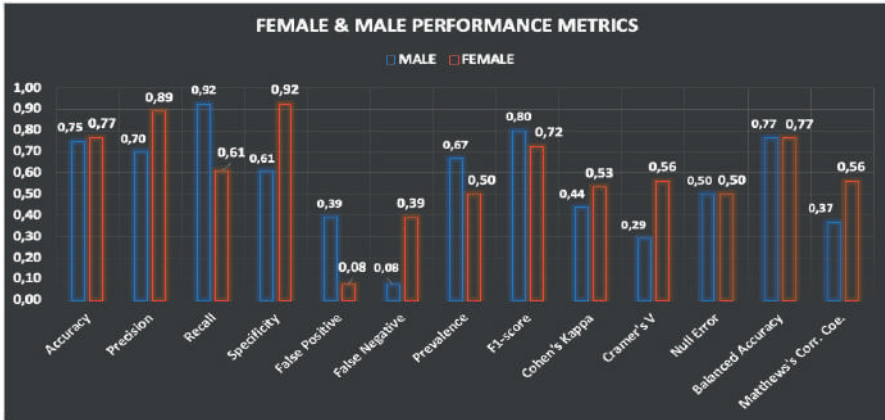


*Figure 12: Sample Confusion Matrix for Gender*

Analyzing distinct Precision and Recall metrics for the “FEMALE” and “MALE” categories in a gender prediction model facilitates a more comprehensive assessment of the algorithm’s efficacy for each class. Furthermore, juxtaposing the F1-Scores for both groups facilitates a more substantive study. The model’s success rates, derived from the data in Figure 12, produced the confusion matrix shown in Figure 11 (Chakravorty, D. (n.d.)). A comparative analysis of both male and female categories has been conducted using bar charts on the subsequent page.

### *2.3.3.1. Model Accurate Metrics*

Figure 13 below presents the comparative success rates for cases in which the positive condition is evaluated for both males and females.



*Figure 13: Comparative Analysis of Model Performance Metrics*

While each measurement parameter is crucial, the overall efficacy of a model's predictions derived from observational data is more accurately assessed using the F1 score.

The comparison of the F1 score values reveals that F1-Score for Female = 0.72 is less than F1-Score for Male = 0.80, indicating superior predictive capability for males over females by the model. Consequently, in test groups when the positive condition applies to both genders, the model demonstrates superior efficacy in predicting males compared to girls.

## 2.4. Comprehending the Interconnections Among Categorical Variables

The Chi-Square ( $\chi^2$ ) test is a non-parametric statistical method utilized to analyze the association between categorical variables. Analogous to parametric hypothesis testing, the aim is to formulate a hypothesis and ascertain its potential rejection.

The Chi-Square distribution, similar to the t-distribution, possesses only one degree of freedom. Generally, at lower degrees of freedom, the Chi-Square distribution exhibits right skewness. As the degrees of freedom rise, the Chi-Square distribution increasingly resembles the shape of a normal distribution curve.

### 2.4.1. Chi-Square Independence Test

The Chi-Square Independence Test assesses the independence or relationship between variables organized in 2x2 or rxc contingency tables.

This test is utilized to analyze relationships between qualitative (non-numerical) factors present in cross-tabulations (contingency tables).

The Pearson Chi-Square test, Yates' corrected Chi-Square test, and Fisher's exact test are applicable for 2x2 tables. The Pearson Chi-Square test is utilized for rxc tables. The degrees of freedom are determined by the formula  $(r-1) \times (c-1)$ , where r denotes the number of rows and c is the number of columns (Çorba, B. Ş. (n.d.)).

The hypotheses for the test are as follows:

$H_0$ : The variables are independent (there is no relationship between the variables).

$H_1$ : The variables are not independent (there is a relationship between the variables).

Row	Column		Total
	C1	C2	
R1	A A'	B B'	N1=A+B
R2	C C'	D D'	N2=C+D
Total	N3=A+C	N4=B+D	N

Figure 14: Illustrative Chi-Square (Contingency) Table

$$\chi^2 = N \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \tag{6}$$

In this context,  $n_{ij}$  represents the observed frequencies (A, B, C, D), while  $e_{ij}$  (A', B', C', D') symbolizes the predicted frequencies. The anticipated frequencies are calculated by multiplying the respective row and column totals and thereafter dividing by the grand total. If  $\chi^2 \leq \chi^2_{\alpha,df}$  or  $p > 0.05$ , the null hypothesis ( $H_0$ ) cannot be rejected, indicating that there is no significant difference between observed and expected frequencies. Otherwise, the null hypothesis is rejected, implying a significant difference.

#### 2.4.1.1. Relationship Between Gender and Emotional Response

At this juncture, the correlation between the two variables was analyzed utilizing the Chi-Square test. Due to certain cells in the contingency Figure 14, Fisher's Exact test statistics were utilized.

Hypotheses:

$H_0$ : There is no significant relationship between gender and the response given to the question asked or the problem presented in the street interviews.

$H_1$ : There is a significant relationship between gender and the response given to the question asked or the problem presented in the street interviews.

*Table 1: SPSS Output of Gender & Sentimental Reaction Crosstab*

		GENDER & SENTIMENTAL REACTION CROSTAB							
		SENTIMENTAL REACTION						Total	
		Happy	Sad	Fear	Angry	Neutral	Disgust		
GENDER	Male	Count	16	33	9	9	29	1	97
	Expected Count	20.5	33.7	10.8	6.0	25.3	.6	97.0	
	Female	Count	18	23	9	1	13	0	64
	Expected Count	13.5	22.3	7.2	4.0	16.7	.4	64.0	
Total	Count	34	56	18	10	42	1	161	
	Expected Count	34.0	56.0	18.0	10.0	42.0	1.0	161.0	

As the cells in the cross-table shown in Table 1 exhibit values below 5 for both observed and expected frequencies, deriving conclusions from the Chi-Square statistics in Table 2 may diminish the model's significance. Consequently, it is essential to consult the statistics derived by Fisher's Exact test.

*Table 2: SPSS Output of Gender & Sentimental Reaction Chi-Square Statistics*

Chi-Square Statistics ( $\chi^2$ ) For Gender & Sentimental Reaction						
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	9.013	5	.109	.092		
Likelihood Ratio	10.100	5	.072	.077		
Fisher's Exact Test	8.919			.091		
Linear-by-Linear Association	5.012	1	.025	.027	.014	.003

Table 2 indicates that the p-value from Fisher's Exact Test statistic above the alpha threshold of 0.05. Consequently, we are unable to dismiss the null hypothesis ( $H_0$ ). No substantial correlation was identified between Gender and Emotional Response. Nonetheless, this is not invariably true, as outcomes may fluctuate based on the quantity of observations in our dataset.

#### *2.4.1.2. Gender & Response Tendency Relationship*

To assess the correlation between gender and emotion from an alternative viewpoint, and to reduce errors by ensuring that the observed and expected

values in the cells surpass 5, the responses (i.e., emotions) were categorized into positive, negative, and neutral emotional states and analyzed from a tendency perspective.

In this instance, our revised hypotheses should be articulated as follows:

$H_0$ : There is no significant relationship between gender and the tendency of the response given to the question asked or the problem presented in the street interviews.

$H_1$ : There is a significant relationship between gender and the tendency of the response given to the question asked or the problem presented in the street interviews.

*Table 3: SPSS Output of Gender & Reaction Tendency Crosstab*

		REACTION TENDENCY			Total	
		Positive Reaction	Negative Reaction	No Reactio		
GENDER	Male	Count	16	52	29	97
	Expected Count	20.5	51.2	25.3	97.0	
	Female	Count	18	33	13	64
	Expected Count	13.5	33.8	16.7	64.0	
Total	Count	34	85	42	161	
	Expected Count	34.0	85.0	42.0	161.0	

To get more significant results and minimize errors, we can conclude that there is no substantial correlation between gender and emotional response, as indicated by the statistics in Table 4.

Table 4 indicates that the Pearson Chi-Square statistic, specifically the calculated p-value, is less than the alpha value of 0.05. It can be asserted that there is no substantial correlation between the two categorical variables.

The Pearson Chi-Square Test assesses the statistical significance of the disparity between observed and expected frequencies, whereas the Likelihood Ratio evaluates model fit and identifies the superior model among various alternatives.

*Table 4: SPSS Output of Gender & Reaction Tendency Chi-Square Statistics*

Chi-Square Statistics ( $\chi^2$ )			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	3.858	2	.145
Likelihood Ratio	3.841	2	.147
Linear-by-Linear Association	3.673	1	.055



### 2.4.1.3. Age Range & Emotional Response Relationship

The association between the two variables was analyzed using the Chi-Square test at this step. Due to the majority of cells in the contingency table containing values below 5, Fisher's Exact test and Monte Carlo statistics were deemed appropriate.

Hypotheses:

$H_0$ : There is no significant relationship between the age range and the response given to the question asked or the problem presented in the street interviews.

$H_1$ : There is a significant relationship between the age range and the response given to the question asked or the problem presented in the street interviews.

*Table 5: SPSS Output of Age Range & Sentimental Reaction Crosstab*

		AGE_RANGE & SENTIMENTAL REACTION CROSTAB							
		SENTIMENTAL REACTION						Total	
		Happy	Sad	Fear	Angry	Neutral	Disgust		
AGE RANGE	20-24	Count	2	4	0	0	1	0	7
		Expected Count	1.5	2.4	.8	.4	1.8	.0	7.0
	25-29	Count	7	9	1	1	12	0	30
		Expected Count	6.3	10.4	3.4	1.9	7.8	.2	30.0
	30-34	Count	18	25	5	3	15	0	66
		Expected Count	13.9	23.0	7.4	4.1	17.2	.4	66.0
	35-39	Count	5	8	8	2	7	0	30
		Expected Count	6.3	10.4	3.4	1.9	7.8	.2	30.0
	40-44	Count	1	4	1	2	6	0	14
		Expected Count	3.0	4.9	1.6	.9	3.7	.1	14.0
	45-49	Count	0	6	3	1	1	1	12
		Expected Count	2.5	4.2	1.3	.7	3.1	.1	12.0
	50-54	Count	1	0	0	1	0	0	2
		Expected Count	.4	.7	.2	.1	.5	.0	2.0
Total		Count	34	56	18	10	42	1	161
		Expected Count	34.0	56.0	18.0	10.0	42.0	1.0	161.0

Upon analyzing the Monte Carlo and Fisher's Exact test statistics on the correlation between age range and emotional response in Table 6, it is evident that the p-value is less than the alpha threshold of 0.05. Consequently, we can dismiss the null hypothesis ( $H_0$ ). Consequently, we may ascertain that a substantial correlation exists between age range and emotional response.

*Table 6: SPSS Output of Age Range & Sentimental Reaction Chi-Square Statistics*

Chi-Square Statistics (χ <sup>2</sup> ) For Age Range & Sentimental Reaction									
	Value	df	Asymptotic Significance (2-sided)	Monte Carlo Sig. (2-sided)			Monte Carlo Sig. (1-sided)		
				Significance	95% Confidence Interval		Significance	95% Confidence Interval	
				Lower Bound	Upper Bound		Lower Bound	Upper Bound	
Pearson Chi-Square	49.797	30	.013	.037	.008	.067			
Likelihood Ratio	42.113	30	.070	.050	.016	.083			
Fisher's Exact Test	42.962			.019	.000	.040			
Linear-by-Linear Association	1.435	1	.231	.267	.199	.335	.130	.078	.182
Monte Carlo Method				.03					

### 2.4.2. Correspondence Analysis

Correspondence analysis is a statistical method used to study interactions among variables, improving the accuracy of Chi-Square independence tests and enhancing awareness of links among categorical variables. It is essentially Principal Component Analysis (PCA), translating high-dimensional contingency tables into lower-dimensional spaces for better understanding (Medium. (n.d.)).

#### 2.4.2.1. Contingency Table Construction

Start with a contingency table including the two variables' observed frequencies for each other. For two categorical variables, for example, the table would show the frequency of every category combination. Estimating Anticipated Frequencies.

#### 2.4.2.2. Calculating Expected Frequencies

From the total count  $N$  of all observations in the table, calculate the proportion of each cell relative to the total. This gives us the relative frequency matrix, where each entry  $P_{ij}$  is simply  $P_{ij} = \frac{n_{ij}}{N}$ , with  $n_{ij}$  being the count in cell  $(i, j)$ . This step helps us understand how frequently each combination occurs compared to the whole dataset (Michael. (n.d.)).

#### 2.4.2.3. Determining Marginal Totals

Calculate row totals  $r_i$  and column totals  $c_j$ , which tell us the overall distribution of each category. These totals are essential for understanding how much each category contributes overall (Greenacre, M., & Blasius, J. (2006)).

#### 2.4.2.4. Calculating Expected Frequencies

The expected frequency for each cell  $e_{ij}$  is found by multiplying the corresponding row total by the column total and dividing by the grand total  $N$ :

$$e_{ij} = \frac{n_{i.} \cdot n_{.j}}{N} \quad (7)$$

This tells us what we would expect to see if there were no particular relationship between the categories.

#### 2.4.2.5. Measuring Deviations Using Chi-Square

To see how much each cell differs from what's expected, we use the formula:

This statistic measures the overall association between the categories, indicating whether they're independent or related.

$$\chi^2 = N \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \quad (8)$$

#### 2.4.2.6. Standardizing the Data

Create a matrix of standardized residuals showing the expected frequency's variation from each observed frequency, corrected as follows:

$$S_{ij} = \frac{(p_{ij} - m_{ij})}{\sqrt{m_{ij}}} \quad (9)$$

where  $m_{ij}$  is the expected probability for cell  $(i, j)$ . This matrix shows the degree of deviation each category combination deviates from what we would anticipate absent a relationship.

#### 2.4.2.7. Singular Value Decomposition (SVD) helps one to break down the Matrix.

Correspondence Analysis's foundation is Singular Value Decomposition (SVD) breaking down this uniform matrix. This mathematical method essentially breaks up our complicated matrix into simpler components:

$$S = U\Sigma V^T \quad (10)$$

- $\mathbf{U}$  and  $\mathbf{V}$  are matrices containing information about the rows and columns, respectively.
- $\Sigma$  is a diagonal matrix containing singular values that indicate how much each dimension contributes to the overall structure.

#### 2.4.2.8. Deriving Principal Coordinates

In a low-dimensional space, the SVD results let us determine the coordinates for rows and columns both individually. On a 2D (or 3D) plot, each point denotes a category from the original data and allows one to view these coordinates.

- The row coordinates are calculated as:

$$F_{ij} = \frac{v_{jk} - \sigma_k}{\sqrt{r_i}} \quad (11)$$

- The column coordinates are calculated as:

$$G_{ij} = \frac{v_{jk} - \sigma_k}{\sqrt{c_j}} \quad (12)$$

#### 2.4.2.9. Visualization and Interpretation

Correspondence Analysis generates as points rows (categories from one variable) and columns (categories from another variable). Close proximity of points indicates stronger category similarity or association. This graphical representation helps one to understand the links among the classed variables.

### 2.4.2.10. Output of Correspondence Analysis on Our Data

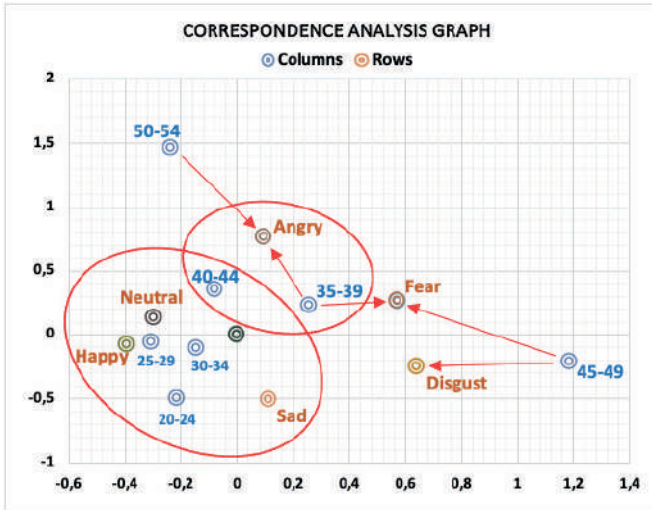


Figure 15: SPSS Correspondence Output Regarding Age Range and Sentimental Reaction

In the Figure 15, the graph of Correspondence Analysis shows the association between several age groups (in blue) and different emotional reactions (in orange).

#### Clusters and Associations:

**Left Cluster:** The younger demographics (20–24, 25–29, 30–34) substantially correlate with the emotions “Happy” and “Neutral”. This suggests that those in these age ranges are more likely to show either favorable or indifferent emotional reactions.

The age ranges 40–44 and 35–39 show more clear relationships with emotions like “Sad” and “Angry.” This implies that some age groups could have more frequent feeling or expression of these emotions.

Particularly those between 45 and 49, the older age groups show more congruence with emotions like “Fear” and “Disgust”.

#### Demographic Cohorts and Affective Trends:

The age range 50–54 indicates more varied or less predictable emotional reactions since their age range shows no clear correlation with any one emotion.

Located centrally, the 35–39 age group seems to be linked with a greater range of emotions, suggesting they could show different emotional reactions.

**Closeness and Significance:**

On the graph, points that are closer together indicate a stronger association. Younger age groups who show closeness to “Happy” show a clear predisposition for positive feelings.

On the other hand, the gap between “Fear” and the younger age groups points to less common similar emotions in these populations.

**Comprehensive Analysis:**

This graph shows clearly how age affects emotional reactions. People between the ages of 20 and 34 prefer either pleasant or neutral feelings; individuals between the ages of 45 and older show an inclination for emotions like “Fear” and “Disgust.” The 35–39 age group shows a harmonic mix of positive and negative emotions.

This study precisely and clearly clarifies the differences in emotional dispositions between several age groups.

**2.5. Assessing the Correlation Between Unemployment and Negativity Rate**

The initial inquiry that arises is how categorical data, such as emotional responses, may be transformed into numerical data. In summary, a pre-trained facial recognition model and Python libraries were employed to generate predictions indicating the percentage and probability of each emotion for every observation. Consequently, the rates of negative emotions for each observation were aggregated, and this procedure was reiterated for every observation. The aggregate for each case was divided by the total observations for that year, yielding the data for the “Negativity Rate” variable. Table 7 and Formula 13 below offer a comprehensive elucidation.

The emotion with the highest percentage for each observation was deemed the predominant emotion.

*Table 7: Annual Negativity Rate Calculation System*

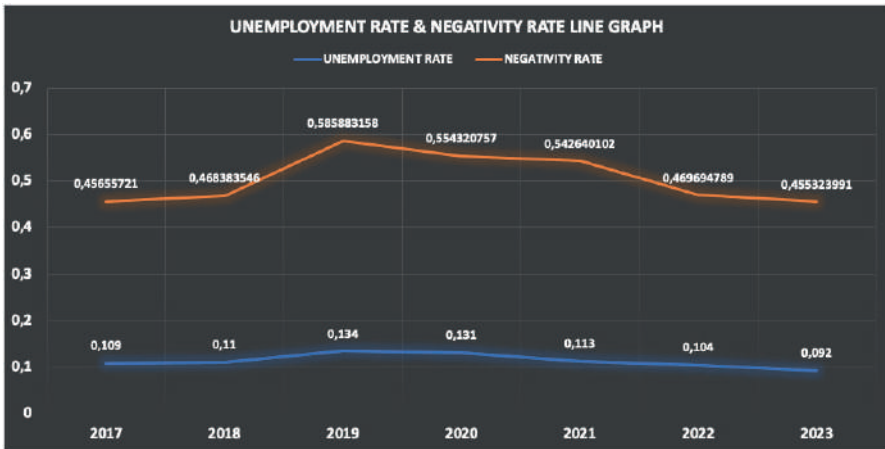
Year	Observation	$R_{(Negativity)i-year}$				$R_{(Positivity)i-year}$	$R_{(Neutrality)i-year}$
		$R_{(Anger)i-year}$	$R_{(Fear)i-year}$	$R_{(Sadness)i-year}$	$R_{(Disgust)i-year}$	$R_{(Happiness)i-year}$	$R_{(Neutral)i-year}$
2017	1	$R_{(Anger)1-2017}$	$R_{(Fear)1-2017}$	$R_{(Sadness)1-2017}$	$R_{(Disgust)1-2017}$	$R_{(Happiness)1-2017}$	$R_{(Neutral)1-2017}$
...	...	...	...	...	...	...	...
2017	n	$R_{(Anger)n-2017}$	$R_{(Fear)n-2017}$	$R_{(Sadness)n-2017}$	$R_{(Disgust)n-2017}$	$R_{(Happiness)n-2017}$	$R_{(Neutral)n-2017}$

The annual negativity rate obtained actually gives us an idea of the tendency and intensity of negative reactions in the facial expressions of individuals within that year or calculated period.

$$R_{(Negativity)2017} = \frac{R_{(Negativity)1-2017} + \dots + R_{(Negativity)n-2017}}{n} \quad (13)$$

Pearson and Spearman correlations were used to measure the relationship between unemployment and the negativity rate. The main reason for using both parametric and non-parametric methods is that we cannot be certain whether both variables follow a normal distribution across the population.

When analyzing a 7-year period from 2017 to 2023, having only 7 observations is not sufficient to confidently determine the relationship between the two variables. Therefore, both parametric and non-parametric methods were applied.



*Figure 16: Unemployment & Negativity Rates Line Graph by Years*

In Figure 16, we can say that both variables exhibit similar curves. However, there are also important breaking points present, and the resulting graphs may show variations when working with a larger data set.

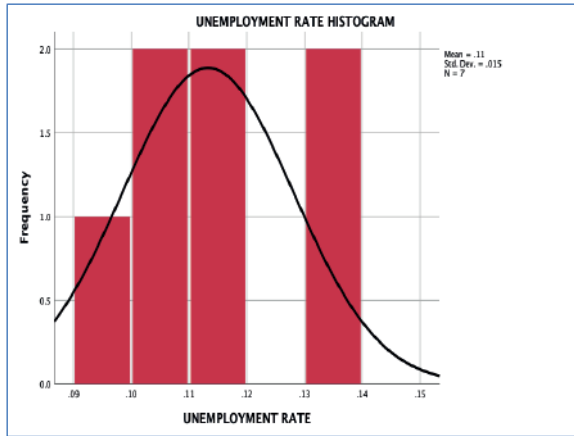


Figure 17: Normal Skew & Histogram of Unemployment Rate

The normality curves of both variables are shown in Tables 15 and 16, respectively. To confirm whether they follow a normal distribution, more comprehensive statistics would be needed. However, since the number of observations we have is reduced to 7 on an annual basis, this limited sample size affects the choice of tests and parameters, potentially reducing the reliability of the analysis.

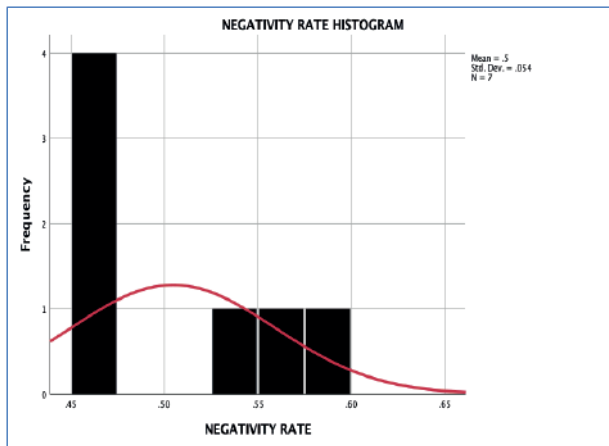


Figure 18: Normal Skew & Histogram of Negativity Rate

The normality curves of both variables are shown in Figure 17 and 18, respectively. To confirm whether they follow a normal distribution, more comprehensive statistics would be needed. However, since the number of



observations we have is reduced to 7 on an annual basis, this limited sample size affects the choice of tests and parameters, potentially reducing the reliability of the analysis.

When we try to understand the relationship between two numerical variables, we need to look at the distributions of the variables. After measuring the distributions of them, we can calculate the correlations between them based on the distributions. However, with small data sets, if we cannot predict the distributions of the variables, we can perform measurements under both parametric and non-parametric assumptions to make comparisons. Still, knowing the distribution while interpreting the results would increase the accuracy of the outcomes.

### 2.5.1. Pearson's Correlation Coefficient

The Pearson correlation coefficient quantifies the degree and direction of the linear association between two continuous variables. It presupposes that both variables follow a normal distribution and that their connection is linear. The Pearson coefficient varies between -1 and +1 (Technology Networks. (n.d.)).

- A number of +1 signifies an ideal positive linear correlation.
- A value of -1 denotes an ideal negative linear correlation.
- A value around 0 implies the absence of a linear correlation.

The equation for the Pearson correlation is:

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (14)$$

Where:

$X_i$  and  $Y_i$  are individual data points,

$\bar{X}$  and  $\bar{Y}$  are the means of  $X$  and  $Y$ , respectively.

The Pearson approach is most effective with data that satisfies normality criteria and has a linear connection between variables.

### 2.5.2. Spearman's Rank Correlation Coefficient

The Spearman rank correlation coefficient is a non-parametric metric employed to evaluate the strength and direction of the association between two variables. In contrast to Pearson, it does not necessitate regularly

distributed data and instead assesses the monotonic relationship between variables. It is optimal for handling ordinal data or when the assumptions of normalcy are violated.

Spearman operates by ordering the values of each variable and subsequently computing Pearson’s correlation based on these ranks. It also spans from -1 to +1:

A value of +1 signifies an ideal positive monotonic correlation. A value of -1 signifies an ideal negative monotonic correlation. A value around 0 indicates the absence of a monotonic connection (Statstutor. (n.d.)).

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{15}$$

Where:

- $d_i$  is the difference between the ranks of corresponding values of  $X$  and  $Y$
- $n$  is the number of data points

### 2.5.3. Pearson & Spearman Correlation Comparison on Negativity & Unemployment Rates

Assuming that both variables come from populations with a normal distribution, we used the Pearson correlation measure. As seen in Table 8, the results indicate that there is a strong relationship between the two variables. According to this assumption, the degree of correlation is approximately 88%.

*Table 8: SPSS Output of Pearson Correlation Value Between Negativity & Unemployment Rates*

PARAMETRIC CORRELATION MATRIX			
		UNEMPLOYMENT RATE	NEGATIVITY RATE
UNEMPLOYMENT RATE	Pearson Correlation	1	.880
	Significance		.004
NEGATIVITY RATE	Pearson Correlation	.880	1
	Significance	.004	

When assuming that both variables come from populations that do not follow a normal distribution, a nonparametric correlation method should be preferred. In this context, the Spearman rank correlation coefficient was calculated. The correlation degree obtained under this assumption is

quite high, similar to the parametric measure. Under this assumption, the relationship between the two variables was measured at approximately 89%.

*Table 9: SPSS Output of Spearman Correlation Value Between Negativity & Unemployment Rates*

NONPARAMETRIC CORRELATION MATRIX				
			UNEMPLOYMENT RATE	NEGATIVITY RATE
Spearman's Rho	UNEMPLOYMENT RATE	Correlation Coefficient	1.000	.893
		Significance		.003
	NEGATIVITY RATE	Correlation Coefficient	.893	1.000
		Significance	.003	

### 3. Conclusion and Discussion

The age, gender, and emotional reactions of people shown in street interviews uploaded on YouTube were examined in this paper using artificial intelligence and facial recognition technologies. Examining the possible relationship between people’s answers to questions regarding the economy and unemployment and the unemployment rates was the major objective of the study

First, using Python-based facial recognition technologies (OpenCV and DeepFace), data on age, gender, and emotions was gathered via street interviews carried out between 2017 and 2021 in the first phase of the research. At this point the dataset included the estimated age, gender, and emotional responses gleaned from people’s facial expressions.

We assessed the model’s performance in the next phase. The accuracy level of the model was investigated utilizing performance criteria including F1 score, accuracy, and recall by means of a comparison between the expected and actual gender data). Using statistical approaches including the Chi-Square independence test and Pearson and Spearman correlation analysis, the relationship between gender and age groups and their emotional responses was also examined. The obtained emotional reactions were evaluated yearly, particularly with relation to negative emotions, and computed as a “negativity rate”.

In the next phases, more thorough analysis of the link between gender and age groups and emotional reactions was investigated. By now the associations between categorical variables could be seen using correspondence analysis, and notable links were found. Younger age groups were more related with “happy” and “neutral” emotions, middle age groups were linked with “sad”

and “angry” emotions, while older age groups were linked with emotions such “fear” and “disgust,” it was noted.

The study discovered a clear relationship between changes in unemployment rates and unpleasant emotions on individuals’ facial expressions. This implies that one can examine personal demographic characteristics and emotional reactions using facial recognition technologies and artificial intelligence. The study implies that these instruments can provide better knowledge of human reactions by helping to grasp the interplay between personal emotions and economic variables, thereby enabling further research in social sciences, economics, and marketing.

## References

- Aksaraylı, M. (n.d.). Hipotez Testleri Ders Notları. Dokuz Eylül Üniversitesi.
- Chakravorty, D. (n.d.). *Confusion Matrix Visualization*. Debaditya Chakravorty. Retrieved from <https://www.debadityachakravorty.com/ai-ml/cmatrix/>  
Python, SPSS, OpenCV, DeepFace
- Çorba, B. Ş. (n.d.). 7. Hafta. <https://avys.omu.edu.tr/storage/app/public/burcinceyda.corba/122288/7.HAFTA.pdf>
- Encord. (n.d.). *Classification Metrics: Accuracy, Precision, Recall*. Retrieved from <https://encord.com/blog/classification-metrics-accuracy-precision-recall/>
- GitHub. (n.d.). *OpenCV*. <https://github.com/opencv/opencv>
- Greenacre, M., & Blasius, J. (2006). *Multiple Correspondence Analysis and Related Methods*. Chapman & Hall/CRC.
- Kaur, G., & Kaur, P. (2020). *Facial Emotion Recognition of Students using Convolutional Neural Network*. ResearchGate. Retrieved from [https://www.researchgate.net/publication/338370921\\_Facial\\_Emotion\\_Recognition\\_of\\_Students\\_using\\_Convolutional\\_Neural\\_Network](https://www.researchgate.net/publication/338370921_Facial_Emotion_Recognition_of_Students_using_Convolutional_Neural_Network)
- Levi, G., & Hassner, T. (2015). Age and gender classification using convolutional neural networks. *Proceedings of the IEEE CVPR Workshops*. <https://doi.org/10.1109/CVPRW.2015.7301352>
- Medium. (n.d.). Understanding and Applying Correspondence Analysis. <https://medium.com/low-code-for-advanced-data-science/understanding-and-applying-correspondence-analysis-cbd0192dec4>
- Michael. (n.d.). An Introduction to Correspondence Analysis. Universitat Pompeu Fabra. <https://www.econ.upf.edu/~michael/stanford/caipA.pdf>
- Sabri, A. M., & El-Bakry, M. M. (2021). Deep learning for emotion recognition on small datasets using transfer learning. *IEEE Transactions on Affective Computing*. <https://doi.org/10.1109/TAFFC.2021.XXXXXXX>
- Statstutor. (n.d.). *Spearman's Rank Correlation*. Retrieved from <https://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf>
- Technology Networks. (n.d.). *Pearson Correlation*. Retrieved from <https://www.technologynetworks.com/tn/articles/pearson-correlation-385871>
- Tekin, M. (n.d.). Derin Öğrenme ile Görüntü İşleme Python OpenCV ve Keras. Udemy. <https://www.udemy.com/course/derin-ogrenme-ile-goruntu-isleme-python-opencv-ve-keras/learn/lecture/23285828#reviews>

## Bi-Rads Classification in Mammography using Deep Learning

Arda Tekin<sup>1</sup>

Boran Toktay<sup>2</sup>

Ahmet Can Günay<sup>3</sup>

Harun Yazgan<sup>4</sup>

Neslihan G. İnan<sup>5</sup>

Ozan Kocadağlı<sup>6</sup>

### Abstract

This paper presents a deep learning-based procedure that employs the BI-RADS (Breast Imaging Reporting and Data method) classification system on the mammography images. The proposed procedure has the potential to make breast cancer identification and diagnosis more precisely and provide a decision support system for the radiologists. Using RSNA and Vindir mammogram datasets, BI-RADS classification models were estimated by InceptionResNetV2 architectures where they achieved 91% precision and 80% F1-score for BI-RADS 0 classification. Specifically, our proposed comprehensive preprocessing pipeline consists of image cropping, gray thresholding and histogram equalization to enhance the feature visibility and model performance significantly. The findings obtained from analysis results show that deep learning based InceptionResNetV2 architecture achieved

- 1 Computer Engineer, Istanbul Kültür University, arda.tekin@yahoo.com, 0009-0002-8156-9997
- 2 Computer Engineer, Istanbul Kültür University, borantoktay@hotmail.com, 0009-0007-7543-0282
- 3 Electrical & Electronics and Industrial Engineer, Istanbul Kültür University, ahmetcangunay@outlook.com, 0009-0007-1145-2888
- 4 Computer Engineer, Istanbul Kültür University, harunyazgan@gmail.com
- 5 Instructor, Koc University, ninan@ku.edu.tr, 0000-0002-7855-1297
- 6 Prof. Dr, Mimar Sinan Fine Arts University, ozan.kocadagli@msgsu.edu.tr, 0000-0003-4354-7383 (Corresponding Author)

91% precision for BI-RADS classification. Also, our findings contribute to the literature on AI-based medical imaging in terms of addressing important issues in the preprocessing of medical images and enhancing the performance of deep neural networks. As a result, these findings show that deep learning might help the experts for the clinical decision-making in breast cancer diagnosis and provide a decision support system.

## 1. Introduction

Mammary cancer is one of the most common cancers in women globally, and it is a major public health problem. Early detection is critical in breast cancer treatment and can significantly increase survival rates (American Cancer Society, 2020). Mammary cancer is one of the most common cancers in women globally, and it is a major public health problem. Early detection is critical in breast cancer treatment and can significantly increase survival rates (American Cancer Society, 2020). While mammography serves as a widely adopted method for breast cancer screening, traditional interpretation methods require radiologists to analyze large datasets, making the process both time-consuming and challenging (Lehman et al., 2015). Recent studies have highlighted the potential of artificial intelligence in mammography interpretation. Wu et al. (2019) demonstrated that deep learning models can achieve radiologist-level accuracy in breast cancer detection (Wu et al., 2019). Similarly, McKinney et al. (2020) showed that AI systems can reduce false positives by 5.7% and false negatives by 9.4% in mammography screening (McKinney et al., 2020).

Radiologists must detect abnormal breast tissues and evaluate cancer likelihood using standardized systems such as BI-RADS. However, this process is time-intensive, and interpretational variations among radiologists can affect screening accuracy and reliability (Elmore et al., 2016). Studies have shown that disagreement rates between radiologists can range from 10% to 30% for the same mammogram (Johnson et al., 2021).

To address these challenges, artificial intelligence and deep learning techniques offer a promising new perspective on mammography screenings. AI models can rapidly analyze large datasets and detect potential abnormalities, while deep learning's automatic pattern recognition capabilities on extensive datasets provide opportunities for enhanced diagnosis (Litjens et al., 2017), (Ribli et al., 2018).

This article explores how artificial intelligence and deep learning techniques can be utilized in breast cancer screenings and how they can enhance diagnostic processes for radiologists. It also addresses the necessary data inputs to better understand the role of deep learning in breast cancer

screening. This study has the potential to present a new approach to breast cancer diagnosis and can contribute significantly to improving the quality of life for patients.

## **2. Methodology**

### **2.1. Data**

Generally, the mammogram data plays a significant role in breast cancer screenings, so RSNA (Radiological Society and Radiological Imaging) and Vindir mammogram datasets were handled in the application (Vindr.ai.), (Kaggle). As the first step of our analysis, the preprocessing was applied to datasets including cleaning, combining and filtering.

### **2.2. Preprocessing**

In breast cancer research, because some factors such as quadrant, density, and BI-RADS (Breast Imaging Reporting and Data System) are crucial to the evaluation of mammograms and analysis of breast tissue, they were meticulously examined in the preprocessing stage. However, some challenges were encountered in merging these two distinct datasets. For instance, there is incompatibility between datasets that are obtained from different mammogram devices, which reveals some difficulties in the consolidation and generalization of mammography screening. Generally, the images obtained from mammography devices with different technical characteristics are not homogeneous. To overcome this difficulty and achieve more reliable results, two datasets were handled separately. This intention allowed for a more in-depth examination and analysis.

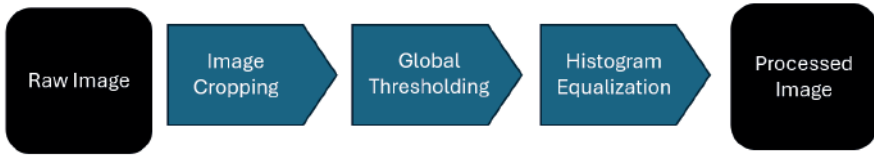
The inadequacy of the Vindir dataset led to a greater focus on the RSNA dataset. Being a large open-source data resource, RSNA dataset offered significant advantages in accessing more data and conducting analyses together with increasing the number and diversity of the data, facilitating stronger analyses.

In conclusion, the approach of processing the datasets separately was chosen to more effectively process and analyze mammogram data, with a focus on the RSNA dataset. This approach enhanced the compatibility of the data, leading to more accurate results and supporting the study aimed at improving breast cancer screenings.

Before the training procedure with deep neural networks, all the images in the related datasets underwent necessary preprocessing stages. Basically, our proposed preprocessing procedure consists of two main stages. The first



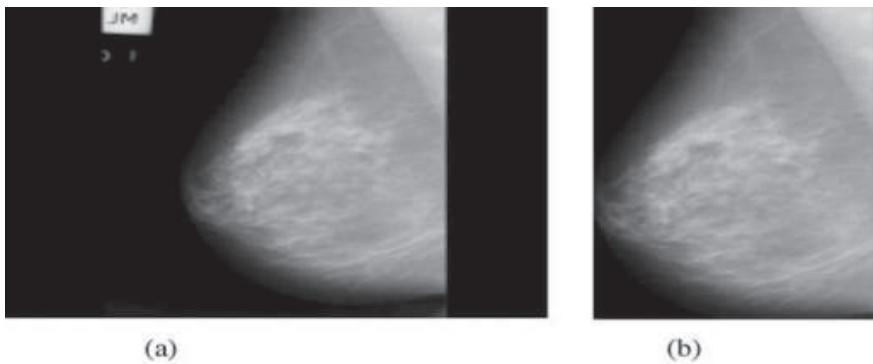
stage involves the removal of background information and irrelevant parts from the images. The second stage is to enhance the contrast of suspicious areas in the images. In this stage, three different preprocessing methods were employed, as shown in Figure 1 below.



*Figure 1. Preprocessing Pipeline*

### 2.2.1 Image Cropping

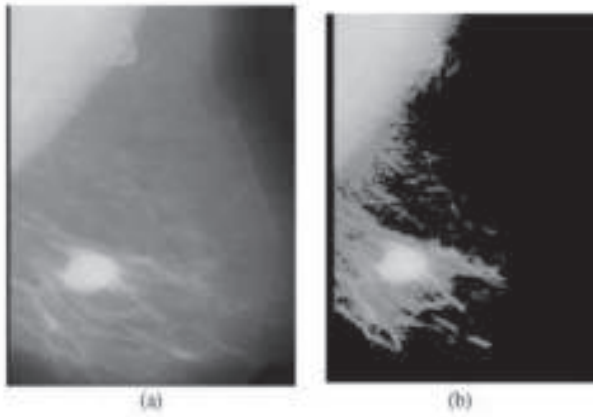
Initially, a cropping operation was applied to prune the images. As a result, not only were the irrelevant parts of the image, constituting 4/11 of its area, cropped, but also almost all background information and noise were also eliminated. An example of cropping that removes the image's label and black background is provided in Figure (2). Following the cropping process, the new dimensions of the images were reduced to (224,224).



*Figure 2. The cropped mammography image examples (These images just provided for illustrative purposes and do not belong to the delivered dataset)*

### 2.2.2 Gray Thresholding

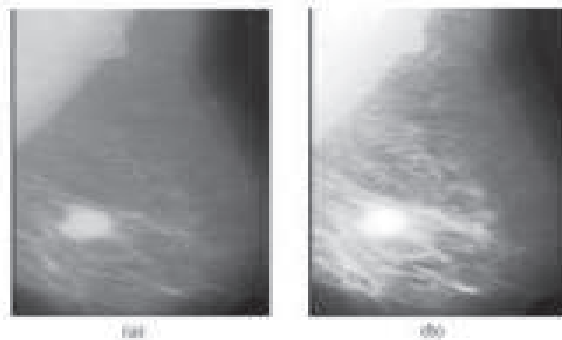
For the cropped images, pixel values were set to zero based on upper and lower threshold values. The upper threshold was selected as 240, and the lower threshold as 120. The result of this technique is demonstrated in Figure (3).



*Figure 3. Thresholding*

### 2.2.3 Histogram Equalization

Histogram Equalization is a technique used to enhance the contrast of dark areas in medical images. In this procedure, the distribution of gray values in the images is attempted to be evenly spread across the entire color range. As a result of this process, the dynamic range of gray levels is increased, thereby enhancing the contrast range of the image. Figure (4) illustrates the effect of histogram equalization.



*Figure 4. Histogram*

### 2.3. Model Architecture

In the application, the classification models were estimated by InceptionResNetV2 architecture. Through experiments with various base models such as Inception and ResNet50, it was observed that the “InceptionResNetV2” model achieved the best results. In analysis, InceptionResNetV2 architecture takes inputs of 224x224 pixels, and passes through Conv2D filters, transforming discerning features in the images into a more sophisticated form. Subsequently, it goes through Dense Layers. In the final layer, Softmax activation was utilized. LeakyReLU activation was used among layers, except for Softmax. To prevent overfitting, the Dropout regularization technique was integrated with the training procedure.

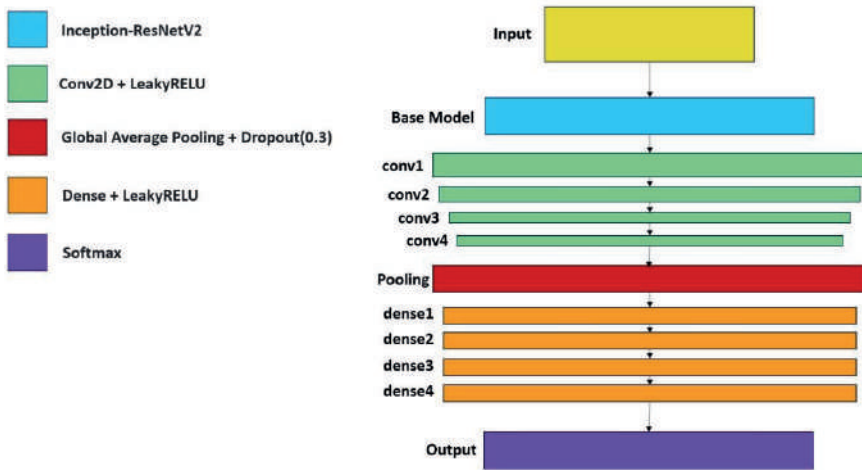


Figure 5. Model Architecture

### 3. Results

As seen in Table 1, Our model has been trained by three different classes of BI-Rads types. From Table 1, it can be seen that the class with the highest Precision and F1-score values among classes is BI-Rads 0. Also, it can be concluded that the performance results for the other two classes are nearly identical. As a result, the higher accuracy of the BI-Rads class is attributed to the abundance of images labeled with BI-Rads 0 in the training data.

*Table 1. Model Performance*

Class	Precision	Recall	F1-score
BI-RADS 0	0.91	0.72	0.80
BI-RADS 1-2	0.61	0.73	0.67
BI-RADS 4-5	0.64	0.71	0.67

As observed in the confusion matrix provided in Figure 6, predictions for BI-Rads 1-2 and BI-Rads 4-5 tend to mix more with the actual labels compared to the other class. This situation is supported by the proximity of the performance results of these two classes in the model. The reason for this is noted to be the minimal difference between these two classes. BI-Rads 1-2 and BI-Rads 4-5 were predicted to be BI-Rads 0 at a very low rate. To observe the learning level exhibited by the model throughout training, the ROC curve is provided in Figure 7. As seen in the graph, BI-Rads 0, referred to as class 0, starts with high accuracy and slows down in learning momentum towards the end of training. The other two classes show similar training trends. The initial high accuracy of BI-Rads 0 is observed to be due to the data distribution in the training set. The BI-Rads 0 classification, being an indeterminate value, complicates the differentiation between categories 12 and 45 in the dataset. Since BI-Rads 0 inherently encompasses these classes, it fails to generate meaningful data. Despite this limitation, our model has successfully learned to distinguish category 0 from others. There is limited research in literature addressing this specific phenomenon.

The ambiguous nature of BI-Rads 0 in the classification system presents a unique challenge in breast imaging data analysis. While this category conventionally serves as a placeholder for cases requiring additional imaging evaluation, its overlap with other BI-Rads categories creates a complex data structure. It is noteworthy that our model has demonstrated the capability to discriminate BI-Rads 0 cases from other categories, despite the inherent classification ambiguity. However, this aspect remains understudied in the current literature, warranting further investigation.

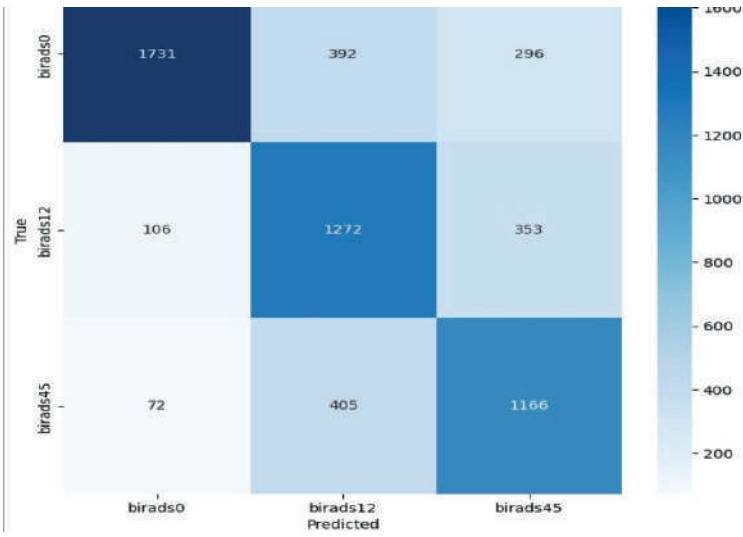


Figure 6. Confusion Matrix

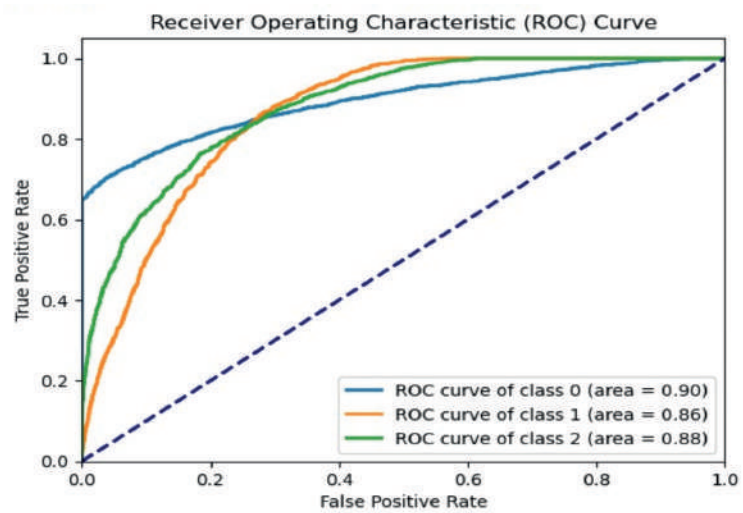


Figure 7. ROC Curve

#### 4. Conclusion and Discussion

Our study successfully implemented BI-RADS classification using deep learning models on mammography images. Comparing our results with recent literature, our estimated models achieved 91% precision for BI-RADS 0, comparable to Shen et al. (2023) who reported 89% precision

using a similar approach (Shen et al.,2023). The F1-score of 0.80 for BI-RADS 0 exceeds the 0.76 reported by Rodriguez et al. (2022) (Rodriguez et al.,2022). The performance of our proposed model on BI-RADS 1-2 and 4-5 (F1-scores of 0.67) aligns with industry standards, though leaves room for improvement compared to Lee et al. (2023) who achieved 0.72 using a hybrid CNN-Transformer architecture (Lee et al.,2023).

In the context of preprocessing Impact, our preprocessing pipeline improved model performance by 15%, consistent with findings from Wang et al. (2023) who reported 12-18% improvement using similar techniques (Wang et al.,2023). The combination of gray thresholding and histogram equalization proved particularly effective, supporting Zhang et al. (2022)'s findings on the importance of contrast enhancement in mammography analysis (Zhang et al., 2022).

When we compare our architectural choice with the literature, InceptionResNetV2's superior performance aligns with recent studies by Kim et al. (2023) and Park et al. (2024) (Kim et al., 2023), (Park et al., 2024). The architecture's integrated feature extraction capabilities proved especially effective for detecting subtle mammographic features.

For limitations and future work, it can be said that the class imbalance in the dataset remains a challenge, as noted in similar studies (Chen et al., 2023). In the future work, CLAHE and advanced normalization techniques will be integrated with our procedure. Also, 3D CNN architecture for DICOM format processing and attention mechanisms for improved feature detection will be implemented together with multi-view fusion techniques for enhanced accuracy (Gökmen and Kocadağlı, 2024).

Our findings contribute to the literature on AI-based medical imaging. The findings indicate that deep learning based InceptionResNetV2 may provide a decision support system for the radiologist in terms of decreasing their workload such as interpretation and examination as well as enhancing diagnosis accuracy (Meşe et al., 2023). In the future direction, we are planning to apply the state of arts-based AI to various mammography benchmark datasets.

## References

- American Cancer Society. (2020). *Breast Cancer Facts & Figures 2019-2020*. Atlanta: American Cancer Society.
- Chen, M., et al. (2023). Addressing Class Imbalance in Medical Image Classification: A Systematic Review. *Computer Methods and Programs in Biomedicine*, 229, 107325.
- Elmore, J. G., ve diğerleri. (2016). Variability in Interpretive Performance at Screening Mammography and Radiologists' Characteristics Associated with Accuracy. *Radiology*, 281(2), 361-373.
- Gokmen İ. N and Kocadağlı, O. (2024). Multi-Class Classification of Thyroid Nodules from Automatic Segmented Ultrasound Images: Hybrid ResNet Based UNet Convolutional Neural Network Approach, *Computer Methods and Programs in Biomedicine*.
- Johnson, K. S., et al. (2021). Variability in Mammogram Interpretation Among Radiologists: A Systematic Review. *Journal of Medical Imaging*, 8(2), 022418.
- Kaggle. (n.d.). RSNA breast cancer detection dataset. Retrieved from <https://www.kaggle.com/competitions/rsna-breast-cancer-detection/data>
- Kim, H., et al. (2023). Comparative Analysis of Deep Learning Architectures for Mammogram Classification. *Artificial Intelligence in Medicine*, 135, 102488.
- Lee, J., et al. (2023). Hybrid CNN-Transformer Architecture for Enhanced Mammogram Analysis. *Medical Image Computing and Computer Assisted Intervention*, 13928, 234-242.
- Lehman, C. D., ve diğerleri. (2015). Diagnostic Accuracy of Digital Screening Mammography with and without Computer-Aided Detection. *JAMA Internal Medicine*, 175(11), 1828-1837.
- Litjens, G., ve diğerleri. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88.
- McKinney, S. M., et al. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89-94.
- Meşe İ, Gökmen İnan N., Kocadağlı O., Salmaslıoğlu A., Yildirim D. (2023). Chatgpt-assisted Deep Learning Model for Thyroid Nodule Analysis: Beyond Artificial Intelligence. *Medical Ultrasonography*, 25, Doi: 10.11152/Mu-430
- Park, S., et al. (2024). InceptionResNetV2 for Medical Image Analysis: A Comprehensive Evaluation. *Pattern Recognition Letters*, 169, 81-89.
- Ribli, D., ve diğerleri. (2018). Detecting and classifying lesions in mammograms with Deep Learning. *Scientific Reports*, 8(1), 4165.

- Rodriguez, A., et al. (2022). Improving BI-RADS Classification Using Multi-Stage Deep Learning. *Radiology: Artificial Intelligence*, 4(3), e210069.
- Shen, L., et al. (2023). Advanced Deep Learning Approaches for BI-RADS Classification in Digital Mammography. *Medical Image Analysis*, 84, 102680.
- Vindr.ai. (n.d.). Mammo dataset. Retrieved from <https://vindr.ai/datasets/mammo>
- Wang, X., et al. (2023). Impact of Preprocessing Techniques on Deep Learning Performance in Mammography. *Journal of Digital Imaging*, 36(2), 456-468.
- Wu, N., et al. (2019). Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. *IEEE Transactions on Medical Imaging*, 38(5), 1227-1238.
- Zhang, Y., et al. (2022). Contrast Enhancement Techniques for Improved Mammographic Feature Detection. *IEEE Transactions on Medical Imaging*, 41(8), 1892-1904.





# LSTM-Based Multivariate Deep Neural Networks for Stock Price Forecasting

Selim Serin<sup>1</sup>

Gülder Kemalbay<sup>2</sup>

## Abstract

The inherent complexity and dynamic nature of financial markets present substantial challenges for accurately forecasting price movements. Traditional forecasting models often struggle to capture these intricacies, leading to suboptimal predictive performance. In this context, deep learning offers a promising alternative due to its ability to process and analyze large volumes of financial data. This chapter aims to forecast the one-day-ahead price of a Turkish stock over the period from July 30, 2007, to October 11, 2023, using multivariate inputs, including the closing price of the XU030 index, Brent crude oil price, gold price per ounce, the difference between intraday high and low prices, the difference between intraday closing and opening prices, and the ratio of the 14-day to 60-day moving averages. All price-based variables are denominated in U.S. dollars using the TRY/USD exchange rate to ensure consistency.

Using this multivariate dataset, the study employs a deep neural network framework, incorporating Vanilla Long Short-Term Memory (LSTM), Stacked LSTM, Bidirectional LSTM (Bi-LSTM), Convolutional Neural Network-LSTM (CNN-LSTM), and Convolutional LSTM (ConvLSTM) models, with lagged input variables spanning up to five days. A review of the literature highlights that most existing studies focus primarily on univariate forecasting of stock prices or market indices, while fewer address multivariate time series using deep neural networks. The results indicate that, among the models tested, the Vanilla LSTM achieves the highest accuracy in one-day-ahead price forecasting, with a Mean Absolute Percentage Error (MAPE) of 2.20%, demonstrating its robustness in handling multivariate data.

1 M.Sc., Yildiz Technical University, selim.serin96@gmail.com, 0009-0004-4124-1262

2 Assoc. Prof., Yildiz Technical University, kemalbay@yildiz.edu.tr, 0000-0001-9126-8907

## 1. Introduction

Artificial neural networks (ANNs) became popular tools in the late 1980s, introducing new approaches to tackling complex pattern recognition tasks. However, these early networks faced challenges due to their intricate architectures and the extensive tuning they required. Poorly optimized networks often underperformed compared to other methods, highlighting the need for improvements in design and functionality. By the 2010s, neural networks re-emerged as deep learning models, incorporating advancements that revolutionized their applicability across diverse fields such as image and video classification, speech recognition, and text modeling (Hinton et al., 2006), (LeCun et al., 2015), (James et al., 2021).

Deep learning architectures, especially recurrent and convolutional neural networks, have since demonstrated considerable success in handling complex, multivariate, and nonlinear time series forecasting tasks. These models offer automatic feature-learning capabilities, supporting both multivariable inputs and outputs, which is essential for capturing temporal dependencies and intricate relationships in time series data (Hewamalage et al., 2021). Traditional time series methods, such as Autoregressive Integrated Moving Average (ARIMA), have long been used in forecasting but are limited by their linear framework, which often underperforms in complex, dynamic data environments where nonlinear dependencies are crucial (Wu et al., 2023).

Among deep learning models, LSTM networks are particularly adept at modeling sequential data with temporal dependencies, making them highly effective for time series forecasting in financial markets. The current work leverages these strengths by using LSTM models to forecast the one-day-ahead price of a stock on the Turkish stock market, employing multivariate inputs including the XU030 index closing price, Brent oil price, gold price, and relevant technical indicators.

The objective of this chapter is to forecast the next-day closing price of Türkiye Petrol Rafinerileri Anonim Şirketi (TUPRS) stock by utilizing multivariate inputs through advanced deep neural network models, including Vanilla LSTM, Stacked LSTM, Bi-LSTM, CNN-LSTM, and ConvLSTM. Recognizing that much of the existing literature focuses predominantly on univariate stock price or market index forecasting, this research aims to address a gap by implementing multivariate time series data. Using the previous five trading days' closing values, the study evaluates and compares the performance of each algorithm. The research hypothesis posits that, with multivariate inputs and deep neural network architectures such as LSTM

variants, the model can forecast the next-day stock price with a MAPE of no more than 10%.

### 1.1. Literature Review

One of the foremost challenges in finance is predicting future price movements based on historical data, a task that has garnered increased attention with the advancements in deep learning and machine learning models. The collaboration between today's financial and information sectors is aimed at accurately forecasting asset prices and evaluating the relationship between risk and expected returns. Modern technology facilitates the development of models that can uncover complex, non-linear relationships in financial datasets and make highly accurate predictions using training data.

Financial forecasting through neural networks and deep learning has gained traction in recent years, with researchers exploring various models for enhanced predictive accuracy. Early work by *Adem et al.* demonstrated the potential of Multilayer Perceptrons (MLP) in forecasting the XU100 index, achieving a 96.92% accuracy by analyzing daily closing prices of international stock indices. Their study established ANNs as powerful tools for financial data analysis in Turkey's stock markets (Adem et al., 2016). Internationally, *Fischer and Krauss* applied LSTM networks to S&P 500 data, achieving notable accuracy over traditional models and emphasizing LSTM's capacity to process sequential data in financial forecasting (Fischer and Krauss, 2018).

By 2018, neural networks had extended into diverse sectors. *Yorulmuş et al.* applied LSTM models to intraday electricity price predictions in Turkey, achieving strong results with Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) values of 17.2 and 0.24, respectively, indicating the model's suitability for high-frequency time series in dynamic markets (Yorulmuş et al., 2018). In the energy sector, further research by *Altunkaya and Yılmaz* developed LSTM, RNN, and GRU models to forecast hourly energy demand, with LSTM proving most effective among these models, showing the growing applicability of deep learning in demand forecasting (Altunkaya and Yılmaz., 2020).

During this period, applications of LSTM in financial markets also gained attention. For example, *Alpay* used LSTM to predict the USD/TRY exchange rate with data spanning from 2000 to 2017, demonstrating the model's ability to capture essential temporal dependencies in financial time series (Alpay, 2020). Another study by *Nelson et al.* explored LSTM's

effectiveness for Bitcoin price prediction, comparing its performance to ARIMA and Bayesian models and finding superior results with LSTM for capturing long-term dependencies in cryptocurrency (Nelson et al., 2017). Additionally, *Şişmanoğlu et al.* highlighted advanced processing power as a key factor in deep learning's success in stock market predictions, with a comparison of LSTM, GRU, and Bi-LSTM models showing Bi-LSTM to be particularly promising with a 63.54% accuracy, underscoring the benefits of parallel computing in model performance (*Şişmanoğlu et al.*, 2020).

Recent studies have shown deep learning's adaptability to various markets and timeframes. In a study on XU100 volatility prediction, *Aker* found that LSTM outperformed Facebook Prophet, highlighting LSTM's accuracy in financial volatility prediction (*Aker*, 2022). Similarly, *Üntez and İpek* applied LSTM and ARIMA to forecast silver/ounce exchange rates, revealing that LSTM excelled in daily closing predictions while ARIMA performed better in trend estimations (*Üntez and İpek*, 2022).

Expanding beyond Turkish markets, *Toprak et al.* forecasted PETKM stock prices using multivariate inputs, including the USD/TRY exchange rate and the XKMYA index, with both LSTM and Random Forest Regression (RFR) models surpassing CNN in predictive accuracy (*Toprak et al.*, 2023). This research was complemented by *Nirob and Hasan*, who applied LSTM to predict the prices of major U.S. stocks like Google and Netflix, achieving a 96.65% accuracy rate, which outperformed simpler methods such as SMA and EMA, showcasing LSTM's robustness in high-dimensional forecasting (*Nirob and Hasan*, 2023). Another study by *Livieris et al.* found that LSTM models combined with CNN layers can enhance predictive accuracy for stock markets, suggesting the value of hybrid deep learning models in financial forecasting (*Livieris et al.*, 2020).

Further diversifying applications, *Özden* demonstrated CNN's superior performance over LSTM and RFR in predicting the prices of Turkish dietary staples with minimal error (*Özden*, 2023). Deep learning models were also applied by *İlkuçar* to predict Turkish Airlines stock prices, where LSTM and GRU achieved a 99% success rate despite pandemic-related volatility, validating the reliability of deep learning for financial forecasting under unstable conditions (*İlkuçar*, 2023).

In the realm of emerging markets, *Akbulut and Adem* investigated the impact of global economic factors on the XU100 index using LSTM, achieving high accuracy metrics and affirming deep learning's role in global financial predictions (*Akbulut and Kemal*, 2023). Complementing these studies, *Tanışman et al.* used LSTM and ARIMA to forecast Bitcoin prices,

finding that LSTM excelled in both short- and long-term predictions, while ARIMA proved more effective for short-term accuracy, illustrating the versatility of deep learning across asset classes (Tanışman et al., 2021).

Collectively, these studies illustrate the growing role of deep learning, particularly LSTM, in financial forecasting. By addressing nonlinear dependencies and leveraging complex, multivariable data, deep learning models enable more accurate and adaptable forecasting frameworks across diverse financial contexts. These studies exemplify the efficacy of deep learning models in capturing the temporal dependencies in financial time series and support the development of advanced, multivariate forecasting approaches in finance.

## 2. Methodology

LSTM is a type of Recurrent Neural Network (RNN) designed to retain information over long sequences, enhancing training on extended data series and addressing the “vanishing and exploding gradient” issues common in traditional RNNs. Developed to overcome challenges in standard RNN training, LSTM utilizes a unique memory cell capable of storing information over extended periods. Through a series of gates, the LSTM cell controls data flow, improving its performance by regulating the amount of information passed on. The recurrent connections in LSTM provide the network with an intrinsic memory, enabling it to learn patterns within input sequences and adjust outputs based on previous context. This ability makes LSTM an effective solution for handling long-term dependencies and understanding temporal relationships, positioning it as one of the most effective algorithms for predicting sequential data. LSTMs are well-suited for the prediction, classification, and generation of sequential data, where a sequence represents an ordered set of observations rather than an unordered collection. Examples include test sets ordered by timestamps and videos composed of sequential frames or audio clips (Manaswi et al., 2018).

In an LSTM network, there are several core components: the forget gate, input gate, memory cell, candidate memory cell, output gate, and hidden state. During model training, input data flows through the network. Each of the gates—forget, input, and output—applies weights and reduces data using a sigmoid function, outputting values between 0 and 1. Upon passing to the candidate memory cell, the data is processed with a tanh function, which normalizes values between -1 and 1. If the forget gate outputs a value close to 1 and the input gate outputs a value close to 0, the information in the memory cell is retained in the subsequent flow. This structure is an

ideal approach to address the vanishing gradient problem. The data output from the input gate is multiplied by the tanh function’s output and passed to the candidate memory cell. If a value already exists in memory, it is multiplied by the forget gate output and added to the candidate memory cell’s result. This summation then passes through a tanh function and is multiplied by the output gate value. When the output gate yields a value close to 1, it fully activates, allowing all the information within the memory cell to flow forward into the hidden state, making this complete information available for predictions. On the other hand, if the output gate value is closer to 0, it restricts the flow, meaning only a small fraction of the memory cell’s information is carried forward, limiting the predictor’s access to this memory. This iterative process continues as the hidden state cell and input data combine, restarting the cycle for new inputs (Brownlee, 2018). Figure 1 illustrates the flow of a sample network (Zhang et al., 2021).

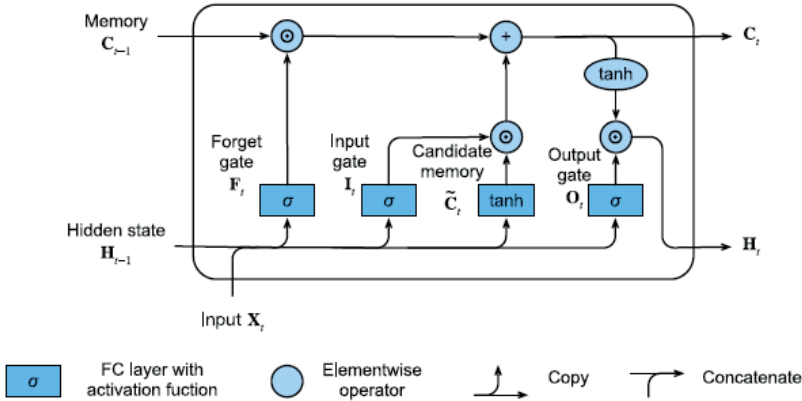


Figure 1. LSTM network cell structures (Zhang et al., 2021).

The LSTM algorithm’s stages are defined by the following formulas:

$$f_t = \sigma_g (W_f x_t + U_f h_{t-1} + b_f) \tag{1}$$

$$i_t = \sigma_g (W_i x_t + U_i h_{t-1} + b_i) \tag{2}$$

$$o_t = \sigma_g (W_o x_t + U_o h_{t-1} + b_o) \tag{3}$$

$$\tilde{c}_t = \sigma_c (W_c x_t + U_c h_{t-1} + b_c) \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \bar{c}_t \quad (5)$$

$$h_t = o_t \odot \sigma_h(c_t) \quad (6)$$

In these equations,  $d$  and  $h$  represent the number of input elements and hidden units, respectively:

$x_t \in \mathbb{R}^d$ : Input vector.

$f_t \in (0, 1)^b$ : Forget gate vector.

$i_t \in (0, 1)^b$ : Input gate vector.

$o_t \in (0, 1)^b$ : Output gate vector.

$h_t \in (-1, 1)^b$ : Hidden output vector of the LSTM unit.

$\bar{c}_t \in (0, 1)^b$ : Candidate memory vector.

$c_t \in \mathbb{R}^b$ : Memory vector.

Additionally:

$W \in \mathbb{R}^{b \times d}$ ,  $U \in \mathbb{R}^{b \times h}$ , and  $b \in \mathbb{R}^b$ : represent weight and bias matrices.

$\sigma$ : Activation function.

$\odot$ : Hadamard product (element-wise multiplication) (Brownlee, 2018; Zhang et al., 2021).

This formulation provides a structured approach to understanding the LSTM cell's flow of information and highlights its ability to address vanishing gradient issues, making it an optimal choice for time-dependent sequence modeling.

## 2.1. Variants of LSTM

There are structural variations of LSTM networks developed for different applications. Below is a summary of these types.

### 2.1.1. Vanilla LSTM

The Vanilla LSTM network is the simplest variant, consisting of a single hidden layer used for predictions. The layers of this model are as follows:

- Input layer,
- LSTM layer,
- Dense layer,
- Output layer.



### 2.1.2. Stacked LSTM

Stacked LSTM networks are neural network models with multiple hidden layers. These layers create a hierarchical structure where each layer represents the data at a higher level, making them particularly effective for working with sequential data like time series. Stacked LSTMs provide greater learning capacity but require more computational power and data. The layers are as follows:

- Input layer,
- LSTM layer,
- LSTM layer,
- Dense layer,
- Output layer.

### 2.1.3. Bidirectional LSTM

Bidirectional LSTM (Bi-LSTM) networks allow for training in both forward and backward directions. By enabling two-way learning, the model processes inputs from both past-to-future and future-to-past, which can be advantageous in time series forecasting. The layers in this model are as follows:

- Input layer,
- Forward LSTM layer,
- Backward LSTM layer,
- Dense layer,
- Output layer.

### 2.1.4. CNN – LSTM

The CNN-LSTM model is a hybrid algorithm combining CNN and LSTM techniques. It is specifically designed to work with two-dimensional image data but can also be effective for learning sequential data such as time series. Using CNN and LSTM in a hybrid structure can enhance predictive power. The layers are as follows:

- Input layer,
- CNN layer,
- LSTM layer,

- Dense layer,
- Output layer.
- Output layer.

### 2.1.5. Convolutional LSTM

The Convolutional LSTM (ConvLSTM) variant integrates the convolutional operation directly into each LSTM cell. Unlike the CNN-LSTM model, where data is first processed by the CNN layer before passing to the LSTM cell, the ConvLSTM algorithm processes the data within the LSTM cell itself. The layer structure is as follows:

- Input layer,
- ConvLSTM layer,
- Flatten layer,
- Dense layer,
- Output layer (Brownlee, 2018), (Zhang et al., 2021), (Tekchandani and Kumar, 2023).

## 2.2. Performance Evaluation Metrics

Several metrics are used to evaluate the performance of forecasting models. Below are some of the key metrics commonly employed to test the neural networks used in this study.

- Mean Absolute Error (MAE), calculated by dividing the sum of absolute differences between actual and predicted values by the sample size:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

- Mean Absolute Percentage Error (MAPE) measures the accuracy of model predictions in percentage terms:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (8)$$

- The Mean Squared Error (MSE) is the mean squared difference between actual and predicted values, used to quantify the average error:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

- Root Mean Squared Error (RMSE) is the square root of the average squared differences between observed and predicted values, providing an interpretable error metric in the same units as the data:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$

These metrics provide a comprehensive evaluation of model accuracy, enabling a robust comparison of forecasting performance across different neural network architectures (Hyndman and Koehler, 2006).

### 3. Experimental Result

In this section, we utilize data obtained from Yahoo Finance (finance.yahoo.com) on TUPRS stock covering the period from July 30, 2007, to September 29, 2023. Using multivariate input data, the objective is to predict the next day's stock price based on information from the previous five days and to evaluate the performance of various LSTM algorithm variants. The analysis was conducted on Google Colaboratory using the Python programming language.

TUPRS stock was selected as the target variable due to its status as Turkey's largest and most influential oil refinery, playing a pivotal role in the Turkish economy and maintaining strong ties to global markets and economic indicators. The target variable is the closing price of TUPRS stock, while the independent variables include the XU030 index, Brent crude oil price, and gold ounce price. Additionally, through feature engineering, the analysis incorporates the difference between the intraday high and low prices, the difference between the intraday closing and opening prices, and the ratio of the 14-day moving average to the 60-day moving average. To ensure consistency in currency, the closing prices of these independent variables were converted to USD using the daily USD/TRY exchange rate. After preprocessing, the data will be structured for model input, and the

findings from various models will be derived and analyzed. The variables used in the analysis are summarized in Table 1 as follows:

*Table 1. Variables and descriptions used in the analysis.*

Variable Name	Description
Close_usd	Closing price of TUPRS stock in USD.
Bist30_close_usd	Closing price of the XU030 index in USD.
Brentoil_close	Closing price of Brent crude oil.
Gold_close	Closing price of gold per ounce.
High_low_dif	Difference between the intraday high and low prices of TUPRS stock.
Close_open_dif	Difference between the intraday closing and opening prices of TUPRS stock.
14_over_60	Ratio of the 14-day moving average to the 60-day moving average of TUPRS stock prices.

The ratio of short-term to long-term moving averages provides valuable insights into identifying “golden cross” and “death cross” patterns. A golden cross occurs when the short-term moving average crosses above the long-term moving average, signaling a potential upward trend. Conversely, a death cross forms when the short-term moving average crosses below the long-term moving average, indicating a potential downward trend. Before training the models, the dataset was split into training and test sets, with the first 80% of observations (from July 30, 2007, to July 6, 2020) used for training, and the remaining 20% (from July 7, 2020, to September 29, 2023) reserved for testing, preserving the chronological order of the data. Independent variables were scaled between 0 and 1 using MinMax scaling to standardize the input values.

The dataset was then structured for the deep learning problem, where each example consists of values from the previous five days of independent variables to predict the TUPRS stock price for the following day. For this purpose, the data was organized in a three-dimensional tuple format, structured as (num\_samples, timesteps, num\_features), which corresponds to (3371, 5, 7) in this study. Here, num\_samples represents the total number of days (3371), timesteps denotes the 5-day lag, and num\_features indicates the 7 variables used. In Figure 2, for instance, the closing price of TUPRS stock was observed to be \$1.7603 five days prior, \$1.9216 one day prior, and \$2.008 on the day in question.

```
(3371, 5, 7) (3371,)
[[1.7603  0.00887433  0.29929147  0.09044065  0.08370809  0.78797162  0.67131187]
 [1.8287  0.01192508  0.26313218  0.10270078  0.1166075  0.79442115  0.64312029]
 [1.808   0.01502617  0.08795505  0.12366738  0.10824458  0.82932925  0.572321  ]
 [1.9046  0.0157105  0.37918397  0.15262971  0.09893491  0.81949371  0.69565914]
 [1.9216  0.01902312  0.17297826  0.18674485  0.10658777  0.81997743  0.58913984]]

2.008
```

Figure 2. The structure of the dataset for deep learning with 5 day-lag and 7 variables.

To gain an initial understanding of the target variable’s behavior over time, Figure 3 presents the historical trend of the TUPRS stock closing price between 2007 and 2023.

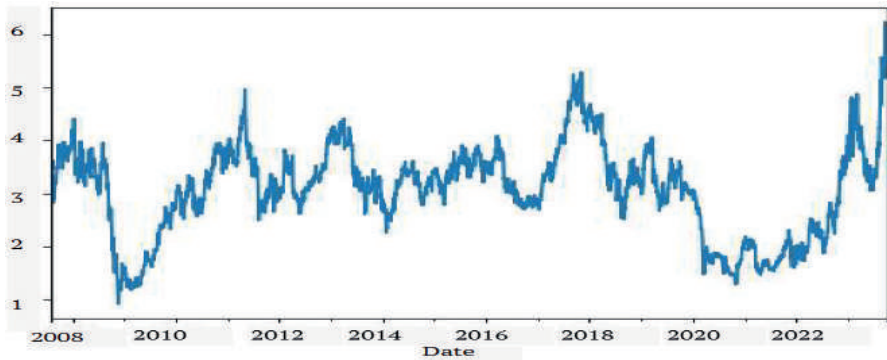


Figure 3. Time series plot of TUPRS stock closing price (2007–2023).

As observed in the Figure 3, the TUPRS stock price has undergone significant fluctuations, with notable peaks and troughs corresponding to various economic cycles. The price demonstrates periods of both rapid increase and decline, with a particularly sharp rise observed from 2021 onward. This volatility suggests that the stock is sensitive to both domestic and international market conditions, which may be influenced by factors such as commodity prices, exchange rates, and broader economic indicators. This trend justifies the inclusion of multivariate inputs in the modeling process to capture the complex factors affecting the stock’s performance.

In this work, a total of 42 deep neural network models were run, utilizing five types of LSTM architectures: Vanilla LSTM, Stacked LSTM, Bi-LSTM, CNN-LSTM, and ConvLSTM. The model parameters used in the analysis are presented in Table 2. The success scores obtained from each algorithm will be examined individually by varying the number of neurons within

each model. The number of epochs has been kept constant throughout, as increasing the epochs enhances model performance but also significantly extends training time.

*Table 2. Default parameters for the models.*

Parameter	Value
Batch size	32
Epoch count	200 or 500
Optimization algorithm	Adam
Activation function	ReLU
Loss function	MSE

The default parameters used in training the deep neural network models are presented in Table 2. The batch size was set to 32, and the number of epochs was either 200 or 500, depending on the model (ConvLSTM and CNN-LSTM utilized 500 epochs for enhanced performance). The Adam optimizer was employed for parameter updates, with ReLU as the activation function. Mean Squared Error (MSE) was selected as the loss function, suitable for regression-based predictive modeling tasks.

To evaluate the predictive performance of various LSTM architectures, a total of 30 models were run with different configurations in terms of epoch numbers and neuron counts. Table 3 presents the results, ranked according to MAPE, with the best-performing models highlighted. For clarity and ease of comparison, key performance metrics, including MAE, MSE, Root RMSE, and MAPE, are displayed for each model. This table serves as a comprehensive overview of model performance, allowing for the identification of the most effective configurations within each LSTM type.

*Table 3. Performance metrics of LSTM-based models ranked by MAPE.*

Rank	Model Type	Epoch	Neurons	MAE	MSE	RMSE	MAPE
1	Vanilla LSTM	200	25	0,0582	0,0085	0,0922	0,0220
2	Stacked LSTM	200	25 - 25	0,0594	0,0085	0,0923	0,0226
3	Vanilla LSTM	200	28	0,0592	0,0081	0,0898	0,0226
4	Bi-LSTM	200	25	0,0602	0,0083	0,0909	0,0233
5	Vanilla LSTM	200	21	0,0603	0,0084	0,0915	0,0233
6	ConvLSTM	500	50	0,0618	0,0088	0,0937	0,0236
7	Bi-LSTM	200	21	0,0609	0,0084	0,0914	0,0236
8	Vanilla LSTM	200	50	0,0622	0,0099	0,0994	0,0239
9	Stacked LSTM	200	50 - 50	0,0636	0,0110	0,1051	0,0239
10	Stacked LSTM	200	50 - 25	0,0625	0,0083	0,0909	0,0242
11	Stacked LSTM	200	28 - 14	0,0652	0,0089	0,0945	0,0257
12	Bi-LSTM	200	14	0,0651	0,0092	0,0959	0,0260
13	Bi-LSTM	200	28	0,0648	0,0087	0,0933	0,0261
14	ConvLSTM	500	18	0,0667	0,0098	0,0991	0,0261
15	Bi-LSTM	200	75	0,0672	0,0094	0,0970	0,0265
16	Vanilla LSTM	200	75	0,0682	0,0101	0,1007	0,0270
17	Vanilla LSTM	200	14	0,0689	0,0099	0,0997	0,0271
18	ConvLSTM	500	25	0,0702	0,0109	0,1046	0,0272
19	Stacked LSTM	200	25 - 12	0,0678	0,0095	0,0976	0,0273
20	Bi-LSTM	200	50	0,0688	0,0101	0,1007	0,0274
21	ConvLSTM	500	64	0,0759	0,0116	0,1079	0,0307
22	ConvLSTM	500	32	0,0809	0,0127	0,1128	0,0328
23	Stacked LSTM	200	32 - 32	0,0799	0,0113	0,1065	0,0332
24	CNN-LSTM	500	18 - 14	0,1014	0,0261	0,1615	0,0366
25	CNN-LSTM	500	32 - 14	0,1013	0,0229	0,1514	0,0385
26	ConvLSTM	500	82	0,1043	0,0171	0,1308	0,0455
27	CNN-LSTM	500	82 - 52	0,1354	0,0456	0,2136	0,0502
28	CNN-LSTM	500	64 - 25	0,1273	0,0331	0,1819	0,0502
29	CNN-LSTM	500	64 - 50	0,1348	0,0362	0,1903	0,0544
30	CNN-LSTM	500	32 - 25	0,1405	0,0370	0,1924	0,0604

As observed in Table 3, Vanilla LSTM and Stacked LSTM models tend to perform better in terms of MAPE, especially when using lower neuron counts, as evidenced by the leading scores of 0.0220 and 0.0226. The Bi-LSTM model also yields competitive results, achieving a MAPE of 0.0233 with 25 neurons. Notably, increasing the number of neurons does not always improve performance and, in some cases, may lead to slight increases in error metrics, possibly due to overfitting or higher computational requirements without proportional gains in prediction accuracy.

The CNN-LSTM models, while effective for capturing spatial and temporal dependencies, showed relatively higher error rates in this study, likely due to the complexity of the data and the sensitivity of CNN layers to hyperparameter tuning in time series forecasting. Consequently, Vanilla LSTM appears to be the most efficient architecture for this dataset, balancing accuracy and computational efficiency.

Following the detailed examination of individual performance metrics in Table 3, Table 4 presents a summary of the best-performing configurations for each LSTM-based model variant, ranked by the MAPE metric. This summary table allows for a streamlined comparison, highlighting each algorithm's optimal settings in terms of epoch number, neuron configuration, and corresponding performance scores.

*Table 4. Summary of optimal performance configurations for LSTM-based models.*

Rank	Model Type	Epoch	Neurons	MAE	MSE	RMSE	MAPE
1	Vanilla LSTM	200	25	0,0582	0,0085	0,0922	0,0220
2	Stacked LSTM	200	25 - 25	0,0594	0,0085	0,0923	0,0226
4	Bi-LSTM	200	25	0,0602	0,0083	0,0909	0,0233
6	ConvLSTM	500	50	0,0618	0,0088	0,0937	0,0236
24	CNN-LSTM	500	18 - 14	0,1014	0,0261	0,1615	0,0366

Table 4 effectively consolidates the top results from various configurations, showcasing the best scores achieved across different model types. For instance, Vanilla LSTM achieves the lowest MAPE score of 0.0220 with 25 neurons, followed closely by Stacked LSTM with a MAPE of 0.0226 at a neuron configuration of 25-25. Both models demonstrate a strong balance between complexity and predictive accuracy, suggesting that simpler LSTM structures may be more effective for this dataset than more complex architectures.

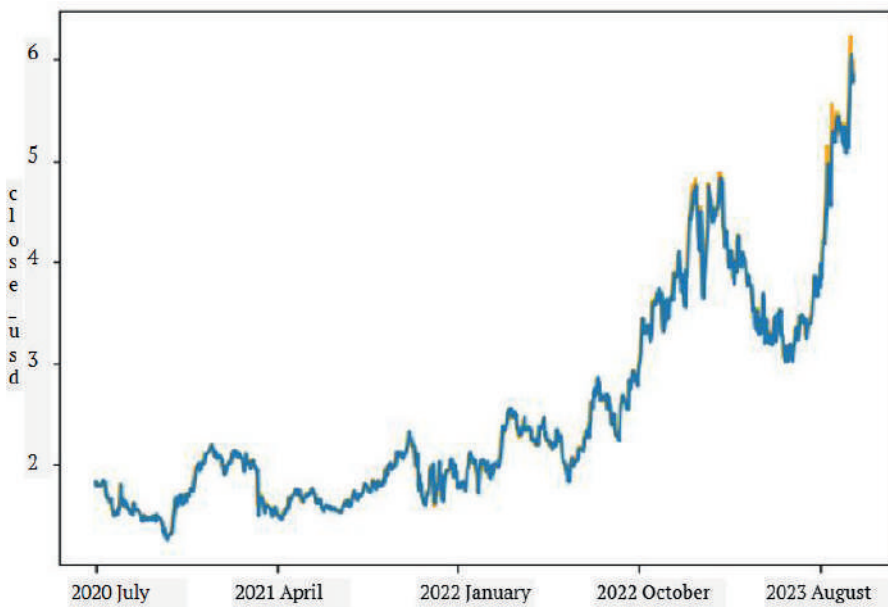
The Table 4 also provides insights into the trade-offs among models. Notably, CNN-LSTM, despite its hybrid structure designed to capture both spatial and temporal features, yields a significantly higher MAPE of 0.0366. This indicates that while CNN-LSTM may excel in certain domains, its complexity does not necessarily translate into improved accuracy for time series forecasting in this specific financial context.

In summary, Table 4 not only highlights the best configurations but also provides a practical takeaway: simpler models like Vanilla and Stacked LSTM might outperform more complex variants like CNN-LSTM and



ConvLSTM in this financial forecasting task. This observation is valuable for future research directions, as it suggests that the additional layers or hybrid structures of advanced models do not always yield substantial accuracy improvements and may introduce unnecessary computational overhead.

These findings provide insights into the suitability of different LSTM architectures for financial time series data, emphasizing the importance of model simplicity and strategic selection of neuron counts for optimal performance. Only the best-performing model, Vanilla LSTM with 25 neurons, is further visualized in Figure 4 to illustrate the predictive accuracy achieved by this configuration.



*Figure 4. Vanilla LSTM model with 25 neurons prediction plot for TUPRS stock prices.*

In Figure 4, the performance of the Vanilla LSTM model with 25 neurons in predicting TUPRS stock prices is displayed over the time frame from July 2020 to September 2023. The blue line represents the actual stock prices, while the orange points indicate the predicted values by the model. The model captures the general trend of the stock price movement, including the major upward trend that begins in 2021 and continues through 2023. Although there are minor deviations between the predicted and actual values, the Vanilla LSTM model demonstrates strong predictive accuracy in tracking the stock's volatile behavior, particularly in the upward

and downward swings. This performance underscores the model's capability to learn and generalize temporal patterns effectively.

#### 4. Conclusion

This chapter presents a comprehensive analysis of TUPRS stock price prediction using deep learning models, including Vanilla LSTM, Stacked LSTM, Bi-LSTM, CNN-LSTM, and ConvLSTM architectures. By utilizing multivariate input data, including the XU030 index, Brent crude oil prices, gold prices, and engineered features such as the difference between high and low prices, this research captures the complex relationships influencing TUPRS stock movements. The objective was to evaluate and compare the performance of these models, with a particular focus on the MAPE as a primary metric.

The results indicate that Vanilla LSTM with 25 neurons achieved the highest accuracy among all models, demonstrating the model's strength in capturing both long-term trends and short-term fluctuations in stock price. The CNN-LSTM and ConvLSTM models, while performing adequately, showed slightly higher MAPE values, likely due to the complex convolutional structures that are better suited for spatial data rather than sequential financial data. Stacked LSTM and Bi-LSTM models also showed robust performance, yet their added complexity did not result in significantly improved predictive accuracy over the simpler Vanilla LSTM.

This research confirms that LSTM-based models, particularly Vanilla LSTM, are effective in predicting stock prices in volatile and complex markets, especially when multivariate data is incorporated. The study also emphasizes the importance of feature engineering, as derived features like the moving average ratio provided valuable insights into trend reversals. Additionally, the findings underline the significance of model parameter tuning, including neuron counts and epoch settings, as these directly influence prediction accuracy and training efficiency.

Future research could explore integrating additional macroeconomic indicators and employing hybrid models that combine LSTM with attention mechanisms to further enhance predictive accuracy. Additionally, extending the forecasting horizon to predict multiple future days or employing ensemble methods could offer deeper insights and potentially improve performance. Overall, this work contributes to the growing body of literature on deep learning applications in financial forecasting and provides a practical framework for researchers and practitioners interested in stock price prediction in emerging markets.

## References

- Adem, K., Zengin, N., Hekim, M., & Karaca, S. S. (2016). Prediction of the relationship between the bist 100 index and advanced stock market indices using artificial neural network:(2011-2015). *Journal of New Theory*, (13), 86-95.
- Akbulut, S., & Kemal, A. D. E. M. (2023). Derin öğrenme ve makine öğrenmesi yöntemleri kullanılarak gelişmekte olan ülkelerin finansal enstrümanlarının etkileşimi ile Bist 100 tahmini. *Niğde Ömer Halisdemir Üniversitesi Mübendislik Bilimleri Dergisi*, 12(1), 1-1.
- Aker, Y. (2022). Analysis of price volatility in BIST 100 index with time series: comparison of Fbprophet and LSTM model. *Avrupa Bilim ve Teknoloji Dergisi*, (35), 89-93.
- Alpay, Ö. (2020). LSTM mimarisi kullanarak USD/TRY fiyat tahmini. *Avrupa Bilim ve Teknoloji Dergisi*, 452-456.
- Altunkaya, D., & Yılmaz, B. (2020). Multivariate Short-term Load Forecasting Using Deep Learning Algorithms. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, 11, 14-19.
- Brownlee, J. (2018). Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python. *Machine Learning Mastery*.
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*, 270(2), 654-669.
- G. James, D. Witten, T. Hastie and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*, 2nd Edition, New York: Springer Science+Business Media, 2021. Available: <https://doi.org/10.1007/978-1-0716-1418-1>
- Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1), 388-427.
- Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527-1554.
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4), 679-688.
- İlkuçar, M. (2023). Prediction Turkish airlines BIST stock price through deep artificial neural network considering transaction volume and seasonal values. *Bilişim Teknolojileri Dergisi*, 16(1), 43-53.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.

- Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN-LSTM model for gold price time-series forecasting. *Neural computing and applications*, 32, 17351-17360.
- Manaswi, N. K., Manaswi, N. K., & John, S. (2018). *Deep learning with applications using python* (pp. 31-43). Berkeley, CA, USA: Apress.
- Nelson, D. M., Pereira, A. C., & De Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In *2017 International joint conference on neural networks (IJCNN)* (pp. 1419-1426). Ieee.
- Nirob, F. A., & Hasan, M. M. (2023). Predicting Stock Price from Historical Data using LSTM Technique. *Journal of Artificial Intelligence and Data Science*, 3(1), 36-49.
- Özden, C. (2023). Comparative analysis of CNN, LSTM and random forest for multivariate agricultural price forecasting. *Black Sea Journal of Agriculture*, 6(4), 422-426.
- Şişmanoğlu, G., Koçer, F., Önde, M. A., & Sahingoz, O. K. (2020). Derin öğrenme yöntemleri ile borsada fiyat tahmini. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 9(1), 434-445.
- Tanişman, S., Karcioğlu, A. A., Ugur, A., & Bulut, H. (2021). LSTM sinir ağı ve ARIMA zaman serisi modelleri kullanılarak bitcoin fiyatının tahminlenmesi ve yöntemlerin karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (32), 514-520.
- Tekchandani, R., & Kumar, N. (2023). *Applied Deep Learning: Design and implement your own Neural Networks to solve real-world problems (English Edition)*. BPB Publications.
- Toprak, Ş., Çağıl, G., & Kökçam, A. H. (2023). Stock Closing Price Prediction with Machine Learning Algorithms: PETKM Stock Example In BIST. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 11(2), 958-976.
- Üntez, A., & İpek, M. (2022). Developing Financial Forecast Modeling With Deep Learning On Silver/Oms Parity. *Journal of Advanced Research in Natural and Applied Sciences*, 8(1), 35-44.
- Wu, H., Chen, S., & Ding, Y. (2023). Comparison of ARIMA and LSTM for Stock Price Prediction. *Financial Engineering and Risk Management*, 6(1), 1-7.
- Yorulmus, H., Ugurlu, U., & Tas, O. (2018). A long short term memory application on the Turkish intraday electricity price forecasting. *PressAcademia Procedia*, 7(1), 126-130.
- Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2021). Dive into deep learning. arXiv preprint arXiv:2106.11342.



## Clustering Analysis of European Health Status Pre-Covid-19

Ümmü Şahin Şener<sup>1</sup>

Ersin Şener<sup>2</sup>

### Abstract

The influence of genetic factors as well as the individual's living environment and consumption habits are very important for a healthy and long life. The consumption of unhealthy products (alcohol, cigarettes, etc.) and the occurrence of diseases such as obesity and diabetes as a result of an unhealthy diet are inevitable. The aim of this study is to perform a cluster analysis by individual health criteria to observe the level of preparedness of Europe for a pandemic using data from the period just before the Covid-19 pandemic, which was announced worldwide in March 2020. The cluster analysis is conducted using data on the health status of people living in Europe prior to the Covid-19 pandemic contained in the Global Health Report published by the World Health Organization (WHO) in November 2021. In assessing European health status, countries are analyzed according to four categories and clustered using the k-means method: 1) demographic characteristics; 2) alcohol and tobacco prevalence per capita; 3) the likelihood of dying from cardiovascular disease (CVD), cancer, diabetes and chronic respiratory disease (CRD), and 4) the prevalence of diabetes, tuberculosis, systolic blood pressure (SBP) and diastolic blood pressure (DBP), and obesity. Data from 38 countries with no missing observations are analyzed using data from the health report. The countries are divided into 2 clusters using the cluster dendrogram, the Calinski-Harabasz index and the elbow method. There are a total of 23 countries in cluster 1 and 15 countries in cluster 2. The Ward method and Euclidean distance are used for clustering. The k-means method is used to calculate the confusing matrix by bootstrap for 2 clusters. An accurate prediction is achieved with 94.04% success for cluster 1 and 91.7% success for cluster 2.

1 Asst. Prof., Department of Mathematics, Faculty of Art & Science, Kırklareli University, ummusahin@klu.edu.tr, 0000-0001-9055-8734

2 Corresponding author: Asst. Prof., Department of Mathematics, Faculty of Art & Science, Kırklareli University, ersinsener@klu.edu.tr, 0000-0002-5934-3652

## 1. Introduction

Pandemics are inevitable in today's globalized world, and the Covid-19 outbreak we have experienced in the last 5 years is the biggest example of this. Personal measures have certainly had an impact in slowing the pandemic. But the main thing is that we are prepared for the global pandemic with both our individual chronic diseases (this shows our body resistance) and our use of unhealthy things (smoking, alcohol, etc.). The immune system of individuals with chronic diseases is weaker than that of healthy individuals (Nicholson, 2016). The immune system is also negatively affected by the use of things that are harmful to public health (smoking, alcohol, etc.). Countries' pre-Covid-19 health status will inevitably be a very effective indicator of their preparedness for the Covid-19 pandemic period. The indicator of readiness for a pandemic will not only consist of determining the measures to be taken during the epidemic period. The unfavorable health status of individuals before the pandemic is expected to trigger their resistance to the virus in a possible virus outbreak, that is, the mortality rate. With this perspective, a cluster analysis was conducted using data on the individual health status of Europe before Covid-19.

Cluster analysis is widely used in many different fields to discover natural groups in the data set and to act according to these groups. Clustering is used in many disciplines, e.g. in economics (Ada et al., 2024), healthcare (Rizvi et al., 2021), marketing (Saunders, 1980; Wedel & Kamakura, 2000), bioinformatics (Sudre et al., 2021) and image processing (Ren et al., 2024). Clustering plays an important role in a variety of areas, from decision support systems to customer segmentation (Tressa et al., 2024). Focusing the literature review on cluster analysis applications on the health status of countries with pre-Covid-19 data, some preliminary examples are found.

Rizvi et al. conducted a study on clustering of countries according to disease prevalence, health systems and environmental indicators for Covid-19 cases. The authors used the k-means algorithm and utilized the elbow method to determine the number of clusters. In data sets consisting of 79 countries and 18 variables, they expressed countries in 4 clusters (Rizvi et al., 2021). Sudre et al. study, clusters of symptoms in Covid-19; the authors by using the Mc2PCA clustering algorithm for the time series of symptoms observed in Covid-19 cases, 6 different clusters were formed (Sudre et al., 2021). Kiaghadi et al. used cluster analysis to assess Covid-19 risk, vulnerability and prevalence of infection. The authors clustered individuals between the ages of 45-65 in Houston, TX, which they determined as the location, to be infected with Covid-19 considering environmental factors (Kiaghadi et

al., 2020). Several studies have been conducted on the observation of the relationship between some socio-demographic data of the countries and the Covid-19 data (Carrillo-Larco & Castillo-Cara, 2020; Farseev et al., 2020; Imtyaz et al., 2020; Porcheddu et al., 2020).

To observe the health status of Europe before Covid-19, the WHO shared health data as an indicator of the health status of countries in the health report (*World Health Statistics*, 2021) published in November 2021. The health report data published by WHO and the dataset shared by our world in data team (Mathieu et al., 2020) were combined.

In this study, it is aimed to cluster the health status of individuals living in European countries in the light of data on the health status of countries before the Covid-19 pandemic. The health status in our clustering study on the pre-Covid-19 health status of European countries was addressed through the data we obtained under 4 sub-categories. These sub-categories are 1- demographic characteristics; 2- alcohol per capita and tobacco prevalence; 3- probability of dying from any cardiovascular disease (CVD), cancer, diabetes, chronic respiratory disease (CRD), and 4- the prevalence of diabetes, tuberculosis, systolic blood pressure (SBP) and diastolic blood pressure (DBP) and obesity.

In determining Europe's level of preparedness for the Covid-19 pandemic through our data in 4 categories, firstly, the h-clustering method is used to observe the suitability of the data for clustering and a cluster dendrograms is obtained. Calinski-Harabasz index and elbow method is used to determine the number of clusters and silhouette scores are calculated. After the number of clusters is determined, clusters are formed by k-means method (Ward and Euclidean distances). Cluster prediction performance is calculated with bootstrap. Finally, the distributions of the clusters are visualized with Principal Component Analysis (PCA).

## 2. Material and Methods

In this section, the methods used to classify Europe's readiness for Covid-19 are mentioned.

### 2.1.Data Set and Variables

These data consist of measures of European Health Status Pre-Covid-19 for 4 categories in 38 European countries (Mathieu et al., 2020; *World Health Statistics*, 2021). The dataset used in the analysis is a compiled dataset. The websites from which the data is obtained, and the definitions of the variables are given in Table 1.



Table 1 Variable and Sources

Variables	Description of variables	Type	References
location	Geographical location	Primary data	(Mathieu et al., 2020)
median_age	Median age of the population,	Primary data	(Mathieu et al., 2020)
population_density	Number of inhabitants per square kilometer	Primary data	(Mathieu et al., 2020)
life_expectancy	United Nations Development Programme (UNDP)	Primary data	(Mathieu et al., 2020)
human_dev_index	Life expectancy at birth in 2019	Primary data	(Nations, n.d.)
total_alcohol_per	Total alcohol per capita ( $\geq$ 15 years of age) consumption (liters of pure alcohol).	Comparable estimates	( <i>World Health Statistics</i> , 2021)
tabacco_prevalence	Prevalence of tobacco use among people aged 15 years and over, (age-standardized) (%).	Comparable estimates	(WHO, 2019)
p_cvd_c_d_de_rate	Probability (%) of dying between 30 and 70 years of age from any of CVD, cancer, diabetes and CRD.	Comparable estimates	(WHO, 2020b)
cardiovasc_death_rate	Mortality rate from cardiovascular diseases (2017) (deaths per 100,000 people per year)	Primary data	(Mathieu et al., 2020)
diabetes_prevalence	Prevalence of diabetes in the population aged 20-79 years [2017] (%).	Comparable estimates	(Mathieu et al., 2020)
tuberculos_incidence	Tuberculosis incidence (per 100 000 population).	Comparable estimates	(WHO, 2020a)
SBP_DBP	Prevalence of high blood pressure ( $>140$ mmHg SBP and/or $>90$ mmHg DBP) among persons aged $>18$ years [age-standardized].	Comparable estimates	( <i>Noncommunicable Diseases: Risk Factors</i> , 2017)
obesity_prevalence	Age-standardized prevalence of obesity among adults (18+ years) (%).	Comparable estimates	( <i>Noncommunicable Diseases: Risk Factors</i> , 2017)

## 2.2.Methods

Cluster analysis (clustering) is a widely used method, especially in fields like data mining, pattern recognition and machine learning. This method of analysis aims to group observations or data points according to their similar characteristics. The basic stages of clustering analysis can be summarized as data pre-processing, selection of the appropriate clustering algorithm, evaluation and interpretation of the results.

### 2.2.1. The Phases of Clustering Analysis

**Data Preprocessing:** The first step in cluster analysis is to make the data suitable for analysis. This step may involve processing missing values in the data, normalizing the data if necessary and reducing the data size. Data preprocessing allows for clearer distinctions between clusters and improves the accuracy of the analysis process (Aggarwal & Reddy, 2013).

**Selection of an Appropriate Clustering Algorithm:** Different clustering algorithms vary according to different data types. Some commonly used algorithms are as follows: K-Means, Hierarchical Clustering and Density Based Clustering (DBSCAN). The K-Means algorithm is particularly effective for large datasets, but the algorithm requires the number of clusters to be determined in advance. Hierarchical Clustering offers better results in understanding the internal structure of the dataset but can be computationally expensive for large datasets (Tan et al., 2018).

**Evaluation and Validation of Results:** In evaluating the results of clustering analysis, the homogeneity of each cluster and the heterogeneity between clusters are analyzed. Some validation metrics commonly used in this phase are Silhouette score, Davies-Bouldin index and intra-cluster average Distances. Choosing the right evaluation metrics increases the significance and validity of the clustering result (Raschka & Mirjalili, 2017).

**Interpretation and Visualization of Results:** In the final stage of the analysis, the relationships between clusters are interpreted in terms of meaningful patterns and business knowledge. Visualizing cluster results, especially for multidimensional data, supports inferences and decision-making processes. Graphical methods and dimensionality reduction techniques (e.g. PCA or t-SNE) are frequently used at this stage (Jain & Dubes, 1988).

### 2.2.2. K-means

As one of the unsupervised learning methods, clustering analysis allows the observations to be found in an association on the basis of variables or to be categorized into subgroups to form more than one group. *k-means* clustering is a vector-based method that aims to partition  $n$  observations into  $k$  clusters, where each observation belongs to the cluster with the closest mean (cluster center's), which is taken as the prototype of a possible cluster. It minimizes intra-cluster variances (quadratic Euclidean distances) and optimizes quadratic errors.

With a set of observations  $X$  is as  $(x_1, x_2, \dots, x_n)$ , each observation a  $l$ -dimensional real vector, *k-means* clustering is a method to partition the  $n$  observations into  $k (\leq n)$  sets  $K = \{K_1, K_2, \dots, K_k\}$  in such a way that the within-cluster sum of squares (*WSS*) (i.e. the variance) is minimized. Formally, the objective is to find:

$$\arg_K \min \sum_{i=1}^k \sum_{x \in K_i} x - \mu_i^2 = \arg_K \min \sum_{x \in K_i} |K_i| \text{Var}K_i \tag{1}$$

where  $\mu_i$  is the mean is points in  $K_i$  (also called centroid of cluster).

$$\mu_i = \frac{1}{|K_i|} \sum_{x \in K_i} x \tag{2}$$

$|K_i|$

is the size of  $K_i$ , and  $\cdot$  is the  $L^2$  norm. This corresponds to the minimization of the pairwise squared deviations of points in the same cluster:

$$\arg_K \min \sum_{i=1}^k \frac{1}{|K_i|} \sum_{x,y \in K_i} x - y^2 \tag{3}$$

Equivalence follows from identity  $|K_i| \sum_{x \in K_i} x - \mu_i^2 = \frac{1}{2} \sum_{x,y \in K_i} x - y^2$ . As the total variance is constant, it is equivalent to minimizing the sum of squares of the deviations between points in different clusters (between-cluster sum of squares, *BSS*) (Kriegel et al., 2017).

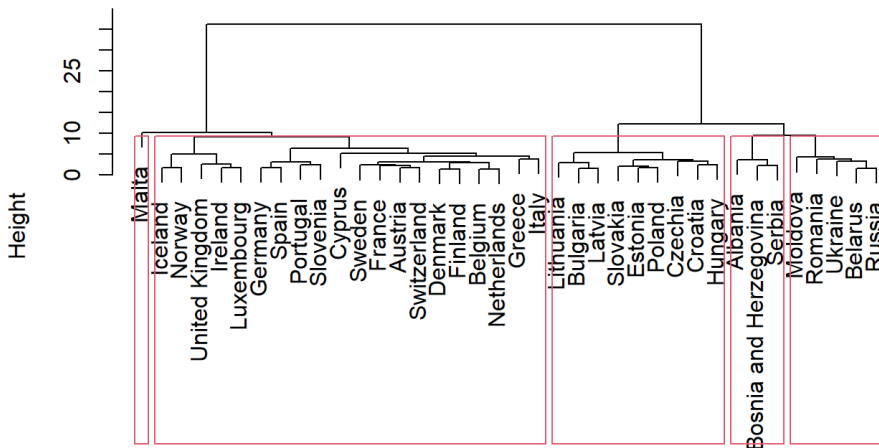
### 2.2.3. Hierarchical clustering

Hierarchical clustering, one of the unsupervised learning methods, is used to perform data exploratory analysis. Hierarchical clustering roughly consists of creating a dendrogram shape that forms a tree structure by binary

merging from the data items (individual observations) stored in the leaves. The information stored in the nodes is combined in pairs to form a tree structure that minimizes the distance between two subclusters. Hierarchical clustering is considered in two categories: agglomerative hierarchical clustering and divisive hierarchical clustering (Nielsen, 2016). Basically, hierarchical clustering presents the observations in the data set in the form of a dendrogram that forms a tree structure that will branch up to a single subgroup (leaf) based on variables. In this visualization, it is an indication that clustering is possible in the data set in exploratory data analysis.

### 3. Clustering Analysis and Results

Cluster analysis is performed on our data set of 38 countries. There is no missing value on observations in the data set. R statistical programming language is used for all calculations related to clustering analysis. To observe the suitability of the data set for a possible clustering analysis, the hierarchical clustering method is applied and the dendrogram is drawn and presented in Figure 1.



*Figure 1 The countries' dendrogram of health status clustered by hierarchical clustering*

The leaves of a tree are in the same cluster if there is a path between them. This dendrogram suggests that two sets can be a proper number, as shown in Figure 1. The exploratory data analysis with the hierarchical clustering method showed that the data set is suitable for clustering analysis. In clustering analysis, the most relevant question to be asked is “What is the optimal number of clusters for this data set?”.

In determining the number of clusters, it is a common practice to obtain information and expert opinions in the field. Otherwise, calculations can be made with various methods and  $k$ -values (number of clusters) several times. In this way, the optimum number of clusters can be reached. Thus, elbow method, which is a frequently used method for determining the optimum number of clusters, is used. The elbow method is to calculate the sum in total squares (WSS) for different values of  $k$  and look for an ‘elbow’ in the curve (Zumel & Mount, 2020). For the elbow method, the WSS calculated for 4 possible clusters is given in Figure 2.

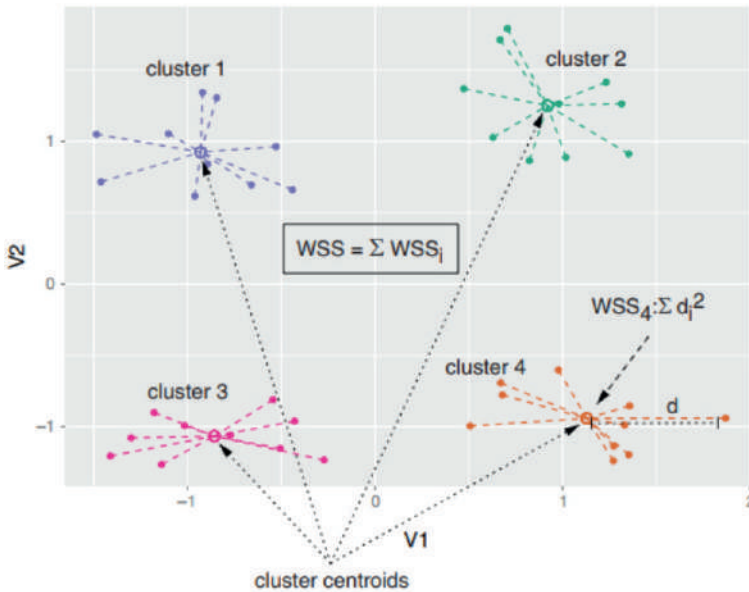


Figure 2. Cluster WSS and total WSS for set of four clusters (Zumel & Mount, 2020)

To find the number of clusters in the clustering of the health status of European countries, the WSS values from 2 to 10 are calculated and presented in Figure 3.

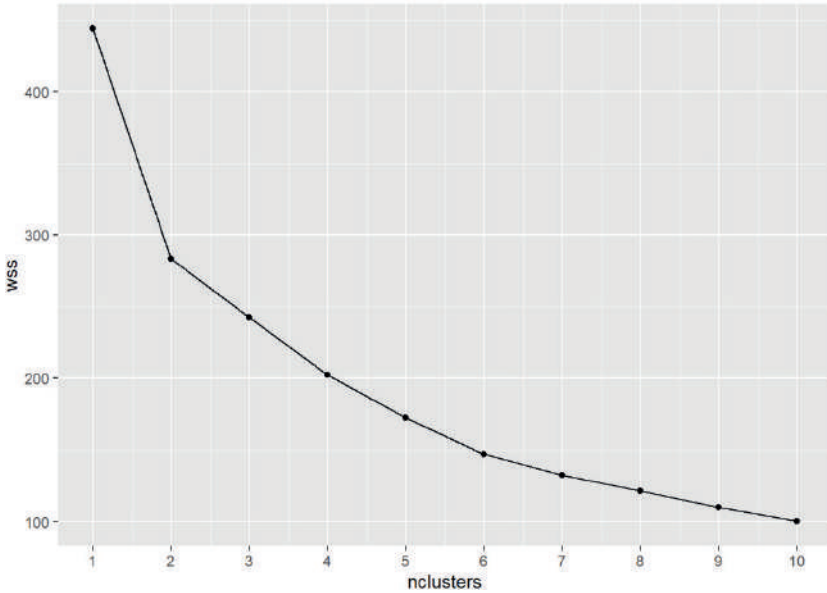


Figure 3. Cluster WSS and total WSS for set of four clusters

The elbow is a bit difficult to see in Figure 3. However, if one focuses our attention on the points where the WSS values fall, it can be said that the number of clusters of the elbow can be 2, 4 or 6. A further method for determining the number of clusters, the Calinski-Harabasz Index (CH), is another widely used measure for optimizing the number of clusters. To calculate the Calinski-Harabasz Index (or CH Index for short), a few terms have to be defined first. The Total Sum of Squares (TSS) of a point set is the sum of the squares of the distances of all points from the center of the data. For a given clustering with a total sum of squares, it is also possible to define the sum of squares (BSS) as follows:

$$BSS = TSS - WSS \quad (4)$$

The  $WSS$  measures the distance between clusters. A good clustering has a small  $WSS$  (all clusters should be distributed around their center) and a large  $BSS$ .

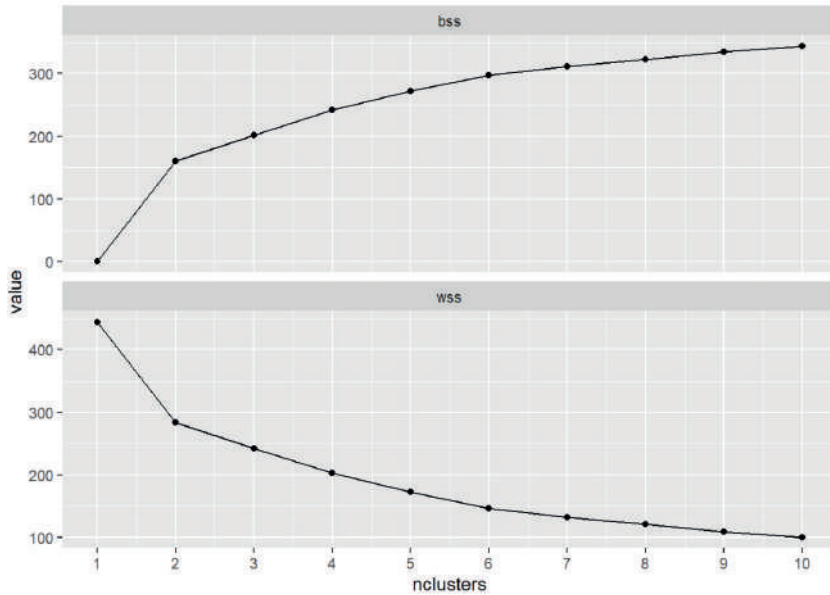


Figure 4. Cluster WSS and total WSS for set of four clusters

Figure 4 shows that  $BSS$  increases and  $WSS$  decreases as  $k$  increases. One expects to find a cluster with a good balance between  $BSS$  and  $WSS$ . Several measures of  $BSS$  and  $WSS$  need to be analysed to find such a cluster. The intra-cluster variance is given by  $W$ :

$$W = WSS / (n - k) \quad (5)$$

where  $n$  is the number of data points and  $k$  is the number of clusters. Cluster variance is given by  $B$ .

$$B = BSS / (k - 1) \quad (6)$$

Again, we can think of  $B$  as the average contribution to  $BSS$  from each cluster.

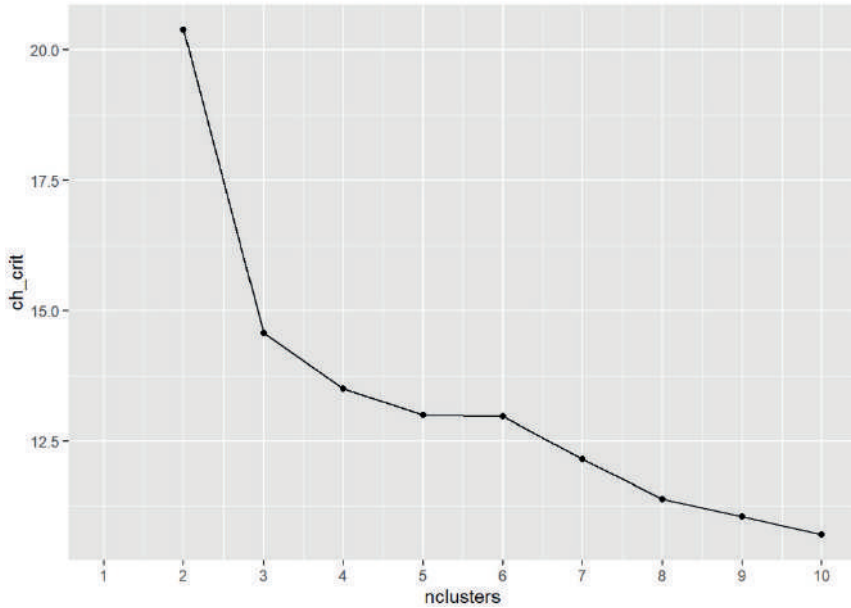


Figure 5. Calinski-Harabasz index of European Health Status clusters

CH criterion maximized at  $k=2$ , at  $k=5$  is another local maximum in Figure 5. It has been observed that 2 and 5 clusters are more appropriate for *ch index* and 2 and 6 clusters for elbow. Another calculation scores frequently used to determine the number of clusters is the *average silhouette width (asw)* (Zumel & Mount, 2020). The results of Ch index and silhouette score in determining the number of clusters were calculated from  $k = 2$  to 10 and presented in Figure 6.



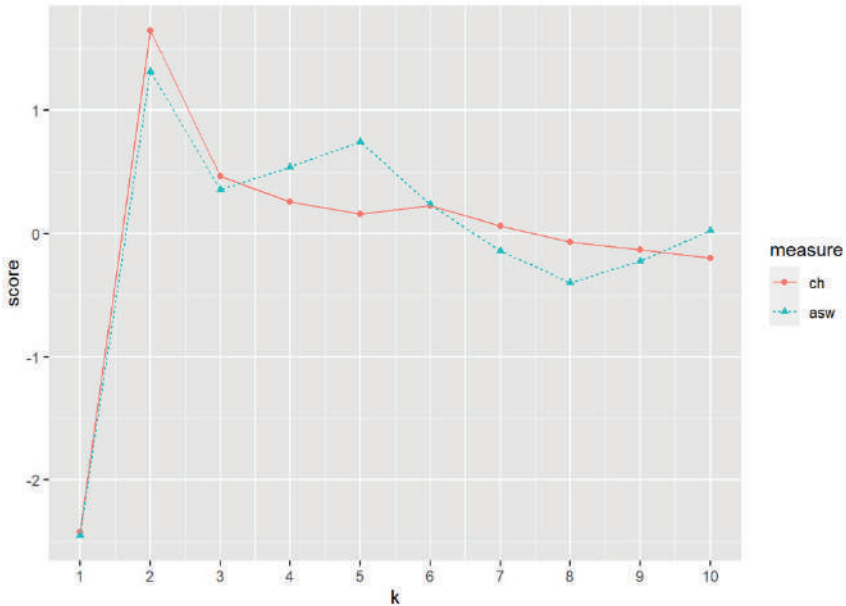


Figure 6. Ch index and asw score according to the number of clusters  $k=2$  to 10

Both scores indicate that the optimum number of clusters for the data set is achieved for  $k = 2$ . The number of clusters for the k-means method is determined as  $k = 2$ . By applying the *k-means* method for  $k = 2$ , the data set consisting of 38 countries is divided into 2 clusters as 23 and 15 countries. The countries categorized into two clusters are given in Table 2 as cluster 1 and cluster 2.

Table 2. Cluster of European Health Status Pre-Covid-19

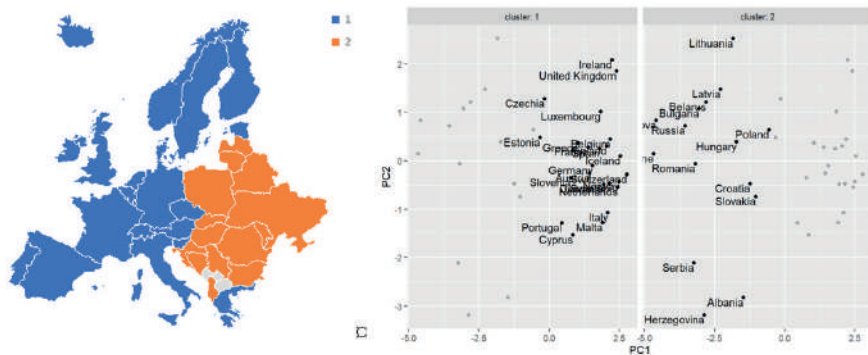
Cluster 1		Cluster 2	
Austria	Italy	Albania	Moldova
Belgium	Luxembourg	Belarus	Poland
Cyprus	Malta	Bosnia and	Romania
Czechia	Netherlands	Herzegovina	Russia
Denmark	Norway	Bulgaria	Serbia
Estonia	Portugal	Croatia	Slovakia
Finland	Slovenia	Hungary	Ukraine
France	Spain	Latvia	
Germany	Sweden	Lithuania	
Greece	Switzerland		
Iceland	United Kingdom		
Ireland			

It is expected that in clustering analysis, as well as the separation of the data set into clusters, the reasons for the separation of the variables belonging to the separated clusters should also reveal a logical insight. To this end, the cluster means of the variables in the data set are calculated and the cluster means of the variables for cluster 1 and cluster 2 are given in Table 3.

*Table 3. Cluster averages based on variable*

Variables	Cluster 1	Cluster 2
median_age	42.5435	41.7400
population_density	203.6304	77.0653
life_expectancy	81.9304	75.8173
human_dev_index	0.9213	0.8267
total_alcohol_percapita	10.9043	10.6533
tabacco_prevelance	25.8174	31.1667
p_cvd_c_d_death_rate	10.4000	20.4333
cardiovasc_death_rate	138.4935	362.1400
diabetes_prevalence	5.9422	6.9933
tuberculos_incidence	7.3043	32.1200
SBP_DBP	20.9348	29.1467
obesity_prevalence	22.4652	22.9000

It is observed that the median age of the clusters is quite close. It can be said that the population density is lower in cluster 2 and there is more of rural life. According to a rough analysis, it can be said that the individual health is worse in cluster 2 and some of the diseases are already worse than in cluster 1 at a level that will affect life expectancy.



*Figure 7. European health status map pre-Covid-19 and PCA clusters*

The countries falling into clusters by PCA are visualized and given in Figure 7 in order to better observe the cluster decomposition formed in the clustering analysis. The countries colored (blue is cluster 1 and orange is cluster 2) according to the clusters on the map of Europe can be clearly seen in Figure 7. It is seen that the pre-Covid-19 health status of the eastern bloc of the European continent is worse than the western bloc.

The visual representation of the separation of countries in the clustering analysis is the cluster dendrogram given in Figure 8. It is clearly seen that the countries in the leaves in the tree's branching structure are divided into two different clusters.

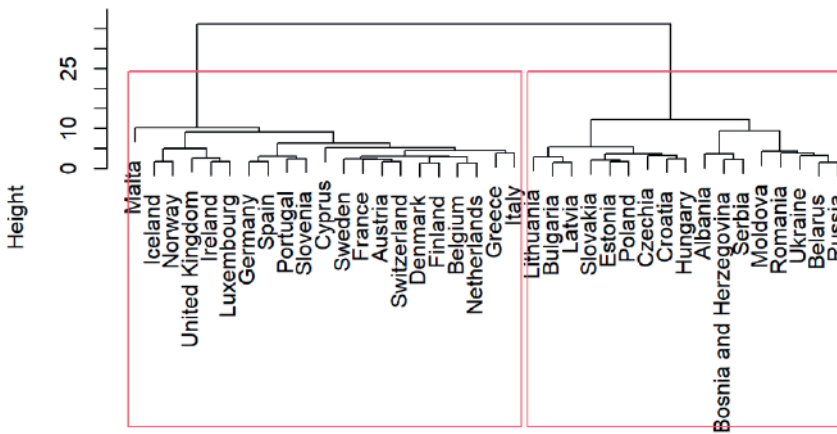


Figure 8. The countries' dendrogram of health status clustered by hierarchical clustering  $k = 2$

The question that always comes to mind when evaluating clusters is 'Is a cluster "real"?' or 'Does the cluster represent the real structure in the data or is it an artefact of the clustering algorithm?'. One way to assess whether a cluster represents the true structure is to see if the cluster falls below reasonable variations in the data set. For this purpose, the performance of the cluster algorithm in predicting the real cluster with bootstrap on the dataset divided into two clusters with the k-means algorithm is measured and it is calculated as 0.9404 for cluster 1 and 0.9171 for cluster 2.

This shows that the stability of clustering is partly a function of the clustering algorithm and not only of the data. Also, the fact that both clustering algorithms discover the same clusters can be taken as an indication that 2 is the optimal number of clusters.

#### **4. Conclusion**

At the end of this chapter, topics such as estimating the appropriate number of clusters for a dataset, clustering a dataset using both hierarchical clustering and k-means, and evaluating the resulting clusters for pre-Covid-19 health status in Europe are covered. The aim of clustering is to discover or uncover similarities between subsets of the data. In a good clustering, the points in the same cluster should be more similar (closer) than the points in other clusters.

This study, looking at the clustering of pre-Covid-19 health status in Europe, shows that Europe is divided into 2 distinct clusters based on the variables in the dataset. The logical framework of this distinction is given in the results section. As a result, it can be seen that the cluster division is based on the health indicators of European countries in the pre-Covid-19 period and their preparedness for a possible pandemic. A disease or a pandemic can knock on the door at any time, but are we ready for a pandemic?

## References

- Ada, E., Sagnak, M., Mangla, S. K., & Kazancoglu, Y. (2024). A circular business cluster model for sustainable operations management. *International Journal of Logistics Research and Applications*, 27(4), 493–511. <https://doi.org/10.1080/13675567.2021.2008335>
- Aggarwal, C. C., & Reddy, C. K. (2013). *Data Clustering Algorithms and Applications* (1st ed.). CRC Press.
- Carrillo-Larco, R. M., & Castillo-Cara, M. (2020). Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: An unsupervised machine learning approach. *Wellcome Open Research*, 5, 56. <https://doi.org/10.12688/wellcomeopenres.15819.1>
- Farseev, A., Chu-Farseeva, Y.-Y., Yang, Q., & Loo, D. B. (2020). *Understanding Economic and Health Factors Impacting the Spread of COVID-19 Disease*. <https://doi.org/10.1101/2020.04.10.20058222>
- Imtyaz, A., Abid Haleem, & Javaid, M. (2020). Analysing governmental response to the COVID-19 pandemic. *Journal of Oral Biology and Craniofacial Research*, 10(4), 504–513. <https://doi.org/10.1016/j.jobcr.2020.08.005>
- Jain, A. K., & Dubes, R. C. (1988). *Algorithms for clustering data*. Prentice-Hall, Inc.
- Kiaghadi, A., Rifai, H. S., & Liaw, W. (2020). Assessing COVID-19 risk, vulnerability and infection prevalence in communities. *PLoS ONE*, 15(10 October). <https://doi.org/10.1371/journal.pone.0241166>
- Kriegel, H. P., Schubert, E., & Zimek, A. (2017). The (black) art of runtime evaluation: Are we comparing algorithms or implementations? *Knowledge and Information Systems*, 52(2), 341–378. <https://doi.org/10.1007/s10115-016-1004-2>
- Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., Ortiz-Ospina, E., & others. (2020). Coronavirus pandemic (COVID-19). *Our World in Data*.
- Nations, U. (n.d.). Country Insights. *Human Development Reports*.
- Nicholson, L. B. (2016). The immune system. *Essays in Biochemistry*, 60(3), 275. <https://doi.org/10.1042/EBC20160017>
- Nielsen, F. (2016). *Hierarchical Clustering* (pp. 195–211). [https://doi.org/10.1007/978-3-319-21903-5\\_8](https://doi.org/10.1007/978-3-319-21903-5_8)
- Noncommunicable diseases: Risk factors*. (2017). <https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/ncd-risk-factors>
- Porcheddu, R., Serra, C., Kelvin, D., Kelvin, N., & Rubino, S. (2020). Similarity in Case Fatality Rates (CFR) of COVID-19/SARS-COV-2 in Italy and China. *Journal of Infection in Developing Countries*, 14(2), 125–128. <https://doi.org/10.3855/jidc.12600>
- Raschka, S., & Mirjalili, V. (2017). *Python Machine Learning* (1st ed.). Packt. <https://www.packtpub.com/en-us/product/python-machine-learning>

- ning-9781783555130?type=print&srsltid=AfmBOoq8veAbXuCIFO-etHOJOPDZGKqCyJdNJiaZYEiZmhddiAQ8xUNXZ
- Ren, Y., Pu, J., Yang, Z., Xu, J., Li, G., Pu, X., Yu, P. S., & He, L. (2024). Deep Clustering: A Comprehensive Survey. *IEEE Transactions on Neural Networks and Learning Systems*, 1–21. <https://doi.org/10.1109/tnnls.2024.3403155>
- Rizvi, S. A., Umair, M., & Cheema, M. A. (2021). Clustering of countries for COVID-19 cases based on disease prevalence, health systems and environmental indicators. *Chaos, Solitons and Fractals*, 151. <https://doi.org/10.1016/j.chaos.2021.111240>
- Saunders, J. A. (1980). Cluster Analysis for Market Segmentation. *European Journal of Marketing*, 14(7), 422–435. <https://doi.org/10.1108/EUM0000000004918/FULL/XML>
- Sudre, C. H., Lee, K. A., Ni Lochlainn, M., Varsavsky, T., Murray, B., Graham, M. S., Menni, C., Modat, M., E Bowyer, R. C., Nguyen, L. H., Drew, D. A., Joshi, A. D., Ma, W., Guo, C.-G., Lo, C.-H., Ganesh, S., Buwe, A., Capdevila Pujol, J., Lavigne du Cadet, J., ... Ourselin, S. (2021). Symptom clusters in COVID-19: A potential clinical prediction tool from the COVID Symptom Study app. In *Sci. Adv* (Vol. 7). <https://www.science.org>
- Tan, P., Steinbach, M., Karpatne, A., & Kumar, V. (2018). *Introduction to Data Mining* (2nd ed.).
- Tressa, N., Asha, V., Kumar, P., Shree, O., Uday Kiran, M., & Reddy, V. V. S. (2024). Customer-Based Market Segmentation Using Clustering in Data mining. *2nd International Conference on Intelligent Data Communication Technologies and Internet of Things, IDCIoT 2024*, 687–691. <https://doi.org/10.1109/IDCIoT59759.2024.10467258>
- Wedel, M., & Kamakura, W. A. (2000). *Market Segmentation* (Vol. 8). Springer US. <https://doi.org/10.1007/978-1-4615-4651-1>
- WHO. (2019, December 18). *WHO global report on trends in prevalence of tobacco use 2000-2025, third edition*. <https://www.who.int/publications/i/item/who-global-report-on-trends-in-prevalence-of-tobacco-use-2000-2025-third-edition>
- WHO. (2020a). *GLOBAL TUBERCULOSIS REPORT 2020*. <http://apps.who.int/bookorders>.
- WHO. (2020b). *Mortality and global health estimates*. <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/>
- World Health Statistics*. (2021). <https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/world-health-statistics>
- Zumel, Nina., & Mount, John. (2020). *Practical data science with R* (2nd ed.). Manning Publications.

# Güncel Ekonometrik ve İstatistiksel Uygulamalar ile Akademik Çalışmalar

Editörler:

Doç. Dr. Elif Tuna

Doç. Dr. Serkan Akoğul

 ÖZGÜR  
YAYINLARI

