

T.C
İNÖNÜ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ



**FAKTÖR ANALİZİ VE KÜMELEME
ANALİZİNİN KARŞILAŞTIRILMASI:
Teori ve Uygulama**

YÜKSEK LİSANS TEZİ

DANIŞMAN
Prof. Dr. Mehmet GÜNGÖR

HAZIRLAYAN
Esen ÇÖP

MALATYA-2022

**T.C.
İNÖNÜ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI**

YÜKSEK LİSANS TEZİ

**FAKTÖR ANALİZİ ve KÜMELEME ANALİZİNİN
KARŞILAŞTIRILMASI: Teori ve Uygulama**

Esen ÇÖP

**Danışman
Prof. Dr. Mehmet GÜNGÖR**

MALATYA, 2022

ONUR SÖZÜ

Prof. Dr. Mehmet GÜNGÖR'ün danışmanlığında yüksek lisans tezi olarak hazırladığım **“FAKTÖR ANALİZİ VE KÜMELEME ANALİZİNİN KARŞILAŞTIRILMASI: Teori ve Uygulama”** başlıklı bu çalışmanın, bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın tarafımdan yazıldığını ve yararlandığım bütün yapıtların hem metin içinde hem de kaynakçada yöntemine uygun biçimde gösterilenlerden oluştuğunu belirtir, bunu onurumla doğrularım.

Esen ÇÖP

İmza



TEŞEKKÜR

Öncelikle danışmanlığımı üstlenen, çalışmam boyunca gerekli olan kaynakları bulmamda bana yardımcı olan, engin görüşleri ve deneyimleriyle yoluma ışık tutan değerli hocam Prof. Dr. Mehmet GÜNGÖR’e teşekkür ederim. Yüksek lisans eğitimim boyunca bana yön gösteren ve benden desteklerini esirgemeyen Ekonometri bölümü hocalarına teşekkür ederim.

Çalışmam boyunca benden maddi ve manevi desteklerini esirgemeyen ablam Ebru ÇÖP ve eşi Arif ÇÖP’e ve son olarak bana gerekli olan çalışma ortamını sağlayan, hayatımın her alanında olduğu gibi, tez çalışmamda da beni destekleyen ve yanımda olan başta annem Sakine ÇÖP ve babam Ali ÇÖP olmak üzere bütün aileme teşekkürü bir borç bilirim.

Esen ÇÖP
MALATYA, 2022

ÖZET

Bu çalışmada, çok değişkenli istatistiksel yöntemlerden olan faktör analizi ve kümeleme analizinin karşılaştırılması amaçlanmıştır. Bu kapsamda, pandemi döneminin makro ekonomik değişkenler üzerindeki etkisi ele alınarak verilere faktör analizi ve kümeleme analizi uygulanmış, elde edilen sonuçlar karşılaştırılmıştır. Çalışmada, ekonomik birimlerden seçilen 15 adet değişken kullanılmıştır.

Faktör analizi sonucunda üç tane faktör belirlenmiştir. Faktörlerden ilki; “döviz, rezerv ve krediler”, ikincisi; “dış ticaret ve turizm”, üçüncüsü ise “taşınmazlar ve ekonomik büyüme” olarak adlandırılmıştır.

Kümeleme analizi sonucunda ise üç tane küme belirlenmiştir.

Faktör analizi ve kümeleme analizinde, kullanılan değişken sayısı ve analiz sonucunda belirlenen küme sayıları birbirine eşit olarak belirlenmiştir. Faktör analizi ve kümeleme analizi sonuçları arasındaki fark ise değişkenlerin kümelere göre dağılımıdır.

Çalışmanın sonucunda, pandemi döneminde uygulanan tedbirlerin ekonomik göstergeler üzerinde etkili olduğu tespit edilmiştir. Dolayısıyla pandemi öncesi ve sonrası ekonomik dinamiklerin farklı olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Faktör Analizi, Kümeleme Analizi, Ekonomik Göstergeler.

ABSTRACT

In this study, it is aimed to compare factor analysis and cluster analysis, which are multivariate statistical methods. In this context, factor analysis and clustering analysis were applied to the data by considering the effect of the pandemic period on macroeconomic variables, and the results were compared. In the study, 15 variables selected from economic units were used.

As a result of factor analysis, three factors were determined. The first of the factors is “currency, reserves and loans”, the second is “Foreign trade and tourism”, and the third is “immovables and economic growth”.

As a result of the cluster analysis, three clusters were determined.

In factor analysis and cluster analysis, the number of variables used and the number of clusters determined as a result of the analysis were determined to be equal to each other. The difference between the results of factor analysis and clustering analysis is the distribution of variables according to clusters.

As a result of the study, it was determined that the measure simple mented during the pandemic period were effective on economic indicators. Therefore, it has been concluded that the economic dynamics before and after the pandemic are different.

Keywords: Factor Analysis, Cluster Analysis, Economic Indicators.

İÇİNDEKİLER

ONUR SÖZÜ	iii
TEŞEKKÜR	iv
ÖZET	v
ABSTRACT	vi
İÇİNDEKİLER	vii
TABLolar LİSTESİ	xi
ŞEKİLLER LİSTESİ	xii
GRAFİKLER LİSTESİ	xiii
KISALTMALAR LİSTESİ.....	xiv
1. GİRİŞ.....	1
2. FAKTÖR ANALİZİ	3
2.1. Faktör Analizi Kavramı.....	3
2.2. Faktör Analizi Modeli	4
2.3. Faktör Analizi Çeşitleri.....	5
2.3.1. Açıklayıcı Faktör Analizi:	5
2.3.2. Doğrulayıcı Faktör Analizi.....	5
2.4. Örneklem Büyüklüğü Seçimi.....	6
2.5. Faktör Analizi Modelinin Uygulama Aşamaları.....	6
2.5.1. Kullanılacak Veri Setinin Uygunluğu	6
2.5.1.1. Bartlett Testi.....	7
2.5.1.2. Kaiser-Meyer-Olkin (KMO) Ölçütü.....	7
2.5.2. Faktör Bulma Yöntemleri	8
2.5.2.1. Temel Bileşenler Analizi (TBA)	8
2.5.2.2. Temel Eksenler Analizi	8
2.5.2.3. Maksimum Olabilirlik Yöntemi	9
2.5.2.4. İmaj Faktör Çıkartma Yöntemi	9
2.5.2.5. Ağırlıklandırılmamış En Küçük Kareler Yöntemi	9

2.5.2.6. Genelleştirilmiş En Küçük Kareler Yöntemi.....	10
2.5.2.7. Alfa Faktör Çıkarma Yöntemi	10
2.5.3. Faktör Sayısının Belirlenmesi	10
2.5.3.1. Özdeğer İstatistiğine Göre Belirleme	11
2.5.3.2. Scree Test (Yamaç Grafiği) Kriterine Göre Belirleme	11
2.5.3.3. Varyans Oranına Göre Belirleme.....	11
2.5.4. Faktör Döndürme Teknikleri.....	11
2.5.4.1. Dik (Orthogonal) Rotasyon.....	12
2.5.4.1.1. Varimax Yöntemi.....	13
2.5.4.1.2. Quartimax Yöntemi	13
2.5.4.1.3. Equamax Yöntemi.....	13
2.5.4.2. Eğik (Oblique) Rotasyon.....	13
2.5.4.2.1. Oblimax Yöntemi	13
2.5.4.2.2. Quartimin Yöntemi	13
2.5.4.2.3. Covarimin Yöntemi.....	14
2.5.4.2.4. Biquartimin Yöntemi	14
2.5.4.2.5. Oblimin Yöntemi.....	14
2.5.4.2.6. Binoramin Yöntemi.....	14
2.5.5. Faktör Skorlarının Elde Edilmesi.....	14
2.5.5.1. Thomson (Regresyon) Yöntemi	14
2.5.5.2. Bartlett Yöntemi.....	15
2.5.5.3. Anderson Rubin Yöntemi	15
2.5.6. Faktörlerin İsimlendirilmesi ve Yorumlanması	15
3. KÜMELEME ANALİZİ	16
3.1. Kümeleme Analizi Kavramı	16
3.2. Kümeleme Analizinin Amacı	16

3.3. Kümeleme Analizinin Adımları.....	17
3.3.1.Verilerin Standartlaştırılması ve Değişken Seçimi.....	17
3.3.2. Benzerlik ve Uzaklık Ölçüleri	17
3.3.2.1. Benzerlik Ölçüleri	17
3.3.2.2. Uzaklık Ölçüleri	17
3.3.2.2.1.Mahalonobis Uzaklığı.....	18
3.3.2.2.2. Minkowski Uzaklığı.....	18
3.3.2.2.3. Öklid Uzaklığı.....	18
3.3.2.2.4. Manhattan (City Blok) Uzaklık Ölçüsü.....	18
3.3.3. Kümeleme Yöntemleri.....	18
3.3.3.1. Hiyerarşik Kümeleme Teknikleri.....	19
3.3.3.1.1. Bağlantı Teknikleri.....	20
3.3.3.1.1.1.Tek Bağlantı Tekniği (En Yakın Komşuluk).....	20
3.3.3.1.1.2.Tam Bağlantı Tekniği (En Uzak Komşuluk).....	20
3.3.3.1.1.3. Ortalama Bağlantı Tekniği	20
3.3.3.1.2. Varyans Teknikleri.....	20
3.3.3.1.2.1.Ward’’s Bağlantı Tekniği.....	20
3.3.3.1.3. Merkez Yöntemi	20
3.3.3.2. Hiyerarşik Olmayan Kümeleme Yöntemleri	21
3.3.3.2.1. k-Ortalama Yöntemi	21
3.3.3.2.2. En Çok Olabilirlik Yöntemi.....	21
3.3.3.3. İki Aşamalı Kümeleme Analizi (Two-Step Cluster Analysis).....	21
3.3.4. Kümelerin Geçerliliği ve Yorumu.....	21
4. UYGULAMA.....	22
4.1. Literatür Taraması	22
4.2. Pandemi Sürecinin Ekonomik Etkileri.....	24

4.3. Veri Seti	25
4.4. Bulgular	26
4.4.1. Faktör Analizi Sonuçları	26
4.4.2. Kümeleme Analizi Sonuçları.....	28
5. SONUÇ	32
KAYNAKÇA	34



TABLÖLAR LİSTESİ

Tablo 2.1. KMO Değerleri ve Yorumu.....	7
Tablo 4.1. Değişkenlerin Tanımlamaları ve Kaynakları.....	26
Tablo 4.2. KMO ve Bartlett Testi İstatistikleri	26
Tablo 4.3. Açıklanan Toplam Varyans Tablosu.....	27
Tablo 4.4. Döndürülmüş Faktör Matrisi	28
Tablo 4.5. Durum Özeti	29
Tablo 4.6. Hiyerarşik Kümeleme Analizi Sonuçları	29
Tablo 4.7. Küme Üyeliği	30



ŞEKİLLER LİSTESİ

Şekil 2.1. Dik ve Eğik Rotasyon Gösterimi.....	12
--	----



GRAFİKLER LİSTESİ

Grafik 2.1. Scree Test Grafiği.....	11
Grafik 4.1. Scree Test Grafiği.....	27
Grafik 4.2. Dendrogram Grafiği.....	31



KISALTMALAR LİSTESİ

TCMB:	Türkiye Cumhuriyet Merkez Bankası
AFA:	Açıklayıcı Faktör Analizi
DFA:	Doğrulayıcı Faktör Analizi
KMO:	Kaiser-Meyer-Olkin
TBA:	Temel Bileşenler Analizi
EVDS:	Elektronik Veri Dağıtım Sistemi



1. GİRİŞ

Günümüzde araştırmaların kapsamının genişlemesi ve teknolojinin gelişmesi ile birlikte karmaşık yapıda veri setleri meydana gelmektedir. Birçok araştırmacı bu veri setlerini daha basit yorumlanabilir hale getirmek için sınıflandırmaya ihtiyaç duymaktadır. Çalışmanın amacına uygun olarak bazen değişkenler sınıflandırılırken bazen de birim veya nesneler sınıflandırılmaktadır. Aşağıda tanımları yer alan faktör analizi ve kümeleme analizi, sınıflandırma yapan çok değişkenli istatistiksel yöntemlerdendir.

Faktör analizi, çok sayıdaki değişkeni daha az sayıya indirgeyen çok değişkenli istatistiksel bir yöntemdir. Faktör analizinde sınıflandırma, birbirleriyle ilişkili değişkenler dikkate alınarak yapılmaktadır.

Kümeleme analizi ise nesne veya birimleri özelliklerine göre sınıflandıran çok değişkenli istatistiksel bir analizdir. Ayrıca kümeleme analizinde değişkenler sınıflandırılabilir. Her iki analizde sınıflandırma yapılırken aynı zamanda boyut indirgeme de yapılmaktadır.

Bu çalışmada, ekonomik birimlerden seçilen değişkenlerin faktör analizi ve kümeleme analizi yöntemlerine göre sınıflandırılması gösterilmiştir. İki yöntemle göre belirlenen sonuçlar karşılaştırılmıştır. Bu kapsamda yapılan çalışma beş bölümden oluşmaktadır.

Çalışmanın birinci bölümünde giriş bölümü yer almaktadır. İkinci bölümde faktör analizi teorik olarak açıklanmıştır. Faktör analizi modeli, çeşitleri ve analizde kullanılacak olan örneklem büyüklüğünün seçimi ayrıntılı bir biçimde ele alınmıştır. Bu bölümde son olarak faktör analizinin aşamaları açıklanmıştır.

Üçüncü bölümde; kümeleme analizi teorik olarak açıklanmıştır. Kümeleme analizinin amaçları ve aşamaları detaylı bir biçimde belirtilmiştir.

Dördüncü bölümde; ilk olarak literatür çalışması yapılmıştır. Çalışmaya dâhil edilen değişkenlerin pandemi öncesi ve sonrası durumu ele alınmıştır. Belirlenen verilere ilk olarak açıklayıcı faktör analizi, daha sonra hiyerarşik kümeleme analizi uygulanmış ve bulgular başlığı altında sonuçlar açıklanmıştır. Her iki analizin sonucuna göre, ekonomik değişkenlerin benzer ve farklı olanları belirlenmiştir.

Beşinci bölümde ise faktör analizi ve kümeleme analizi yöntemlerinin sonuçları karşılaştırılıp, iki analiz arasındaki benzerlikler ve farklılıklar açıklanmıştır. Ayrıca

ekonomik deęişkenlerin pandemi döneminde iki analiz yöntemine göre gruplanmasındaki deęişimi literatür dikkate alınarak açıklanmıştır.



2. FAKTÖR ANALİZİ

2.1. Faktör Analizi Kavramı

Faktör analizinin temelleri sosyal tutum ve davranışlara dayanmaktadır. İlk olarak psikoloji bilimi alanında kullanılmıştır. Charles Spearman, 20. Yüzyılın başlarında Amerikan psikoloji dergisinde yayımlanan bir çalışmada ilk defa bu analizi kullanmıştır. Çalışmada zekâ kavramı üzerine durulmuştur. Faktör analizi üzerine yapılan çalışmalar 1925 yılından sonra hız kazanmıştır.

Karl Pearson, 1930 yılında yaptığı çalışma ile temel eksen faktörü yöntemini bu alana dâhil etmiştir. Pearson tarafından planlanan yöntem 1933 yılında Hotelling tarafından geliştirilmiştir. Garnett 1919 yaptığı çalışmalarla çoklu faktör analizini keşfetmiştir. Daha sonra Thurstone yaptığı çalışmalarla bu analizi geliştirmiştir (Harman, 1976: 4). Lawley 1940 yılında maksimum olabilirlik yöntemine ilişkin çalışmalar yapmıştır. Lawley'in sunduğu çalışma faktör sayısının artması durumunda kullanılamamıştır. Jöreskop 1967'de yaptığı çalışma ile bu sorunu çözmüştür (Bektaş, 2017: 20).

1900'lerin ortasına gelindiğinde teknoloji alanında yaşanan gelişmelerle birlikte analizin kullanım alanı da genişlemiştir. Bilim insanları bu yıllarda yaptıkları çalışmalarla faktör analizini açıklayıcı (açımlayıcı) ve doğrulayıcı faktör analizi olmak üzere iki ana başlığa ayırmıştır.

İlk defa psikoloji alanında kullanılan faktör analizi ilerleyen yıllarda iktisat, işletme, coğrafya, tıp, eğitim gibi birçok alanda kullanılmaya başlanmıştır. Analize özellikle sosyal bilimlerde kullanılan soyut kavramlar konu olabileceği gibi somut kavramlar da konu olabilir. Analizde kullanılacak konu ile ilgili herhangi bir kısıtlama bulunmamaktadır.

Güncel araştırmaların artması ve kapsamının genişlemesi sonucunda bu analize olan ihtiyaç artmaktadır. Araştırmacılar doğru bilgiyi elde etmek ve çalışmalarının güvenilirliği için veri sayısını ve değişken sayısını artırmaktadır. Değişken sayısının artmasına paralel olarak çalışmalarda birbirleriyle bağlantılı olan değişkenler ortaya çıkmaktadır.

Bu konu ile ilgili aşağıda belirtilen iki sorunun üzerinde durulması gerekmektedir (Albayrak, 2006: 108);

- Orijinal değişken seti yerine daha az sayıda değişkenden oluşan bir alt değişken seti kullanılabilir mi?
- Ölçülen değişken setinin temel boyutları (özellikleri) nelerdir?

Bu iki önemli sorunun cevabı ise çok değişkenli istatistiksel analizlerden biri olan faktör analizinin tanımı ve amaçlarıyla bağlantılıdır. Faktör analizinin tanımı ve amaçları aşağıdaki gibidir;

Faktör analizi p boyutlu bir uzayda birbiri ile ilişkili değişkenleri bir araya getirerek, az sayıda yeni birbiriyle bağlantısız değişkenin bulunduğu çok değişkenli analiz yöntemlerinden birisidir (Tatlídil, 2002: 167). Analizde dikkat edilmesi gereken konu en az bilgi kaybıyla daha az sayıda faktör oluşturulmasıdır.

Faktör analizinin temel amacı, analiz sonucunda gizil değişkenleri ortaya çıkarmaktır. Değişkenlerin gözlenemeyen ilişkileri sonucu ortaya çıkan değişkene gizil değişken denir. Faktör analizinin diğer amaçları ise boyut indirgeme ve faktörleri hazır veri haline getirmektir. Analiz sonucunda bulunan faktörler diğer çok değişkenli analizlerde kullanılmak üzere veri olarak kullanılabilir.

2.2. Faktör Analizi Modeli

Bağımlı ve bağımsız değişken ayrımının olmadığı matematiksel model aşağıdaki gibidir (Bektaş, 2017: 29);

$$X_i = A_{i1} F_1 + A_{i2} F_2 \dots + A_{im} F_m + U_i \quad (2.1)$$

Denklem 2.1’de yer alan simgelerin açıklaması aşağıdaki gibidir;

X_i = standartlaştırılmış i. değişken

F = ortak faktör

A_{ij} = ortak j faktöründe standartlaştırılmış çoklu regresyon sayısı

m = ortak faktör sayısı

U_i = i. faktör için eşsiz (unique) faktör sayısıdır.

Modelde i sayıda değişken arasındaki ilişkiler m tane faktör tarafından açıklanmaktadır. Faktör analizinde değişkenlerin aynı birimde olması için standartlaştırma yapılmaktadır. Eşsiz faktörler, ortak faktörlerce açıklanamayan kısmı ifade etmektedir. Eşsiz faktörlerin hem ortak faktörlerle hem de birbirleriyle ilişkili olmadığı farz edilmektedir. F_m ’nin tahmini gözlenen değişkenlerle yapılır. Bu tahminler gözlenen değerlerin doğrusal bir birleşimidir ve i’inci faktör olan F ’nin eşitliği aşağıdaki gibidir;

$$F_i = W_{i1}X_1 + W_{i2}X_2 \dots + W_{ik}X_k \quad (2.2)$$

Denklem 2.2’de W_i vektör skoru, k ise faktör sayısıdır.

2.3.Faktör Analizi Çeşitleri

Analiz, açıklayıcı faktör analizi yöntemi ve doğrulayıcı faktör analizi yöntemi olmak üzere iki başlık altında ele alınmaktadır.

2.3.1. Açıklayıcı Faktör Analizi

Bu yöntemde değişkenler arasındaki doğrusal yapı ortaya koyularak boyut indirgemesi yapılmaktadır. Değişkenlerin ilişkisi araştırmacı tarafından belirlenmektedir. Çok değişkenli istatistiksel yöntemler arasında boyut indirgeme konusunda önemli bir yeri olan yöntem, güvenilirlik ve ölçek geliştirme analizinde de sıklıkla kullanılmaktadır. İstatistikî olarak güvenilirlik, aynı kişileri kapsayacak şekilde aynı ve farklı birimlerde (yer veya zaman) ölçüm sonuçlarının benzer olmasıdır. Belirli bir zaman geçtikten sonra veya farklı mekânlarda uygulanan testin ölçüm sonuçlarının benzer olması gerekmektedir.

2.3.2. Doğrulayıcı Faktör Analizi

Bu yöntem ile kuramsal bir yapı yönünde geliştirilen ölçme aracından elde edilen verilerle, modeldeki yapının doğrulanıp doğrulanmadığı test edilmeye çalışılmaktadır. Yöntem özellikle kuram geliştirme ve var olan kuramların geçerliliğini belirlemede sıklıkla kullanılmaktadır (Erkuş, 2003’ten aktaran Çokluk vd., 2012: 177). Bu yöntemde, önceden belirlenen bir model üzerinde çalışılmaktadır. Burada amaç, modelin uyumunu test etmektir. Çalışmayı yapan araştırmacının açıklayıcı faktör analizi hakkında da bilgisinin olması gerekmektedir. Çünkü açıklayıcı faktör analizi, doğrulayıcı faktör analizinin temelini oluşturmaktadır.

İki model arasında benzer ve farklı durumlar mevcuttur. Benzer durumlar her iki modelin doğrusal istatistik modellerine dayalı olması, değişkenlerin normal dağılıma uyumu, modellerin gizil yapıyı ortaya koymaları olarak sayılabilir. Açıklayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) arasındaki farklılıklar ise DFA’da hazır ve faktör sayıları belli bir model üzerinde çalışılırken, AFA’da varyansı maksimum tutarak modeldeki yapı ortaya çıkartılmaktadır. AFA’da yeni bir model keşfedilmektedir (Suhr, 2006: 4-5).

2.4. Örneklem Büyüklüğü Seçimi

Örneklem büyüklüğünün belirlenmesi, analizin güvenilirliği açısından son derece önemlidir. Nitekim küçük örneklemeler tarafından belirlenen korelasyon sayısı düşüktür. İstatistikte genel bir kanıya göre anakütleden çekilen küçük örneklerle yapılan tahminler tutarsızdır. Bu yüzden faktör analizinde örneklem büyüklüğünün seçimi önemli bir konudur. Faktör analizinde kullanılacak olan örneklem büyüklüğü seçimi hususunda literatürde birçok yaklaşım söz konusudur. Literatürde bu konuya dair bazı araştırmacıların fikirleri aşağıdaki gibidir;

Gorsuch (1983) örneklem büyüklüğünün 100'den az olmaması gerektiğini savunmuştur. Gorsuch ayrıca ölçülen değişken başına beş katılımcı olması gerektiğini vurgulamaktadır. Nunally (1978) ve Everitt (1975) ise bu oranın bire on olması gerektiğini savunmuşlardır (Fabrigar vd., 1999: 274). Genel olarak 50 sayısından az örneklerle yapılan araştırmalar faktör analizine uygun değildir. Hatta mümkünse bu oran 100 ve üzeri olarak alınmalıdır (Hair vd., 2010: 101). Bazı araştırmacılar örneklem büyüklüğü seçiminin bu oranların yanı sıra değişkenlerin özelliklerine dayandığını savunmuşlardır. Macculum vd. (1999) yazdıkları bir makalede bu özelliklerden biri olan ortak varyans terimi üzerinde durmuşlardır. Yaptıkları çalışmalar sonucunda örneklem büyüklüğü seçiminin ortak varyansa bağlı olduğunu belirtmişlerdir (Albayrak, 2004: 58). Yukarıda açıklanan yaklaşımlardan anlaşılacağı üzere bu konu hakkında kabul edilmiş bir kural söz konusu değildir.

2.5. Faktör Analizi Modelinin Uygulama Aşamaları

2.5.1. Kullanılacak Veri Setinin Uygunluğu

Analize başlamadan önce sonuçların geçerliliği ve doğru yorumun yapılması açısından kullanılacak verinin bazı özelliklerinin bilinmesi gerekmektedir. Verilerin istatistiksel olarak normallik, çoklu doğrusal bağlantı ve doğrusallık gibi parametrik varsayımları sağlaması gerekmektedir. Verilerin ölçümünde ölçek tipi olarak en azından likert ölçeği kullanılmalıdır.

Bunun yanı sıra gözlenen değişkenler arasında korelasyon düzeyinin yeterliliğine bakılması gerekmektedir. Çünkü analiz sonucu bir veya daha fazla anlamlı faktörün oluşması, değişkenler arasındaki korelasyonun yüksek olmasına bağlıdır. Kısmi korelasyon katsayısı incelendiğinde oluşan değerlerin 0,30'dan büyük olması gerekmektedir.

Ayrıca verilerin analiz için uygunluğunun değerlendirilmesinde, Bartlett testine ve Kaiser-Meyer-Olkin (KMO) ölçütü değerine bakılması gerekmektedir.

2.5.1.1. Bartlett Testi

Bartlett testinde, korelasyon matrisindeki korelasyonların anlamlı olup olmadığı bulunmaktadır. Bir başka ifadeyle korelasyon matrisinin birim matris olup olmadığı araştırılmaktadır. Testin hipotezi ve istatistiği aşağıdaki gibidir (Albayrak, 2006: 130-131);

H_0 : Korelasyon matrisi birim matristir.

H_1 : Korelasyon matrisi birim matris değildir.

$$X^2 = -[(N - 1) - (2p + 5) / 6] \ln|R| \quad (2.3)$$

Denklem 2.3'te yer alan N, örneklem büyüklüğü, p, değişken sayısı, |R| ise korelasyon matrisinin determinantıdır.

Testin serbestlik derecesi ise $(p(p - 1) / 2)$ formülü ile hesaplanmaktadır. Analize devam edilmesi için Ki-kare test istatistiği sonucunun belirlenen kritik değerden büyük olması gerekmektedir. Bir başka deyişle yokluk hipotezinin reddedilmesi gerekmektedir.

2.5.1.2. Kaiser-Meyer-Olkin (KMO) Ölçütü

Bu ölçüt, veri setinin analiz için elverişliliğini test etmek için kullanılmaktadır. Gözlenen korelasyon katsayıları büyüklüğü ile kısmi korelasyon katsayılarının büyüklüğünü karşılaştıran bir ölçüttür (Kalaycı, 2010: 322).

Ölçüt için gereken aralık ve yorumlar aşağıdaki tabloda belirtilmiştir;

Tablo 2.1.KMO Değerleri ve Yorumu	
KMO değeri	Yorum
0,90 ve üzeri	Mükemmel
0,80	Çok iyi
0,70	İyi
0,60	Orta
0,50	Zayıf
0,50'den aşağısı	Kabul edilemez

KMO değerinin 1'e yakınlığı veri setinin faktör analizi için daha iyi oranda uygunluğunu gösterir. Bu değer 0,50 oranının altında çıkması durumunda ise faktör analizi yapılamaz.

Uygulanan bu adımlar neticesinde araştırmada kullanılacak olan analizin faktör analizi olduğu sonucuna varıldıktan sonra ilerleyen adımlara geçilir. Bu adımlar sırasıyla aşağıdaki gibidir;

- Faktör türetmek için uygun yöntemlerin seçimi,
- Araştırma sonucunda elde edilmesi gereken faktör sayısının belirlenmesi,
- Basit bir yapının elde edilmesi ve yorumlanabilir faktörler elde etmek için rotasyon adımı,
- Araştırmacının bilgi ve deneyimlerine uygun olarak oluşan yeni faktörleri isimlendirmesi ve sonuçları yorumlamasıdır.

2.5.2. Faktör Bulma Yöntemleri

Faktör bulma yöntemlerinde amaç, korelasyon matrisini yeniden oluşturmaktır.

2.5.2.1. Temel Bileşenler Analizi (TBA)

Faktör çıkartma yöntemleri arasında en sık kullanılan analiz olan temel bileşenler analizinde amaç, her bir bileşen ile veri setindeki maksimum varyansı ortaya çıkarmaktır. Analizde ilk temel bileşen maksimum varyansa sahiptir ve bu bileşenin toplam varyansa katkısı en yüksek düzeydedir. Ardışık olarak ilerleyen diğer bileşenlerin varyansları ise azalan bir yapıdadır. Yöntem matematiksel olarak tek bir çözüme sahiptir. TBA modeli en kısa haliyle aşağıdaki gibi yazılabilir (Albayrak, 2006: 133);

$$X_j = b_{j1}F_1 + b_{j2}F_2 \dots + b_{jp}F_p \quad j=1,2,\dots,p \quad (2.4)$$

Denklem 2.4'te yer alan X_j =standartlaştırılmış j. değişkendir. Ayrıca modeldeki p sayıdaki değişken, bağımsız p tane temel bileşenle temsil edilmektedir.

2.5.2.2. Temel Eksenler Analizi

Bu yöntem bazı kaynaklarda faktör analizi, temel faktör analizi veya ortak faktör analizi olarak ifade edilmektedir. Yöntemin temel amacı, ardışık faktörler ile veri setinden maksimum sayıda ortogonal varyansı çıkarmaktır. Bu yöntemin avantajları, yöntemin yaygın kullanımı, eşsiz (unique) ve hata varyansının kaldırılarak ortak varyansın bulunmasıdır. Bu durum ise faktör analizi yapısına uymaktadır. Bu yöntem, ortak varyans terimi üzerinde durmaktadır. Ortak faktör, bir araştırmadaki gözlemlenen değişkenleri etkileyen ve değişkenler arasındaki ilişkiyi açıklayan yapı olarak tanımlanmaktadır. Ortak varyans ise ortak faktörler tarafından açıklanan varyanstır. Bu

yöntem diğer faktör bulma yöntemleri ile karşılaştırıldığında bazen korelasyonu yeniden türetme konusunda başarılı değildir (Tabachnick ve Fidell, 2007: 636).

2.5.2.3. Maksimum Olabilirlik Yöntemi

Özellikle doğrulayıcı faktör analizinde kullanılan bu yöntemde, evren parametreleri tahmin edilmeye çalışılmaktadır. Bu tahmin aşamasında örneklem korelasyon matrisinin olasılığını maksimum yapan faktör yükleri kullanılmaktadır. Yöntemin en önemli avantajlarını bir modelin uyumluluk indeksini hesaplamada kullanımı ve faktörler için manidarlık testi uygulanarak araştırmanın başında belirlenen faktör sayısı ile kurulan hipotezi test etmesi oluşturmaktadır (Tabachnick ve Fidell, 2007: 636).

Maksimum olabilirlik yöntemi, çoklu normallik varsayımının sağlandığı veri grubu için en uygun olan yöntemdir. Maksimum olabilirlik yöntemi bu varsayımın sağlandığı modelde gözlem verilerine dayalı korelasyon katsayılarının alabildiği en yüksek değerin olmasını sağlamaktadır (Şencan, 2005: 374).

2.5.2.4. İmaj Faktör Çıkartma Yöntemi

Bu yöntem temel eksenler analizindeki eksik yönlerin tamamlanması amacıyla oluşturulmuştur. Temel eksenler analizi eşsiz ve ortak faktörler arasındaki ilişkiyi açıklama konusunda yeterli bir yöntemdir. Ancak bir veri grubunda bu iki faktörün nasıl belirleneceği konusunda yeterli bir yöntem değildir. Bu sebeple Guttman (1953) tarafından imaj faktör çıkartma yöntemi belirlenmiştir. Yöntem başka değişkenler tarafından “yansıtılan” gözlenen bir değişkenin varyansı sonucunda oluşan faktörler arasındaki dağılımı yapmasıdır. Bu sebeple de imaj faktör ismini almıştır (Çokluk vd., 2012: 199). Bu yöntem, temel bileşenler analizi ve temel eksenler analizi ile benzer durumlara sahiptir. TBA ile benzerliği korelasyon matrisinin pozitif köşegeninde sabit değerler olması sebebiyle matematiksel olarak tek bir çözüme sahip olmasıdır. Temel eksenler analizi ile olan benzerliği ise köşegen değerlerinin eşsiz ve hata değişkenliğinden arındırılmış bölge (ortak varyans) olmasıdır (Tabachnick ve Fidell, 2007: 636).

2.5.2.5. Ağırlıklandırılmamış En Küçük Kareler Yöntemi

Comrey (1962) tarafından keşfedilip; Harman ve Jones (1966) tarafından tasarlanmıştır. Literatürde en düşük kalıntı olarak da kullanılan yöntemde amaç, gözlenen ve yeniden üretilen korelasyon matrisleri arasındaki farkın karelerini en düşük

seviyeye indirmektir. Yöntemde ortak varyanslar çözümden sonra belirlenmektedir. Bunun sebebi ise bu yöntemde korelasyon matrisinin köşegen dışı elemanlarının dikkate alınması ve bunun sonucunda ortak varyansın çözümden türetilmesidir (Tabachnick ve Fidell, 2007: 636-637).

Her bir yöntemin kendine has bazı avantajları ve dezavantajları vardır. Veri grubunda normal dağılımdan sapmanın büyük olduğu durumlarda bu yöntemin kullanımı, yöntemin bilinen en önemli avantajıdır. Yöntemin sonuçlarının, değişkenlerin ölçek tipi değişimine duyarlı olması bu yöntemin bilinen bir dezavantajıdır. Bu dezavantaj neticesinde varyans-kovaryans matrisi yerine korelasyon matrisi kullanılarak yapılan çalışmalarda farklı sonuçlara sebep olmaktadır (Bektaş, 2017: 68).

2.5.2.6. Genelleştirilmiş En Küçük Kareler Yöntemi

Bu yöntemin amacı, ağırlıklandırılmamış en küçük kareler yönteminde olduğu gibi gözlenen ve yeniden üretilen korelasyon matrisleri arasındaki farkın karelerini en düşük seviyeye indirmektir. Bu iki yöntem arasındaki fark ise genelleştirilmiş en küçük kareler yönteminde değişkenlere ağırlıklandırılma yapılmasıdır. Yöntemde ağırlıklandırılma, ortak varyansa sahip olan değişkenler eşsiz varyansa sahip olan değişkenlere göre daha çok olacak şekilde yapılmaktadır. Yani modeldeki diğer değişkenlerle ilişkili olmayan değişkenlerin çözümde önemli bir yeri yoktur (Tabachnick ve Fidell, 2007: 637).

2.5.2.7. Alfa Faktör Çıkarma Yöntemi

Bir diğer faktör türetme yöntemi olan alfa faktör çıkarma yönteminde amaç, ortak faktörlerin güvenilirliğini maksimize etmektir (Alpar, 2011: 290).

Yöntem bir değişken popülasyonundan tekrarlanan değişken örnekleri alındığında hangi ortak faktörlerin tutarlı bir şekilde bulunduğunu araştıran psikometri araştırmaları sonucunda ortaya çıkmıştır. Bu yöntem grup farklarının güvenilirliğinden daha çok ortak faktörlerin güvenilirliği ile ilgilenmektedir. Yöntemde süreç yinelemeli bir şekilde devam eder (Tabachnick ve Fidell, 2007: 637).

2.5.3. Faktör Sayısının Belirlenmesi

Bu aşamanın amacı, değişkenler arasındaki ilişkileri en iyi derecede ölçebilecek daha az sayıda anlamlı faktör ya da faktörler elde etmektir (Kalaycı, 2010: 322). Bu faktörleri belirlemek için aşağıda bahsedilen yöntemler ele alınmaktadır.

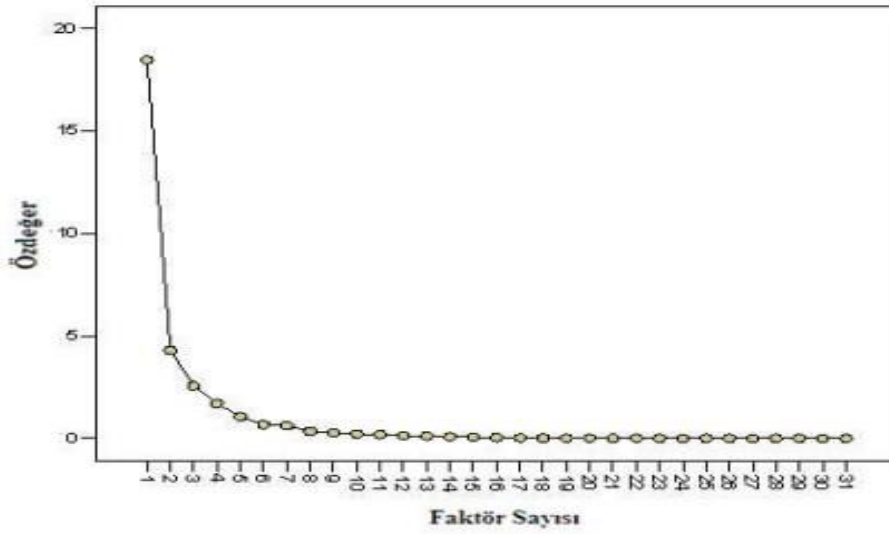
2.5.3.1. Özdeğer İstatistiğine Göre Belirleme

Özdeğer, bir faktörün açıkladığı toplam varyanstır. Özdeğer istatistiği 1'den büyük olan faktörler kabul edilir. Ayrıca bu konuda dikkat edilmesi gereken bir diğer kriter Joliffe kriteridir. Bunun için belirlenen kısıt ise 0.7'nin üstünde olan faktörlerin modele dâhil edilmesidir (Kalaycı, 2010: 322).

2.5.3.2. Scree Test (Yamaç Grafiği) Kriterine Göre Belirleme

Bu kriterde grafik üzerinde özdeğer sayısı (dikey eksen) ve faktör numaraları (yatay eksen) belirtilerek uygun faktör sayısına karar verilmektedir. Grafik 2.1'de bulunan eğrinin yüksek özdeğer sayısı ile başlayıp gittikçe azalan bir yapıya büründüğü görülmektedir. Aynı zamanda bu azalmanın belli bir süreden sonra durağanlaştığı ve düz bir çizgi halini aldığı da görülmektedir. Grafiğin düz bir çizgi halini almadan önceki alan bize bu yöntemle belirlenen faktör sayısını vermektedir. Aşağıdaki grafik incelendiğinde yamaç grafiği kriterine göre 5 adet faktör türetilmiştir.

Grafik 2.1.Scree Test Grafiği



Kaynak: Yılmaz, 2011: 54

2.5.3.3. Varyans Oranına Göre Belirleme

Açıklanan varyansın toplam varyans üzerindeki gücü faktörlerin temsil yeteneğini belirlemektedir. Literatürdeki çalışmalar dikkate alındığında açıklanan varyans oranının %50'yi geçiyor olmasının yeterli olduğu görülmektedir.

2.5.4. Faktör Döndürme Teknikleri

Araştırmada uygun faktörlere karar verildikten sonra bir sonraki aşama olan döndürme (rotasyon) aşamasına geçilir. Modelden daha basit bir yapı elde etmek ve yorumlanabilir bir faktör matrisi elde etmek için rotasyon adımı uygulanır.

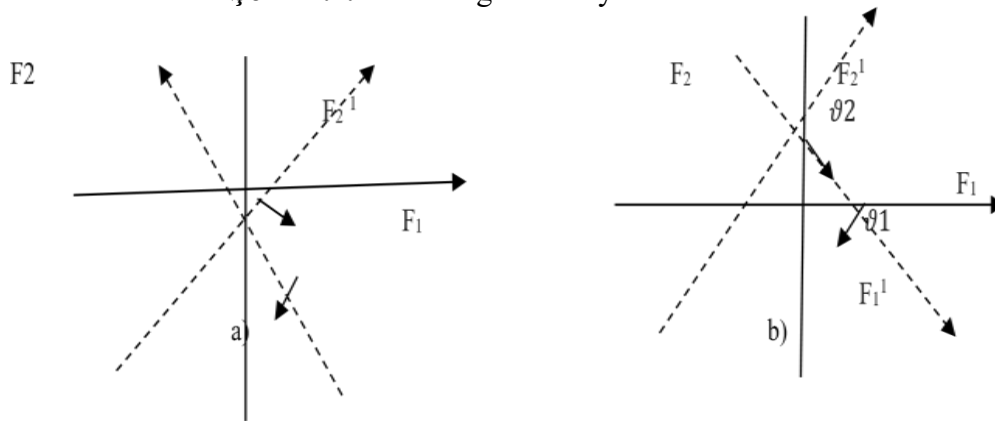
Çözümün basit bir yapı olabilmesi için Thurstone tarafından belirlenen beş şart aşağıdaki gibidir (Alpar, 2011: 291-292);

1. Faktör yükleri matrisindeki her satır ya da faktör yükleri matrisi en azından 1 adet sıfır değeri içermelidir.
2. Dikkate alınacak faktör sayısı k ise her bir kolonda (faktörde) en az k tane sıfır değeri bulunmalıdır.
3. Yükler matrisinin her kolon çiftinin birinde yükler görülürken diğerinde görülmemelidir.
4. Faktör sayısının dört ya da fazla olduğu durumlarda her faktör çifti için değişkenlerin büyük çoğunluğunun yükleri sıfır olmalıdır.
5. Dikkate alınan faktör sayısı dört ya da daha fazla ise, yükler matrisindeki her faktör çiftinde az sayıda değişkenin yük değeri olmalıdır.

Araştırmacılar döndürülmemiş çözümlerin yeterli olmadığını düşünmektedir. Çözümde rotasyon adımı uygulanarak ön analize eşlik eden bazı belirsizlikler azaltılarak yorumlama kolaylaştırılmaktadır (Hair vd., 2010: 114).

Faktör rotasyonu dik (orthogonal) ve eğik (oblique) rotasyon olmak üzere ikiye ayrılır. Şekillerden ilki (a) dik rotasyonu, ikincisi (b) ise eğik rotasyonu belirtmektedir.

Şekil 2.1. Dik ve Eğik Rotasyon Gösterimi



Kaynak: Eşmekaya, 2019: 31

2.5.4.1. Dik (Orthogonal) Rotasyon

Bu yöntemde eksenlerin konumu değiştirilmeden döndürme yapılır. Bir başka ifadeyle 90 derecelik açı ile döndürme işlemi yapılır. Bu yöntemle elde edilen faktörler birbirinden bağımsızdır (Kalaycı, 2010: 322). Bu özellik dik ve eğik döndürmeyi ayırt etmemizde belirleyicidir. Dik döndürmede faktörler arasında korelasyon yok iken eğik döndürmede faktörler arasında korelasyon vardır. Bir başka ifadeyle faktörler birbirine

bağımlıdır. Üç adet dik rotasyon yöntemi vardır. Bu yöntemler varimax, quartimax ve equamaxtır.

2.5.4.1.1. Varimax Yöntemi

Kaiser tarafından geliştirilen yöntemde amaç faktör varyanslarını maksimum yaparak basit bir yapı elde etmektir. Dik döndürme yöntemleri arasında en sık kullanılan bu yöntemde faktör matrisinin sütunları ile işlem yapılır. Yöntemde her sütundaki bazı yük değeriyle 1'e yaklaştırılırken geriye kalan yük değerleri 0'a yaklaştırılır (Tatlıdil, 2002: 180).

2.5.4.1.2. Quartimax Yöntemi

İki faktörlü modellerden en iyi sonucu veren bu yöntemde basit bir yapı elde etmek için faktör matrisinin satırları kullanılır. Yöntemde her satırdaki herhangi bir değer büyütülüp 1'e yaklaştırılırken, öteki değerler küçültülerek 0'a yaklaştırılır (Tatlıdil, 2002: 180). Varimax yaklaşımında faktör merkezli işlem yapılırken; quartimax yaklaşımında değişken merkezli işlem yapılır.

2.5.4.1.3. Equamax Yöntemi

Bu yaklaşım varimax ve quartimax yöntemlerinin kesiştiği alanı ifade eder. Basit bir yapı elde etmek amacıyla faktörler ve değişkenler aynı zamanda kullanılır. Bu yöntem genel olarak kabul görmemiştir ve nadiren kullanılmaktadır (Hair vd., 2010: 114).

2.5.4.2. Eğik (Oblique) Rotasyon

90 derecelik açı ile döndürme işleminin yapılmadığı yöntemdir. Bu döndürme tekniği sonuçların veriye en iyi şekilde uyumu öncelikliyse ve elde edilen faktörler arasında ilişkinin varlığına ait bir öngörü varsa kullanılır (Alpar, 2011: 293). Oblimax, quartimin, covarimin, biquartimin, oblimin, binoramin yöntemleri en sık kullanılan eğik döndürme yöntemleridir.

2.5.4.2.1. Oblimax Yöntemi

Yöntem basıklık katsayısının maksimize edilmesi üzerine kuruludur (Tatlıdil, 2002: 183).

2.5.4.2.2. Quartimin Yöntemi

Bu yaklaşımda basit bir yapı elde etmek için faktör matrisindeki kare yüklerin çaprazlarının toplamını en aza indirmek amaçlanır. Faktörler arasında oldukça yüksek

korelasyonun mevcut olduğu durumda bu yöntem kullanılabilir (Tabachnick ve Fidell, 2007: 639).

2.5.4.2.3. Covarimin Yöntemi

Caroll tarafından önerilen yaklaşımda faktör matrisindeki kare faktör yüklerinin kovaryanslarının toplamının minimum edilmesi amaçlanır (Gorsuch, 2014: 200).

2.5.4.2.4. Biquartimin Yöntemi

Yöntemde döndürme sonucunda faktör yükü karesi en yüksek olan tek bir faktör oluşur. Geriye kalan faktörlerin değişken yükü kareleri ise sıfıra yakın değerlerden oluşur (Şencan, 2005: 399).

2.5.4.2.5. Oblimin Yöntemi

Yöntemde faktör yüklerinin çaprazlarını minimize ederek basit bir yapı oluşturmak hedeflenmiştir.

2.5.4.2.6. Binoramin Yöntemi

Yöntemde E ile gösterilen fonksiyonun minimum olması hedeflenmektedir.

$$\min E = \sum_{1 \leq q=1}^m \left[\sum_{j=1}^p (v_{j1}^2 / h_j^2) (v_{jp}^2 h_j^2) \left[\sum_{j=1}^p (v_{j1}^2 / h_j^2) (\sum_{j=1}^p (v_{j1}^2 / h_j^2)) \right] \right] \quad (2.5)$$

2.5.5. Faktör Skorlarının Elde Edilmesi

Analizin amaçlarından biri de temsili değişken oluşturmaktır. Bu amaç doğrultusunda faktör skorları elde edilmektedir. Faktör skorları, orijinal değişkenleri temsil eden ve onların yerine kullanılabilecek olan boyutu indirgenmiş değişkenlerdir. Araştırmanın amacına bağlı olarak faktör skorları hesaplanabilir. Bunun için analizin aşamaları tamamlandıktan sonra faktörleri isimlendirme ve yorumlama yapılır veya faktörler, faktör skorları hesaplanarak diğer analizler için hazır veri haline getirilir.

Her bir faktör için elde edilecek faktör skorları değişik yöntemlerle elde edilebilir. Orijinal değişken değerleriyle faktör yüklerinin çarpımı veya Z ile standartlaştırılmış değişken değerleriyle faktör yüklerinin çarpımıyla faktör skorları elde edilebilir. (Alpar, 2011: 274). Regresyon, Bartlett ve Anderson Rubin yöntemi olmak üzere üç adet faktör skoru yöntemi vardır.

2.5.5.1. Thomson (Regresyon) Yöntemi

Yöntemde hesaplanan faktör değerlerinin ortalaması sıfır, varyansı ise gerçek faktör değerleri ile tahmin edilen faktör değerleri arasındaki çoklu korelasyon katsayısının karesine eşittir. Yöntemde hesaplanan değerler birbiriyle ilişkili olabilir. Faktörler birbiriyle ilişkisizdir (Albayrak, 2006: 174).

2.5.5.2. Bartlett Yöntemi

Bu yöntemde faktör skorları yalnızca kendi faktörleriyle ilişkilidir ve faktör skorları tarafsızdır (Tabachnick ve Fidell, 2007: 650-651). Yöntemi diğer yöntemlerden ayıran özellik gerçek faktör puanlarının yansız tahminlerini üretmesidir. Bunun sebebi Bartlett puanlarının en çok olabilirlik yaklaşımı kullanılarak elde edilmesidir (DiStefano vd., 2009: 4).

Bu yöntemde bulunan faktör skorları regresyon yöntemi ile aynı ortalama ve standart sapmaya sahiptir.

2.5.5.3. Anderson Rubin Yöntemi

Bu yöntem Bartlett yöntemiyle benzerlik göstermektedir. Hesaplama şekli Bartlett yönteminden daha karışıktır. Gorsuch (1983) tarafından önerilen yaklaşımda faktör skorları 0 ortalama ve 1 standart sapmaya sahiptir. Ayrıca tahmin edilen faktörler birbirleriyle ilişkili olmasına rağmen faktör skorları birbirleriyle ilişkisizdir (Tabachnick ve Fidell, 2007: 651).

2.5.6. Faktörlerin İsimlendirilmesi ve Yorumlanması

Analiz yukarıda belirtilen istatistikî kısıtlamalar ve kurallar dikkate alınarak araştırmacı tarafından yapılmalıdır. Analiz sonucu ortaya çıkan daha az sayıdaki faktörler iki şekilde kullanılabilir. Birinci olarak, oluşan faktörler araştırmanın amacı dikkate alınarak yorumlanmalıdır. İkinci olarak ise oluşan faktörler, faktör skorları hesaplanarak hazır veri haline getirilebilir. Temsili değişkenler diğer çok değişkenli istatistiksel yöntemlerde, belirtilen özellikleri taşıması halinde kullanılabilir.

3. KÜMELEME ANALİZİ

3.1. Kümeleme Analizi Kavramı

Çok değişkenli istatistiksel yöntemlerden kümeleme analizi, birincil amacı nesneleri sahip oldukları özelliklere göre gruplandırmak olan çok değişkenli teknikler grubudur (Hair vd., 2010: 481). Nesneler gruplanırken sahip oldukları özellikler dikkate alınır.

Kümeleme analizinde çok değişkenli istatistiksel yöntemlerin varsayımlarının pek bir önemi yoktur. Sadece varsayımlardan biri olan çoklu doğrusal bağlantı bu analiz için önemlidir. Analizde değişkenler arasında çoklu doğrusal bağlantı olmaması gerekmektedir. Bunun yanı sıra analizde belirtilen uzaklık değerlerinin normal dağılıma uyması gerekmektedir.

Kümeleme analizi tanımlayıcı bir analizdir. Analiz, gözlemlerin yapısal özelliklerini belirlemektedir. Bu yüzden analizde kullanılacak olan örneklem sayısı anakütleyi tanımlayacak sayıda olmalıdır. Analizde kullanılan örneklem büyüklüğü ile ilgili bir kural yoktur.

Kümeleme analizi sonucu meydana gelen daha az sayıdaki kümenin elemanları, kendi arasında benzer yapıda oluşurken diğer grup elemanları ile benzer yapıda değildir.

3.2. Kümeleme Analizinin Amacı

Kümeleme analizinin asıl amacı, nesneleri veya birimleri belirlenen özelliklerine göre sınıflandırmaktır. Araştırmacı birbirine benzeyen birim veya nesneleri bir araya getirerek sonuçların daha kolay ve anlaşılır bir şekilde yorumlanmasını sağlamaktır. Analiz sonucu boyutu indirgenen küme sayısının yorumu büyük sayılardan oluşan küme sayısının yorumuna göre daha kolaydır. Kümeleme analizi bazı özel amaçlar içinde kullanılabilir. Bu amaçlar aşağıdaki gibidir (Tatlídil, 2002: 252-253);

- Gerçek tiplerin (cinslerin- ırkların) belirlenmesi,
- Model uydurmanın kolaylaştırılması,
- Gruplar için ön tahmin,
- Hipotezlerin testi,
- Veri yapısının netleştirilmesi,
- Veri indirgemesi (veriler yerine kümeler değerlendirilmesi),
- Aykırı değerlerin (outliers) bulunmasıdır.

3.3. Kümeleme Analizinin Adımları

Kümeleme analizi dört adımdan oluşmaktadır. Bu adımlar sırasıyla aşağıdaki gibidir;

- Verilerin standartlaştırılması ve değişken seçimi,
- Benzerlik veya uzaklık ölçülerinin belirlenerek benzerlik-uzaklık matrisinin oluşturulması,
- Araştırmaya uygun kümeleme yönteminin seçimi ve küme sayısının belirlenmesi,
- Analiz sonucu oluşan kümelerin yorumlanmasıdır.

3.3.1. Verilerin Standartlaştırılması ve Değişken Seçimi

Değişken seçimi, çoğu analizde olduğu gibi kümeleme analizinde de önemlidir. Kümeleme analizinde nesnelerin özelliklerini içeren değişkenlerin tanımının tutarlı bir şekilde yapılması gerekmektedir. Nitekim önemli bir değişkenin analiz dışında bırakılması kümeleme sürecinin yanlış olmasına sebep olabilir.

Analizde aynı ölçekte olmayan değişkenlerin standartlaştırılması gerekmektedir. En yaygın kullanılan standartlaştırma türü Z puanlarıdır (Hair vd., 2010: 497). Z puanları dönüşümü aşağıdaki gibi hesaplanmaktadır:

$$Z = \frac{x - \bar{x}}{s} \quad (3.1)$$

Denklem 3,1'de; X= ham veri, $\bar{x}$ = ortalama ve S= standart sapmadır.

3.3.2. Benzerlik ve Uzaklık Ölçüleri

Araştırmacı tarafından veri girişi ve değişken seçimi yapıldıktan sonra benzerlik veya uzaklık ölçülerinin belirlenmesi aşamasına geçilir. Çalışmada kullanılacak yöntem kümeleme analizi ise birey veya nesnelerin gruplandırılması bu iki ölçü dikkate alınarak çalışmanın yapılması gerekmektedir.

3.3.2.1. Benzerlik Ölçüleri

Benzerlik büyük bir sayı ise iki gözlem birbirine yakın olup, benzerlik küçük bir sayı ise iki gözlem birbirine uzaktır. Bir başka ifadeyle benzerlik ve uzaklık zıttır.

3.3.2.2. Uzaklık Ölçüleri

Uzaklık ölçüleri, gözlemlerin birbirine olan yakınlığını ölçmektedir. En sık kullanılan uzaklık ölçüleri; Mahalonobis, Minkowski, Öklid ve Manhattan uzaklıklarıdır.

3.3.2.2.1. Mahalonobis Uzaklığı

Bu uzaklık ölçüsü ile uç değerler hesaplanmaktadır. Bu özellik mahalonobis uzaklık ölçüsünü, diğer uzaklık ölçülerinden ayıran en temel avantajıdır. Mahalonobis uzaklığı aşağıdaki gibi hesaplanmaktadır (Tatlıdil, 2002: 254);

$$d(x_i, x_j) = \sqrt{(x_i - x_j)^T S^{-1} (x_i - x_j)} \quad (3.2)$$

Denklem 3.2’de yer alan S, örneklem kovaryans matrisidir.

3.3.2.2.2. Minkowski Uzaklığı

Uzaklık ölçüsü aşağıdaki gibi tanımlanmıştır:

$$d_{ij} = (\sum_{k=1}^p |x_{ik} - x_{jk}|^m)^{1/m} \quad (3.3)$$

Denklem 3.3’de; m=1 için Manhattan uzaklık ölçüsünü, m=2 için Öklid uzaklık ölçüsünü vermektedir (Alpar, 2011: 165).

3.3.2.2.3.Öklid Uzaklığı

Bu uzaklık yönteminin kullanımı gözlem sayısının 100’den fazla olduğu durumlarda uygundur. En sık kullanılan uzaklık yöntemidir. Öklid uzaklığında iki nokta arasındaki uzaklık, Pisagor teoremi yardımı ile bulunabilir. Bu teoreme göre bir dik üçgenin dik kenarları uzunlukları bilindiğinde hipotenüs (d_{ij}) elde edilebilir. İki nokta arasındaki (x_{i1}, x_{i2}) ve (x_{j1}, x_{j2}) uzaklık, Öklid uzaklığı aşağıdaki gibi hesaplanır (Alpar, 2011: 165-166);

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (3.4)$$

Denklem 3.4’te;

X_{ik} =i. gözleme ait k. değişken değeri,

X_{jk} =j. gözleme ait k. değişken değeri, p =değişken sayısıdır.

Kareli Öklid uzaklığı, Öklid uzaklığının karesinin alınması ile elde edilmektedir.

3.3.2.2.4. Manhattan (City Blok) Uzaklık Ölçüsü

Bu uzaklık ölçüsü, iki gözlemin arasındaki farkın mutlak değerleri toplamının hesaplanması ile bulunur (Çokluk vd., 2012: 150). Kesikli sayısal verilerin kullanımı için uygun olan uzaklık ölçüsünün formülü aşağıdaki gibidir;

$$d_{ij} = \sum_{k=1}^p |x_{ik} - x_{jk}| \quad (3.5)$$

3.3.3. Kümeleme Yöntemleri

Benzerlik ve uzaklık ölçüleri belirlendikten sonra kümeleme yöntemlerinin belirlenmesi ve küme sayısına karar verme aşamasına geçilir. Kümeleme yöntemleri

temelde hiyerarşik ve hiyerarşik olmayan kümeleme analizi yöntemi olmak üzere iki yöntem olarak ele alınabilir. Bu iki yöntem dışında sadece SPSS programında bulunan iki aşamalı kümeleme analizi (Two Step Cluster Analysis) yöntemi bulunmaktadır. Bu yöntemler arasından nesneleri ya da birimleri gruplayacak olan uygun yöntem seçilmelidir.

3.3.3.1. Hiyerarşik Kümeleme Teknikleri

Hiyerarşik kümeleme yöntemi, başlangıçta kaç tane küme sayısının oluşacağını bilinmediği kümeleme yöntemidir. Hiyerarşik tekniklerin özellikleri aşağıdaki gibi sıralanabilir (Akat, 2007'den aktaran Çokluk vd., 2012: 147);

- Hiyerarşik kümeleme teknikleri, kümeleri artarda birleştirme sürecidir.
- Hiyerarşik kümeleme yöntemleri, veri matrisindeki birimlerin/değişkenlerin başlangıçta kaç küme oluşturduğuna ve küme elemanlarını başlangıçta hangi ölçütün seçildiğine göre iki temel gruba ayrılırlar. Bu gruplar birleştirici (toplamalı) hiyerarşik kümeleme yöntemleri ve ayırıcı (bölünmeli) hiyerarşik kümeleme yöntemleridir.
- Birleştirici tekniklerde n adet nesne ardıl olarak birleştirilir.
- Ayırıcı tekniklerde n adet nesne ardıl olarak daha küçük kümelere ayrılır.
- Birleştirici tekniklerin kullanımı ayırıcı tekniklere göre daha fazladır.
- Hiyerarşik kümeleme analizi uçdeğerlere oldukça duyarlıdır. Bu yüzden farklı hiyerarşik kümeleme teknikleri uygulanıp sonuçları karşılaştırılmalıdır. Sonuçlar arasında tutarlılık söz konusu ise kümeleme işlemi uygulanmalıdır.
- Büyük örneklerde hiyerarşik kümeleme tekniklerinin işlemlerinin uzun sürmesinden dolayı hiyerarşik kümeleme yöntemi küçük örnekler için uygundur.
- Hiyerarşik yöntemle yapılan gruplandırmalar geri alınamaz. Bir başka deyişle bir grup, diğeri ile bir kez birleştirildikten (ayrıldıktan) sonra, daha sonraki adımlarda ayrılamaz (birleştirilemez).
- Hiyerarşik tekniklerin ağaç diyagramları ile gösterilen sonuçlarına dendrogram denir.

En çok kullanılan hiyerarşik kümeleme teknikleri birleştirici tekniklerdir. Bu yöntemlerden en sık kullanılan teknikler ise; bağlantı teknikleri (tek bağlantı, tam

bağlantı ve ortalama bağlantı teknikleri), varyans teknikleri (Ward's bağlantı tekniği) ve merkez yöntemidir.

3.3.3.1.1. Bağlantı Teknikleri

3.3.3.1.1.1. Tek Bağlantı Tekniği (En Yakın Komşuluk)

Bu teknik Florek vd. (1951) tarafından keşfedilip, Sneath (1957) ve Johnson (1967) tarafından geliştirilmiştir. Teknikte ilk olarak, uzaklıklar matrisindeki en küçük uzaklığa (en yakın) sahip iki gözlem dikkate alınır. Bu birinci küme olarak belirlenir. Daha sonra uzaklık matrisinden bir sonraki en küçük uzaklık bulunur. Bu aşamada, yeni gözlem ya ilk iki gözlemin oluşturduğu kümeye katılır ya da iki elemanlı başka bir küme oluşturur. Bu süreç, tüm gözlemler tek bir kümede toplanana kadar tekrarlayarak devam eder (Alpar, 2011: 316).

3.3.3.1.1.2. Tam Bağlantı Tekniği (En Uzak Komşuluk)

Bu tekniğin küme yapısını bulma algoritması, uzaklık matrisindeki gözlemler arasındaki en büyük uzaklığın (en az benzerlik) dikkate alınması hariç tek bağlantı tekniği ile aynıdır (Anderberg, 1973: 41-42). Tekniğin en uzak komşuluk olarak da adlandırılmasının sebebi budur.

3.3.3.1.1.3. Ortalama Bağlantı Tekniği

Sokal ve Michener (1958) tarafından önerilen bu teknikte, herhangi iki kümenin benzerliği, bir kümedeki tüm bireylerin diğer bir kümedeki tüm bireylerle ortalama benzerliğidir. Uç değerlerden en az etkilenen tekniklerden biri olan bu teknikte, küçük küme içi değişkenliğe sahip kümeler üretilmektedir (Hair vd., 2010: 504).

3.3.3.1.2. Varyans Teknikleri

3.3.3.1.2.1. Ward's Bağlantı Tekniği

Bu teknikte gruplandırmanın her aşamasında küme içi varyansı en aza indirmek amaçlanmaktadır. Yöntemde küme oluşturulurken her aşamada hata kareler toplamı minimum olan kümeler birleştirilir. Yöntemde küme içi homojenlik optimize edilir. Uzaklık ölçüsü olarak kare öklid uzaklığı kullanılır (Borgen ve Barnett, 1987: 464-465).

3.3.3.1.3. Merkez Yöntemi

Yöntemde iki küme merkezi belirlenerek, iki gözlem arasındaki benzerlik bulunur. Bu yöntem ile birleştirilmiş bir kümenin merkezi, iki ayrı kümenin merkezlerinin ağırlıklı bir kombinasyonudur. Bu yöntemin dezavantajı kümelerin birleştirildiği uzaklığın bir adımdan diğerine azalmasıdır. Bu kümelemede istenmeyen

bir özelliktir. Sonraki aşamalarda birleştirilen kümeler, erken aşamada birleştirilen kümelerden daha farklı olduğundan bu istenmeyen bir durum olmaktadır (Norusis, 2011: 388).

3.3.3.2. Hiyerarşik Olmayan Kümeleme Yöntemleri

Hiyerarşik olmayan kümeleme yönteminde, analize başlamadan önce küme sayısına karar verilir. En sık kullanılan hiyerarşik olmayan kümeleme yöntemleri; k-ortalama ve en çok olabilirlik yöntemleridir.

3.3.3.2.1. k-Ortalama Yöntemi

Bu yöntem, verileri kullanıcı tarafından belirlenen sayıda kümeye bölerek, ardından bazı sayısal kriterler karşılanana kadar yinelemeli olarak gözlemleri kümelere yeniden atayarak çalışır. Bu yöntemde amaç, bir kümedeki gözlemlerin birbirinden uzaklığını en aza indirmek ve kümeler arasındaki mesafeyi en üst düzeye çıkarmaktır. Yöntem, hiyerarşik olmayan kümeleme tekniklerinden en sık kullanılanıdır (Hair vd., 2010: 507).

3.3.3.2.2. En Çok Olabilirlik Yöntemi

Bu yöntemde, gözlemlerden her biri en büyük olabilirlik değeri verecek biçimde daha önceden belirlenen kümelere atanmaktadır.

3.3.3.3. İki Aşamalı Kümeleme Analizi (Two-Step Cluster Analysis)

Çalışmada kullanılacak olan veri tipi sürekli veya kategorik veri ise hiyerarşik ve hiyerarşik olmayan kümeleme yöntemleri kullanılamaz. Bu tip çalışmalarda iki aşamalı kümeleme yöntemi kullanılmaktadır. Bu yöntem çalışmada kullanılacak olan örneklem sayısının çok büyük olduğu durumlar için uygundur. Ayrıca iki aşamalı kümeleme analizi, tüm değişkenlerin bağımsız olduğu, sürekli değişkenlerin normal dağılıma sahip olduğu ve kategorik değişkenlerin ise çok terimli bir dağılıma sahip olduğu modellerde en iyi sonucu verir (Norusis, 2011: 394).

3.3.4. Kümelerin Geçerliliği ve Yorumu

Geçerlilik örneklemin anakütleyi ne kadar iyi temsil ettiğini göstermektedir. Kümeleme analizinde geçerlilik birkaç yöntemle yapılmaktadır. Bu yöntemlerin bazıları, çözümlerinin uzun olması nedeniyle tercih edilmemektedir. En iyi geçerlilik yöntemi ise örneği iki gruba ayırmaktır. Ayrılmış olan grupların her birine kümeleme analizi yöntemlerinden biri uygulanır ve sonuçlar karşılaştırılır (Kalaycı, 2010: 362).

4. UYGULAMA

Çalışmada, faktör analizi yöntemlerinden açıklayıcı faktör analizi yöntemi kullanılmıştır. Verinin faktör analizine uygunluğunu test etmek için KMO ve Bartlett test istatistiği uygulanmıştır. Faktör bulma yöntemi olarak, temel bileşenler analizi kullanılmıştır. Temel bileşenler analiziyle veri setindeki maksimum varyans ortaya çıkartılarak faktör türetme amaçlanmıştır. Faktör sayısına karar vermek için sırasıyla özdeğer istatistiği, toplam varyans oranı ve scree plot grafiği incelenmiştir. Döndürme yöntemi olarak, dik döndürme yöntemlerinden varimax yöntemi seçilmiştir.

Kümeleme yöntemi olarak hiyerarşik kümeleme yöntemi seçilmiştir. Veriler, tek bağlantı tekniği uygulanarak analiz edilmiştir. Uzaklık ölçüsü olarak kareli Öklid uzaklığı kullanılmıştır.

4.1. Literatür Taraması

Atilla ve Arıkan (2001) yaptıkları çalışmada; Antalya ili platosunda yer alan yeraltı sularını, benzerlik ve farklılıklarına göre sınıflandırmayı amaçlamışlardır. Bu kapsamda yeraltı sularına ait kimyasal ve çevresel izotop verilerine faktör analizi ve kümeleme analizi uygulanmıştır. Kümeleme ve faktör analizi yöntemlerinin sonuçlarının benzer olduğu ve aynı tür gruplamanın her iki yöntem ile elde edildiğine karar verilmiştir.

Pamuk (2005), Türkiye'nin 81 iline ait GSYİH verileri ile belirlenen 10 sektörü faktör analizi ve kümeleme analizi yöntemlerini kullanarak gruplandırmayı amaçlamıştır. Ayrıca çalışmada 81 il gruplandırılmıştır. Faktör analizi ve kümeleme analizi sonuçları dikkate alındığında değişkenlerin, gruplara dağılımında farklılıklar olduğu saptanmıştır.

Kümeleme analizi ve faktör analizi yöntemlerinin kullanıldığı bir diğer çalışma ise Günden ve Miran (2008) tarafından yapılmıştır. Çalışmada, İzmir ili Torbalı ilçesi bölgesinden seçilen çiftçilerin çevreye karşı tutumlarını ve çevre duyarlılığını ölçmek için 13 soruluk bir anket düzenlenmiştir. Analiz sonucunda “ekolojik denge”, “çevre ve canlılar”, “doğal kaynaklar” ve “doğa” olarak adlandırılan dört tane faktör belirlenmiştir. Faktör analizi sonucunda oluşan faktörler kullanılarak ise kümeleme analizi yapılmıştır. Bu analiz sonucunda ise bölgedeki çiftçiler “duyarlılar”, “ılımlılar” ve “düşük ılımlılar” olarak adlandırılan üç tane kümeye göre gruplanmıştır.

Doğan ve Başokçu (2010) tarafından yapılan “İstatistik Tutum Ölçeği İçin Uygulanan Faktör Analizi ve Aşamalı Kümeleme Analizi Sonuçlarının Karşılaştırılması” isimli makalede faktör analizi ve kümeleme analizi yöntemlerinin benzer sonuç verip vermediği, yapıya ilişkin kuramsal tutarlığı sağlayıp sağlamadıkları ve açıkladıkları yapı için elde edilen doğrulayıcı faktör analizi sonuçlarının benzer olup olmadığı amaçlanmıştır. Araştırmaya, Hacettepe Üniversitesi’nden seçilen 297 kişi dâhil edilmiştir. Bu çalışmada, faktör analizi ve hiyerarşik kümeleme analizi yöntemlerinin sonuçlarının benzerlik gösterdiği belirlenmiştir. İki analizin sonuçları arasındaki farklılık ise boyutlardaki maddeler ve madde sayısının farklı olmasıdır.

Yılmaz (2011) tarafından yapılan “Türkiye’de İllerin Sosyoekonomik Gelişmişlik Düzeylerinin Faktör Analizi ve Kümeleme Analizi ile İncelenmesi” isimli çalışmada 2008 yılına ait verilerle, 81 ilin sosyo-ekonomik gelişmişlik seviyelerini tespit etmek ve bu tespit doğrultusunda gelişmişlik sıralamalarını yapmak amaçlanmıştır. Bu doğrultuda sosyal ve ekonomik birimlerden seçilen 31 değişken kullanılmıştır. Belirlenen verilere açıklayıcı faktör analizi yöntemi ve hiyerarşik olmayan kümeleme analizi yöntemi uygulanmıştır. Çalışmada faktör analizi sonucunda, demografik ve sağlık göstergeleri, tarım göstergeleri, istihdam göstergeleri, ekonomi göstergeleri ve eğitim göstergeleri şeklinde isimlendirilen beş faktör elde edilmiştir. İllerin, elde edilen bu faktörlerdeki faktör yüklerine göre sosyal ve ekonomik gelişmişlik açısından yerleri belirlenmiştir. Kümeleme analizi sonucunda ise, iller üç gruba ayrılmıştır. Birinci grup en gelişmiş illeri, ikinci grup gelişmiş illeri ve üçüncü grupta gelişmemiş illeri göstermektedir.

Temel ve Uzundumlu (2015) yaptıkları çalışmada, Rize ilinde ikamet eden 115 tane hane halkı ile yüz yüze anket yaparak, bu ildeki balık tüketimini belirleyen tutum ve davranışları belirlemek ve balık tüketimini etkileyen faktörleri belirlemeyi amaçlamışlardır. Bu kapsamda 2013 yılına ait veriler kullanılmıştır. Çalışmanın sonucunda likert ölçeğine göre belirlenen 30 soruya faktör analizi uygulanarak 9 faktöre indirgenmiştir. Faktör skorları belirlenen çalışmada, hiyerarşik olmayan kümeleme yöntemi uygulanarak tüketiciler 4 kümeye indirgenmiştir. Ayrıca her bir tüketici grubun ekonomik, sosyal ve demografik özellikleri dikkate alınarak tüketicilerin daha fazla önem verdiği değişkenlere yönelik çalışmalar yapılarak tüketici memnuniyeti artırılabilceği sonucuna ulaşılmıştır.

Özgür Güler ve Veysikarani (2018) faktör analizi ve kümeleme analizi yöntemlerini kullanarak OECD ülkelerini, inovasyon göstergelerine göre karşılaştırmıştır. Sonuç olarak, benzer ve farklı ülkeler saptanıp gruplanmıştır. Çalışmada 12 değişken kullanılmıştır. Faktör analizi sonucunda değişkenler dört faktör çatısı altında toplanmıştır. Belirlenen faktör skorlarına ise kümeleme analizi uygulanmıştır. Analiz sonucunda beş adet küme belirlenmiştir.

Han vd. (2022) yaptıkları çalışmada, pandemi sürecinin Türkiye ekonomisine etkisini incelemişlerdir. Bu bağlamda araştırmada ekonomik birimlerden seçilen 15 adet değişken kullanılmıştır. Verilere, bağımsız iki örnek t testi ve Mann-Whitney U testi uygulanmıştır. Veriler 2019 ve 2021 yılına ait olarak seçilmiştir. Sonuç olarak pandemi döneminin ekonomik göstergeler üzerinde etkili olduğu belirlenmiştir.

Bu tez çalışmasında literatürden farklı olarak; pandemi döneminin Türkiye ekonomisine etkisi, verilere faktör analizi ve kümeleme analizi uygulanarak sonuçları karşılaştırılmıştır.

4.2. Pandemi Sürecinin Ekonomik Etkileri

Çin'in Wuhan kentinde 2019 yılının son günlerinde ortaya çıkan salgın, hızlı bir şekilde diğer ülkelere yayılmıştır. Türkiye'de ise ilk vaka 2020 yılının mart ayında görülmüştür. Vaka sayılarının artması ile birlikte her ülke kendi izlediği politikalar neticesinde bir dizi önlemler almaya başlamıştır. Türkiye'de ilerleyen dönemlerde salgının seyrinin kötüye gitmesinden dolayı önlemler almaya başlamıştır. Alınan önlemlerin etkisi ise kısa bir süre içerisinde eğitim, ulaşım, turizm, ekonomi gibi birçok alanda hissedilmiştir. Salgın dolayısı ile tedavi gören veya gözlem altında tutulan insan sayısının çok olması, birçok ülkenin sınırlarını kapatması, uzaktan çalışma modeline geçilmesi gibi durumlar ülke ekonomisini doğrudan etkilemiştir. Bu durumlar neticesinde, salgın öncesi duruma göre üretim kapasitesinde azalma meydana gelmiş, fiyat istikrarsızlığı oluşmuş ve tedarik zincirinde bozulmalar meydana gelmiştir. Bu süreçte temel ekonomik göstergelerden olan enflasyonun artması ise birçok ülkede olduğu gibi, Türkiye açısından da büyük bir problem olmuştur. Türkiye'de enflasyon salgın öncesi döneme göre artma eğilimi göstermiştir. Birçok sektörde yaşanan uzun süreli kapanmalar sonucunda ise işten çıkarma oranlarında artış meydana gelmiştir.

Dış ticaret rakamları dikkate alındığında da salgın öncesi duruma göre düşüş yaşanmıştır. Ülkelerin sınırlarını kapatması, üretim kapasitesinin azalması bu durumun sebeplerindendir.

Alınan tedbirler büyüme rakamlarına da yansımıştır. Özellikle işletmelerin uzun süreli kapalı olması büyüme oranlarında düşüşe sebep olmuştur. Bu durum ilk etapta

GSYİH'nın düşmesine, daha sonra ise üretimin ve dolaylı olarak ihracatın azalmasına sebep olmuştur (Maliszewska, 2020: 6).

Döviz rezervlerindeki düşüş ve Türk lirasının yabancı para birimleri karşısında değer kaybetmesi ise pandemi sürecinde hızını artırmıştır. Rezervlerdeki düşüş ve Türk lirasının değer kaybının bir süre daha devam edeceği uzmanlar tarafından tahmin edilmektedir.

Toplam kredi miktarında salgın döneminde artış meydana gelmiştir. Firmalar ve kişiler istek ve ihtiyaçlarını karşılayabilmek için bankalardan kredi çekmişlerdir. Harcamalar ve tüketici kredilerinde ise artış meydana gelmiştir.

Kişilerin alınan tedbirler dolayısı ile seyahatlerini ertelemeleri veya iptal etmesinden dolayı seyahatten elde edilen gelirler düşmüştür. Hükümetin uyguladığı teşvik ve yatırımlar neticesinde ise konut ve taşıt satışlarında bir dönem artış sağlanmıştır.

4.3. Veri Seti

Çalışmada, 2019:01 ile 2021:01 dönemlerine ait verilerle pandemi sürecinin ekonomik etkilerinin faktör analizi ve kümeleme analizi ile incelenmesi amacıyla makro ekonomik değişkenler ele alınmıştır. Makro ekonomik değişken bağlamında; toplam ödemeler, dolar kuru, euro kuru, ekonomik büyüme, toplam krediler, toplam taşıt sayısı, konut satış sayısı, tüketici kredileri, döviz rezervleri, seyahat gelirleri, işsizlik, enflasyon, ithalat, harcama tutarı ve ihracat değişkenleri ele alınmıştır. İncelenen verilere ait özet bilgiler aşağıdaki gibidir;

Tablo 4.1. Değişkenlerin Tanımlamaları ve Kaynakları

Değişkenler	Açıklaması	Kaynak
Dolar	Dolar	
Euro	Euro	
lnrezerv	Döviz Rezervleri (Milyon Dolar)	
lntopödeme	Toplam Ödemeler (Milyon Dolar)	
lntopkredi	Toplam Krediler (Bin TL)	
lntüketicikredi	Tüketici Kredileri (Bin TL)	
lnharcama	Banka Kartı ve Kredi Kartı Harcama Tutarı (Bin TL)	EVDS, https://evds2.tcmb.gov.tr
lntasit	Taşıt Sayısı	
lnkonut	Konut Satış	
lnseyahat	Seyahat Gelirleri (Milyon Dolar)	
İssizlik	İşsizlik %	
Enflasyon	Enflasyon (Tüfe)	
lnbüyüme	Sanayi Üretim Endeksi	
lnithalat	İthalat (milyon dolar)	
lnihracat	İhracat (Milyon Dolar)	

Çalışmadaki bazı değişkenlerin logaritması alınmıştır. Bunun sebebi varyansı sabitleştirmek ve uçdeğerlerin etkilerini azaltmaktır (Türe ve Akdi, 2005: 6). Çalışmada yer alan değişkenler standartlaştırılmıştır.

Çok değişkenli parametrik test varsayımlarının sağlandığı bilinmektedir.

4.4. Bulgular

Bu başlık altında faktör analizi ve kümeleme analizi sonuçlarına yer verilmiştir.

4.4.1. Faktör Analizi Sonuçları

Verilerin faktör analizine uygunluğunu test etmek için yapılan KMO ve Bartlett test istatistik sonuçları aşağıdaki gibidir;

Tablo 4.2. KMO ve Bartlett Testi İstatistikleri

KMO Uygunluk Ölçütü		0,695
	Yaklaşık Ki-kare	686,519
Bartlett Küresellik Testi	df	105
	Anlamlılık	0,00

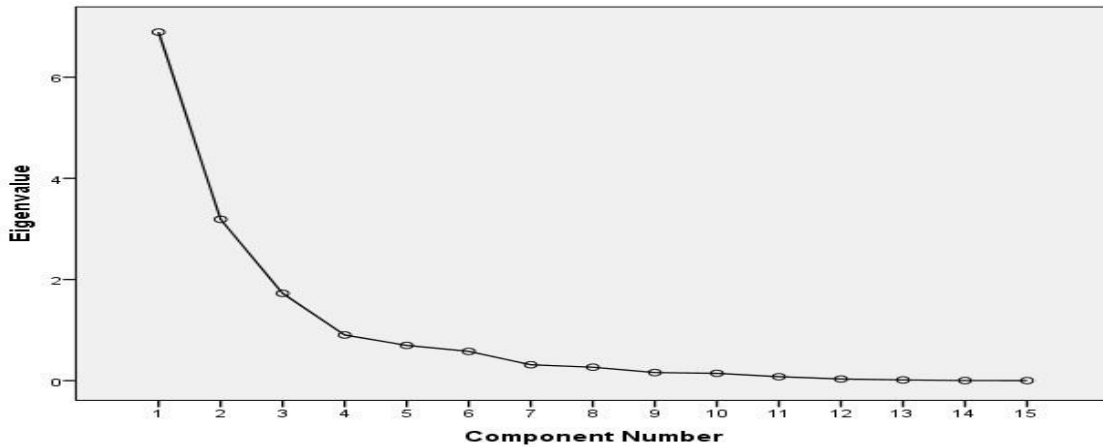
KMO ve Bartlett testi sonuçları ile çalışmada kullanılan veri setinin faktör analizi için uygun olduğu belirlenmiştir.

Açıklanan toplam varyans tablosu Tablo 4.3'te ifade edildiği gibidir;

Tablo 4.3. Açıklanan Toplam Varyans Tablosu

Faktör	İlk Özdeğerler			Türetilen Kareli Ağırlıklar Toplamı			Çevrilmiş Kareli Ağırlıklar Toplamı		
	Toplam	Varyans (%)	Kümülatif (%)	Toplam	Varyans (%)	Kümülatif (%)	Toplam	Varyans (%)	Kümülatif (%)
1	6,892	45,946	45,946	6,892	45,946	45,946	6,711	44,741	44,741
2	3,189	21,263	67,21	3,189	21,263	67,21	2,92	19,466	64,207
3	1,726	11,505	78,714	1,726	11,505	78,714	2,176	14,508	78,714
4	0,903	6,022	84,736						
5	0,697	4,648	89,384						
6	0,579	3,857	93,241						
7	0,314	2,096	95,337						
8	0,266	1,773	97,11						
9	0,16	1,068	98,178						
10	0,145	0,969	99,147						
11	0,078	0,521	99,669						
12	0,032	0,216	99,885						
13	0,014	0,096	99,98						
14	0,002	0,015	99,996						
15	0,001	0,004	100						

Tablo 4.3'te özdeğer istatistiği 1'den büyük üç adet faktör bulunmaktadır. Elde edilen faktörlerden varyansın toplamını açıklama oranı olarak en yüksek etkiye birinci faktör sahiptir. Birinci faktörün varyansı açıklama oranı % 44,74'tür. Birinci ve ikinci faktörün birlikte varyansı açıklama oranı ise % 64'tür. Üç adet faktörün toplam varyansı açıklama oranı ise % 78,71'dir. Açıklanan varyans oranının %50'yi geçmesi yeterlidir. Dolayısıyla özdeğer istatistiği ve açıklanan varyans oranları dikkate alındığında çalışmada üç adet faktör belirlenmiştir. Faktör türetme metodu olarak temel bileşenler analizi kullanılmıştır. Faktör sayısını belirlemede kullanılan bir diğer yöntem ise Scree test grafiğidir. Aşağıda scree test grafiği bulunmaktadır;

Grafik 4.1. Scree Test Grafiği

Grafik 4.1'e göre faktör sayısının 3 olduğu söylenebilir.

Elde edilen faktörlere döndürme işlemi uygulanmıştır. Döndürme işlemi yapılan faktörlere ait tablo aşağıdaki gibidir;

Tablo 4.4. Döndürülmüş Faktör Matrisi

	Faktörler		
	1	2	3
Lntopkredi	0,982		
Dolar	0,980		
Euro	0,980		
Lntüketicikredi	0,958		
Lntasit	0,783		
Lnharcoma	0,745		0,338
Lnrezerv	-0,732	-0,334	
İssizlik	-0,687		
Lntopödeme	0,613		
Lnithalat		0,916	
Lnihracat		0,820	
Enflasyon		0,755	
Lnseyahat		0,656	
Lnkonut			0,895
Lnbüyüme		0,328	0,740

Tablo 4.4 dikkate alındığında her bir değişkenin bulunduğu faktörler mevcuttur.

Varimax yöntemine göre her bir faktörde bulunan değişkenlerin değeri 1'e yaklaştırılmıştır. Elde edilen birinci faktördeki değişkenler; toplam krediler, dolar ve euro kuru, tüketici kredileri, toplam taşıt sayısı, toplam harcama tutarı, döviz rezervleri, işsizlik ve toplam ödemelerdir. Belirlenen birinci faktör “döviz, rezerv ve krediler” olarak isimlendirilmiştir. İkinci faktördeki değişkenler; ithalat ile ihracat değişkenleri, enflasyon ve seyahat gelirleridir. İkinci faktör “dış ticaret ve turizm” olarak isimlendirilmiştir. Üçüncü faktördeki değişkenler ise; toplam konut sayısı ve ekonomik büyümedir. Üçüncü faktör “taşınmazlar ve ekonomik büyüme” olarak isimlendirilmiştir.

4.4.2. Kümeleme Analizi Sonuçları

Kümeleme yöntemi olarak hiyerarşik kümeleme yöntemi seçilmiş olup sonuçları aşağıdaki gibidir;

Tablo 4.5. Durum Özeti

Durumlar					
Geçerli veri		Kayıp veri		Toplam	
N	%	N	%	N	%
32	100	0	0	32	100

Tablo 4.5 dikkate alındığında, eksik veri olmadığı belirtilmiştir. Çalışmada verilerin tamamı kullanılmıştır.

Çalışmada kümeleme yöntemi olarak hiyerarşik kümeleme yöntemi seçilmiştir. Aşağıdaki tabloda bu yöntemin sonuçları verilmiştir;

Tablo 4.6. Hiyerarşik Kümeleme Analizi Sonuçları

Aşama	Küme Birleştirme		Katsayı	İlk Görünen Küme Aşaması		Sonraki Aşama
	Küme 1	Küme 2		Küme 1	Küme 2	
1	1	2	0,529	0	0	3
2	5	6	0,635	0	0	3
3	1	5	2,444	1	2	4
4	1	8	13,280	3	0	6
5	14	15	13,743	0	0	8
6	1	7	16,547	4	0	9
7	9	13	23,619	0	0	12
8	10	14	24,254	0	5	10
9	1	4	33,605	6	0	12
10	10	12	37,166	8	0	13
11	3	11	41,226	0	0	13
12	1	9	43,478	9	7	14
13	3	10	50,914	11	10	14
14	1	3	72,944	12	13	0

Tablo 4.6'ya göre ilk aşamada birinci kümede 1., ikinci kümede ise 2. gözlem birbirine en yakın iki gözlemdir. Dolayısıyla çalışmada belirtilen dolar ve Euro kuru değişkenleri birleşmiştir. Katsayılar bölümünde ise 1. ve 2. gözlem arasındaki mesafe bulunmaktadır. İki gözlem arasındaki uzaklık 0,529 olarak hesaplanmıştır.

Çalışmada uzaklık ölçüsü olarak kareli Öklid uzaklık ölçüsü kullanılmıştır. Kümeleme yöntemi olarak tek bağlantı tekniği kullanılmıştır. Kümelerin ilk görüldüğü bölümde, bir kümenin hangi aşamada meydana geldiği belirtilmektedir. İlk küme oluşumu 3. aşamada gerçekleşmiştir. Sonraki aşama bölümünde ise, 1. ve 2. gözlemin hangi aşamada diğer gözlemlerle birleşerek küme meydana getireceği belirtilmektedir. Çalışmada ilk aşamadan sonra 3. aşamaya geçileceği belirtilmektedir. Bu aşamada 5.

gözlem olan toplam kredi değişkeni ile dolar ve euro kuru değişkeni birleşir. Kümeleme süreci 14. aşamada tüm gözlemlerin tek bir küme halini almasına kadar bu şekilde devam eder.

Aşağıda yer alan tabloda, çalışmada yer alan değişkenlerin kümelere göre dağılımı gösterilmektedir;

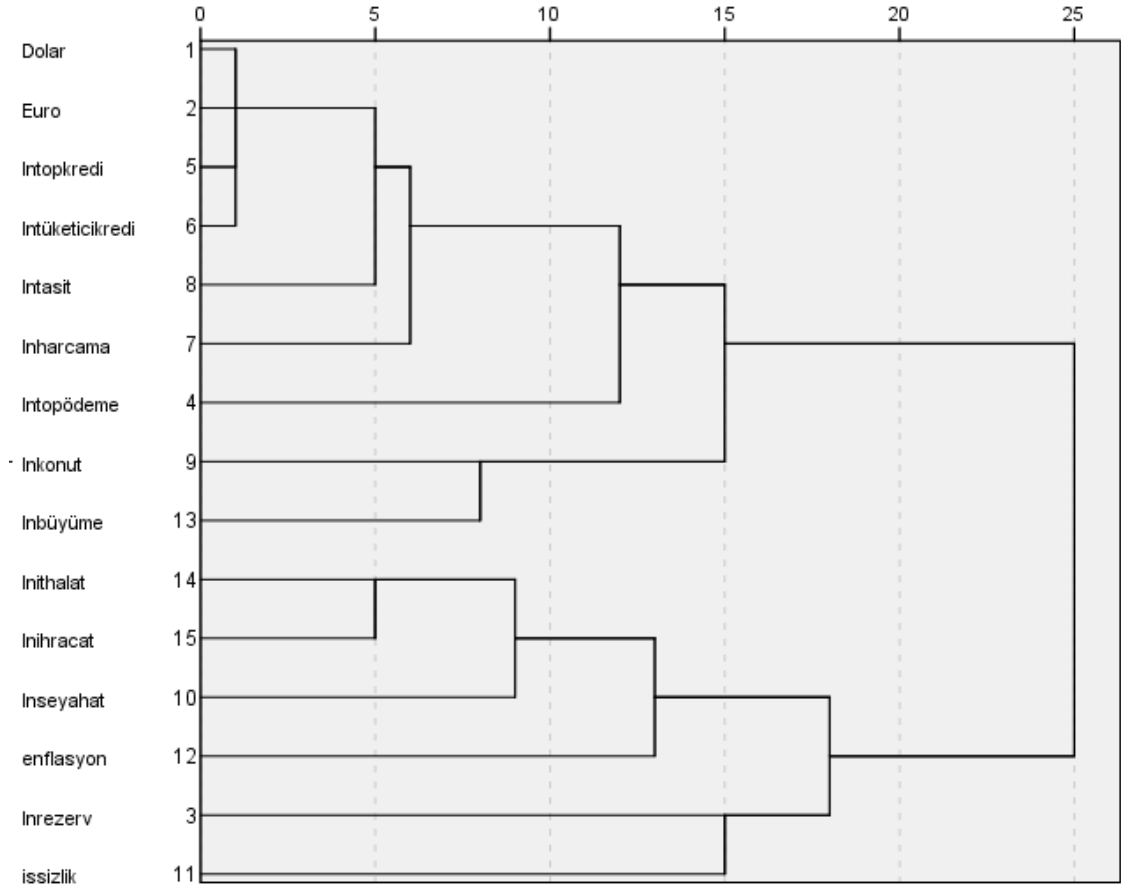
Tablo 4.7. Küme Üyeliği

Değişken	3 Küme
Dolar	1
Euro	1
Lnrezerv	2
Lntopödeme	1
Lntopkredi	1
Lntüketicikredi	1
Lnharcoma	1
Lntasit	1
Lnkonut	1
Lnseyahat	3
İssizlik	2
Enflasyon	3
Lnbüyüme	1
Lnithalat	3
Lnhracat	3

Tablo 4.7'ye göre çalışmada küme sayısı üç olarak belirlenmiştir. Birinci kümede yer alan değişkenler; dolar ve euro kuru, toplam ödeme, toplam kredi miktarı, tüketici kredisi, toplam harcama tutarı, toplam taşıt sayısı, toplam konut sayısı ve ekonomik büyümedir. İkinci kümede yer alan değişkenler; döviz rezervleri ve işsizliktir. Üçüncü kümede yer alan değişkenler; seyahat gelirleri, enflasyon, ithalat ve ihracattır.

Hiyerarşik kümeleme sonuçları dendrogram grafiği yardımı ile de açıklanabilir. Aşağıda çalışmanın dendrogram grafiği gösterilmektedir;

Grafik 4.2. Dendrogram Grafiği



Grafik 4.2'ye göre değişkenler uzaklıklarına göre sınıflandırılmıştır. Birbirine benzeyen değişkenler aynı küme içinde yer almıştır.

5. SONUÇ

Faktör analizi, birbirleriyle ilişkili değişkenleri sınıflandıran çok değişkenli istatistiksel bir yöntemdir. Faktör analizi aynı zamanda bir boyut indirgeme olup amaç, daha kolay yorumlanabilen faktörler elde etmektir.

Faktör analizinin amaçları; boyut indirgeme, faktörleri diğer çok değişkenli analizlerde kullanmak üzere hazır veri haline getirmek ve değişkenler arasındaki gizil yapıyı ortaya çıkarmaktır. Faktör analizi yöntemleri, açıklayıcı faktör analizi ve doğrulayıcı faktör analizi olmak üzere ikiye ayrılır.

Kümeleme analizi; birimleri, özelliklerine göre gruplayan çok değişkenli istatistiksel bir yöntemdir. Birimler gruplanırken uzaklık veya benzerlik ölçüleri dikkate alınmaktadır. Kümeleme analizinin amaçları; boyut indirgeme, aykırı değerlerin bulunması ve gruplar için ön tahmindir. Kümeleme analizi yöntemleri; hiyerarşik kümeleme, hiyerarşik olmayan kümeleme ve iki aşamalı kümeleme yöntemidir.

Çalışmada, pandemi döneminin Türkiye ekonomisi üzerindeki etkisi, verilere faktör analizi ve kümeleme analizi uygulanarak incelenmiştir.

Faktör analizi ve kümeleme analizi sonucunda, 15 değişken 3 faktör altında toplanmıştır. Faktör analizi ve kümeleme analizi sonuçları karşılaştırıldığında, her iki analizde de boyut indirgemesi gerçekleşmiştir. İki analizde kullanılan değişken sayısı ve küme sayısı aynıdır. Faktör analizi ve kümeleme analizi sonuçları arasındaki farklılık ise değişkenlerin kümelerle göre dağılımıdır. Faktör analizi sonucunda rezerv ve işsizlik değişkenleri birinci kümede, kümeleme analizinde ise bu değişkenler, ikinci kümede yer almıştır. Konut ve büyüme değişkenleri ise faktör analizi sonucunda üçüncü kümede yer alırken, kümeleme analizi sonucunda bu değişkenler birinci kümede yer almıştır. İncelenen diğer değişkenler iki analiz sonucunda da aynı kümelerde yer almıştır. Faktör analizi çok değişkenli istatistiksel varsayımların sağlandığı durumlarda kullanılıyorken; kümeleme analizi bu varsayımların sağlanmadığı durumlarda kullanılmaktadır.

Literatüre bakıldığında Yılmaz (2011) tarafından yapılan çalışmada, 2008 yılına ait verilere faktör analizi uygulanması sonucunda işsizlik, ithalat ve ihracat değişkenleri aynı grupta yer almıştır. Pandemi dönemine ait ekonomik değişkenlerin ele alındığı bu çalışmada ise faktör analizi sonucunda ithalat ve ihracat değişkenlerinin aynı grupta olduğu, işsizlik değişkeninin ise başka bir grupta olduğu sonucuna ulaşılmıştır.

Pandemi döneminde uygulanan tedbirler ile değişkenlerde meydana gelen değişim sonucunda bu farklılığın ortaya çıktığı tespit edilmiştir. Elde edilen bu sonuçlarla pandemi öncesi ve sonrası dönemlerin ekonomik dinamiklerinin uyuşmadığı sonucuna ulaşılmıştır. Sonuçların 2008 yılına ait verilerle karşılaştırılmış olduğuna dikkat edilmesi ve pandemi öncesi tüm dönemler için bir genelleme yapılmaması gerekmektedir. Aynı zamanda her iki çalışmada kullanılan değişken ve değişken sayısının farklı olması da bu sonuçların oluşmasında etkili olmuştur.

Han vd. (2022) yaptıkları çalışmada, pandemi döneminin aynı ekonomik göstergelerin etkinliğini incelemişlerdir. Yapılan bu çalışmada benzer sonuçlar elde edilmiştir.



KAYNAKÇA

- Albayrak, A. S. (2006), “*Uygulamalı Çok Değişkenli İstatistik Teknikleri*”, (1. baskı), Asil Yayın Dağıtım, Ankara.
- Albayrak, A. S. (2004), “Türkiye’de İllerin Sosyoekonomik Gelişmişlik Düzeylerini Belirleyen Hipotetik Yapıların Faktör Analiziyle İncelenmesi”, *Yönetim Dergisi*, Haziran, (48), 53-79.
- Alpar, R. (2011), “*Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*”, (3. baskı), Detay Yayıncılık, Ankara.
- Anderberg, M.R. (1973), “*Cluster Analysis Applications*”, (1.baskı), Academic Press, New York.
- Atila, A. Ö. ve Alparsan, A. (2001), “Antalya Liaverten Platosu Yeraltı Sularının Kümeleme ve Faktör Analizi ile Sınıflandırılması”, *Jeoloji Mühendisliği Dergisi*, 1(23), 41-54.
- Bektaş, H. (2017), “*Açıklayıcı Faktör Analizi*”, (1. baskı), Beta Basım, İstanbul.
- Borgen, F.H. ve Borgen, D. C. (1987), “Applying Cluster Analysis in Counseling Psychology Research”, *Journal of Counseling Psychology*, 34(4), 456-458.
- Çokluk, Ö, Şekercioğlu G ve Büyüköztürk, Ş. (2012), “*Sosyal Bilimler için Çok Değişkenli İstatistik SPSS ve LISREL Uygulamaları*”, (2. baskı), Pegem Akademi, Ankara.
- DiStefano, C., Zhu, M. ve Mindrila, D. (2009), “Understanding and Using Factor Scores: Considerations for the Applied Researcher”, *Practical Assessment, Research and Evaluation*, 20(14).
- Erkuş, A. (2003), “*Psikometri Üzerine Yazılar*”, (1. baskı), Türk Psikologlar Derneği Yayınları, Ankara.
- Eşmekaya, E. (2019), “Faktör Analizi (Factor Analysis)”, *YBS Ansiklopedi Dergisi*, 7 (1).
- Fabrigar, L. R., Weegener, D. T., MacCallum, R. C. ve Strahan, E. J. (1999), “Evaluating the Use of Exploratory Factor Analysis in Psychological Research”, *Psychological Methods*, 4(3), 272-299.
- Hair, J. F., Black, W. C., Babin, B. J. ve Anderson, R. E. (2010), *Multivariate Data Analysis*, (7. basım), Pearson, New York.

- Han, A., Pehlivan, C., ve Konat, G. (2022), “Covid-19 Pandemi Sürecinin Türkiye Ekonomisine Etkilerinin Ampirik Analizi”, *Sağlık ve Sosyal Refah Araştırmaları Dergisi*, 1(4), 32-49.
- Harman, H. H. (1976), “*Modern Factor Analysis*”, (3. baskı), The University of Chicago Press, Chicago.
- Gorsuch, R. L. (2014), “*Factor Analysis*”, (2. baskı), Psychology Press., New York.
- Günden, C., ve Miran, B. (2008), “Yeni Çevresel Paradigma Ölçeğiyle Çiftçilerin Çevre Tutumunun Belirlenmesi: İzmir İli Torbalı İlçesi Örneği”, *Ekoloji Dergisi*, 18, (69), 41-50.
- Kalaycı, Ş. 2010, “*SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri*”, (5. baskı), Asil Yayın Dağıtım, Ankara.
- Maliszewska, M., Aaditya, M. ve Mensbruggha, V. D. (2020), “*The Potential Impact of Covid-19 on GDP and Trade: A Preliminary Assessment*”, World Bank Policy Research Working Paper.
- Norusis, M. J. (2011), “*Cluster Analysis in: IBM SPSS Statistics 19 Advanced Statistical Procedures Companion*”, Prentice Hall.
- Özgür Güler, E., ve Veysikarani, D. (2018), “OECD Ülkelerinin İnovasyon Göstergeleri Açısından Çok Değişkenli İstatistiksel Analizlerle Karşılaştırılması”, *Çukurova Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 2(19), 157-168.
- Pamuk, M. (2005), “Türkiye’de İllerin GSYİH’nin Sektör Paylarına Göre Gruplandırmasına İstatistiksel Bir Yaklaşım”, *Maliye Araştırma Konferansları*, (47), 142-156.
- Suhr, D. D. (2006), “*Exploratory or Confirmatory Factor Analysis?*”, SAS Institute, 1-17.
- Şencan, H. (2005), “*Sosyal ve Davranışsal Ölçümlerde Güvenilirlik ve Geçerlilik*”, (1. baskı), Seçkin Yayınları, Ankara.
- Tabachnick, B. G. ve Fidell, L. S. (2007), “*Using Multivariate Statistics*”, (5. baskı), Pearson Education Limited, Washington.
- Tatlıdil, H. (2002), “*Uygulamalı Çok Değişkenli İstatistiksel Analiz*”, (2. baskı), Ziraat Matbaacılık, Ankara.

- Temel, T., A. ve Uzundumlu, S. (2015), “Rize İlinde Hanelerin Balık Tüketimi Üzerine Etkili Olan Faktörlerin Belirlenmesi”, *Menba Kastamonu Üniversitesi Su Ürünleri Fakültesi Dergisi*, 1(82), 14-22.
- Türe, H. ve Akdi, Y. S. (2015), “Mevsimsel Kointegrasyon: Türkiye Verilerine Bir Uygulama”, 7.Ulusal Ekonometri ve İstatistik Sempozyumu, Düzenleyen: İstanbul Üniversitesi, 26-27 Mayıs, 1-19.
- Yılmaz, Ü. (2011), “Türkiye’de İllerin Sosyoekonomik Gelişmişlik Düzeylerinin Faktör Analizi ve Kümeleme Analizi ile İncelenmesi”, *Yüksek Lisans Tezi*, Karadeniz Teknik Üniversitesi, Ekonometri Anabilim Dalı, Trabzon.

