



T.C.

AKDENİZ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

EĞİTİM BİLİMLERİ ANA BİLİM DALI

YÜKSEK
LİSANS
TEZİ

EĞİTİM VERİLERİ ÜZERİNDE TEMEL
BİLEŞENLER ANALİZİ VE REGRESYON
ANALİZİNİN R PROGRAMI ÜZERİNDE
UYGULANMASI VE YORUMLANMASI

NEFİN YAŞAR

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME
TEZLİ YÜKSEK LİSANS PROGRAMI

Antalya, 2022

T.C.
AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME
TEZLİ YÜKSEK LİSANS PROGRAMI

EĞİTİM VERİLERİ ÜZERİNDE TEMEL BİLEŞENLER ANALİZİ
VE REGRESYON ANALİZİNİN R PRORAMI ÜZERİNDE
UYGULANMASI VE YORUMLANMASI

YÜKSEK LİSANS TEZİ
Nefin YAŞAR

Danışman
Doç. Dr. Alper SİNAN

Antalya, 2022

DOĐRULUK BEYANI

Yüksek lisans tezi olarak sunduđum bu çalıřmayı, bilimsel ahlak ve geleneklere aykırı düřecek bir yol ve yardıma başvurmaksızın yazdıđımı, yararlandıđım eserlerin kaynakçalardan gösterilenlerden olduđunu ve bu eserleri her kullandıđımda alıntı yaparak yararlandıđımı belirtir; bunu onurumla dođrularım. Enstitü tarafından belli bir zamana bađlı olmaksızın, tezimle ilgili yaptıđım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara katlanacađımı bildiririm.

28/06/2022

Nefin YAŐAR

AKDENİZ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Nefin YAŞAR'ın bu tez çalışması tarihinde jürimiz tarafından Eğitim Bilimleri Ana Bilim Dalı Eğitimde Ölçme ve Değerlendirme (Tezli Yüksek Lisans Programında Yüksek Lisans Tezi) olarak oy birliği/oy çokluğu ile kabul edilmiştir.

İmza

Başkan :.....

Üye :.....

Üye :.....

YÜKSEK LİSANS TEZİNİN ADI: Eğitim Verileri Üzerinde Temel Bileşenler Analizi ve Regresyon Analizinin R Pprogramında Uygulanması Ve Yorumlanması

ONAY: Bu tez, Enstitü Yönetim Kurulunca belirlenen yukarıdaki jüri üyeleri tarafından uygun görülmüş ve Enstitü Yönetim Kurulunun tarihli ve sayılı kararıyla kabul edilmiştir.

Doç. Dr. Güçlü ŞEKERCİOĞLU

Enstitü Müdürü

TEŐEKKÜR

Tez alıőmamın her anında bilgi ve gürüşlerini benden esirgemeyen, sabır ve hoşgörü ile bana her zaman destek olan, akademik kimlięi ve başarılarıyla her daim rol model bir çizgiye sahip pek deęerli danışmanım Do. Dr. Alper SİNAN' a, sonsuz saygı ve teşekkürlerimi sunarım.

Bu süreçte, her anıma őahitlik eden, benimle gülen, aęlayan, kimi zaman sadece susan deęerli eőim, hayat arkadaőım Recep Gürkan YAŐAR'a yanımdaydım, bana olan inancını hiç kaybetmedięi ve her zaman gülümseyerek hatırlayacaęım anılar biriktirmemi sağladığı için teşekkür ederim.

ÖZET

EĞİTİM VERİLERİ ÜZERİNDE TEMEL BİLEŞENLER ANALİZİ VE REGRESYON ANALİZİNİN R PROGRAMINDA UYGULANMASI VE YORUMLANMASI

YAŞAR, Nefin

Yüksek Lisans, Eğitim Bilimleri Ana Bilim Dalı,

Tez Danışmanı: Doç. Dr. Alper SİNAN

June 2022, 66 sayfa

İnsan hayatının her anında bir karar verme süreci içerisindeyiz. Aldığımız kararların, seçtiğimiz yolların birçok sebebi vardır. Ama hepimiz biliriz ki onlarca sebep arasından bazıları diğerlerine göre daha önemlidir ve düşüncelerimizi en çok onlar şekillendirir. Bunu bilerek insan beyni bilgiyi de aynı mantıkla işleme ihtiyacı duymaktadır.

Bilgi yığılmalı bir şekilde birikmeye devam etmekte ve gün geçtikçe daha kaotik bir hal almaktadır. Teknolojinin de gelişmesiyle artık hayatımızda çok boyutlu verilerin varlığı yadsınamaz bir gerçektir. Bu büyük verilere ulaşmak günümüzde oldukça kolaylaşmış olsada, ulaştığımız bu veriyi işlemek çok daha karmaşık bir süreç haline gelmiştir. Özellikle son dönemlerde tüm dünyada eğitim yöntemlerine ve eğitim başarısının sebeplerini araştırmaya yönelik artan ilgi aynı oranda elde edilen bilginin de artmasına sebep olmuştur.

Bilimsel araştırmaların odak noktalarından biri veri ve bu verilerin doğru bir şekilde işlenerek ve analiz edilmesidir. Veri analizi işlemleri özellikle çok boyutlu verilerde birtakım zorlukları da beraberinde getirir. Verinin boyutu, depolama zorlukları, betimsel istatistiksel, regresyon, korelasyon ve benzeri yapılması istenen tüm işlemler veri analiz sürecinin içerisinde yer almaktadır.

Eğitim yöntemlerini ve başarılarını uluslararası boyutta karşılaştırmak için, Ekonomik Kalkınma ve İşbirliği Örgütü (OECD) tarafından ilk olarak 1997 yılında uygulanan Uluslararası Öğrenci Değerlendirme Programı (PISA), üç yılda bir on beş yaşındaki öğrencilerin başarısını test etmektedir. Bu karşılaştırma ise eğitim başarısında, hangi faktörlerin daha etkili olduğu sorusunu akla getirmektedir.

PISA (Programme for International Student Assessment-Uluslararası Öğrenci Değerlendirme Programı), çok sayıda değişkene ilişkin veri toplayan geniş ölçekli bir sınavdır. Uluslararası düzeyde uygulanan bu sınav sayesinde sadece sınav puanına değil aynı zamanda sınav puanını etkileyebilecek faktörlere ilişkin verilere de ulaşmak mümkündür.

Gerçekleştirilen bu tez çalışmasında ise büyük verilerle çalışırken en çok kullanılan boyut indirgeme ve anlamlandırma yöntemi olan Temel bileşenler analizi ve regresyon analizi yöntemi ile R programı üzerinde PISA 2018 sınav puanları ve o döneme ait farklı değişkenlere ait veriler işlenerek alınan puanlar üzerindeki en etkili değişkeni belirlemek ve aralarındaki ilişkiyi gözlemlemek amaçlanmıştır.

Anahtar Sözcükler: Temel Bileşenler Analizi, Regresyon Analizi, R Programı, PISA 2018

ABSTRACT

APPLICATION AND INTERPRETATION OF BASIC COMPONENT ANALYSIS AND REGRESSION ANALYSIS ON EDUCATIONAL DATA ON THE R PROGRAM

YAŞAR, Nefin

Master of Arts, Department of Educational Sciences,

Supervisor: Assoc. Prof. Dr. Alper SİNAN

May 2022, 66 pages

A person is in a decision-making process at every moment of his life. There are many reasons for the decisions we make, the paths we choose. But we all know that among dozens of reasons, some are more important than others and they shape our thoughts the most. By knowing this, the human brain needs to process information with the same logic.

Information continues to accumulate in a cumulative manner and is becoming more chaotic day by day. With the development of technology, the existence of multidimensional data in our lives is an incontrovertible fact. Although it has become quite easy to reach this big data today, processing this data we have reached has become a much more complex process. Especially in recent years, the increasing interest in education methods and researching the reasons for educational success all over the world has caused an increase in the knowledge obtained at the same rate.

One of the focal points of scientific research is data and the correct processing and analysis of this data. Data analysis processes bring some difficulties, especially in multidimensional data. One of the focal points of scientific research is data and the correct processing and analysis of this data. Data analysis processes bring some difficulties, especially in multidimensional data. The size of the data, storage difficulties, descriptive statistics, regression, correlation and all similar operations are included in the data analysis process. The Program for International Student Assessment (PISA), which was first implemented by the Organization for Economic Cooperation and Development (OECD) in 1997, in order to compare educational methods and achievements internationally, tests the

success of fifteen-year-old students every three years. This comparison brings to mind the question of which factors are more effective in educational success.

PISA (Programme for International Student Assessment) is a large-scale exam that collects data on many variables. Thanks to this exam, applied at an international level, it is possible to access not only the exam score, but also data on the factors that may affect the exam score.

In this thesis study, it was aimed to determine the most effective variable on the scores obtained by processing the PISA 2018 exam scores and the data of different variables of that period on the R program with the Principal Components analysis and regression analysis method, which are the most used dimension reduction and interpretation methods when working with big data.

Keywords: PISA 2018, R Program, Principal Component Analysis, Regression Analysis

İÇİNDEKİLER

DOĞRULUK BEYANI.....	III
TEŞEKKÜR.....	İİ
ÖZET.....	İİİ
ABSTRACT	V
İÇİNDEKİLER.....	Vİİ
TABLolar LİSTESİ	X
KISALTMALAR LİSTESİ	Xİ

BÖLÜM I

GİRİŞ

1.1. Problem Durumu	1
1.2. Araştırmanın Amacı ve Problemleri.....	2
1.2.1. Araştırmanın Alt Problemleri	2
1.3. Araştırmanın Önemi	2
1.4. Araştırmanın Sayıltıları	3
1.5. Araştırmanın Kapsam ve Sınırlılıkları.....	3
1.6. Tanımlar	3

BÖLÜM II

KURAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

2.1. İstatistiksel Temel Bilgiler	4
2.1.1. Kovaryans Matris	4
2.1.2. Özvektörler ve Özdeğerler	5
2.1.3. Tekil Değer Ayrışımı.....	6
2.1.4. Dikey Bazı Matrisler	8
2.1.5. Dik İzdüşüm	8
2.2. Temel Bileşenler Analizi	9
2.2.1. Temel Bileşenler Analizinin Özellikleri ve Sağladığı Yararlar.....	12
2.3. Regresyon Analizi	13
2.4. Çoklu Regresyon Modeli.....	14
2.4.1.1. Çoklu Regresyon Modeli Varsayımları	15
2.5. Bağımsız Değişkenler Arasında İlişki Olmamalıdır.....	15
2.5.1. Çoklu Bağlantı Sorununun Meydana getirdiği Etkiler	16
2.5.2. Çoklu Bağlantı Sorununun Belirlenmesi.....	16
2.5.3. Çoklu Bağlantı Sorununun VIF Değerleri İle İncelenmesi	16
2.6. Temel Bileşenler Analizi ve Regresyon Analizi	17
2.7. PISA Sınavı	17
2.8. R Programı.....	19

BÖLÜM III

YÖNTEM

3.1. Araştırma Modeli.....	20
3.2. Araştırmanın Materyali.....	20

3.3. Veri Çözümleme Yöntemi.....	21
3.4. Veri Toplama Aracı.....	21
3.5. Veri İşleyiş Süreci	21

BÖLÜM IV

BULGULAR VE YORUMLAR

4.1. Değişkenlere İlişkin Regresyon Analizi Sonuçları.....	23
4.2. Temel Bileşenler Analizine İlişkin Sonuçlar.....	29
4.3. Temel Bileşenlere İlişkin Regresyon Analizi Sonuçları.....	33

BÖLÜM V

SONUÇ VE ÖNERİLER

5.1. Sonuç ve Öneriler	37
------------------------------	----

KAYNAKÇA.....	39
----------------------	-----------

EKLER	41
--------------------	-----------

Ek-1: Temel Bileşenler Analizi ve Regresyon Analizine İlişkin R Kodları.....	41
--	----

Ek-2: Temel Bileşenler Tablosu	45
--------------------------------------	----

Ek-2: Temel Bileşenler Tablosu (Devam).....	47
---	----

ÖZGEÇMİŞ	49
-----------------------	-----------

BİLDİRİM.....	50
----------------------	-----------

İNTİHAL RAPORU	51
-----------------------------	-----------

TABLolar LİSTESİ

Tablo 4.1. *Tez Çalışmasına Dahil Edilen Ülkeler*

Tablo 4.2. *Tez Çalışmasına Dahil Edilen Bağımlı ve Bağımsız Değişkenler*

Tablo 4.3. *Ülkelerin PISA Puanlarına Ait Dağılım Grafiği*

Tablo 4.4. *Okuma Puanı Bağımlı Değişkenine Ait Regresyon Sonuçları*

Tablo 4.5. *Matematik Puanı Bağımlı Değişkenine Ait Regresyon Sonuçları*

Tablo 4.6. *Fen Puanı Bağımlı Değişkenine Ait Regresyon Sonuçları*

Tablo 4.7. *Kaiser-Meyer-Olkin (KMO) Testi Sonuçları*

Tablo 4.8. *Okuma Puanı Değişkenine Ait Temel Bileşenler Regresyonu Sonuçları*

Tablo 4.9. *Matematik Puanı Değişkenine Ait Temel Bileşenler Regresyonu Sonuçları*

Tablo 4.10. *Fen Puanı Değişkenine Ait Temel Bileşenler Regresyonu Sonuçları*

KISALTMALAR LİSTESİ

PCA: Temel bileşenler analizi

O: Okuma puanı

M: Matematik okuryazarlığı puanı

F: Fen okuryazarlığı puanı

X1: Nüfus

X2: Gayri safi yurt içi hasıla

X3: İşsizlik oranı:

X4: Kadın iş gücüne katılım oranı(15-24 yaş)

X5: Erkek iş gücüne katılım oranı(15-24 yaş)

X6: Örgün çocuk bakımı ve okul öncesi eğitime kayıt(0-2yaş)

X7: Kişi başı oda sayısı

X8: Eğitimde geçen yıllar

X9: Eğitimde ve istihdamda yer almayan erkekler(15-19yaş)

X10: Eğitimde ve istihdamda yer almayan kadınlar(15-19yaş)

X12: Eğitimde ve istihdamda yer almayan erkekler(15-29yaş)

X11: Eğitimde ve istihdamda yer almayan kadınlar(15-29yaş)

X13: Eğitimde ve istihdamda yer almayan erkekler(20-24yaş)

X14: Eğitimde ve istihdamda yer almayan kadınlar(20-24yaş)

X16: Eğitimde ve istihdamda yer almayan erkekler(25-29yaş)

X15: Eğitimde ve istihdamda yer almayan kadınlar(25-29yaş)

X17: Kaba evlilik oranı

BÖLÜM I

GİRİŞ

1.1. Problem Durumu

Bu tez çalışması, R Programı üzerinde çok değişkenli istatistiksel analiz yöntemi olan Temel Bileşenler Analizi (PCA) ve Regresyon Analizi ile PISA 2018 sınav puanlarının ve o döneme ilişkin çeşitli değişkenlerin incelenmesi ve anlamlandırılması çalışmasıdır.

Eğitime verilen önemin artması ile birlikte bu alanda farklı çalışmalar yapılmakta ve birçok soruya cevap aranmaktadır. Eğitim bir süreçtir ve bu sürecin sonucunda elde edilen başarının bireyden bireye, bölgeden bölgeye ve ülkeden ülkeye farklılıklar göstermesi akıllarda soru işaretleri bırakmaktadır. Bu başarıyı olumlu veya olumsuz etkileyen faktörlerin varlığı ve etkileme oranları yıllarca birçok araştırmacı tarafından inceleme konusu haline gelmiştir.

Eğitim ve birçok alandaki veriyi işleme isteği beraberinde kullanılması gereken program ve analizler neler olmalıdır sorusunu da karşımıza çıkarmaktadır.

Özellikle eğitimdeki başarının artırılma çabaları ışığında 2018 PISA verileri incelenerek başarıyı en çok etkileyen faktörler problemine ve bu analizlerin R programı üzerinde nasıl yapılacağına ilişkin sorulara cevap aranmıştır.

Bu ve buna benzer soruların cevabını aramaya başlayan birçok araştırmacı elde ettiği çok değişkenli ve büyük verilerle işlem yapmanın zorluğuyla karşı karşıya kalmıştır. İlerleyen teknolojik gelişmeler sayesinde verileri bilgisayarlar ve paket programlar üzerinden analiz etmek araştırmacılara gerek zamandan gerek enerjiden tasarruf sağlamıştır. Bu noktada ise hangi programı, hangi paket uygulamalarını kullanmalıyım soruları analiz sürecinde araştırmacıları düşündüren konulardan olmuştur.

Temel bileşenler analizinin bu çalışmada kullanılmasının amacı verideki değişkenler arasındaki ilişki yapısını yok ederek regresyon modeline dahil etmektir.

Yine bu tez çalışması açık kaynak kodlu R Programı üzerinde temel bileşenler ve regresyon analizlerini yaparak araştırmacılara bu programın uygulama kolaylığını göstermeyi amaçlamıştır. Aynı zamanda bu tez çalışması bir araştırmacı için verinin nasıl R programına çekildiği, ön işleme süreçleri, analizlerin hangi adımlarla uygulandığı sorularını da cevaplar niteliktedir.

1.2. Araştırmanın Amacı ve Problemleri

Bu tez çalışmasının amacı açık kaynak kodlu R programı üzerinde PISA 2018 verilerinin Temel Bileşenler Analizi ve Regresyon Analizi yöntemleri kullanılarak incelenmesi ve yorumlamasının yapılmasıdır.

1.2.1. Araştırmanın Alt Problemleri

- i. R Programında çok boyutlu veri setleri nasıl işlenir?
- ii. Çok boyutlu veri setleri için R programında Temel Bileşenler Analizi nasıl uygulanır?
- iii. Büyük veri setleri için R programında Regresyon Analizi nasıl uygulanır?
- iv. R Programında uygulanan Temel Bileşenler Analizi ve Regresyon Analizi ile ulaşılan veriler nasıl yorumlanır?
- v. Eğitim başarısı nelerden etkilenir?
- vi. Eğitim başarısını en fazla etkileyen faktör nedir?

1.3. Araştırmanın Önemi

Bu tez çalışması eğitim alanında henüz kullanımı yaygınlaşmamış olan açık kaynak kodlu R programının büyük veri setlerine nasıl kolayca uygulanabildiğinin örnekleri arasında yer almaktadır. Temel bileşenler analizi ve regresyon analizinin R programında uygulanma aşamalarını göstermektedir. Ayrıca bu analizlerin yorumlama ve anlamlandırma çalışmalarının nasıl yapılacağı sorularını da cevaplar niteliktedir.

1.4. Araştırmanın Sayıtları

Kullanılacak veri setlerinin uzmanlar tarafından doğru bir şekilde toplandığı varsayılacaktır. Sınava katılan öğrenciler, öğrenci anketlerindeki maddeleri gerçeğe uygun cevaplamışlardır. Elektronik ortamdan elde edilen verilerin doğru olduğu varsayılacaktır.

1.5. Araştırmanın Kapsam ve Sınırlılıkları

Araştırma kapsamında çok boyutlu veri setlerinin eğitim ile ilgili istatistiksel verileri kullanılacaktır.

Yapılan tahmin ve yorumlamalar PISA öğrenci anketlerinde yer almış özellikleri kapsamaktadır.

1.6. Tanımlar

Eğitim başarısı: Kısa veya uzun vadeli, öngörülen süre içerisinde kişi veya kurumların hedeflenen eğitim seviyesine ulaşmasıdır.

R: Genellikle istatistiksel analizlerde araştırmacılar tarafından kullanılan açık kaynak kodlu bir yazılım programıdır.

Temel bileşenler analizi: Genel anlamı ile çok boyutlu ve büyük veri setlerinde boyut indirgeyerek toplam varyansı en çok açıklayan daha az sayıda değişken oluşturma işlemidir.

Regresyon analizi: En basit tanımı ile değişkenler arasındaki ilişkinin var olup olmadığını veya var olan ilişkiyi modelleyerek tahmin ve öngörü için kullanılır.

PISA sınavı: Üç yıllık periyotlarla matematik okuryazarlığı, fen okuryazarlığı, okuma yazma alanlarında yapılan ulusal ölçekli bir sınavdır.

Ortogonal matris: Transpozesi ile tersi birbirine eşit olan matristir.

BÖLÜM II

KURAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

2.1. İstatistiksel Temel Bilgiler

2.1.1. Kovaryans Matris

Kovaryans, iki değişkenin hareketlerinin birbirini ne kadar etkiler durumda olduğunu göstermektedir. Diğer bir anlamda kovaryans iki değişkenin birlikte ne kadar değiştiklerinin ölçüsüdür. Kovaryans aşağıdaki formül aracılığı ile hesaplanabilir.

$$x = x_1, x_2, \dots, x_n$$

$$y = y_1, y_2, \dots, y_n \quad \text{olsun,}$$

$$\mathbf{Cov} = \frac{\sum(x-\bar{x})(y-\bar{y})}{N-1} \quad (2.1)$$

\bar{x} : x vektörüne ait verilerin ortalaması

\bar{y} : y vektörüne ait verilerin ortalaması

N: Toplam veri sayısı

Kovaryans matris ise değişkenlerin hesaplanan kovaryans sonuçlarını veren matristir.

$$Cov(x, y) = Cov(y, x) \quad (2.2)$$

Kovaryans matris yukarıdaki eşitlikten anlaşılacağı gibi simetrik bir özelliğe sahiptir.

Ayrıca x ve y birbirinden bağımsız değişkenler ise;

$$Cov(x, y) = 0 \quad (2.3)$$

yani değişkenler arasında bir ilişki yoktur.

Kovaryans birlikte değişmek anlamındadır ve değişkenler arasındaki ilişkinin yönünü analiz etmek için kullanılır. Kovaryans formülü kullanılarak elde edilen sonuçlara baktığımızda; eğer değer pozitif ise pozitif bir ilişki olduğu anlamına gelir, negatif ise negatif bir ilişki olduğu anlamındadır.

Genel anlamıyla kovaryans hesaplamasından elde edilen sonucun pozitif veya negatif ya da sıfır olması durumu önemlidir. Elde ettiğimiz sonucun büyüklüğü bizim için önemli değildir. Çıkan sonuçlara baktığımızda; negatif değerler ters yönde bir ilişkiyi, pozitif değerler aynı yönde bir ilişkinin varlığını doğrularken sıfır değeri ise değişkenler arasında bir ilişki varlığının olmadığını göstermektedir.

Kovaryans matris istatistiksel analizler için oldukça önemlidir. Yapısı itibari ile kovaryans matris köşegenlerinde değişkenlere ait varyans değerleri ile değişkenlerin yayılımını verirken, kovaryans değerleri ile değişkenler arasındaki karşılıklı değişimi yani kovaryans değerini verir.

2.1.2. Özvektörler ve Özdeğerler

Bir matris hakkında en önemli bilgileri matrisin özelliklerine bakarak anlarız. Yani özdeğerler, özuzaylar ve özvektörlere bakarak matris hakkında önemli bulgular edinebiliriz. Aynı zamanda bu özellikleri matrisleri çarpanlarına ayırırken de kullanabiliriz.

Bir matris, bir vektör üzerine etkidiğinde genellikle vektörün büyüklüğünü ve doğrultusunu değiştirir. Ancak bazı durumlarda vektör üzerine uygulanan matris, vektörün

sadece büyüklüğünü değiştirir. Doğrultusunu etkilemez fakat vektörün yönünü ters çevirebilir. Temel referans noktamız doğrultusunun değişmemesidir. Matrisin, vektör üzerine uygulanmasında doğrultusu değişmeyen vektörlere o matrisin özvektörleri adı verilir. Özvektörler sadece kare matrislerden elde edilebilir. Eğer bir özvektöre ve dolayısıyla özdeğere ihtiyacımız var ise kovaryans matris kullanmamız gerekmektedir.

Bir matrisin özdeğeri ise, özvektör üzerine etkiyerek özvektörün büyüklüğünü bir çarpan değerinde değiştirir. Bu çarpan ise o özvektöre ait özdeğer olarak adlandırılır. . Eğer bu çarpan negatif bir değer ise doğrultusunu değiştirmeden vektörün sadece yönünü değiştirir. Çarpan pozitif bir değer almış ise yine doğrultusu sabit kalmak koşuluyla vektörün yönü de değişmez. Yani çarpan pozitif ve negatif olsun vektörün doğrultusu aynı kalır.

Aynı özdeğere sahip bütün özvektörlerin oluşturduğu kümeye ise özuzay denir.

2.1.3. Tekil Değer Ayrışımı

Bilindiği üzere gerekli şartları sağladığında matrisler bir çarpan kadar genişletilebilir. Yine aynı şekilde matrisler karmaşık bir yapıyı daha basit ve daha kolay işlenebilir bir şekle getirmek amacıyla çarpanlarına ayrılabilir. Daha önceleri sadece lineer sistem analizi uygulamalarında kullanılan matris ayrışımı 1960'tan sonraki yıllarda regresyon analizi, , kuantum, yazılım, öneri sistemleri, sayısal hava tahminleri, temel bileşenler analizi ve matris dönüşümlerinde de kullanılmaktadır. SVD (Singular Value Decomposition), LU, QR, CS ayrışımı, Pivoted, QLP ayrışımı, QR ve Pivoted QLP ayrışımı gibi yöntemler bir matrisin ayrışımında kullanılır.

Matris ayrışımında en çok matrisin köşegenleştirilmesi mantığına dayanan Tekil Değer Ayrışımı (Singular Value Decomposition-SVD) kullanılmaktadır. Bir matrisin çarpanlarına ayrıştırılma yöntemlerinden biri olan Tekil Değer Ayrışımı diferansiyel geometriciler tarafından geliştirilmiştir. Eugenio Beltrami (1873) ve Camille Jordan'ın (1874) çalışmalarından bağımsız olarak James Joseph Sylveste 1889 yılında kare matrisler için tekil

değer ayrışımı bularak tekil değerleri A matrisinin “kanonik çarpanları” şeklinde isimlendirdi.

Emile Picard 1910 yılında ilk defa singular değer terimini kullanmıştır. Tekil Değer Ayrışımı birbirinden bağımsız araştırmacıların da içinde yer aldığı Erjand Schmidt (1907), Autonne (1915), Carl Eckart ve Gale J. Young (1936), Kogbetliantz (1954-1955), Hestenes (1958) gibi birçok araştırmacının üzerinde çalışıp yer yer geliştirdiği bir konu olmuştur. Gene Golub ve William Kahan 1965 yılında singular değer ayrışımını hesaplamak için Golub/Kahan adlı algoritmayı yayınlamışlardır. Golub ve Christian Reinsch 1970 yılında günümüzde sıkça kullanılmakta olan Golub/Kahan algoritmasının farklı bir biçimini yayınlamışlardır.

Tekil Değer Ayrışımı (Singular Value Decomposition), büyük boyuttaki veri kümelerinde boyut azaltma, sınıflama ve kümeleme için kullanılan bir yöntemdir. Bu yöntem ile matrisler anlamlı bir vektör uzayı elde etmek için ayrıştırılır. Tekil Değer Ayrışımı bir matrisi ortonormal bir matris, köşegen bir matris ve ortonormal bir matris olmak üzere 3 parçaya ayırıp ve yine bu üç ayrı matris ile aynı matrise tekrar ulaşmayı hedefleyen bir algoritmadır. Matrisler bu yöntem kullanılarak anlamsal bir vektör uzayı oluşturmak için ayrıştırılmaktadır. Bu yöntemle M matrisi üç farklı matris olarak ifade edilir.

$$M = U\Sigma V \quad (2.4)$$

U matrisi, üniter matris (unitary matrix) olmaktadır. Tersyüz eşleniği ile kendisi çarpıldığında birim matris elde edilir.

$$U^{-1}.U = I \quad (2.5)$$

V matrisi, M matrisinin birimlik özelliklerini taşıyan ortogonal matristir.

Σ matrisi ise tekil deęer (singular values) zelliklerini taşıyan bir kşegen (diagonal) matrisi olmaktadır.

Σ , n. basamaktan bir kare matris olmak zere;

$i \neq j$ iin $\Sigma_{ij} = 0$ ise Σ kşegen matristir.

2.1.4. Dikey Bazı Matrisler

$$A = U \Sigma V^H = \sum_{i=1}^n \sigma_i u_i v_i^H \quad (2.6)$$

Bu formlde A, n. dereceden bir matris, 1. dereceden ve birbirine dikey n tane matrisin doęrusal birleřimi (lineer kombinasyonu) biiminde yazılır. A matrisine en ok katkıda bulunan matris matrisi, nk en byk tekil deęer matrislerinin Frobenius normlarıdır.

Eldeki matrisi eřit byklkteki farklı legolardan oluřmuř bir oyuncak gibi dřnlrse, bir biimde bu oyuncakta kullanılan paraları bulunur ve paraların hangi oranda kullanıldıęı belirlenir.

2.1.5. Dik İzdřm

$$P[A] = AA^+ \quad (2.7)$$

A'nın stun vektrlerinin gerdięi $[A] \subset \mathbb{R}^{n \times 1}$ alt uzayı zerine dik izdřm matrisi AA^+ matrisine eřittir.

$$[A] \perp (R^n = [A] \oplus [A] \perp) \quad (2.8)$$

A'nın sütun vektörlerinin gerdiği uzaya dik olan alt uzay olmak üzere, bu alt uzay üzerine dik izdüşüm matrisi aşağıdaki gibidir.

$$P[A] \perp = I - AA^+ \quad (2.9)$$

2.2. Temel Bileşenler Analizi

Temel Bileşenler Analizi (Principal Component Analysis) Karl Pearson (1901) tarafından bilim dünyasına tanıtılmıştır. Hotelling'in 1933'lü yıllarda, Rao'nun ise 1964'lü yıllarda yaptığı çalışmalar ve katkılarıyla temel bileşenler analizinin uygulama alanları da genişletilmiştir. Genel olarak bu tekniğin günümüzde kullanılan tanımının Pearson ve Hotelling tarafından yapıldığı kabul edilmektedir.

İstatistiksel tekniklerin kökeni hakkında net bilgiler ve tarihler söylemek oldukça zordur. Ancak birbirlerinden bağımsız olarak Beltrami, Jordan ve Preisendorfer ile Mobley'in sırasıyla 1873, 1874 ve 1988 yıllarındaki tekil değer ayrışımı ile ilgili çalışmalarının temel bileşenler analizi yönteminin temelini oluşturduğu düşünülmektedir.

Hotelling görüşlerinin temelini faktör analizine dayandırmasına rağmen faktör analizinden belirgin şekilde farklar ortaya koyduğu görülmektedir. Hotelling p sayıdaki değişkenin değerine karşılık gelebilecek k sayıda değişkenler olabileceği düşüncesini savunmuştur. Faktör teriminin bu doğrultuda k'yı tam olarak tanımladığını varsayarak bu değişkenlerin "bileşen" kavramı ile isimlendirilebileceği düşüncesini ortaya atmıştır. Hotelling'in düşüncelerinin temelinde göre bileşenleri toplam varyansa en çok katkı sağlayacak şekilde belirler. Belirlediği bu bileşenlere ise "temel bileşenler" adını verir.

Temel bileşenler analizi ile ilgili çalışmaların yer alacağı listeye Preisendorfer ve Mobley'in (1988) kitabını da eklemek gereklidir. Bu kitap temel bileşenler analizine ilişkin

içeriğiyle Rao'nun düşüncelerine eşdeğer tutulmaktadır. Büyük bir kısmını Preisendorfer'in yazdığı kitabı, ölümünden sonra Mobley düzenleyerek basıma hazır hale getirmiştir.

Temel Bileşenler Analizi tek başına kullanılan bir analiz olduğu gibi değişkenler arasındaki bağımlılık yapısının yok edilmesi veya boyut indirgeme amacıyla ve başka analizler için de veri hazırlama tekniği olarak kullanılabilir. (Arnold, 1981; Ludwig ve Reynolds, 1988; Sharma, 1996). Sınıflama ve kümeleme amaçlı çalışmaların tamamı temel bileşenler analizini kullanmaktadır. Temel bileşenler analizi çok sayıda değişken değerini referans alıp örnekleri guruplara ayırıp kümeleyerek sınıflandırabilir. Çok değişkenli bir veri setinde değişkenlerin bağımlı ve bağımsız olarak ayrılması amacıyla regresyon analizi uygulanmak istenebilir. Değişkenler arasındaki birlikte değişimin anlamlandırılması ihtiyacıyla da temel bileşenler analizine başvurulmaktadır.

Çok değişkenli istatistiksel analizlerde n tane bireye ait p tane değişkenden bahsedilir. Veri setinin büyüklüğü günümüzde bir sorun olarak görünmesede çok sayıdaki değişkene ait analiz sonuçlarının anlamlandırılması ve yorumlanması zaman almakta ve kafa karıştırabilmektedir. Aynı zamanda değişkenlerin genelde çoğunluğunun birbirine bağımlı olması da analizler için sorun oluşturmaktadır. Örneğin bu tez çalışmasında ele alınan Pisa sınav puanlarını etkileyen değişkenler gibi veri setlerindeki değişkenlerden bazıları genellikle birbiri ile ilişkilidir. Bu ilişki değişkenlerin bağımsızlığı kuralını zedeler.

Böyle durumlarda en çok başvurulan temel bileşenler analizi yöntemi, veri içerisinde incelenen çok sayıdaki değişkeni, aralarındaki bağımlılık yapısını yok ederek doğrusal kombinasyonlarından oluşturulmuş daha az sayıdaki bileşene indirger. Varyans-kovaryans yapısını referans alarak verinin boyutunu indirgeyerek yorumlama kolaylığı sağlayan çok değişkenli bir istatistiksel analiz yöntemidir.

x 'in n birey ve p değişkenden oluşan bir veri matrisi olduğu düşünülürse x 'in uzaydaki görünümü her bireyin bir noktayı temsil ettiği çok sayıda noktadan oluşan bir topluluk olarak ifade edilir. Bu değişkenler arasında bir bağımlılık yapısı olacağı için bu topluluk bir bulut şeklinde nitelendirilebilir. Bu geometrik şeklin eksekleri birbirine dik olmadığı için tanımının ve yorumlanmasının yapılması mümkün olmayacaktır. Bu yüzden bu noktaları eksenleri birbirine dik bir elipsoid şeklinde incelemek hem açıklayıcı hem de anlamlandırılması

mümkün bir hale dönüştürecekler. Bu dönüştürmede, noktaların sahip olduđu toplam varyans deđişmeden kalır ve eksenler birbirine dik hale gelir. Böylelikle analize uygun bir yapıya dönüşür.

Harold Hotelling' e göre x ham veri matrisi ($x_{n \times p}$) doğrudan da kullanılabilir. Eğer ham veri matrisi kullanılacak ise temel bileşenleri bulmak için varyans-kovaryans matrisinden faydalanılır. Bu yolun kullanılabilmesi için deđişkenlerin ölçü birimlerinin aynı ve varyanslarının birbirine yakın olması gerekmektedir.

Diđer bir bileşen elde etme tekniđi ise $z_{n \times p}$ biçimindeki standartlaştırılmış deđerler matrisini kullanmaktır. Standartlaştırılmış veri matrisi kullanıldığında temel bileşenleri elde etmek için korelasyon matrisinden faydalanılır. Bu teknikte deđişkenlerin ölçü birimleri farklıdır.

Temel bileşenler x_1, x_2, \dots, x_p 'nin varyans kovaryans matrisi veya x_1, x_2, \dots, x_p 'nin korelasyon matrisine bađlıdır. Bu bileşenler çok deđişkenli normal dağılım varsayımını gerektirmez. Fakat çok deđişkenli normal dağılım kitlesinden çıkarılan temel bileşenler sabit yoğunluklu elipsoidler bakımından yararlı yorumlar yapmamızı sađlar. Bununla birlikte, eđer kitle çok deđişkenli normal dağılımlı ise çıkarsamalar örnek bileşenlerden yapılabilir (Johnson and Wichern, 1998: 357).

Temel bileşenler analizinde, en basit anlatımıyla, p sayıdaki orijinal deđişkeni temel bileşenler olarak bilinen k sayıdaki yeni deđişkenlere dönüştürmektedir. Aslında temel bileşen sayısı verimizdeki deđişken sayısına eşittir fakat yorumlama aşamasında genellikle ilk iki veya üç bileşen incelenir. Bu sayı, bileşenlerin varyansı açıklama miktarına göre deđişiklik gösterir. Temel bileşen sayısını belirlerken referans noktamız kümülatif varyansı % 67'den büyük yapan bileşenleri gözlemlemektir. Dikkate alınan bu bileşenler verinin büyük bir kısmının açıklanmasını ve yorumlanmasını sađlar.

Genel olarak p tane bađımsız deđişkendeki bilgiyi birinci temel bileşen en yüksek oranda; sonuncu bileşen ise en az oranda açıklar.

Tüm bunlar değişkenler arasındaki korelasyonun varlığı savına dayanarak yapılır. Değişkenler arasında ilişki olmaması halinde temel bileşenler analizi uygulamak anlamsız bir işlem haline gelir. Temel bileşenler analizinin uygulanmasının gerekli olup olmadığını incelemek için veri setine ait korelasyon matrisi, Kaiser-Meyer-Olkin (KMO) testi veya Bartlett testi kullanılabilir.

Bartlett testi; değişkenlere ilişkin korelasyon matrisinin birim matrisi ile kıyaslanarak test edilmesi ilkesine dayanır. Bartlett testi aynı zamanda korelasyon matrisinin anlamlılık testidir (Tatlıdil, 1996).

Bartlett sonucunda çıkan değer 0,05'ten küçük olması verilerin temel bileşenler analizine uygun olduğunu gösterir.

Kaiser-Meyer-Olkin (KMO) testi sonuçları ise 0 ile 1 arasında değişen değerler alabilmektedir ve 0,50 alt sınır kabul edilmektedir.

$$H_0: \Sigma = I$$

$$H_1: \Sigma \neq I$$

Biçimindeki hipotezlerden H_0 hipotezinin reddedildiği durumlarda temel bileşenler analizinin kullanımının uygun olduğu söylenebilir.

2.2.1. Temel Bileşenler Analizinin Özellikleri ve Sağladığı Yararlar

- P boyutlu bir veri için belirlenen ilk k tane temel bileşen toplam varyansın büyük bir bölümünün varyansını karşılayabiliyor ise geriye kalan p-k tane temel bileşenin varlığı ihmal edilebilir. Bu durum az bir bilgi kaybıyla, $k < p$ koşulunu sağlayarak p boyutlu bir verinin m boyutlu bir uzaya indirgenmesidir.
- Ham veri matrisi x ve standartlaştırılmış z matrisi içerisindeki değişkenler için bağımlılık yapısı mevcut iken, konum olarak bağımsız fakat ölçeğe bağımlı olan ve

geometrik olarak dik eksenler referans alınarak elde edilen y_j vektörleri eğik eksenler üzerinde bulunan z_j değerlerinden farklı olarak birbirinden bağımsızdır.

- Boyut indirgeme olmaksızın p tane y_j bileşenin kullanılması p tane bağımsız değişken elde edilmesini sağlar.
- y_j vektörleri birbirine dik olduğu için kovaryans terimi bulunmaz ve dağılım yalnızca varyansla yorumlanabilir.
- y_j eksenlerinin varyansları büyükten küçüğe doğru sıralanmaktadır.
- Değişkenleri örnek olarak tahminleme işlemleri yapabilir.
- Çok boyutlu verilerin görselleştirilmesinde kolaylıklar sağlar.
- Veri sıkıştırma işlemi için de kullanılabilir.
- Basitleştirme ve boyut indirgeme görevi görür.
- Birimlerin sınıflandırılması görevini üstlenebilir
- Değişkenler arasındaki bağımlılık yapısını ortaya çıkarır.
- Hipotez oluşturma amacıyla kullanılabilir
- Sıralama ve ölçekleme işlevi vardır.

Tekniğin görünüşteki basitliğine rağmen, temel bileşenler analizi alanında birçok araştırma halen yapılmakta ve yaygın bir şekilde kullanılmaktadır.

2.3. Regresyon Analizi

İngiliz istatistikçi Francis Galton, çocuğun boyu ile anne ve babanın boyu arasındaki ilişkiyi incelerken, araştırmasına dayanarak “geri dönüş” anlamında önce reversion kelimesini kullanmış fakat sonraki çalışmalarında regression sözcüğünü kullanmayı tercih etmiş ve regresyon yöntemini geliştirmiştir. En bilindik isimlerden Karl Pearson, Undy Yüle ve R.A Fisher gibi araştırmacıların da katkılarıyla kullanım alanı genişletilmiştir.

Regresyon analizi iki ya da daha çok deęişken arasındaki ilişkiyi incelememize olanak sağlayan bir yöntemdir. Bu analiz ile varsa deęişkenler arasındaki ilişki ve bu ilişkinin gücü hakkında bilgi sahibi olabiliriz. Hem tanımlayıcı hem de çıkarımsal bir istatistik sağlar.

Regresyon analizi, bağımlı deęişken ile bağımsız deęişken arasındaki ilişkiyi modellemek için kullanılan istatistiksel bir metottur. Daha genel bir tanım olarak regresyon analizi, bağımlı bir deęişkenin bağımsız deęişkenlerdeki deęişikliklere göre nasıl deęiştiğini inceler.

Regresyon analizi ile gelecek tahmini yaparken belirtme katsayısından faydalanılır. Bu katsayı deęerinin 0,5 ile 0,99 arasında olması beklenir. Bu deęer aralığında yer alan katsayı deęişkenlerin birbiri ile pozitif bir ilişki içerisinde olduğunu niteler.

Regresyon analizi yöntemi genellikle, pazarlama, üretim, ilaç, finans gibi sektörlerde SPSS, R ve Minitab programlarıyla birlikte uygulanmaktadır.

2.4. Çoklu Regresyon Modeli

Günümüzde birçok alanda ve özellikle eğitim alanında bağımlı bir deęişken tek bir bağımsız deęişken ile açıklanamaz. Eğitim modelleri genellikle birden fazla sebebin oluşturduğu sonuçlardır. Bu modellerde sonuçlar çok sayıda bağımsız deęişkenin bağımlı deęişkeni etkilemesi ile ortaya çıkar. Bazı durumlarda bu bağımsız deęişkenler birbirini de etkilemektedir. Birden fazla bağımsız deęişkenin bulunduğu bu durumlarda çoklu regresyon analizi (multiple regression analysis) kullanılır.

Y bağımlı deęişken ve $x_1, x_2, x_3, \dots, x_p$ bağımsız deęişken olmak üzere çoklu regresyon denklemi aşağıdaki şekilde verilmiştir.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_{ij} \quad (2.10)$$

Denklem de yer alan $\beta_j, j=0,1,\dots,p$ deęerlerine regresyon katsayıları denir. β_j deęerleri, $i \neq j$ olmak üzere tüm x_i bağımsız deęişkenleri sabit olduğunda, x_j deki her bir

birimlik deęişime karşılık y bağımlı deęişkenindeki beklenen deęişimi gösterir. Bu nedenle β_j deęerlerine kısmi regresyon katsayıları da denir

Ayrıca bağımlı deęişkenimiz Y , $n \times 1$ vektör ve bağımsız deęişken kümemiz X , $n \times (p+1)$ boyutlu matris, β $(p+1) \times 1$ boyutlu katsayılar vektörü ve ε , $(n \times 1)$ boyutlu hata vektörü olmak üzere regresyon denklemi aşığıdaki şekilde yazılabilir.

$$Y = X\beta + \varepsilon \quad (2.11)$$

2.4.1. Çoklu Regresyon Modeli Varsayımları

Çoklu regresyon modelinin varlığından bahsedebilmek için birtakım varsayımların varlığını saptamak önemlidir. Bu varsayımlar aşığıdaki gibidir;

- ❖ Hata terimlerinin hesaplanan aritmetik ortalaması sıfır bulunmalıdır.
- ❖ Hata terimleri normal bir dağılım göstermelidir.
- ❖ Hata terimlerinin varyansının sabit olması gerekmektedir.
- ❖ Hata terimleri birbirinden bağımsız nitelikte olmalıdır.
- ❖ Bağımlı deęişken sürekli ve normal dağılıma sahiptir.
- ❖ Parametre sayısı gözlem sayısından küçük olmalıdır.
- ❖ Bağımlı ve bağımsız deęişkenler arasında doğrusal bir ilişkinin varlığı söz konusu olmalıdır.
- ❖ Bağımsız deęişkenler arasında ilişki olmamalıdır.

2.5. Bağımsız Deęişkenler Arasında İlişki Olmamalıdır.

Oluşturulan regresyon modellerinde bağımsız deęişkenler arasında bir ilişki olmaması beklenir. Deęişkenlerden elde edilen ikili korelasyon deęerleri sıfır ise deęişkenlerin ortogonal olduğu söylenebilir. Ancak deęişkenler ortogonal yani dik deęil ise aralarında ilişkinin varlığı söz konusu olur. Bağımsız deęişkenlerin arasında bir ilişkinin var olması çoklu bağlantı sorunu veya deęişkenlerin ilişkili olması sorunu olarak nitelendirilebilir.

2.5.1. Çoklu Bağlantı Sorununun Meydana getirdiği Etkiler

Oluşturulan modellerde çoklu bağlantı sorunu varlığının etkileri aşağıdaki gibidir.

- ❖ En küçük kareler yöntemi ile elde edilen katsayıların varyans ve kovaryans değerleri olması gerekenden büyük çıkacaktır.
- ❖ Regresyon katsayıları gerçek katsayılarından farklı olacağı için bağımlı değişken kestirimlerinin standart hatası büyür.
- ❖ H_0 hipotezinin reddedilmesi zorlaşır.
- ❖ Z ve t testi istatistiklerinin sonuçlarının yanlış olmasına sebep olur.

2.5.2. Çoklu Bağlantı Sorununun Belirlenmesi

Regresyon analizinde değerlendirilmesi gereken en önemli noktalardan biri verinin çoklu bağlantı sorununa sahip olup olmadığıdır. Bu sebeple çoklu bağlantı sorununun varlığının ve derecesinin belirlenmesi amacıyla kullanılan en popüler yöntemler aşağıda listelenmiştir.

- ❖ Çoklu bağlantı sorununun $X'X$ korelasyon matrisiyle belirlenmesi
- ❖ Çoklu bağlantı sorununun açıklayıcılık katsayısı ile incelenmesi
- ❖ Çoklu bağlantı sorununun kısmi korelasyon katsayıları ile incelenmesi
- ❖ Çoklu bağlantı sorununun tolerans değerleri ile incelenmesi
- ❖ Çoklu bağlantı sorununun $X'X$ matrisinin özdeğerleri ile belirlenmesi
- ❖ Çoklu bağlantı sorununun korelasyon matrisinin determinant değeri ile incelenmesi
- ❖ Çoklu bağlantı sorununun VIF değerleri ile incelenmesi

2.5.3. Çoklu Bağlantı Sorununun VIF Değerleri İle İncelenmesi

Glauber ve Farrar 1967 yılında ilk kez çoklu bağlantı sorununun varlığını incelemek amacıyla bu değerleri kullanmıştır. Ancak terimsel olarak VIF adı Marquardt tarafından 1970 yılında isimlendirilmiştir.

Bağımsız değişkenlerin oluşturduğu korelasyon matrisinin tersi alındıktan sonra köşegen üzerindeki elemanların varyans şişme faktörleri VIF değerleri olarak nitelendirilir.

Webster' e göre $VIF \geq 10$ olduğu durumlarda çoklu bağlantı sorununun varlığından söz edilir.

2.6. Temel Bileşenler Analizi ve Regresyon Analizi

Çok değişkenli bir veri setinde bağımlı ve bağımsız değişkenlerin varlığının söz konusu olduğu durumlarda regresyon analizine tekniklerine ihtiyaç duyulabilir. Bağımlı ve bağımsız değişkenler arasında ilişki ise çoklu bağlantı sorununu oluşturduğu için yapılan regresyon analizi sonuçlarının anlamlılığını zedeler. Bu gibi durumlarda regresyon analizinin değişkenlerin ilişkisiz olma varsayımını sağlamasına ihtiyaç duyulur. Çoklu bağlantı sorununu gidermek için ise bağımsız değişkenlere temel bileşenler analizi uygulanır. Böylelikle bağımsız değişkenler ortogonal yani birbirine dik hale getirilir. Boyut azaltma yapılmaksızın değişken sayısını değiştirmeden ilişkisiz bileşenler oluşturulur. Temel bileşenler analizinin kullanım amacı bağımsızlaştırma yapmaktır. Bu bileşenlere uygulanan regresyon analizi daha doğru ve anlamlı sonuçlar elde edilmesini sağlar.

2.7. PISA Sınavı

1961 yılında kurulan Ekonomik Kalkınma ve İşbirliği Örgütü (OECD) ekonomik gelişmeler için eğitimin iyi ve sistemli olması gerektiğini ön görmüştür. PISA sınavı, bu çerçevede kapsamında OECD tarafından ilk olarak 1997 yılında geliştirilen bir uygulamadır.

İlk olarak 2000 yılında 32 farklı ülkeden toplam 265.000 öğrenciye uygulanmıştır. Türkiye ise o yıl ilk defa yapılan sınavlara dâhil olmamıştır.

PISA sınavı (Uluslararası Öğrenci Değerlendirme Programı) uluslararası boyutta 2000 yılından itibaren üçer yıl arayla yapılmaktadır. Zorunlu eğitimi bitiren ve bitirme çağındaki 15 yaş grubu öğrencilerin okuldaki öğrenimlerini günlük hayatta uygulayabilme becerisini ölçmeyi hedefleyen bir tarama çalışmasıdır.

2000 yılından itibaren üçer yıl arayla yapılan bu sınav, öğrencileri 3 temel alanda değerlendirmeye almaktadır. 15 yaş grubundaki öğrenciler; Matematik okuryazarlığı, Fen Bilimleri okuryazarlığı ve Okuma Becelerileri olmak üzere üç alanda puansal olarak değerlendirilmişlerdir. Fakat aynı süreç içerisinde yapılan anketlerden elde edilen veriler bizlere puanların ötesinde bilgiler sunmaktadır. Elde edilen veriler; gençlerimizin istek ve beklentilerini, motivasyonlarını, görüşlerini, öğrenme şekillerini ve buldukları maddi ve manevi çevresel faktörlerini daha anlamlı ve faydalı bir şekilde değerlendirmemize olanak sağlamıştır.

Üç yılda bir yapılan bu sınavda dokuzar yıllık periyotlar söz konusudur. Yani o yıl yapılan sınav belirlenmiş bir alana yoğunluk verilir. 2000 yılında yapılan sınav okuma becerilerine, 2003 yılında yapılan sınav matematik okuryazarlığına, 2006 yılında yapılan sınav ise fen bilgisi okuryazarlığına ağırlık vermiştir. 2009 yılından itibaren tekrar okuma becerilerine ağırlık vermesiyle yeni bir periyota başlamış ve 2009, 2012, 2015 yıllarındaki sınavların yapılması ile ikinci periyotunu tamamlamış ve 2018 yılı itibari ile yeni bir periyota başlamıştır.

PISA sınavında katılımcı ülkelerin okul sistemleri bölgelere ve okul türlerine göre sınıflandırılmakta ve daha sonra okullar ve öğrenciler tabakalı seçkisiz (tesadüfi) örnekleme yoluyla belirlenmektedir. Öğrenciler bilgisayar tabanlı sınav uygulamasına katılırlar. Sınavda açık-kapalı uçlu, çoktan seçmeli ve karmaşık çoktan seçmeli gibi soru tipleri kullanılmaktadır. Değerlendirme uygulamasının ardından öğrenciler anket uygulamasına katılarak süreci sonlandırılırlar.

PISA sınavı uygulamasından elde edilen veriler ulusal bir rapor haline getirilmektedir. Bu rapor eğitim sistemlerinin karşılaştırılması ve geliştirilmesi kapsamında tüm ülkeler açısından oldukça faydalı bir veri olmaktadır.

Bu araştırmada ise 2018 yılında yapılan uygulama verileri kullanılmıştır. Sınava 79 ülke ve ekonomiden 600 öğrenci katılmıştır. Türkiye ise 186 okul ve 6890 öğrenci ile uygulamaya dahil olmuştur.

Araştırmaya konu olan 2018 PISA verileri <http://www.pisa.oecd.org> adresinden erişime açılmaktadır. Ülkemize ait uygulama sonuç ve verileri ise <http://pisa.meb.gov.tr> adresinde Raporlar başlığında yayınlanmıştır.

2.8. R Programı

Gelişen teknolojinin de etkisiyle veri bilimi günümüzde oldukça popülerleşmiş ve yığılarak büyüyen bilgiyi incelemek için çeşitli alternatifler karşımıza çıkmaya başlamıştır. Yapılması gereken istatistiksel hesaplamaları ise uygulama aşamasında akla gelen en güçlü programlama dili R'dır.

R, en basit tanımıyla istatistiksel analiz uygulamalarında sıkça kullanılan açık kaynak kodlu bir programlama dilidir. Tarihsel gelişimi ise 1900'lü yıllara uzanmaktadır. Yeni Zelanda, Auckland Üniversitesi İstatistik Bölümü'nden Robert Gentleman ve Ross Ihaka grafik gösterimleri, veri işleme ve istatistik hesaplamaları için adını geliştiricilerinin baş harflerinden alan R programlama dilini tasarlanmıştır. Böylelikle 1993 yılında istatistik ve veri bilimi dünyasına giriş yapmıştır. 1997 yılında "R Core Team" adını alan ve bünyesinde farklı ülkelerden araştırmacıların olduğu bir grup tarafından geliştirilmiştir. Araştırmacıların da katkısıyla bir metin editöründen çalışma ortamına dönüşmüş ve günümüzde iki milyondan fazla kullanıcı potansiyeline ulaşmıştır.

R programlama dilinin en bariz avantajı ücretsiz hizmet sunmasıdır. Çapraz bir platformdur ve hemen hemen bütün işletim sistemlerinde çalıştığı için kullanım kolaylığı sağlamaktadır.

R programı Google, Microsoft, Twitter, Mozilla, Uber, HP ve benzeri birçok tanınmış firma tarafından aktif olarak kullanılmaktadır.

R programının avantajları şunlardır:

- ❖ Büyük verileri işleme ve saklama için çeşitli fonksiyonlara sahiptir.

- ❖ Ücretsizdir.
- ❖ Tablolar üzerinde işlem yapma olanağı sağlar.
- ❖ Veri analizleri için sürekli geliştirilen paketlere sahiptir.
- ❖ Hemen hemen bütün işletim sistemlerinde çalışır.
- ❖ Diğer birçok programla kıyaslandığında görselleştirmeleri ileri düzeydedir.
- ❖ Basit ve etkili bir programlama diline sahiptir.

BÖLÜM III

YÖNTEM

Bu bölümde araştırmanın modeli, araştırma materyali, veri çözümleme yöntemi veri toplama aracı ve veri işleyiş sürecine ilişkin bilgiler verilmiştir.

3.1. Araştırma Modeli

Bu araştırma PISA 2018 puanlarını ve değişkenlerini R programı üzerinde Temel Bileşenler Analizi ve Regresyon Analizi ile incelerken sayısal veriler üzerinden yorumlama ve uygulama yapmaktadır. Aynı zamanda birden fazla değişken arasındaki ilişkinin varlığını ve gücünü incelemektedir. Dolayısıyla bu araştırma bağıntısal nicel bir yaklaşımı benimsemiştir.

3.2. Araştırmanın Materyali

Araştırma materyalini PISA 2018 uygulamasında elde edilen punlar ve değişkenlerin o yılları kapsayan sayısal verileri oluşturmaktadır.

PISA 2018 sınavına 79 ülke ve ekonomiden toplam 600.000 öğrenci katılmıştır. Bu öğrenciler 15 yaş grubu 32 milyon öğrenciyi temsil etmektedir. Bu tez çalışmasında ise 29 ülke ve bu ülkelere ait 3 bağımlı değişken, 17 bağımsız değişken kullanılmıştır.

3.3. Veri Çözümleme Yöntemi

Bu çalışmada alt problemler de göz önünde bulundurularak öncelikle analizler için gerekli olan varsayımlar test edilmiştir. Daha sonra PISA 2018 verileri kullanılarak R programı (versiyon: 4.0.5) üzerinde Temel Bileşenler Analizi ve Regresyon analizi uygulanmıştır.

3.4. Veri Toplama Aracı

Araştırmada kullanılan veriler OECD ve World Bank tarafından oluşturulan veri tabanlarından yararlanılarak elde edilmiştir. Ayrıca yine kullanılan değişkenler ve sayısal veriler bu kuruluşların elektronik ortamda sunduğu bilgilerden yararlanılarak oluşturulmuştur. Dolayısıyla bu verilerin doğru olduğu varsayılarak analizler yapılmıştır.

3.5. Veri İşleyiş Süreci

Araştırma için kullanılan veriler OECD'nin resmi internet sayfası olan <https://stats.oecd.org/> adresinden elde edilmiştir. Bu veriler excel formatına dönüştürülerek R programı üzerinde incelenmiştir.

BÖLÜM IV

BULGULAR VE YORUMLAR

Tezin bu bölümünde araştırmanın amaçları doğrultusunda, PISA 2018 puanları ve ilgili yıla ait bağımsız değişkenlerin oluşturduğu veri setine R programı üzerinde uygulanan Temel Bileşenler Analizi ve Regresyon Analizine ilişkin analiz sonuçları ve bu sonuçlardan elde edilen çıkarımlar yer almaktadır.

Tez çalışmasında kullanılan veriler <https://stats.oecd.org/> sitesinden elde edilerek excel dosyasına dönüştürülmüştür. Araştırmaya dahil edilen ülkeler aşağıdaki gibi listelenmiştir.

Tablo 4.1. *Tez Çalışmasına Dahil Edilen Ülkeler*

1) ABD	11) Finlandiya	21) Norveç
2) Almanya	12) Fransa	22) Polonya
3) Avustralya	13) Hollanda	23) Portekiz
4) Avusturya	14) İspanya	24) Rusya
5) Belçika	15) İsrail	25) Slovakya
6) Birleşik Krallık	16) İsveç	26) Slovenya
7) Brezilya	17) Letonya	27) Türkiye
8) Çekya	18) Lüksemburg	28) Yeni Zelanda
9) Danimarka	19) Macaristan	29) Yunanistan
10) Estonya	20) Meksika	

Tablo 4.2. *Tez Çalışmasına Dahil Edilen Bağımlı ve Bağımsız Değişkenler*

Bağımlı Değişkenler	Bağımsız Değişkenler
O: Okuma Puanı F: Fen Okuryazarlığı Puanı M: Matematik Puanı	X1: Nüfus X2: Gayri safi yurt içi hasıla X3: İşsizlik oranı X4: Kadın iş gücüne katılım oranı(15-24 yaş) X5: Erkek iş gücüne katılım oranı(15-24 yaş) X6: Örgün çocuk bakımı ve okul öncesi eğitime kayıt(0-2yaş) X7: Kişi başı oda sayısı X8: Eğitimde geçen yıllar X9: Eğitimde ve istihdamda yer almayan erkekler(15-19yaş) X10: Eğitimde ve istihdamda yer almayan kadınlar(15-19yaş) X12: Eğitimde ve istihdamda yer almayan erkekler(15-29yaş) X11: Eğitimde ve istihdamda yer almayan kadınlar(15-29yaş) X13: Eğitimde ve istihdamda yer almayan erkekler(20-24yaş) X14: Eğitimde ve istihdamda yer almayan kadınlar(20-24yaş) X16: Eğitimde ve istihdamda yer almayan erkekler(25-29yaş) X15: Eğitimde ve istihdamda yer almayan kadınlar(25-29yaş) X17: Kaba evlilik oranı

Tez çalışmasında 29 ülkeye ait 2018 yılı okuma, matematik ve fen okuryazarlığı olmak üzere 3 bağımlı değişken puanı ve ülkelere ait ilgili yıllardaki Tablo 4.2.'de belirtilen 17 bağımsız değişken incelenmiştir.

4.1. Değişkenlere İlişkin Regresyon Analizi Sonuçları

Tez çalışmasının bu bölümünde bağımlı değişkenlere uygulanan regresyon analizine ilişkin parametrelere yer verilmiştir.

Tablo 4.1.1. *Okuma Puanı Değişkenine Ait Regresyon Analizine İlişkin Model Özeti*

MODEL ÖZETİ						
R²	Düzeltilmiş R²	Standart Hatalar	df1	df2	f İstatistiği	p
,815	,530	17,46517	17	11	2,860	,040

Tablo 4.1.2. Okuma Puanı Bağımlı Değişkenine Ait Regresyon Analizi Parametreleri

	β	Standart Hata	t	f	VIF
Sabit	3.172e+02	8.574e+01	3.700	0.0035	
X1	-2.585e-08	6.739e-08	-0.384	0.7086	2.309880
X2	-1.500e-04	2.806e-04	-0.535	0.6035	4.347970
X3	-1.495e+00	2.513e+00	-0.595	0.5641	8.229422
X4	2.755e-01	1.728e+00	0.159	0.8762	56.254889
X5	-3.299e-01	1.732e+00	-0.190	0.8524	39.596205
X6	9.559e-02	5.943e-01	0.161	0.8751	10.964877
X7	2.279e+01	1.665e+01	1.369	0.1983	5.531632
X8	6.921e+00	5.191e+00	1.333	0.2094	5.259813
X9	1.471e+00	8.741e+00	0.168	0.8695	69.971638
X10	-3.658e+00	4.094e+00	-0.894	0.3907	44.208329
X11	1.059e+01	1.579e+01	0.670	0.5164	1273.314983
X12	-6.703e+00	2.570e+01	-0.261	0.7990	810.821626
X13	6.487e+00	7.986e+00	0.812	0.4339	130.157882
X14	-4.478e+00	5.280e+00	-0.848	0.4145	187.139746
X15	-3.062e+00	5.937e+00	-0.516	0.6162	324.202782
X16	-9.787e-01	9.795e+00	-0.100	0.9222	205.390866
X17	3.309e+00	5.172e+00	0.640	0.5354	2.438585

Yukarıdaki Tablo 4.1.1. ve Tablo 4.1.2. incelendiğinde model sabitinin 3.17 olduğu gözlenmektedir. Değişken sayısı adedince elde edilen regresyon katsayısı yani B değeri her

bir deęişkenin okuma puanını hangi ölçüde azaltıp arttırdığı hesaplanmıştır. Buna göre $X_1, X_2, X_3, X_5, X_{10}, X_{12}, X_{14}, X_{15}, X_{16}$ deęişkenlerinin okuma puanını negatif yönde; $X_4, X_6, X_7, X_8, X_9, X_{11}, X_{13}, X_{17}$ deęişkenlerinin ise okuma puanını pozitif yönde etkilediğı söylenebilir.

Düzeltilmiş R^2 deęeri modelin genellenebilirlik durumunu göstermektedir ve bu deęer 0 ile 1 arasında bir deęer alır. Modelin R^2 deęerinin 0.30'dan yüksek olması beklenir. Oluşturulan modelde R^2 deęerinin 0.815 olduğı ve model uyumunun iyi olduğı gözlemlenmektedir.

Sig. deęerinin 0.05'ten küçük olması beklenir ve bu deęer varyansların homojen olup olmadığını göstermektedir.

VIF deęerlerinin ise 10'dan büyük çıkması çoklu bağlantı sorunun varlığını göstermektedir ve bu deęerler temel bileşenler analizi yapılarak deęişkenler arası bağımsızlaştırmanın gerekliliğini kanıtlamıştır.

Tablo 4.1.3. Matematik Puanı Deęişkenine Ait Regresyon Analizine İlişkin Model Özeti

MODEL ÖZETİ						
R^2	Düzeltilmiş R^2	Standart Hatalar	df1	df2	f İstatistiğı	p
,863	,653	17,98	17	11	4,107	,010

Tablo 4.1.4. Matematik Puanı Bağımlı Değişkenine Ait Parametreler

	β	Standart Hata	t	f	VIF
Sabit	3.369e+02	8.828e+01	3.816	0.00286	
X1	-6.249e-08	6.939e-08	-0.901	0.38707	2.309880
X2	-2.014e-04	2.889e-04	-0.697	0.50007	4.347970
X3	-8.417e-01	2.588e+00	-0.325	0.75111	8.229422
X4	-1.915e+00	1.779e+00	-1.077	0.30472	56.254889
X5	2.120e+00	1.783e+00	1.189	0.25964	39.596205
X6	6.151e-01	6.118e-01	1.005	0.33635	10.964877
X7	1.020e+01	1.714e+01	0.595	0.56384	5.531632
X8	6.198e+00	5.345e+00	1.160	0.27075	5.259813
X9	1.713e+00	9.000e+00	0.190	0.85249	69.971638
X10	-6.745e+00	4.215e+00	-1.600	0.13786	44.208329
X11	2.104e+01	1.626e+01	1.294	0.22216	1273.314983
X12	-1.124e+01	2.646e+01	-0.425	0.67922	810.821626
X13	6.679e+00	8.222e+00	0.812	0.43382	130.157882
X14	-9.384e+00	5.437e+00	-1.726	0.11228	187.139746
X15	-6.987e+00	6.113e+00	-1.143	0.27730	324.202782
X16	2.118e+00	1.009e+01	0.210	0.83747	205.390866
X17	4.366e+00	5.325e+00	0.820	0.42969	2.438585

Yukarıdaki Tablo 4.1.3. ve Tablo 4.1.4. incelendiğinde model sabitinin 3.17 olduğu gözlenmektedir. Değişken sayısı adedince elde edilen regresyon katsayısı yani B değeri her bir değişkenin okuma puanını hangi ölçüde azaltıp arttırdığı hesaplanmıştır. Buna göre X₁, X₂,

$X_3, X_4, X_{10}, X_{12}, X_{14}, X_{15}$ deęişkenlerinin matematik puanını negatif yönde; $X_5, X_6, X_7, X_8, X_9, X_{11}, X_{13}, X_{16}, X_{17}$ deęişkenlerinin ise matematik puanını pozitif yönde etkiledięi söylenebilir.

Düzeltilmiş R^2 deęeri modelin genellenebilirlik durumunu göstermektedir ve bu deęer 0 ile 1 arasında bir deęer alır. Modelin R^2 deęerinin 0.30'dan yüksek olması beklenir. Oluşturulan modelde R^2 deęerinin 0.863 olduęu ve model uyumunun iyi olduęu gözlemlenmektedir.

VIF deęerlerinin ise 10'dan büyük çıkması çoklu bağlantı sorunun varlığını göstermektedir ve bu deęerler temel bileşenler analizi yapılarak deęişkenler arası bağımsızlaştırmanın gereklilięini kanıtlamıştır.

Tablo 4.1.5. *Fen Puanı Deęişkenine Ait Regresyon Analizine İlişkin Model Özeti*

MODEL ÖZETİ						
R^2	Düzeltilmiş R^2	Standart Hatalar	df1	df2	f İstatistięi	p
,838	,589	17,46	17	11	3,365	,022

Tablo 4.1.6. Fen Puanı Bağımlı Değişkenine Ait Parametreler

	β	Standart Hata	t	f	VIF
Sabit	3.182e+02	8.572e+01	3.713	0.00342	
X1	-2.417e-08	6.738e-08	-0.359	0.72654	2.309880
X2	-2.628e-04	2.805e-04	-0.937	0.36890	4.347970
X3	-4.787e-03	2.513e+00	-0.002	0.99851	8.229422
X4	-7.275e-01	1.727e+00	-0.421	0.68174	56.254889
X5	7.749e-01	1.732e+00	0.448	0.66319	39.596205
X6	2.190e-01	5.941e-01	0.369	0.71936	10.964877
X7	2.664e+01	1.664e+01	1.600	0.13781	5.531632
X8	7.090e+00	5.190e+00	1.366	0.19918	5.259813
X9	3.952e+00	8.739e+00	0.452	0.65994	69.971638
X10	-4.684e+00	4.093e+00	-1.145	0.27667	44.208329
X11	1.402e+01	1.579e+01	0.888	0.39358	1273.314983
X12	-1.614e+01	2.569e+01	-0.628	0.54261	810.821626
X13	8.953e+00	7.984e+00	1.121	0.28600	130.157882
X14	-6.029e+00	5.279e+00	-1.142	0.27770	187.139746
X15	-4.547e+00	5.936e+00	-0.766	0.45976	324.202782
X16	1.871e+00	9.793e+00	0.191	0.85197	205.390866
X17	3.569e+00	5.170e+00	0.690	0.50427	2.438585

Yukarıdaki Tablo 4.1.5. ve Tablo 4.1.6. incelendiğinde model sabitinin 3.17 olduğu gözlenmektedir. Değişken sayısı adedince elde edilen regresyon katsayısı yani B değeri her bir değişkenin okuma puanını hangi ölçüde azaltıp arttırdığı hesaplanmıştır. Buna göre X_1 , X_2 , X_3 , X_4 , X_{10} , X_{12} , X_{14} , X_{15} değişkenlerinin fen okuryazarlığı puanını negatif yönde; X_5 , X_6 , X_7 , X_8 , X_9 , X_{11} , X_{13} , X_{16} , X_{17} değişkenlerinin ise fen okuryazarlığı puanını pozitif yönde etkilediği söylenebilir.

Düzeltilmiş R^2 değeri modelin genellenebilirlik durumunu göstermektedir ve bu değer 0 ile 1 arasında bir değer alır. Modelin R^2 değerinin 0.30'dan yüksek olması beklenir. Oluşturulan modelde R^2 değerinin 0.838 olduğu ve model uyumunun iyi olduğu gözlemlenmektedir.

VIF değerlerinin ise 10'dan büyük çıkması çoklu bağlantı sorunun varlığını göstermektedir ve bu değerler temel bileşenler analizi yapılarak değişkenler arası bağımsızlaştırmanın gerekliliğini kanıtlamıştır.

4.2. Temel Bileşenler Analizine İlişkin Sonuçlar

Tez çalışmasının bu bölümünde verinin temel bileşenler analizine uygunluğunu test eden Kaiser-Meyer-Olkin testi ve bağımsız değişkenlerin aralarındaki ilişki yapısını yok etmek amacıyla uygulanan temel bileşenler analizine ilişkin parametrelere yer verilmiştir.

Bu tez çalışmasında temel bileşenler analizi değişkenler arasındaki bağımlılık yapısını yok etmek için kullanılmıştır. Bu bağımlılık yapısını regresyon analizi uygulayarak gözlemleyebiliriz. R programında O, M ve F bağımlı değişkenlerine ait regresyon analizi sonuçları elde edilmiştir. Bu sonuçlar bize değişkenler arasındaki bağımlılık yapısını göstermektedir. Bu tablolara bakarak Temel Bileşenler Analizi uygulayarak değişkenler arasındaki ilişkiyi yok etmenin gerekliliğine karar verilmiştir.

Temel bileşenler analizi için veri excel formatında oluşturulmuştur. Veri bu formatta R programına aktarılmıştır. Öncelikle değişkenler arasındaki ilişkinin önemi ve verinin temel

bileşenler analizine uygun olup olmadığına karar vermek için Kaiser-Meyer-Olkin (KMO) testi uygulanmıştır. Uygulanan teste ilişkin sonuçlar Tablo 4.2.1.' de gösterilmiştir.

Tablo 4.2.1. Kaiser-Meyer-Olkin (KMO) Testi Sonuçları

MSA	X1	X2	X3	X4	X5	X6
0.61	0.60	0.78	0.83	0.60	0.45	0.57
X7	X8	X9	X10	X11	X12	X13
0.76	0.52	0.56	0.73	0.63	0.57	0.57
X14	X15	X16	X17			
0.65	0.60	0.51	0.28			

KMO testi sonuçlarında alt limit 0.50 kabul edilmektedir. Analiz sonuçlarına bakıldığında değerimiz 0.61 çıkmaktadır. Bu sonuç veri setimizin temel bileşenler analizine uygun olduğunu göstermektedir.

Veri setiminin temel bileşenlere uygunluğunun kontrolünün ardından veriye R programı üzerinde değişkenler arasındaki ilişkiyi yok etmek amacıyla temel bileşenler analizi uygulanmıştır. Uygulanan temel bileşenler analizine ilişkin parametreler Tablo 4.2.2. ve Tablo 4.2.3.'te gösterilmiş; elde edilen temel bileşenler ise EK-2: Temel Bileşenler Tablosunda verilmiştir.

Tablo 4.2.2. Temel Bileşenler Analizine İlişkin Parametreler

	BOYUT1	BOYUT2	BOYUT3	BOYUT4	BOYUT5	BOYUT6
Özdeğer	7.14666	3.29241	2.46853	1.20710	0.87446	0.63321
Varyans Yüzdesi	42.03922	19.36717	14.52076	7.1006	5.14390	3.724816
Kümülatif Varyans	42.03922	61.40640	75.92716	83.02778	88.17168	91.89649
	BOYUT7	BOYUT8	BOYUT9	BOYUT 10	BOYUT 11	BOYUT 12
Özdeğer	0.51298	0.27153	0.18687	0.13937	0.11555	0.05845
Varyans Yüzdesi	3.01755	1.59727	1.09925	0.81986	0.67976	0.34384
Kümülatif Varyans	94.91405	96.51132	97.61058	98.43045	99.11021	99.45405
	BOYUT 13	BOYUT 14	BOYUT 15	BOYUT 16	BOYUT 17	
Özdeğer	0.05326	0.02850	0.009595	0.000913	0.000532	
Varyans Yüzdesi	0.31331	0.16768	0.056443	0.005372	0.003130	
Kümülatif Varyans	99.76737	99.93505	99.99150	99.99687	100.0000	

Tablo 4.2.3. Temel Bileşenler Analizi Standartlaştırılmış Özet Parametreleri

	PC1	PC2	PC3	PC4	PC5	PC6
Standart Sapma	2.6733	1.8145	1.5712	1.09868	0.93513	0.79575
Varyans Oranı	0.4204	0.1937	0.1452	0.07101	0.05144	0.03725
Kümülatif Varyans	0.4204	0.6141	0.7593	0.83028	0.88172	0.91896
	PC7	PC8	PC9	PC10	PC11	PC12
Standart Sapma	0.71623	0.52109	0.43229	0.3733	0.3399	0.24177
Varyans Oranı	0.03018	0.01597	0.01099	0.0082	0.0068	0.00344
Kümülatif Varyans	0.94914	0.96511	0.97611	0.9843	0.9911	0.99454
	PC13	PC14	PC15	PC16	PC17	
Standart Sapma	0.23079	0.16884	0.09796	0.03022	0.02307	
Varyans Oranı	0.00313	0.00168	0.00056	0.00005	0.00003	
Kümülatif Varyans	0.99767	0.99935	0.99991	0.99997	1.00000	

Tablo 4.2.2. ve Tablo 4.2.3. incelendiğinde varyans oranı her bir bileşenin tek başına toplam varyansın yüzde kaçını karşıladığını gözlemlenebilir. Kümülatif varyans ise bileşenlerin birlikte açıkladığı toplam varyans yüzdesini göstermektedir. PC1 tek başına toplam varyansın %42'sini açıklamaktadır. PC2 tek başına toplam varyansın %19'unu açıklarken PC1 ve PC2 birlikte toplam varyansın %61'ini açıklamaktadır.

4.3. Temel Bileşenlere İlişkin Regresyon Analizi Sonuçları

Tezin bu bölümünde temel bileşenler analizi uygulanarak bağımsızlaştırılmış değişkenlere uygulanan regresyon analizine ilişkin parametreler verilmiş ve yorumlanmıştır.

Tablo 4.3.1. *Okuma Puanı Bağımlı Değişkenine Ait Temel Bileşenler Regresyonu Parametreleri*

	β	Standart Hata	t	f	VIF
Sabit	485.4483	3.2432	149.682	< 2e-16	
X1	6.5563	1.2346	5.310	0.000248	1
X2	2.0828	1.8190	1.145	0.276502	1
X3	-0.2944	2.1008	-0.140	0.891096	1
X4	-6.4883	3.0041	-2.160	0.053726	1
X5	-2.7090	3.5296	-0.768	0.458927	1
X6	0.7433	4.1478	0.179	0.861036	1
X7	4.7844	4.6083	1.038	0.321455	1
X8	16.6396	6.3340	2.627	0.023532	1
X9	-9.5880	7.6352	-1.256	0.235218	1
X10	-2.8648	8.8409	-0.324	0.751989	1
X11	-6.0356	9.7094	-0.622	0.546853	1
X12	7.2831	13.6518	0.533	0.604308	1
X13	23.2408	14.3013	1.625	0.132427	1
X14	-11.6295	19.5490	-0.595	0.563953	1
X15	1.6058	33.6948	0.048	0.962843	1
X16	-6.0685	109.2194	-0.056	0.956687	1
X17	98.0460	143.0635	0.685	0.507320	1

Tablo 4.3.2. Matematik Puanı Bağımlı Değişkenine Ait Temel Bileşenler Regresyonu Parametreleri

	β	Standart Hata	t	f	VIF
Sabit	486.8966	3.3392	145.811	2e-16	
X1	8.7843	1.2712	6.910	2.55e-05	1
X2	0.6912	1.8729	0.369	0.7191	1
X3	-5.0566	2.1629	-2.338	0.0393	1
X4	-6.7697	3.0931	-2.189	0.0511	1
X5	0.1843	3.6341	0.051	0.9605	1
X6	3.0925	4.2706	0.724	0.4841	1
X7	3.7616	4.7447	0.793	0.4447	1
X8	13.5193	6.5216	2.073	0.0624	1
X9	-6.0447	7.8612	-0.769	0.4581	1
X10	4.8260	9.1027	0.530	0.6065	1
X11	0.7186	9.9968	0.072	0.9440	1
X12	-4.6537	14.0560	-0.331	0.7468	1
X13	20.9435	14.7247	1.422	0.1827	1
X14	-4.6875	20.1278	-0.233	0.8201	1
X15	-39.9650	34.6924	-1.152	0.2738	1
X16	-22.7062	112.4531	-0.202	8437	1
X17	196.9086	147.2991	1.337	0.2083	1

Tablo 4.3.3. Fen Puanı Bağımlı Değişkenine Ait Temel Bileşenler Regresyonu Parametreleri

	β	Standart Hata	t	f	VIF
Sabit	486.9655	3.2425	150.183	2e-16	
X1	7.2739	1.2344	5.893	0.000104	1
X2	1.0593	1.8186	0.582	0.572008	1
X3	-1.2558	2.1003	-0.598	0.562008	1
X4	-6.3381	3.0035	-2.110	0.058551	1
X5	-1.0592	3.5288	-0.300	0.769644	1
X6	0.3116	4.1469	0.075	0.941447	1
X7	6.1354	4.6073	1.332	0.209906	1
X8	19.9584	6.3326	3.152	0.009214	1
X9	-8.0664	7.6335	-1.057	0.313300	1
X10	-1.1084	8.8390	-0.125	0.902472	1
X11	-1.1692	9.7072	-0.120	0.906302	1
X12	9.2840	13.6488	0.680	0.510434	1
X13	22.5361	14.2981	1.576	0.143292	1
X14	-6.9119	19.5446	-0.354	0.730288	1
X15	-18.0237	33.6873	-0.535	0.603276	1
X16	18.9216	109.1952	0.173	0.865577	1
X17	142.8342	143.0317	0.999	0.339440	1

BÖLÜM V

SONUÇ VE ÖNERİLER

5.1. Sonuç ve Öneriler

Bu tez çalışmasında, öncelikle çok değişkenli büyük veri setlerinin açık kaynak kodlu R programına nasıl aktarılacağı ve analizlerin nasıl yapılacağı sorularına cevap aranmıştır. Ayrıca ilgili kısımlarda temel bileşenler analizi, regresyon analizi ve istatistiksel terimlere yönelik teorik açıklamalar yapılmıştır. 2018 PISA okuma, matematik ve fen okuryazarlığı bağımlı değişken puanları ve ilgili yıla ait 17 adet bağımsız değişkene ait veriler kullanılarak bir uygulama yapılmıştır.

İlk olarak değişkenlere yapılan regresyon analizlerinde değişkenler arasında çoklu bağlantı sorununun varlığı ispatlanmıştır. VIF değerlerinin 10'un üzerinde çıkması çoklu bağlantı probleminin varlığını göstermektedir. Çoklu regresyon analizi için değişkenlerin ilişkisiz olma varsayımı sağlanamadığı için değişkenlerin bağımsızlaştırılması amacıyla temel bileşenler analizinin uygulanmasının gerekliliği belirtilmiştir. Verinin temel bileşenler analizine uygunluğunu test etmek amacıyla KMO testi uygulanmış ve veri temel bileşenler analizi yöntemine uygun bulunmuştur. Temel bileşenler analizi bu uygulamada boyut indirgeme amacıyla değil; değişkenler arasındaki ilişki yapısını yok etmek amacıyla uygulanmıştır. Temel bileşenler analizi uygulanırken korelasyon matrisinden yararlanılmıştır.

Temel bileşenler analizi uygulaması ile değişkenler arasındaki ilişki yok edilmiştir. Elde edilen yeni VIF değerlerinin 1 olması değişkenlerin bağımsızlaştırıldığını ve çoklu bağlantı sorununun giderildiğini göstermektedir.

Tablo 4.3.1. incelendiğinde okuma puanını pozitif yönde etkileyen değişkenler çoktan aza doğru sırasıyla X_{17} , X_{13} , X_8 , X_{12} , X_1 , X_7 , X_2 , X_{15} ve X_6 dir. Okuma puanını negatif yönde etkileyen değişkenlerin çoktan aza doğru sırasıyla X_3 , X_5 , X_{10} , X_{11} , X_{16} , X_4 , X_9 ve X_{14} dir.

Tablo 4.3.2. incelendiğinde matematik puanını pozitif yönde etkileyen değişkenler çoktan aza doğru sırasıyla X_{17} , X_{13} , X_8 , X_1 , X_{10} , X_7 , X_6 , X_{11} , X_2 ve X_5 tir. Matematik puanını

negatif yönde etkileyen değişkenler çoktan aza doğru sırasıyla X_{12} , X_{14} , X_3 , X_9 , X_{12} , X_{16} ve X_{15} dir.

Tablo 4.3.3. incelendiğinde fen okuryazarlığı puanını pozitif yönde etkileyen değişkenler çoktan aza doğru sırasıyla X_{17} , X_{13} , X_8 , X_{16} , X_{12} , X_1 , X_7 , X_2 ve X_6 dir. Fen okuryazarlığı puanını negatif yönde etkileyen değişkenler çoktan aza doğru sırasıyla X_5 , X_{10} , X_{11} , X_3 , X_4 , X_{14} , X_9 , ve X_{15} dir.

Genel olarak bakıldığında okuma, matematik ve fen okuryazarlığı puanlarını etkileme oranı farklı olsa da X_{17} , X_{13} , X_8 bağımlı değişkenleri pozitif yönde en çok etkileyen bağımsız değişkenlerdir. Negatif yöndeki etkilerde benzer yapıya rastlanmamıştır. Tablo 4.3.2., Tablo 4.3.2. ve Tablo 4.3.3.' e bakıldığında PC1, PC4 ve PC8'in modele anlamlı katkıları olduğunu söylemek mümkündür.

Bilginin kaotik olarak biriktiği günümüzde ülke ve toplumların eğitim başarısı ve bu başarıyı etkileyen faktörleri anlamlandırma çabasının giderek arttığı bilinmektedir. Son dönemlerde ulusal düzeyde yapılan sınavlar ise bu çabayı destekler niteliktedir. Eğitim durumları ve eğitim başarısı Eğitimde Ölçme ve Değerlendirme alanının odak noktasıdır.

Gelişen teknolojinin de etkisiyle bilgiye ulaşmak kolaylaşmış fakat bilginin işleme süreci zorlaşmıştır. Bu problem ilgili alanda yeni alternatif analiz programları arayışını beraberinde getirmiştir. R programı ilgili alan için büyük boyutlu verileri depolama, analiz ve görselleştirme aşamalarında diğer geleneksel yöntemlere kıyasla birçok avantaj sağlamaktadır. Düşünüldüğünün aksine R proramında analiz yapmak hızlı ve kolaydır. Elde edilen sonuçlar ise diğer analiz programları gibi güvenilirdir. Bu nedenle açık kaynak kodlu R programının kullanılması özellikle Eğitimde Ölçme ve Değerlendirme alanındaki araştırmacılara tavsiye edilir.

KAYNAKÇA

- Ürük, Ersan. (2007). *İstatistiksel uygulamalarda lojistik regresyon analizi*. (Yüksek Lisans Tezi). Marmara Üniveristesi, İstanbul.
- Alkan, Ömer. (2008). *Temel bileşenler analizi ve bir uygulama örneği*. (Yüksek Lisans Tezi). Atatürk Üniversitesi, Erzurum.
- Çatalbaş, Mehmet. (2014). *Temel bileşenler analizi ve kanonik korelasyon analizi ile imge tanıma ve sınıflandırma*. (Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Yazar, Işıl (2008). *Temel bileşen analizi ve bağımsız bileşen analizi yöntemlerini temel alan bazı görüntü tanıma uygulamaları ve karşılaştırmaları*. (Yüksek Lisans Tezi). Eskişehir Osmangazi Üniversitesi, Eskişehir.
- Sinan, Alper. (2003). *Lineer regresyonda ridge yahren edicileri ve bir uygulama*. (Yüksek Lisans Tezi). Selçuk Üniversitesi, Konya
- Özmuşul, Mustafa. (2003). *PISA 2009 verilerine göre Türkiye’de liselerin okul politika ve uygulamaları ile bunların öğrenme çıktıları üzerindeki etkileri*. (Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Tunç, Zeynep.(2018). *En küçük kareler ve temel bileşenler regresyon Analizlerinin Karşılaştırılması*. (Yüksek Lisans Tezi). İnönü Üniversitesi, Malatya.
- Koçak, İsmail. (1998). *Temel bileşenler analizi ve bir uygulama*. (Yüksek Lisans Tezi). İnönü Üniversitesi, Malatya.
- Şimşek, Sinem. (2010). *Matris denklemlerinin singüler değer ayrışımı ile yaklaşık simetrik çözümleri*. (Yüksek Lisans Tezi). Sakarya Üniversitesi, Sakarya.
- Zhaoyang, Li. (2006). *Matris ayrışımı*. (Yüksek Lisans Tezi). İstanbul Üniversitesi, İstanbul.
- Tıgılı, Nur E. (2020). *Doğrusal olmayan yapısal eşitlik modeli yaklaşımları ve R programında bir uygulama*. (Yüksek Lisans Tezi). Düzce Üniversitesi, Düzce.

- Tekin, Gülce G. (2021). *Pisa 2018 sonuçları temelinde ergenlerin iyi oluşunun psikososyal değişkenler açısından yordanması*. (Yüksek Lisans Tezi). Başkent Üniversitesi, Ankara.
- Ceyhan, Eda. (2019). *Pisa 2012 okuma becerileri ölçeğinin uygulama dili doğrultusunda belirlenen ülkeler arasında ölçme değişmezliğinin incelenmesi*. (Yüksek Lisans Tezi). Akdeniz Üniversitesi, Antalya.
- Aksu, Nurşah. (2019). *Farklı ülkelerden pisa sınavına katılan öğrencilerin matematik okuryazarlığını etkileyen faktörlerin tahmin edilmesi*. (Yüksek Lisans Tezi). Aydın Adnan Menderes Üniversitesi, Aydın.
- Kara, Fahriye A. (2019). *Din ve okul başarısı - pisa verileri kapsamında uluslararası bir karşılaştırma*. (Yüksek Lisans Tezi). Çukurova Üniversitesi, Adana.
- Ertürk, Gözde. (2015). *R programı kullanılarak bağımlı ve bağımsız gerçek ve yaratılmış verilerde yarışan risklerin değerlendirilmesi*. (Yüksek Lisans Tezi). Çukurova Üniversitesi, Adana.
- Yüzük, Feyyaz. (2019). *Çoklu regresyon analizi ve yapay sinir ağları ile türkiye enerji talep tahmini*. (Yüksek Lisans Tezi). Sivas Cumhuriyet Üniversitesi, Sivas.
- Ağlarıcı, Ali V. (2017). *Basit ve çoklu regresyon modellerinde parametre tahminlerinin karşılaştırılması*. (Yüksek Lisans Tezi). Süleyman Demirel Üniversitesi, Isparta.
- Süzenler, Olcay. (2006). *Dalgıçların hava tüketimini etkileyen değişkenlerin analizi ve bir çoklu regresyon modeli uygulaması*. (Yüksek Lisans Tezi). Deniz Harp Okulu, İstanbul.

EKLER

Ek-1: Temel Bileşenler Analizi ve Regresyon Analizine İlişkin R Kodları

```
n<-read.table("C:/Users/Casper/Desktop/data15.csv",sep=";",header=TRUE,row.names=1)

n

n1<-n[,-c(2,3)]

n2<-n[,-c(1,3)]

n3<-n[,-c(1,2)]

n1

n2

n3

modn<-lm(O~.,data=n1)

modn

summary(modn)

car::vif(modn)

modn2<-lm(M~.,data=n2)

modn2

summary(modn2)

car::vif(modn2)

modn3<-lm(F~.,data=n3)

modn3

summary(modn3)
```



```
car::vif(modn3)

data<-read.table("C:/Users/Casper/Desktop/data15.csv",sep=";",header=TRUE,row.names=1)

data

dat<-data[,-c(1,2,3)]

corr=cor(dat,method="pearson")

KMO(corr)

scatterplotMatrix(dat,diagonal="histogram")

dim(dat)

mydata<-scale(dat)

mydata

scatterplotMatrix(mydata,diagonal="histogram")

res.cov<-cov(mydata)

round(res.cov,2)

res.cor<-cor(mydata)

round(res.cor,2)

corrplot(res.cor,method="number")

corrplot(res.cor)

res.eig<-eigen(res.cor)

res.eig

res.pca<-prcomp(mydata)

res.pca

plot(res.pca)
```

```
get_eig(res.pca)
summary(res.pca)
res.pca$rotation[,1:17]
newdata<-res.pca$x[,1:17]
newdata
newdata<-as.data.frame(newdata)
attach(newdata)
newdata
data<-
read.table("C:/Users/Casper/Desktop/tba.csv",sep=";",dec=".",header=TRUE,row.names=1)
data
str(data)
mod<-lm(O~.,data=data)
mod
summary(mod)
plot(mod)
car::vif(mod)
mod2<-lm(M~.,data=data)
mod2
summary(mod2)
plot(mod2)
car::vif(mod2)
mod3<-lm(F~.,data=data)
```

```
mod3
```

```
summary(mod3)
```

```
plot(mod3)
```

```
car::vif(mod3)
```

Ek-2: Temel Bileşenler Tablosu

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
ABD	0,3994	-0,0213	2,5992	0,1019	-3,9174	-1,0705	0,4009	0,2608	0,0179
ALMANYA	2,2543	-0,8587	0,3974	-0,4379	-0,6397	-0,3562	0,2884	-0,3463	0,0318
AVUSTRALYA	2,5475	0,9493	2,3178	-1,6747	0,7633	-0,6354	0,1620	0,5081	0,0927
AVUSTURYA	1,1548	-0,1316	0,5795	-0,5642	-0,2737	0,8304	-0,4636	0,0843	0,6721
BELCIKA	0,7948	1,6754	-1,9050	0,3237	-0,0926	-0,9157	0,2698	0,2792	-1,0974
BIRLESİK KRALLIK	0,7111	0,6270	1,9330	1,2447	0,3949	0,6907	0,1379	0,5410	-0,2703
BREZİLYA	-6,5027	1,5881	2,7233	1,2967	0,3025	1,0021	0,8001	-0,5296	0,4809
CEKYA	1,3096	-3,4835	-1,7237	-0,1966	0,1810	-0,5625	0,3580	0,6028	0,0930
DANIMARKA	1,8977	1,8072	0,7780	-1,6546	-0,1481	-0,2511	-0,6485	-1,0832	-0,3337
ESTONYA	-0,1436	-0,9592	0,3996	0,4481	0,0186	1,5046	-0,7758	1,2668	-0,0877
FINLANDIYA	1,2591	1,1470	0,4222	-1,1687	0,6899	-0,3175	0,3837	0,2213	0,3649
FRANSA	-1,4979	2,4530	-0,7795	1,1344	-0,5777	-0,1041	0,2876	-0,0969	-0,5180
HOLLANDA	3,7346	0,3303	1,600	0,3014	1,2766	0,2469	0,4778	-0,5022	0,3588
İSPANYA	-3,0138	3,6613	-1,0688	-0,4578	-0,0563	-0,1309	0,3041	0,7836	0,1816
İSRAIL	-0,1378	-0,0222	0,1856	0,6491	-0,4640	1,6389	-1,5224	-0,7900	-0,7592
İSVEC	2,2745	0,4454	0,7660	-0,5919	0,2673	0,5100	-0,2844	-0,0275	0,3122
LETONYA	0,5135	-1,1902	-1,5415	-1,9537	-0,8613	0,8344	-0,3462	-0,4635	0,0490
LUKSEMBURG	2,8852	0,4409	-1,6533	3,2797	0,1528	-1,3786	-1,3586	-0,1031	0,8151
MACARİSTAN	-0,6136	-2,6611	-1,0357	-0,0641	0,2434	0,0681	-0,1744	0,0011	-0,3552
MEKSIKA	-3,9356	-3,7438	1,5856	1,1346	0,5926	-0,9469	-0,2622	-0,3433	-0,3535
NORVEC	3,2418	1,0473	0,2127	0,4499	0,1929	-0,5900	-0,4508	-0,4354	0,0995
POLONYA	-0,0970	-2,1420	-1,2798	-0,7686	-0,0138	0,0421	0,6898	-0,2079	0,0401

PORTEKIZ	0,1234	1,2369	-1,8845	0,5000	0,2526	0,5869	1,1190	0,0316	-0,3973
RUSYA	-0,6227	-1,8680	-0,9733	0,8886	-0,7449	0,5871	1,5222	-0,3435	0,4254
SLOVAKYA	-1,5319	-1,9662	-1,3506	-0,5428	-0,4250	0,5655	-0,7171	0,4974	0,3548
SLOVENYA	1,5095	-0,2682	-1,5057	0,1990	1,0024	-0,0610	1,1643	-0,5005	-0,0762
TURKIYE	-6,2919	-1,2951	1,5491	-1,2420	1,1592	-1,5170	-0,5760	-0,0485	-0,12948
YENI ZELANDA	2,3649	0,3194	1,6270	0,2570	0,8133	0,0452	-0,0601	0,7809	-0,4959
YUNANISTAN	-4,5875	2,8823	-2,9748	-0,8910	-0,0890	-0,3154	-0,7256	-0,0376	0,4836

Ek-2: Temel Bileşenler Tablosu (Devam)

	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17
ABD	-0,0807	0,1220	0,0487	-0,1042	0,0191	0,0607	0,0006	-0,0038
ALMANYA	-0,1232	0,1630	0,0782	-0,0487	-0,1443	-0,1615	0,0294	0,0070
AVUSTRALYA	-0,4196	-0,2348	-0,1777	0,2229	0,2635	-0,0595	0,0394	-0,0230
AVUSTURYA	1,1069	-0,2566	0,1093	-0,0350	-0,1946	-0,0289	-0,0194	-0,0171
BELCIKA	-0,1240	-0,3240	-0,1900	-0,0685	-0,2887	-0,0945	-0,0466	0,0247
BIRLESİK KRALLIK	0,0527	0,0487	-0,1035	-0,0075	0,0163	0,0486	-0,0418	0,0284
BREZILYA	-0,1711	-0,3222	-0,1795	0,0038	-0,0222	0,0243	-0,0137	0,0173
CEKYA	-0,2901	0,0260	-0,2762	0,1865	-0,1748	0,1393	0,0199	-0,0020
DANIMARKA	0,4079	0,0530	-0,5334	-0,0833	0,1230	0,1163	0,0170	0,0427
ESTONYA	-0,1275	0,2009	-0,1145	0,1121	-0,1889	-0,1136	0,0465	0,0392
FINLANDIYA	0,1114	-0,3996	0,1543	0,2274	-0,1945	0,1207	0,0095	-0,0190
FRANSA	0,0853	0,1514	-0,0251	0,3875	0,1166	-0,1757	0,0214	-0,0394
HOLLANDA	-0,1473	0,7900	-0,0590	0,0164	-0,0948	-0,0212	0,0285	0,0026
ISpanya	0,1378	-0,0703	-0,0570	-0,2209	0,1461	0,0436	0,0140	-0,0119
ISRAİL	-0,4102	-0,1919	0,0795	0,3031	0,0095	0,0923	0,0095	-0,0364
ISVEC	-0,9297	-0,4758	0,0608	-0,5787	-0,1470	-0,0664	-0,0344	-0,0281
LETONYA	-0,0987	0,1416	0,5652	-0,0674	0,1446	-0,0999	-0,0084	0,0295
LUKSEMBURG	-0,1097	-0,3149	0,1458	0,0498	0,2041	0,0109	0,0086	0,0174
MACARİSTAN	-0,1698	-0,2654	-0,0688	-0,3277	0,1784	0,1796	0,0213	0,0039
MEKSIKA	0,3869	0,3744	-0,0491	-0,3791	-0,1514	-0,0290	0,0180	-0,0372

NORVEC	0,3816	-0,0835	0,0198	0,1031	-0,2248	-0,0250	-0,0175	0,0029
POLONYA	0,1319	-0,1958	-0,0796	0,4592	-0,0489	0,0542	-0,0452	-0,0178
PORTEKIZ	0,5172	-0,3514	0,3766	-0,2758	-0,0092	0,0566	0,0708	0,0020
RUSYA	-0,2802	0,0632	0,0626	0,2111	0,0660	0,0326	-0,0091	0,0222
SLOVAKYA	0,4176	-0,0270	-0,4816	-0,0863	0,2840	-0,1310	-0,0250	-0,0132
SLOVENYA	-0,0229	0,1860	-0,1236	-0,1826	0,2308	-0,1210	-0,0270	-0,0072
TURKIYE	-0,0704	-0,2850	0,3848	0,2186	0,0408	-0,0807	-0,0012	0,0312
YENI ZELANDA	0,2101	0,6142	0,4409	-0,0011	0,2319	0,1295	-0,0556	-0,0064
YUNANISTAN	-0,3717	0,8639	-0,0076	-0,0346	-0,1906	0,0986	-0,0095	-0,0082

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Nefin YAŞAR

Eğitim Durumu

Lisans Öğrenimi : Akdeniz Üniversitesi, Eğitim Fakültesi, Sınıf Öğretmenliği Bölümü.

Yüksek Lisans Öğrenimi : Akdeniz Üniversitesi, Eğitim Bilimleri Enstitüsü, Eğitimde Ölçme ve Değerlendirme Tezli Yüksek Programı.

Bildiği Yabancı Diller : İngilizce.

Tarih : 28/06/2022

BİLDİRİM

Hazırladığım tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kâğıt ve elektronik kopyalarının Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

■ Tezimin/Raporumun tamamı her yerden erişime açılabilir.

□ Tezim/Raporum sadece Akdeniz Üniversitesi yerleşkelerinden erişime açılabilir.

□ Tezimin/Raporumun 1 yıl süreyle erişime açılmasını istemiyorum.

Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

28/06/2022

Nefin YAŞAR

İNTİHAL RAPORU

EĞİTİM VERİLERİ ÜZERİNDE TEMEL BİLEŞENLER ANALİZİ VE REGRESYON ANALİZİNİN R PROGRAMI ÜZERİNDE UYGULANMASI VE YORUMLANMASI

ORJİNALLIK RAPORU

% 13	% 12	% 2	% 5
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	acikbilim.yok.gov.tr İnternet Kaynağı	% 3
2	dspace.akdeniz.edu.tr:8080 İnternet Kaynağı	% 2
3	9lib.net İnternet Kaynağı	% 1
4	Submitted to Akdeniz University Öğrenci Ödevi	<% 1
5	ijaedu.ocerintjournals.org İnternet Kaynağı	<% 1
6	tez.sdu.edu.tr İnternet Kaynağı	<% 1
7	ismailari.com İnternet Kaynağı	<% 1
8	openaccess.inonu.edu.tr:8080 İnternet Kaynağı	<% 1

issuu.com