T.C.

HACETTEPE ÜNİVERSİTESİ

FEN FAKÜLTESİ

İSTATİSTİK BÖLÜMÜ

ÇOK DEĞİŞKENLİ İSTATİSTİKSEL YÖNTEMLER

PROJE RAPOR DOSYASI

Kerem DİKBAŞLI

Ocak, 2023

**İçindekiler**

|  |  |  |
| --- | --- | --- |
| [**Giriş**](#page8) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | | **3** |
| **Bölüm 1: Veri Seti Tanımı** . . . . . . . . . . . . . . . . . . . . . . . . . . . | | **5** |
| [1.1](#page10) | [Değişkenlerin Tanımı](#page10) . . . . . . . . . . . . . . . . . . . . . . . . . . . | 5 |
| [1.2](#page11) | [Tanımlayıcı İstatistikler](#page11) . . . . . . . . . . . . . . . . . . . . . . . . . . | 6 |
| [1.3](#page12) | [Korelasyon Katsayısı](#page12) . . . . . . . . . . . . . . . . . . . . . . . . . . . | 7 |
|  | [1.3.1 Normallik Sınaması](#page13) . . . . . . . . . . . . . . . . . . . . . . . . | 8 |
| [**Bölüm 2: Temel Bileşenler Analizi (PCA)**](#page14) . . . . . . . . . . . . . . . . . | | **9** |
| [2.1](#page14) | [KMO ve Bartlett Testi](#page14) . . . . . . . . . . . . . . . . . . . . . . . . . . | 9 |
| [2.2](#page15) | [Anti Image Matris](#page15) . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 10 |
| [2.3](#page15) | [TBA Analizi](#page15) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 10 |
| [2.4](#page16) | [Communalities](#page16) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 11 |
| [2.5](#page17) | [TBA Graph of Individuals](#page17) . . . . . . . . . . . . . . . . . . . . . . . . | 12 |
| [2.6](#page18) | [TBA Graph of Variables (Fviz)](#page18) . . . . . . . . . . . . . . . . . . . . . | 13 |
| [2.7](#page19) | [Temel Bileşenler](#page19) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 14 |
| [2.8](#page20) | [Scree Plot](#page20) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 15 |
| [2.9](#page21) | [Temel Bileşenlerin Matematiksel Oluşumu](#page21) . . . . . . . . . . . . . . . | 16 |
| **Bölüm 3: Faktör Analizi** . . . . . . . . . . . . . . . . . . . . . . . . . . . . | | **19** |
| [3.1](#page25) | [Artıklar Matrisi](#page25) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 20 |
| [3.2](#page26) | [Faktör Matrisi](#page26) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 21 |
|  | [3.2.1 Factor Transformation Matrisi](#page27) . . . . . . . . . . . . . . . . . . | 22 |
| **Bölüm 4: Diskriminant Analizi** . . . . . . . . . . . . . . . . . . . . . . . . | | **23** |
|  | [4.0.1 Başlarken](#page28) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 23 |
| [4.1](#page29) | [Group Statistic Tablosu](#page29) . . . . . . . . . . . . . . . . . . . . . . . . . | 24 |
| [4.2](#page30) | [Test of Equaltiy of Group Means Tablosu](#page30) . . . . . . . . . . . . . . . . | 25 |
| [4.3](#page31) | [Kovaryans Matrisleri](#page31) . . . . . . . . . . . . . . . . . . . . . . . . . . . | 26 |
| [4.4](#page32) | [Box’s Test of Equality of Covariance Matrices](#page32) . . . . . . . . . . . . . | 27 |
| [4.5](#page33) | [Kanonik Diskriminant Fonksiyonlarının Özeti](#page33) . . . . . . . . . . . . . | 28 |
| [4.6](#page34) | [Structure Matrix ve Coefficentlar](#page34) . . . . . . . . . . . . . . . . . . . . | 29 |
| [4.7](#page35) | [Centroids](#page35) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 30 |
| [4.8](#page36) | [Gruplandırma İstatistikleri](#page36) . . . . . . . . . . . . . . . . . . . . . . . . | 31 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 4 |  |  | İçindekiler | |
|  |  |  | |  |
|  | [4.8.1](#page36) | [Gruplar İçin Öncül Olasılıklar ve Gruplandırma Fonksiyon Co-](#page36) | |  |
|  |  | [efficentları](#page36) . . . . . . . . . . . . . . . . . . | ........... | 31 |
|  | [4.8.2](#page37) | [Casewise Statistics](#page37) . . . . . . . . . . . . . | ........... | 32 |
| **Bölüm 5: Kümeleme Analizi** . . . . . . . . . . . . . . | | | ........... | **35** |
| [5.1](#page41) | [Two Step Cluster](#page41) . . . . . . . . . . . . . . . . . . | | ........... | 36 |
|  | [5.1.1](#page41) | [Model Özeti ve Grup Kalitesi](#page41) . . . . . . . | ........... | 36 |
|  | [5.1.2](#page42) | [Grup Boytuları ve Oranları](#page42) . . . . . . . . | ........... | 37 |
|  | [5.1.3](#page43) | [Grup İçin Daha Spesifik Bilgiler](#page43) . . . . . . | ........... | 38 |
| [5.2](#page44) | [Quick Cluster ile Kümeleme Analizi](#page44) . . . . . . . . | | ........... | 39 |
|  | [5.2.1](#page44) | [Initial Cluster Centers](#page44) . . . . . . . . . . . | ........... | 39 |
|  | [5.2.2](#page45) | [Anova](#page45) . . . . . . . . . . . . . . . . . . . . | ........... | 40 |
| [5.3](#page46) | [Hieararchical Cluster](#page46) . . . . . . . . . . . . . . . . | | ........... | 41 |
| [**Bölüm 6: Çok Boyutlu Ölçeklendirme Analizi**](#page48) . . . | | | ........... | **43** |
| [6.1](#page48) | [Proximities - Proxscal](#page48) . . . . . . . . . . . . . . . | | ........... | 43 |
|  | [6.1.1](#page48) | [Stress and Fit Measures](#page48) . . . . . . . . . . | ........... | 43 |
|  | [6.1.2](#page49) | [Decomposition of Normalised Raw Stress](#page49) . | ........... | 44 |
|  | [6.1.3](#page50) | [Object Points](#page50) . . . . . . . . . . . . . . . . | ........... | 45 |
| [**Sonuç**](#page52) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | | | ........... | **47** |
| [**Kaynaklar**](#page54) . . . . . . . . . . . . . . . . . . . . . . . . . . | | | ........... | **49** |

**Şekil Listesi**

|  |  |  |
| --- | --- | --- |
| [1.1](#page11) | [Tanımlayıcı istatistikler](#page11) . . . . . . . . . . . . . . . . . . . . . . . . . . | 6 |
| [1.2](#page12) | [Korelasyon grafiği](#page12) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 7 |
| [1.3](#page13) | [Shapiro Wilk](#page13) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 8 |
| [1.4](#page13) | [Determinant Değeri](#page13) . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 8 |
| [2.1](#page14) | [KMO-Bartlett test](#page14) . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 9 |
| [2.2](#page15) | [Anti Image Köşegen Değerleri](#page15) . . . . . . . . . . . . . . . . . . . . . . | 10 |
| [2.3](#page16) | [Anti Image Köşegen Değerleri](#page16) . . . . . . . . . . . . . . . . . . . . . . | 11 |
| [2.4](#page17) | [PCA Graph of Individuals](#page17) . . . . . . . . . . . . . . . . . . . . . . . . | 12 |
| [2.5](#page18) | [Graph of Variables-Fviz](#page18) . . . . . . . . . . . . . . . . . . . . . . . . . | 13 |
| [2.6](#page19) | [Total variance explained](#page19) . . . . . . . . . . . . . . . . . . . . . . . . . | 14 |
| [2.7](#page20) | [Scree Plot](#page20) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 15 |
| [2.8](#page21) | [Component Matrix](#page21) . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 16 |
| [3.1](#page25) | [Reproduced Correlations Matrix](#page25) . . . . . . . . . . . . . . . . . . . . . | 20 |
| [3.2](#page25) | [Reproduced Correlations Matrix - Residuals](#page25) . . . . . . . . . . . . . . | 20 |
| [3.3](#page26) | [Normal faktör matrisi ve varimax rotasyonu yapılmış faktör matrisi](#page26) . | 21 |
| [3.4](#page27) | [Factor transformation matrix](#page27) . . . . . . . . . . . . . . . . . . . . . . | 22 |
| [4.1](#page29) | [Group Statistics Tablosu](#page29) . . . . . . . . . . . . . . . . . . . . . . . . . | 24 |
| [4.2](#page30) | [Test of Equality of group Means](#page30) . . . . . . . . . . . . . . . . . . . . . | 25 |
| [4.3](#page31) | [Test of Equality of group Means](#page31) . . . . . . . . . . . . . . . . . . . . . | 26 |
| [4.4](#page32) | [Log determinant tablosu ve Box’s testinin sonuç tablosu](#page32) . . . . . . . | 27 |
| [4.5](#page33) | [Öz değerler ve Wilk’s Lambda test sonuçu tabloları](#page33) . . . . . . . . . . | 28 |
| [4.6](#page34) | [SCAN ve USCAN tabloları](#page34) . . . . . . . . . . . . . . . . . . . . . . . . | 29 |
| [4.7](#page35) | [Centroids](#page35) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 30 |
| [4.8](#page36) | [Classification Function ve Prior Probabilites Tabloları](#page36) . . . . . . . . . | 31 |
| [4.9](#page37) | [Casewise Statistics](#page37) . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 32 |
| [4.10](#page38) | [Classification Results](#page38) . . . . . . . . . . . . . . . . . . . . . . . . . . . | 33 |
| [5.1](#page41) | [Two step cluster Model Summary](#page41) . . . . . . . . . . . . . . . . . . . . | 36 |
| [5.2](#page42) | [cluster Size](#page42) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 37 |
| [5.3](#page43) | [Cluster dağılımı](#page43) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 38 |
| [5.4](#page44) | [Initial cluster Centers](#page44) . . . . . . . . . . . . . . . . . . . . . . . . . . . | 39 |
| [5.5](#page45) | [Anova](#page45) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 40 |
| [5.6](#page46) | [Proximity Matrix](#page46) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 41 |

|  |  |  |  |
| --- | --- | --- | --- |
| Şekil Listesi | | | 1 |
|  |  |  |  |
| [5.7](#page46) | | [Case - Number of Clusters](#page46) . . . . . . . . . . . . . . . . . . . . . . . . | 41 |
| [5.8](#page47) | | [Case - Number of Clusters](#page47) . . . . . . . . . . . . . . . . . . . . . . . . | 42 |
| [6.1](#page48) | | [Stress and Fit Meausers](#page48) . . . . . . . . . . . . . . . . . . . . . . . . . | 43 |
| [6.2](#page49) | | [Raw Stress ve Final Coordinates tabloları](#page49) . . . . . . . . . . . . . . . | 44 |
| [6.3](#page50) | | [Object Points](#page50) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 45 |

**Giriş**

İstatistiksel ilişkiler bir bağımsız değişkenin bir bağımlı değişkeni etkilemesi temeline dayanmaktadır. Oysa ekonomik olaylardan sağlık alanındaki olaylara kadar değişken veya değişkenleri etkileyen birden çok faktör vardır. [Baydemir (2020)](#page54) Tek değişkenli istatistiksel analizlerin eksikliği ve sınırlı olayları açıklayabilmesi, araştırmalarda çok değişkenli istatistiksel analiz yöntemlerin kullanılmasını gerektirmiştir. Böylece, tek değişkenli istatistiklerde varsayılan kısıtlamalar ortadan kalktığından araştırmalarda daha objektif ve tutarlı sonuçlar elde edilir. [Akkuş et al. (2006)](#page54)

Proje süresince çok değişkenli istatistiksel yöntemler dersinde öğrenilen analiz yöntem-leri kullanılarak karmaşık olan veri seti basitleştirilmiştir. Proje konusuna bağlı olarak problemi etkileyen birçok değişken vardır ve tüm değişkenler dikkate alınarak yapılan analizler sonucunda elde edilen veriler ile uygulama yapılmıştır. Çözümlemeye başla-madan önce değişkenlerin yapılarının açık bir şekilde tanımlanması yöntem seçiminde doğru bir sonuç elde edilmesini sağlayacaktır.

**Bölüm 1**

**Veri Seti Tanımı**

Veri setinde Türkiye’nin illerinden oluşan 81 adet deney birimi ve kadın cinayetlerini etkileyen 16 adet değişken vardır. 81\*15 boyutlu matris elde edilmiştir ve toplamda 1215 adet gözlem bulunmaktadır. Her bir deney birimi tüm 16 değişken tarafından ölçülmüştür.

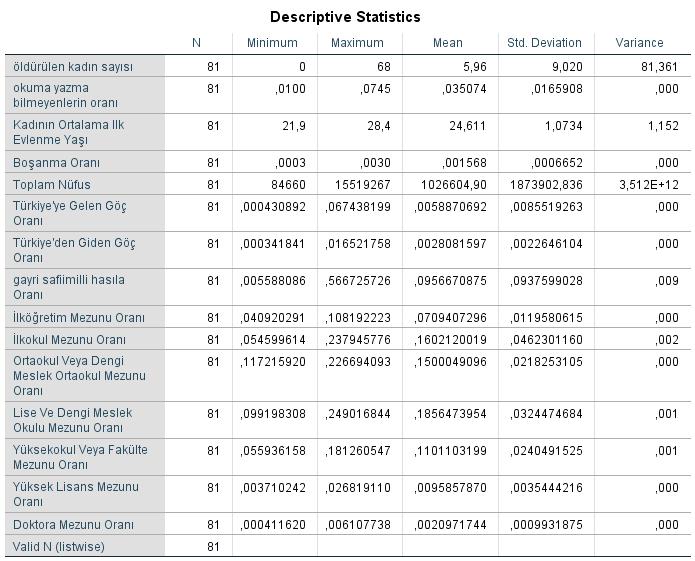
**1.1 Değişkenlerin Tanımı**

* *X*1 : Öldürülen kadın sayısı
* *X*2 : Okuma yazma bilmeyenlerin oranı
* *X*3 : Kadının ortalama ilk evlenme yaşı
* *X*4 : Boşanma oranı
* *X*5 : Toplam nüfus
* *X*6 : Türkiye’ye gelen göç oranı
* *X*7 : Türkiye’den giden göç oranı
* *X*8 : Gayri safii milli hasıla oranı
* *X*9 : İlköğretim mezunu oranı
* *X*10 : İlkokul mezunu oranı
* *X*11 : Ortaokul veya dengi meslek ortaokul mezunu oranı
* *X*12 : Lise ve dengi meslek okul mezunu oranı
* *X*13 : Yüksekokul veya fakülte mezunu oranı
* *X*14 : Yüksek lisans mezunu oranı
* *X*15 : Doktora mezunu oranı

6 Bölüm 1. Veri Seti Tanımı

**1.2 Tanımlayıcı İstatistikler**

Veri setini tanımlamak amacıyla çıkarılan bu tabloda her bir değişkenin "minimum, maksimum değerleri, ortalaması, standart sapması ve varyans değerleri" hesaplanmıştır.



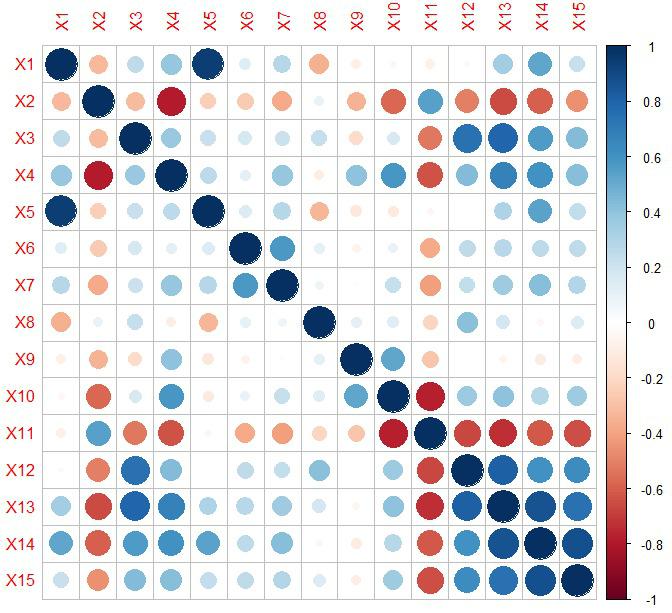
Şekil 1.1: Tanımlayıcı istatistikler

Şekil 1.1 e bakıldığında minimum maksimum değerleri arasındakı farkın az olduğu görülmektedir. Bunun sebebi değişkenlerin toplam nüfusa oranlanmış olmasıdır. Değiş-kenlerin örneklem genişliği beklendiği gibi her değişkende 81’dir ve veri setinde kayıp gözlem bulunmamaktadır. Değişkenlerin standart sapmalarına bakıldığında en küçük sapma 0,00099318 iken en büyük sapma 1873902,836’dır. Bu da analizlerde korelasyon matrisinin kullanılması gerektiği sonucunu vermektedir.

|  |  |  |
| --- | --- | --- |
| 1.3. Korelasyon Katsayısı | | 7 |
|  |  |  |

**1.3 Korelasyon Katsayısı**

R paket programı yardımıyla oluşturulan korelasyon grafiği şekil 1.2 de gözükmektedir.



Şekil 1.2: Korelasyon grafiği

Korelasyon grafiğine bakıldığında sırasıyla şu yorumlar yapılabilir;

*X*1 değişkeni ile *X*5 değişkeni arasında çok güçlü pozitif yönlü bir ilişki vardır.

*X*7 değişkeni ile *X*9 değişkeni arasında herhangi bir ilişki yoktur.

*X*9 değişkeni ile *X*11 değişkeni arasında çok güçlü, negatif yönlü bir ilişki vardır.

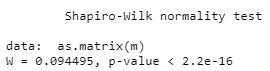
Değişkenler arasındaki ilişki yönleri pozitif ve negatif olarak farklılıklar göstermektedir. Ayrıca bazı değişkenlerin birbiri ile ilişkisinin olmadığını görülmektedir. Determinant değerinin 0’a çok yakın olup 4,920413\*e-08 çıkması seçilen veri setinin çok değişkenli istatistiksel yöntemler ile analiz edilmeye uygun olduğunu göstermektedir.

8 Bölüm 1. Veri Seti Tanımı

**1.3.1 Normallik Sınaması**

*H*0 Veri seti çok değişkenli normal dağılıma sahiptir.

*H*1 veri seti çok değişkenli normal dağılıma sahip değildir.



Şekil 1.3: Shapiro Wilk

* P değeri < 0.05 olduğundan yokluk hipotezi reddedilir ve veri setinin çok değişkenli normal dağılıma sahip olmadığı söylenebilir.



Şekil 1.4: Determinant Değeri

* Determinant değeri "4,301065\*e-20" bulunmuştur. Bu bulunan determinant değerinin 0’dan büyük, 0’a yakın olması beklenmektedir.

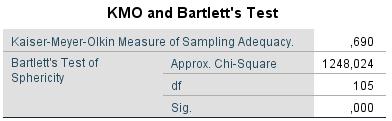
**Bölüm 2**

**Temel Bileşenler Analizi (PCA)**

Türkçesi “temel bileşenler analizi” olan PCA, tanıma, sınıflandırma, görüntü sıkıştırma alanlarında kullanılan, bir değişkenler setinin varyans kovaryans yapısını, bu değişkenle-rin doğrusal birleşimleri vasıtasıyla açıklayarak, boyut indirgenmesi ve yorumlanmasını sağlayan, çok değişkenli bir istatistik yöntemidir. [Yıldız et al. (2010)](#page54)

**2.1 KMO ve Bartlett Testi**

Bartlett testi Temel bileşenler analizi için korelasyon matrisinde değişkenler arasında yeterli korelasyon olup olmadığını test eder. KMO (Kaiser Mayer Olkin) testi ise veri setinin faktör analizi için uygun olup olmadığını test etmek için kullanılır.



Şekil 2.1: KMO-Bartlett test

* *H*0 *p* = 1 **korelasyon matrisi birim matrise eşittir.**
* *H*1 *p* ̸= 1 **korelasyon matrisi birim matrise eşit değildir.**

Bartlett testinin sonucu olarak çıkan sigma değeri 0,5’ten küçük olduğu için (*p <* 0*.*5) korelasyon matrisinin birim matrise eşit olmadığını destekleyecek yeterli kanıt vardır ve bu veri setinde çoklu doğrusal bağlantı sorunu vardır denebilir. Bu yargı hem veri setinin çok değişkenli istatistiksel yöntemlerle analiz edilebileceğini bir kez daha kanıtlarken hem de temel bileşenler analizine uygunluğunu kanıtlamaktadır. Veri setinin faktör analizine uygunluğunun testi için yapılan KMO testi sonucunu 0,690 olarak görünmektedir. bu değer 0,5 in üzerinde olduğu için veri seti faktör analizine uygundur.

10 Bölüm 2. Temel Bileşenler Analizi (PCA)

**2.2 Anti Image Matris**

Ters Görüntü Korelasyon Matrisi (Anti image correlation matrix) Bir matrisin köşegen dışı elemanlarının negatifi alınmış kısmi korelasyon katsayılarından oluşur. Köşegen elemanları ise her değişkenin örneklem uygunluğu değerini verir. Köşegenlerde bulunan değerlerin 0,60 dan yüksek olması beklenir, aksi takdirde örneklem hacmini arttırmak gerekebilir.



Şekil 2.2: Anti Image Köşegen Değerleri

Toplam nüfüs, Türkiye’ye gelen göç oranı, Türkiye’den giden göç oranı, Gayri safi milli hasıla, İlköğretim mezunu oranı değişkenlerinin 0,60 değeri altında olduğu gö-rünmektedir. 0,60 değerinin altında bir çok değişken olduğu için örneklem hacmi arttırılmalıdır.

**2.3 TBA Analizi**

Temel bileşenler analizinin veri setine uygulanabileceğini önceki bölümlerde kanıtla-dıktan sonra, bu bölümde veri setinin öz değerlerinin veri setini açıklama oranlarına bakılacak ve scree plot’ın da yardımıyla temel bileşenler analizine girecek olan değiş-kenlere karar verilecektir.

|  |  |
| --- | --- |
| 2.4. Communalities | 11 |
|  |  |

**2.4 Communalities**

Communalities tablosu her bir faktörün toplam faktörün ne kadarını etkilediğini gösterir. Bu tabloda extraction değerlerinin .30’un üzerinde olması beklenir. Genellikle, 0.30’un altındakiler çıkarılarak işlem tekrar edilir. Bazen .10’a kadar tolere edilebilir. Sorunlu madde var ise, atılıp denenebilir ve atmaya değer mi diye kontrol edilir. Tavsiye edilen, sadece bu tabloya göre hemen madde atmamak ve ilerleyip diğer tablolarda sorun olup olmadığına bakmaktır.Çil [(2014)](#page54)

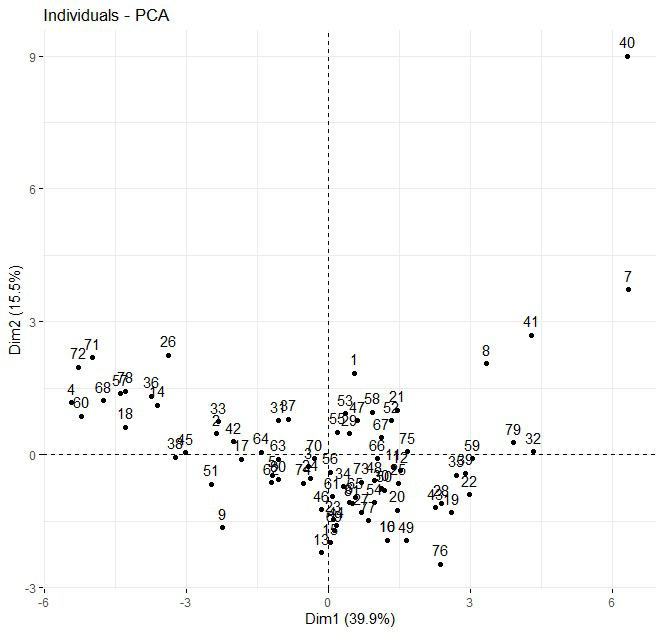


Şekil 2.3: Anti Image Köşegen Değerleri

şekil 2.3’te her bir faktörün toplam faktörü ne kadar etkilediği extraction kısmında yazmaktadır. 0,30’un altında değerimiz olmaması herhangi bir faktörü analizden çıkarmamıza gerek kalmadığını destekler.

12 Bölüm 2. Temel Bileşenler Analizi (PCA)

**2.5 TBA Graph of Individuals**

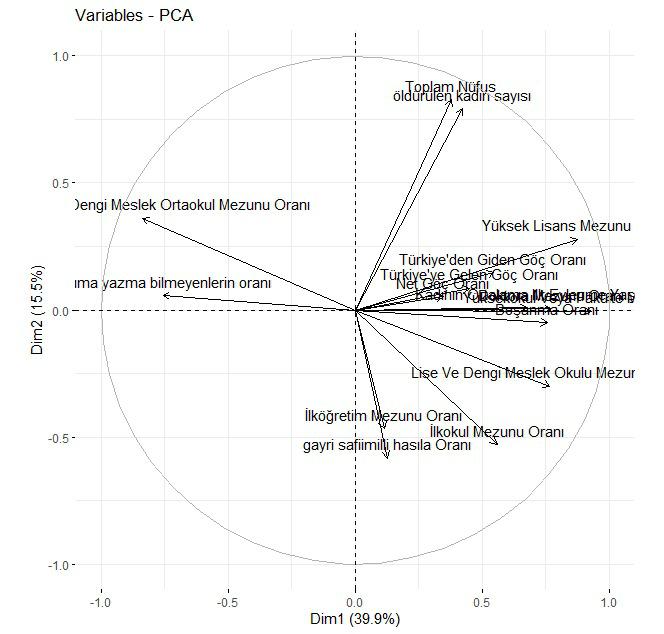


Şekil 2.4: PCA Graph of Individuals

Her bir deney birimin bileşen skorlarının grafiksel gösterimi yukarıdadır. Buradan 40, 7 ve 41 numaralı deney birimlerinin bileşenlerin skorlarına en uzak katkıyı verdiği söylenebilir.

|  |  |
| --- | --- |
| 2.6. TBA Graph of Variables (Fviz) | 13 |
|  |  |

**2.6 TBA Graph of Variables (Fviz)**

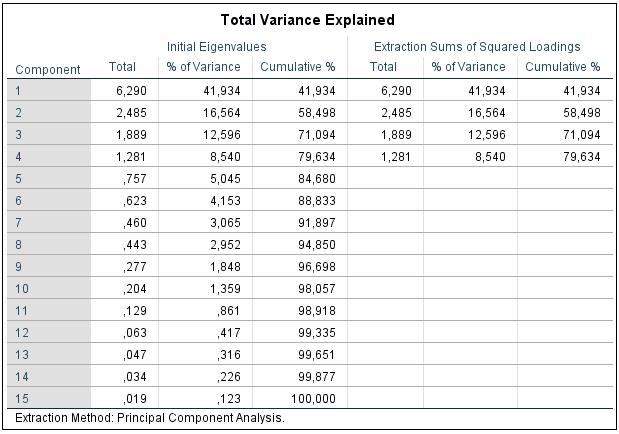


Şekil 2.5: Graph of Variables-Fviz

*X*11 ve *X*2 ’ nİn bileşenlere anlamlı bir katkı vermediği ve Birinci bileşeni ağırlıklı olarak *X*1*, X*5*, X*14*, X*7*, X*6*, X*3*, X*4*, X*15 ’İn oluşturduğu görülmektedir.

14 Bölüm 2. Temel Bileşenler Analizi (PCA)

**2.7 Temel Bileşenler**



Şekil 2.6: Total variance explained

Temel bileşenler analizinin sonucunda şekil 2.4’te "Initial Eigenvalues" kısmında ki "Total" sütununa bakıldığında 1’den büyük değerler faktör analizine girmeye uygun özdeğerlerdir. Bu durumda 4 adet temel bileşen faktör analizine girecektir.

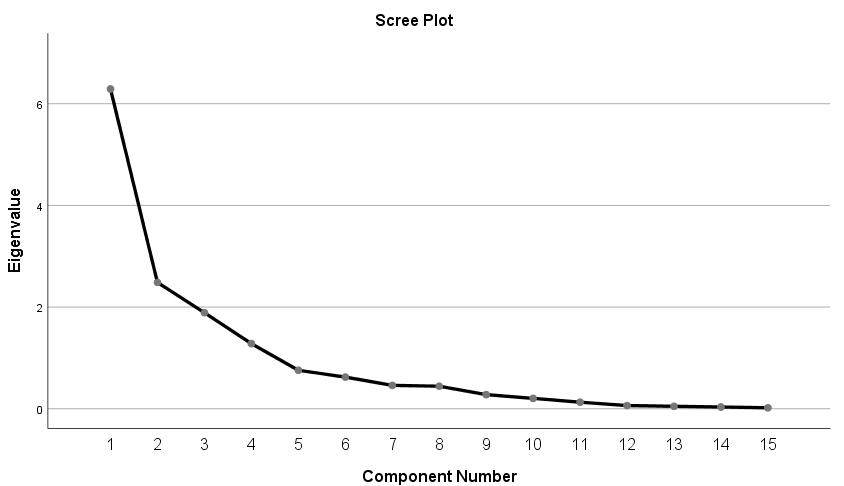
* 1.temel değişkenin toplam değişimin %41,934’ünü,
* 2.temel değişkenin toplam değişimin %16.564’ünü,
* 3.temel değişkenin toplam değişimin %12,596’sını,
* 4.temel değişkenin ise toplam değişimin %8,540’ını açıklar.

Bu bahsi geçen 4 temel değişken toplam değişimin %79,634’ünü açıklar. Toplam açık-lanma oranı %80’e çok yakın olduğu için, alınan temel bileşen sayısının yeterli olduğunu söyleyebiliriz. Bu şaibeyi tam olarak ortadan kaldırmak için scree plot’tan yardım alarak faktör analizine gidecek olan toplam temel bileşen sayısını kesinleştirilebilir.

|  |  |
| --- | --- |
| 2.8. Scree Plot | 15 |
|  |  |

**2.8 Scree Plot**

Scree plot’da bükülmenin olduğu kısma kadar olan değişkenler temel bileşenleri vermektedir.

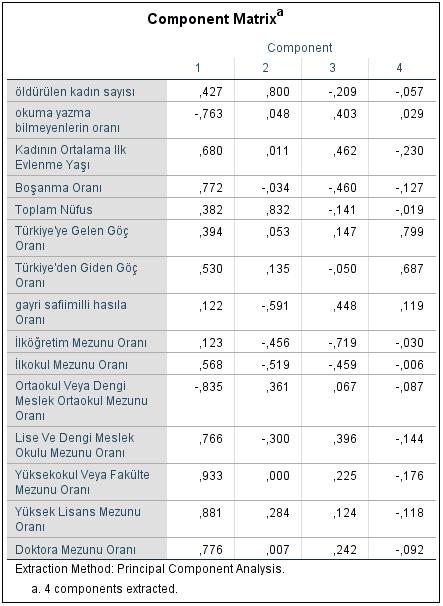


Şekil 2.7: Scree Plot

Şekil 2.5’te görüldüğü üzere 5. bileşene kadar bir bükülme yaşanmış fakat 5. bileşenden sonra bükülme tamamlanıp grafik düzleşmiştir. Bu yüzden 4. bileşene kadar olan bileşenler Temel bileşen olarak adlandırılacaktır.

16 Bölüm 2. Temel Bileşenler Analizi (PCA)

**2.9 Temel Bileşenlerin Matematiksel Oluşumu**



Şekil 2.8: Component Matrix

*Y*1 = (0*,* 427)*X*1 + (−0*.*763)*X*2 + (0*,* 680)*X*3 + (0*,* 722)*X*4 + (0*,* 382)*X*5 + (0*,* 394)*X*6 + (0*,* 530)*X*7 + (0*,* 122)*X*8 + (0*,* 123)*X*9 + (0*,* 568)*X*10 + (−0*,* 835)*X*11 + (0*,* 766)*X*12 + (0*,* 933)*X*13 + (0*,* 881)*X*14 + (0*,* 766)*X*15

*Y*2 = (0*,* 800)*X*1 + (0*.*048)*X*2 + (0*,* 011)*X*3 + (−0*,* 034)*X*4 + (0*,* 832)*X*5 + (0*,* 053)*X*6 + (0*,* 135)*X*7 + (−0*,* 591)*X*8 + (−0*,* 456)*X*9 + (−0*,* 519)*X*10 + (−0*,* 361)*X*11 + (−0*,* 300)*X*12 + (0*,* 000)*X*13 + (0*,* 284)*X*14 + (0*,* 007)*X*15

*Y*3 = (−0*,* 209)*X*1 + (0*.*403)*X*2 + (0*,* 462)*X*3 + (−0*,* 460)*X*4 + (−0*,* 141)*X*5 + (0*,* 147)*X*6 + (−0*,* 050)*X*7 + (0*,* 448)*X*8 + (−0*,* 719)*X*9 + (−0*,* 459)*X*10 + (0*,* 067)*X*11 + (0*,* 396)*X*12 + (0*,* 225)*X*13 + (0*,* 124)*X*14 + (0*,* 242)*X*15

*Y*4 = (−0*,* 057)*X*1 +(0*.*029)*X*2 +(−0*,* 230)*X*3 +(−0*,* 127)*X*4 +(−0*,* 019)*X*5 +(0*,* 799)*X*6 + (0*,* 687)*X*7 + (0*,* 119)*X*8 + (−0*,* 030)*X*9 + (−0*,* 006)*X*10 + (0*,* 087)*X*11 + (−0*,* 144)*X*12 + (−0*,* 176)*X*13 + (−0*,* 118)*X*14 + (−0*,* 092)*X*15

|  |  |
| --- | --- |
| 2.9. Temel Bileşenlerin Matematiksel Oluşumu | 17 |
|  |  |

*Y*1 temel değişkeni bünyesinde en çok ± 0,33’den büyük değer barındıran temel bileşen olduğu için bu değişkene baskın temel bileşen denir.

Baskınlık derecesine göre büyükten küçüğe doğru sıralayacak olursak;

*Y*1 = *X*13*, X*14*, X*11*, X*12*, X*15*, X*2*, X*3*, X*10*, X*7*, X*1*, X*6*, X*5

*Y*2 = *X*5*, X*1*, X*8*, X*10*, X*9*, X*11

*Y*3 = *X*9*, X*3*, X*4*, X*10*, X*8*, X*2*, X*12

*Y*4 = *X*6*, X*7

**Bölüm 3**

**Faktör Analizi**

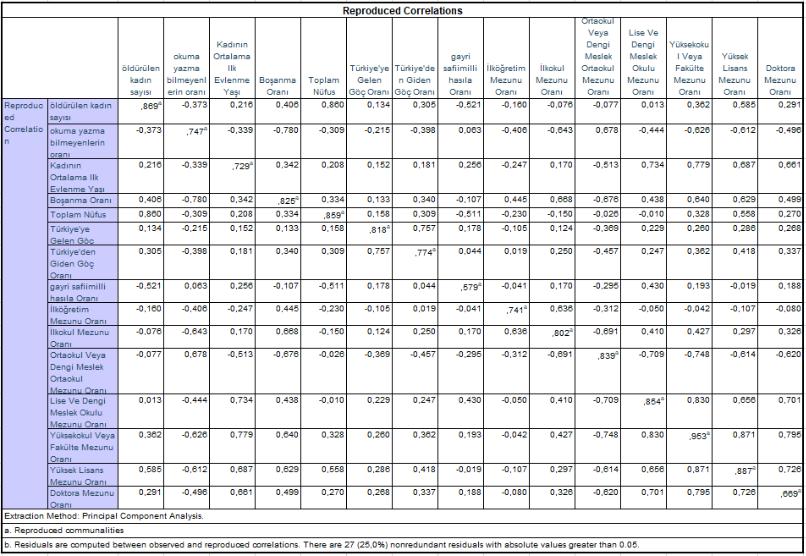
Faktör analizi, aynı yapıyı ölçen çok sayıda değişkenden, az sayıda ve tanımlanabilir nitelikte anlamlı değişkenler elde etmeye yönelik çok değiskenli bir istatistiktir. Davranış bilimlerinde duyuşsal bir özelliği, kişilik ve gelişim gibi pek çok özellikleri ölçmek amacıyla geliştirilen araçların yapı geçerliği faktör analizi kullanılarak [incelenebilir.Büyüköztürk (2002)](#page54)

Temel bileşenler analizinde elde edilmiş olan analizler sonucu veri setinin faktör analizine uygunluğu kanıtlanmış olup çıktı olarak 4 faktör edinilmiştir. Bu 4 fak-tör, faktör analizinde kullanılacaktır. Söz konusu analizde Unweighted least squares yöntemi kullanarak 15 değişkeni (şehir değişkeni hariç) 4 adet faktör ile açıklamak amaçlanmıştır.

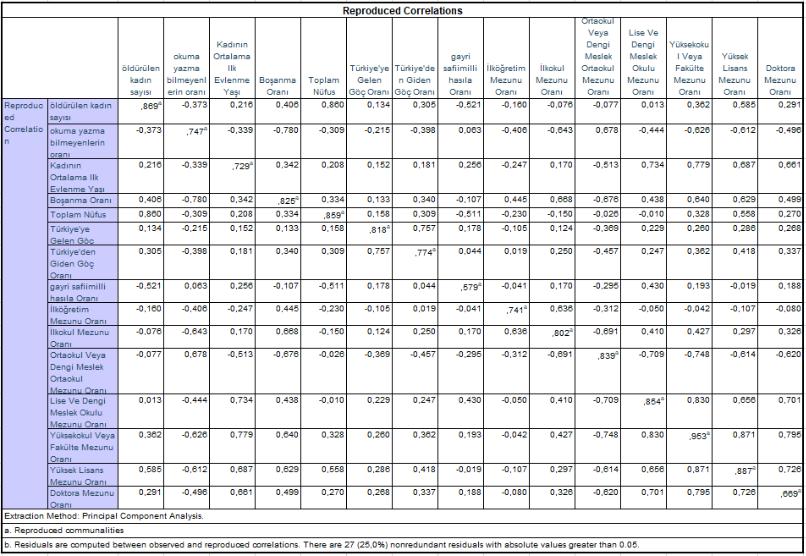
20 Bölüm 3. Faktör Analizi

**3.1 Artıklar Matrisi**

Spss tarafından faktör analizi içerisinde oluşturulan modele göre değişkenler arasındaki korelasyon ve artıklar matrisi üst üste, Reproduced Matrix’ de verilmektedir. Artıkların değeri 0,05’ten az olmalıdır. Değeri 0,05’ten fazla olan artıkların sayısı tüm matrise oranla çeyreğinden (0.25) fazla olmamalıdır. Üzerinde çalıştığımız veri setinde faktör analizi ile oluşturulan modele ait artıklar matrisinde 0.05’in üzerinde 27 adet artık tespit edilmiştir. bu artıklar tüm matrise istinaden, matrisin % 25’ini oluşturmaktadır, kabul edilebilir düzeydedir.



Şekil 3.1: Reproduced Correlations Matrix

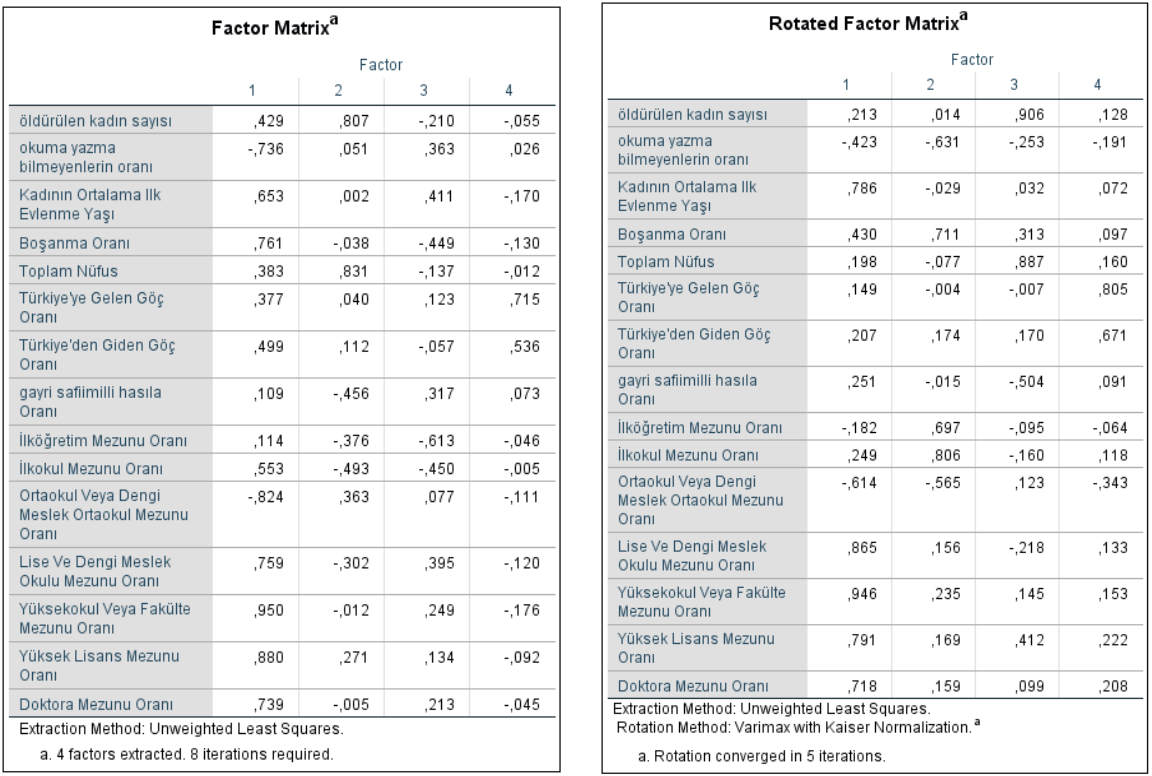


Şekil 3.2: Reproduced Correlations Matrix - Residuals

|  |  |
| --- | --- |
| 3.2. Faktör Matrisi | 21 |
|  |  |

**3.2 Faktör Matrisi**

Faktör Matrisi, değişken ile faktör arasındaki korelasyonlar olan, döndürülmemiş faktör yüklerini içerir. Bunlar korelasyon olduğundan, olası değerler -1 ile +1 arasındadır. Matrsiteki değerlerin yalnızca 0.33’ten yüksek veya daha düşük olan bağıntıları anlam ifade eder. -0.33 ile 0.33 arasındaki değerler analiz için yeterli değildir, bu yüzden bu değerler hesaba katılmaz ve muhtemelen zaten anlamlı olmayan düşük korelasyon yığınını ortadan kaldırarak çıktının okunmasını kolaylaştırır. [Bruin (2011)](#page54)



(a) Faktör matrisi (b) Rotasyonlu faktör matrisi

Şekil 3.3: Normal faktör matrisi ve varimax rotasyonu yapılmış faktör matrisi

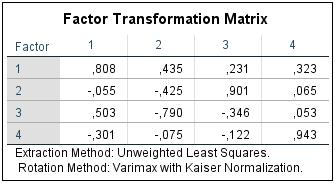
Şekil 3.3.b’de Kaiser normalizasyonu kullanılarak Varimax rotasyon yapılmış faktör matrisine bakıldığında, rotasyon yapılmamış faktör matrisinden daha çok verinin "-0,333<x<0,333" aralığı dışında olduğu görülmektedir. Veriler "-0,333<x<0,333" aralığı dışında oldukça anlamlı bir şekilde yorumlanabilir. Varimax rotasyonlu faktör matrisi tablosunu kullanarak yapılmış olan faktör isimlendirmesi;

"*F*1; Yüksek okur yazarlık, *F*2; Düşük okur yazarlık, *F*3; Düşük modernite ve yoksulluk, *F*4; Göç" şeklinde olmuştur.

22 Bölüm 3. Faktör Analizi

**3.2.1 Factor Transformation Matrisi**

Faktör dönüştürme matrisi, döndürülmüş faktör matrisini elde etmek için döndürül-memiş faktör matrisini çarptığınız matristir.



Şekil 3.4: Factor transformation matrix

Faktör analizinden elde edilen bu 4 faktör, mevcut verisetinin yerine geçecek ve devamında gelecek analizler sadece bu 4 faktörün bulunduğu veri seti aracılığı ile uygulanacaktır.

**Bölüm 4**

**Diskriminant Analizi**

Diskriminant analizi, diskriminant fonksiyonları aracılığı ile gruplar arası ayırıma en fazla etki eden ayırıcı değişkenleri belirlemede ve hangi gruptan geldiği bilinmeyen bir bireyin hangi gruba dahil edileceğini belirlemede kullanılır.

Genel anlamda ayırma olup, bireylere ait p tane özellikten yararlanarak ait oldukları grupları (kütle) belirlemede veya mevcut grupları birbirinden ayıracak en iyi fonksiyonu bulmada kullanılan çok değişkenli istatistik tekniklerinden birisidir. [Cangül (2006)](#page54)

**4.0.1** **Başlarken**

Discriminant analizini gerçekleştirmek için bir önceki faktör analizinden elde edilen yeni veri setİ kullanılacaktır.

Grup değişkeni olarak göç faktörü üzerinden binary değişken oluşturulmuştur. Göç Grup değişkeninin "0" a

24 Bölüm 4. Diskriminant Analizi

**4.1 Group Statistic Tablosu**



Şekil 4.1: Group Statistics Tablosu

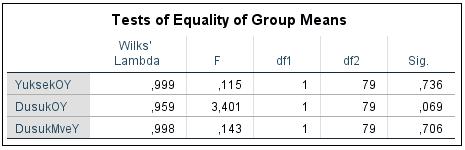
Grup istatistikleri tablosu yeni veri setinin tanınmasında önemli rol oynayan bir tablodur. Bu tablodan şu çıkarımlar elde edilebilir;

* Göç veren şehirlerin yüksek okur yazarlık oranı ortalaması %2.7’dir ve standart sapması 0,9642’dir.
* Göç veren şehirlerin düşük okur yazarlık oranı ortalaması %13.5’dir ve standart sapması 1.033’tür.
* Göç alan şehirlerin yüksek okur yazarlık oranı %5.8’dir ve standart sapması 1.0340’dır.
* Göç alan şehirlerin düşük okur yazarlık oranı %27’dir ve standart sapması 0.685’dir

|  |  |
| --- | --- |
| 4.2. Test of Equaltiy of Group Means Tablosu | 25 |
|  |  |

**4.2 Test of Equaltiy of Group Means Tablosu**

Grup ortalamalarının eşitliğini temsil eden bu analiz, yeni veri setinin karşılaştırma faktörü ile kıyaslanması ve grup ortalamarının varlığı ve yokluğuna dayanarak ista-tistiksel olarak anlamlı fark olup olmadığını, faktörlerin sigma değerlerinin 0.05 ten küçük veya büyük olmasına göre belirler.



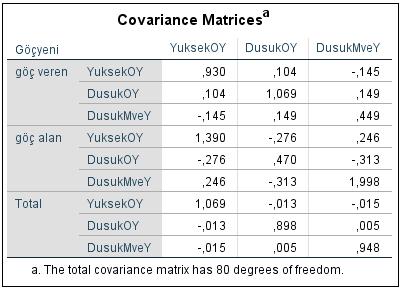
Şekil 4.2: Test of Equality of group Means

Tabloya göre şu çıkarımlar elde edilebilir;

* Şehirlerin **yüksek okur yazarlık** oranları arasındaki fark istatistiksel olarak anlamsızdır.
* Şehirlerin **düşük okur yazarlık** oranları arasındaki fark istatistiksel olarak anlamsızdır.
* Şehirlerin **düşük modernite ve yoksulluk** oranları arasındaki fark istatistiksel olarak anlamsızdır.

26 Bölüm 4. Diskriminant Analizi

**4.3 Kovaryans Matrisleri**

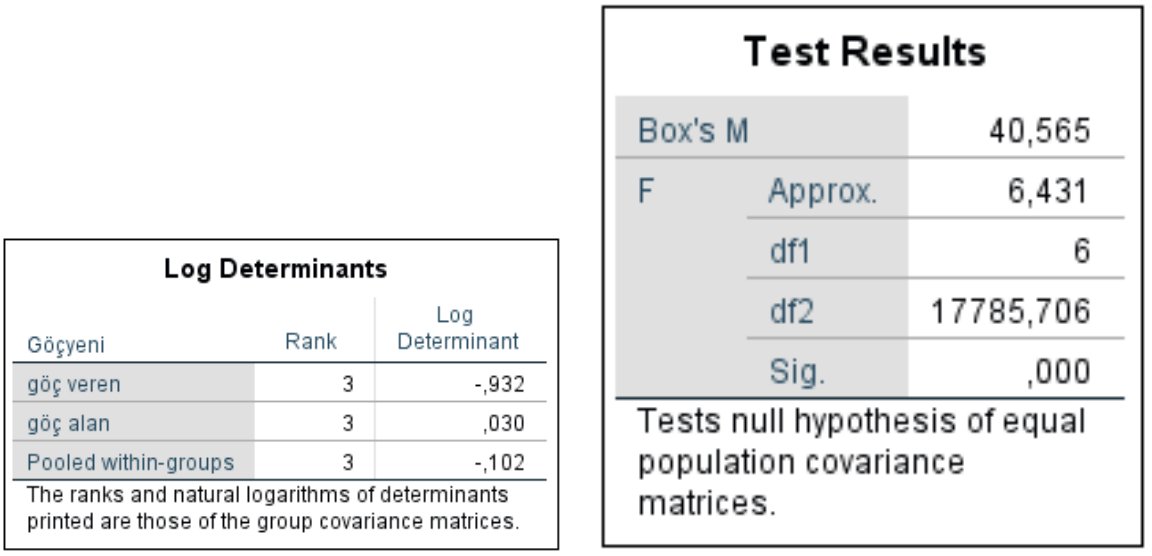


Şekil 4.3: Test of Equality of group Means

* Göç veren şehirlerin **yüksek okur yazarlık**, **düşük okur yazarlık** ve **düşük modernite ve yoksulluk** arasında negatif yönlü bir değişim ve etkileşimbulunmaktadır.
* Göç alan şehirlerin **yüksek okur yazarlık**, **düşük okur yazarlık** ve **dü-şük modernite ve yoksulluk** arasında pozitif yönlü bir değişim ve etkileşimbulunmaktadır.
* Göç veren şehirlerin **yüksek okur yazarlık**, **düşük okur yazarlık** ve **düşük modernite ve yoksulluk** arasında negatif yönlü bir değişim ve etkileşimbulunmaktadır.

|  |  |
| --- | --- |
| 4.4. Box’s Test of Equality of Covariance Matrices | 27 |
|  |  |

**4.4 Box’s Test of Equality of Covariance Matrices**



(a) Log Determinants (b) Box’s Test Results

Şekil 4.4: Log determinant tablosu ve Box’s testinin sonuç tablosu

Box testi, varyans kovaryans matrisinin eşitliğini aşağıdaki hipotezler yardımı ile test etmektedir.

*H*0 : **göç alan ve göç veren şehirler açısından açısından varyans-kovaryans matrisleri eşittir.**

*H*1 : **göç alan ve göç veren şehirler açısından açısından varyans-kovaryans matrisleri eşit değildir.**

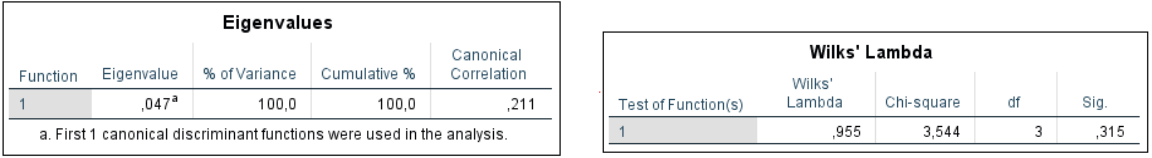
Box’s M testine göre p-değeri 0.05’ten küçük olduğundan dolayı sonuç olarak h0’ red edilmiştir. Göç alan ve göç veren şehirler açısından iki grubun varyans-kovaryans matrisleri arasında fark vardır. Bu durumda doğrusal ayrıştırma analizine devam etmek mümkün değildir. Bu durumu münkün kılmak için;

* Örneklem hacmi arttırılabilir.
* Faktör analizine yeni bir faktör alınabilir.
* Gruplayıcı değişken değiştirilebilir.

Yalnızca gerekli düzenlemeler yapıldıktan sonra tekrar Box’s M testinin sonuçlarına göre ayrıştırma analizine geçilebilir.

28 Bölüm 4. Diskriminant Analizi

**4.5 Kanonik Diskriminant Fonksiyonlarının Özeti**



(a) Eigenvalues (b) Wilks Lambda

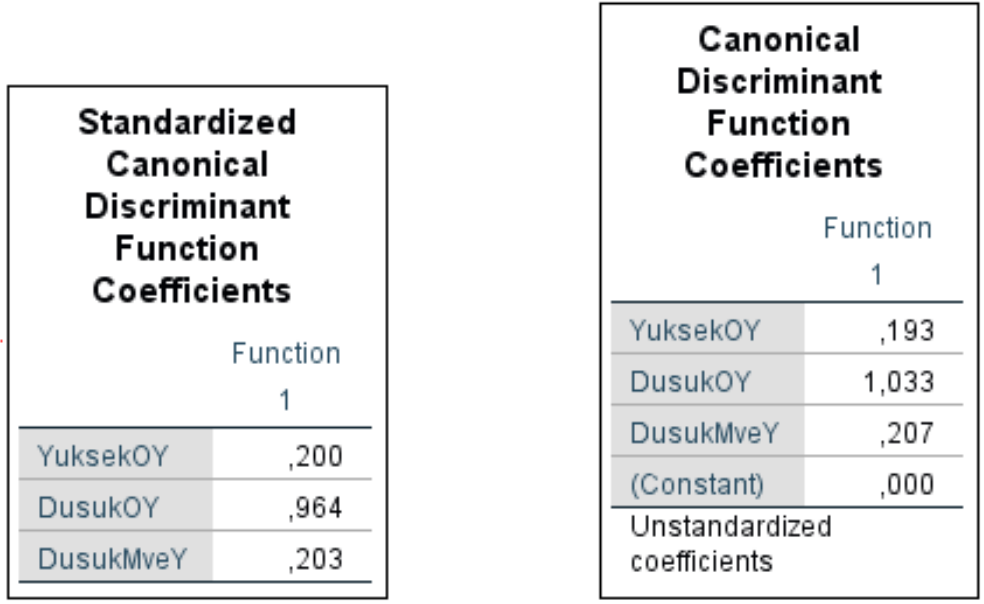
Şekil 4.5: Öz değerler ve Wilk’s Lambda test sonuçu tabloları

Kanonik ayrıştırma fonksiyonunun **eigen values** (Şekil 4.5.a) tablosunda tek bir fonksiyon olması, tek boyutlu uzay sisteminde ayrıştırma analizi yapılabileceği sonucunu verir. Kanonik korelasyon sonucu ne kadar 1 değerine ayrıştırmada o kadar açıklayıcılığı artar. fakat bu bahsedilen durum Wilk’s Lambda için geçerli değildir. Wilk’s lambda’nın sigma değeri olabildiğince küçük olmalıdır. sigma değeri 0’a yaklaştıkça istatistiksel anlamlılık artar.

* Diskriminant fonksiyonu bu veri seti için 0.211’lik kanonik korelasyona sahiptir ve istatistiksel olarak anlamlıdır.
* Wilk’s lambda bu veri seti için 0.315’lik sigma değerine sahiptir ve istatistiksel olarak anlamlıdır.

|  |  |
| --- | --- |
| 4.6. Structure Matrix ve Coefficentlar | 29 |
|  |  |

**4.6 Structure Matrix ve Coefficentlar**



(a) Standirzed Canonical Discriminant Func- (b) Unstandirzed Canonical Discriminant

tion Coefficents Function Coefficents

Şekil 4.6: SCAN ve USCAN tabloları

Standartlaştırılmış kanonik diskriminant fonksiyonunun matematiksel ifadesi aşağıdaki gibidir;

* *SCAN*1 = 0,200 YuksekOY + 0,964 DusukOY + 0,203 DusukMveY

Şehirlere göre göçü iki gruba ayırma durumunda faktörlerin katkısını gözlemlemekteyiz

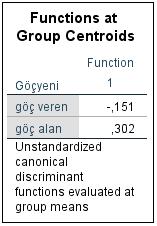
Standartlaştırılmamış kanonik diskriminant fonksiyonunun matematiksel ifa-desi aşağıdaki gibidir;

* *USCAN*2= 0.00 + 0,193 YuksekOY + 1.033 DusukOY + 0.207 DusukMveY Eğer standartlaştırma yapılmazsa faktörlerimiz olmadığında, ortalama olarak şehirlere göre göçün olup olmaması ortalama olarak düeylerindeki farkı verir. Bu fark Unstandardized coefficent tablosunda "constant" satırının karşılığıdır.

Faktörler 1 birim arttırıldığında; faktörler, tablodaki söz konusu faktörün karşısındaki değer kadar anlamlılığa katkı vermiş olur.

30 Bölüm 4. Diskriminant Analizi

**4.7 Centroids**



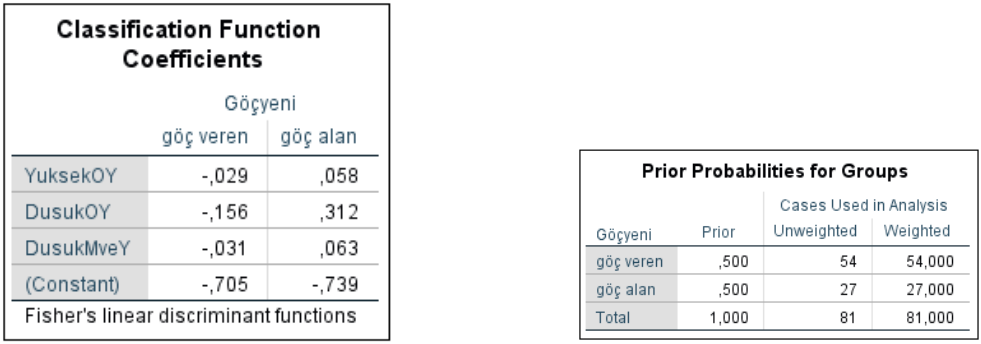
Şekil 4.7: Centroids

Önceki bölümlerde tek boyutlu uzayda işlemlerin yapılacağı analizler ile desteklenmiştir. Buna istinaden Functions at Group Centroids tablosu X ekseninde merkezi konum oluşturur ve bunları listeler. Diskriminant analizi belrilenen bu merkezi konuma göre yapılır. Merkezi değerler birbirine ne kadar yakınsa diskriminant o kadar başarısız olur.

|  |  |
| --- | --- |
| 4.8. Gruplandırma İstatistikleri | 31 |
|  |  |

**4.8 Gruplandırma İstatistikleri**

**4.8.1 Gruplar İçin Öncül Olasılıklar ve Gruplandırma Fonk-siyon Coefficentları**



(a) Classification function coefficents (b) Prior Probabilities for groups

Şekil 4.8: Classification Function ve Prior Probabilites Tabloları

*F*1 = (-0.705) + (-0.156) \* DusukOY + (-0.031) \* DusukMveY + (-0.029) \* YuksekOY (Göç veren şehirler)

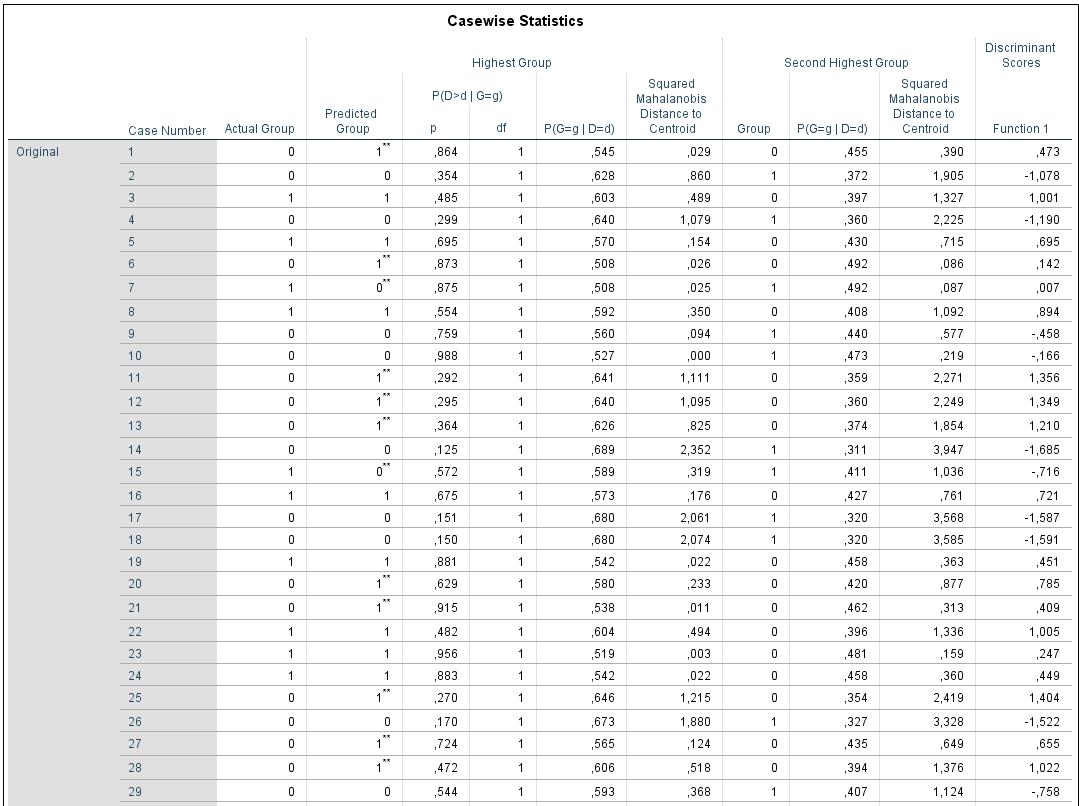
*F*2 = (-0.739) + (0.312) \* DusukOY + (-0.063) \* DusukMveY + (-0.058) \* YuksekOY (Göç alan şehirler)

Yeni deney biriminin Göç alıp verme açısından hangi gruba koyulacağını belirlemek için sırasıyla yukarıda belirtilen denkleme yazılır, denklemde büyük değeri alan deney birimi ilgili gruba yerleştirilir.

Prior Probabilities for Groups tablosuna (şekil 4.8.b) bakıldığında "Göç veren" isimli ilk grubun öncül olasılığı 0.666 iken ikinci grubun öncül olasılığı 0.333’tür

32 Bölüm 4. Diskriminant Analizi

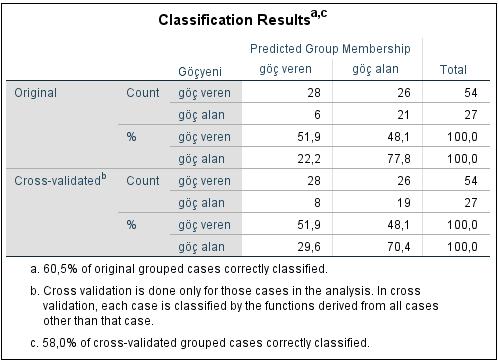
**4.8.2** **Casewise Statistics**



Şekil 4.9: Casewise Statistics

Casewise Statistics tablosu biz araştırmacı ve analizcilere değişkenin şu anda hangi grupta olduğunu, ve olması gerektiğini tahminlediği grubu vermektedir. Bu tahminleme işlemini değerlerin diskriminant skorlarını tek tek Fisher’s LDF denkleminde yerine yazarak, değişkenleri tüm gruplarda dener. değişkenler, en yüksek katkıyı yaptıkları gruba dahil olur ve böylece tahminleme işlemi tamamlanmış olur. Olduğu grup ve olması gerektiği grup değerleri uyuşmayan tablo değerlerinin yanında "\*\*" çift yıldız sembolü yerleştirir.

|  |  |
| --- | --- |
| 4.8. Gruplandırma İstatistikleri | 33 |
|  |  |



Şekil 4.10: Classification Results

Classification Results tablosu, Casewise statistics tablosunda yapılan yanlış gruplan-dırmaların sayısını ve yüzdesini, orjinal veri seti ve crossvalidated veri seti şeklinde özetleyen bir tablodur.

**Bölüm 5**

**Kümeleme Analizi**

Kümeleme analizi bir veri setinin farklı gruplar içerip içermediğini belirlemek ve eğer içeriyorsa bu grupları tespit etmek için kullanılan çok değişkenli istatistiksel bir yöntemdir.

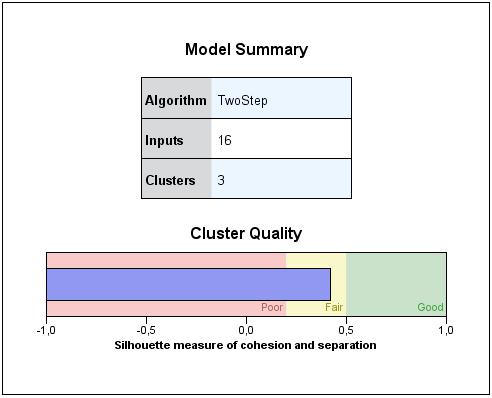
Çok boyutlu uzayda verilerin özetlenmesi ve tanımlanmasında yol gösterici bir araştırma yöntemi olan kümeleme analizi; heterojen olan farklı gruplardaki gözlem yapılarını ya da homojen olan benzer gruplardaki gözlemleri uygun yöntemlerle gruplamayı sağlayan bir yöntemdir.

Diğer çok değişkenli istatistiksel yöntemlerde önemli bir yer tutan normallik, doğrusallık ve homojenlik varsayımları bu yöntemde prensipte kalmakta ve uzaklık değerlerinin normalliği yeterli görülmektedir. [Aydın (2007)](#page54)

36 Bölüm 5. Kümeleme Analizi

**5.1 Two Step Cluster**

**5.1.1 Model Özeti ve Grup Kalitesi**

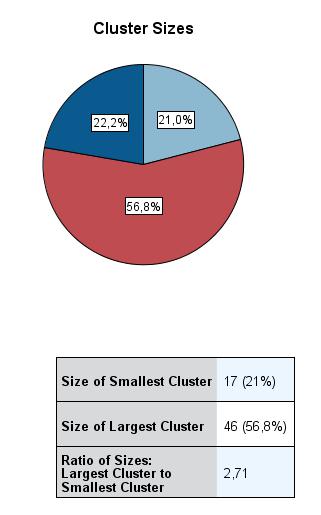


Şekil 5.1: Two step cluster Model Summary

Kümeleme analizi için oluşturulan cluster’ların ne kadar etkili olduğunu görebilmek adına, Model Summary ve Cluster Quality bölümlerine bakılabilir. Bu tablolara dayanarak, kullanılan algoritmanın "TwoStep" olduğu, 16 adet girişi 3 ayrı cluster’da topladığı görülmektedir. Ayrıca yapılan gruplandırmanın kalitesinin çok iyi olmasada yeterli seviyede olduğu çıkarımı da yapılabilir.

|  |  |
| --- | --- |
| 5.1. Two Step Cluster | 37 |
|  |  |

**5.1.2 Grup Boytuları ve Oranları**



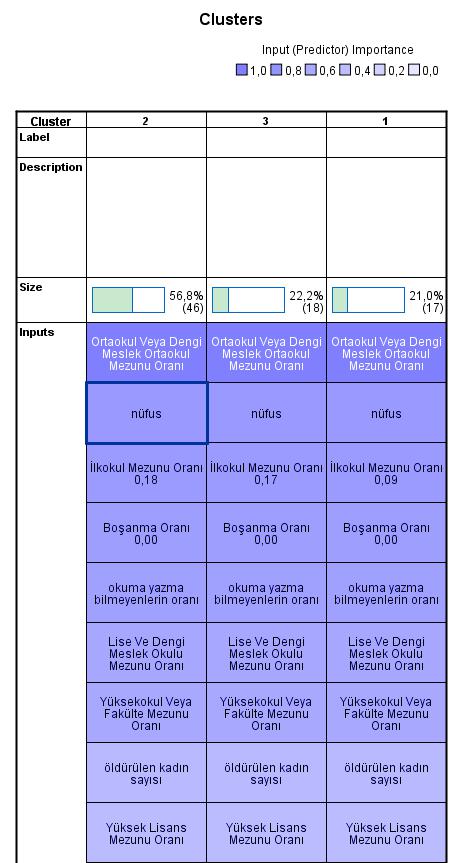
Şekil 5.2: cluster Size

Cluster size grafiğinde ve tablosunda, grupların spesifik özelliklerine yer verilmektedir. Söz konusu tablo ve grafikten aşağıdaki bilgiler elde edilebilir.;

* Seçilecek herhangi bir grubun tüm değişkenlerin % kaçını kapsadığını
* En küçük grubun tüm değişkenlerin % kaçını kapsadığını
* En büyük grubun tüm değişkenlerin % kaçını kapsadığını
* Büyük grubun büyüklüğünün küçük gruba oranı

38 Bölüm 5. Kümeleme Analizi

**5.1.3 Grup İçin Daha Spesifik Bilgiler**



Şekil 5.3: Cluster dağılımı

Şekil 5.3’teki Cluster tablosu, değişkenlerin ne kadar kısmının hangi grupta olduğunu, grup büyüklüklerini ve hangi değişkenin grup için daha önemli olduğu bilgisini veren daha detaylı bir tablodur.

|  |  |
| --- | --- |
| 5.2. Quick Cluster ile Kümeleme Analizi | 39 |
|  |  |

**5.2 Quick Cluster ile Kümeleme Analizi**

**5.2.1 Initial Cluster Centers**

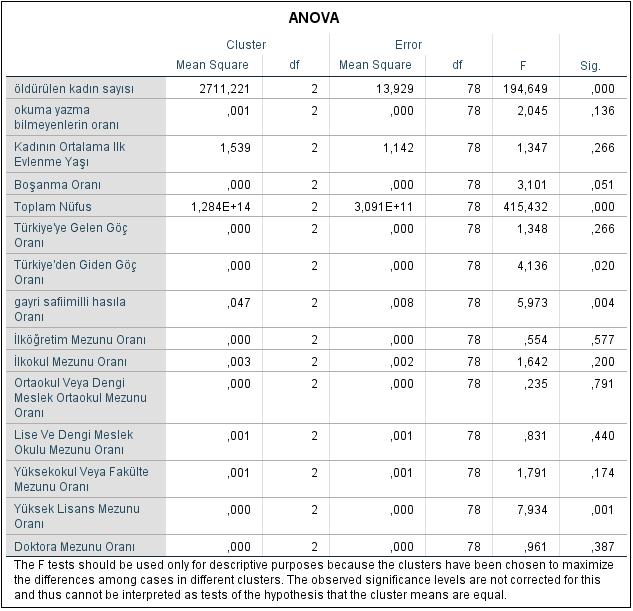


Şekil 5.4: Initial cluster Centers

Inital cluster centers tablosu, grupların merkezlerini verirken aynı zamanda birbirinden çıkarıldığında uzaklıklarını da vermiş olur. Gruplar birbirinden ne kadar uzaksa gruplandırma o kadar başarılı yapılmış demektir.

40 Bölüm 5. Kümeleme Analizi

**5.2.2 Anova**



Şekil 5.5: Anova

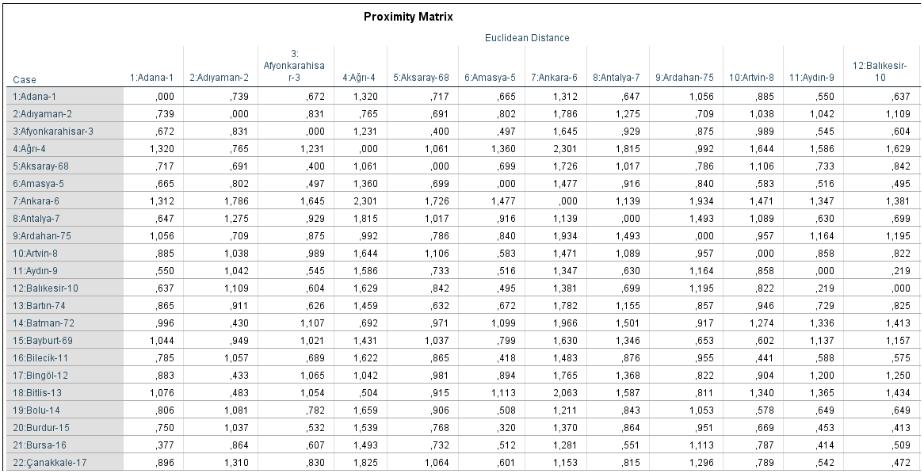
Şekil 5.5’te görünen Anova tablosu içerisindeki sigma değerleri 0’a yaklaştıkça o değişkenin kümeleme açısından önemi artmaktadır.

Veri setindeki sigma değerlerine bakıldığı zaman, **Ortaokul** **veya Dengi Meslek Ortaokul Mezunu Oranı, İlk Öğretim Mezun Oranı, Lise ve Dengi Meslek Okulu Mezunu Oranı ve Doktora Mezunu Oranı** değişkenlerinin sigma

değerlerinin 0’dan çok büyük olduğunu gözlemlenmiştir. Bu değişkenlerin kümeleme açısından önemi diğer değişkenlere göre daha azdır.

|  |  |
| --- | --- |
| 5.3. Hieararchical Cluster | 41 |
|  |  |

**5.3 Hieararchical Cluster**

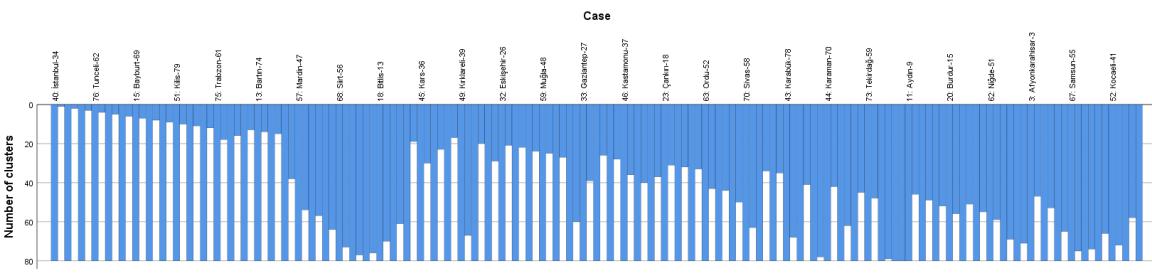


Şekil 5.6: Proximity Matrix

Şekil 5.6’da gözüken Proximity Matrix, 15 değişken açısından şehir değişkenlerinin birbirine olan istatistiksel uzaklıklarını hesaplayıp listeler. Uzaklık ne kadar fazlaysa farklılık o kadar fazla olmaktadır.

Bu tabloya bakarak şu çıkarımlara ulaşılabilir;

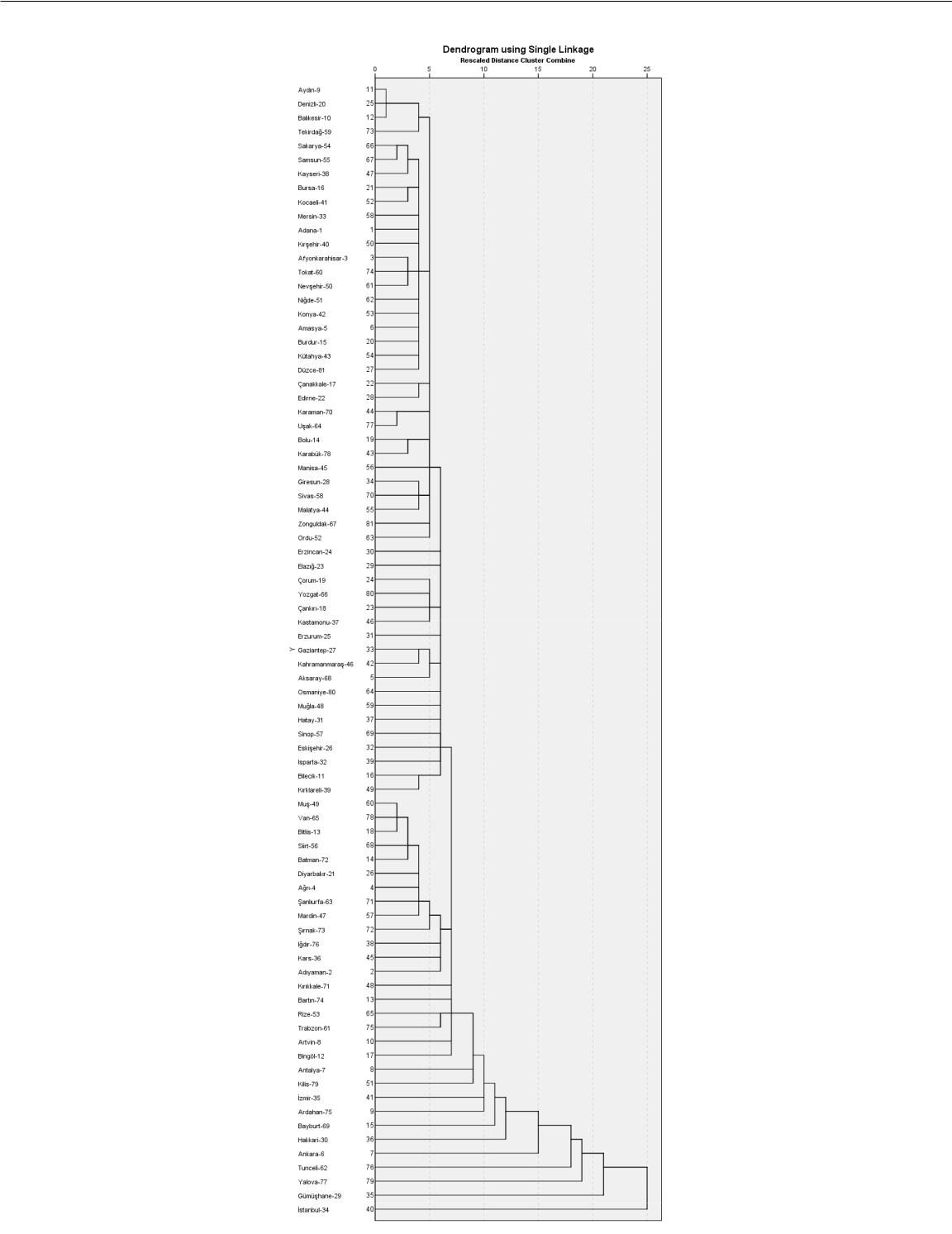
* Adana ve Adıyaman arasında 15 değişken açısından 0,739 birim uzaklık vardır.
* Balıkesir ve Ağrı arasında 15 değişken açısından 1,629 birim uzaklık vardır.



Şekil 5.7: Case - Number of Clusters

Şekil 5.7’de Number of cluster - Cases histogramında şehirlerin hangi grupta kaç tane elemanı olduğu açıkça görülmektedir.

42 Bölüm 5. Kümeleme Analizi



Şekil 5.8: Case - Number of Clusters

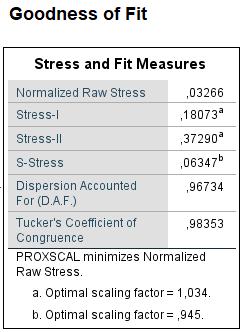
Dendogram grafiğine bakıldığında şehirlerin istatistiksel açıklayıcılıkarına göre sırayla eşleştirildiği görülmektedir. Üretilen dendrogramın X-ekseninde kümeler arasındaki uzaklık görülmektedir. Y-ekseninde ise kümeleri oluşturan veri noktaları bulunmakta-dır.

**Bölüm 6**

**Çok Boyutlu Ölçeklendirme Analizi**

**6.1 Proximities - Proxscal**

**6.1.1 Stress and Fit Measures**

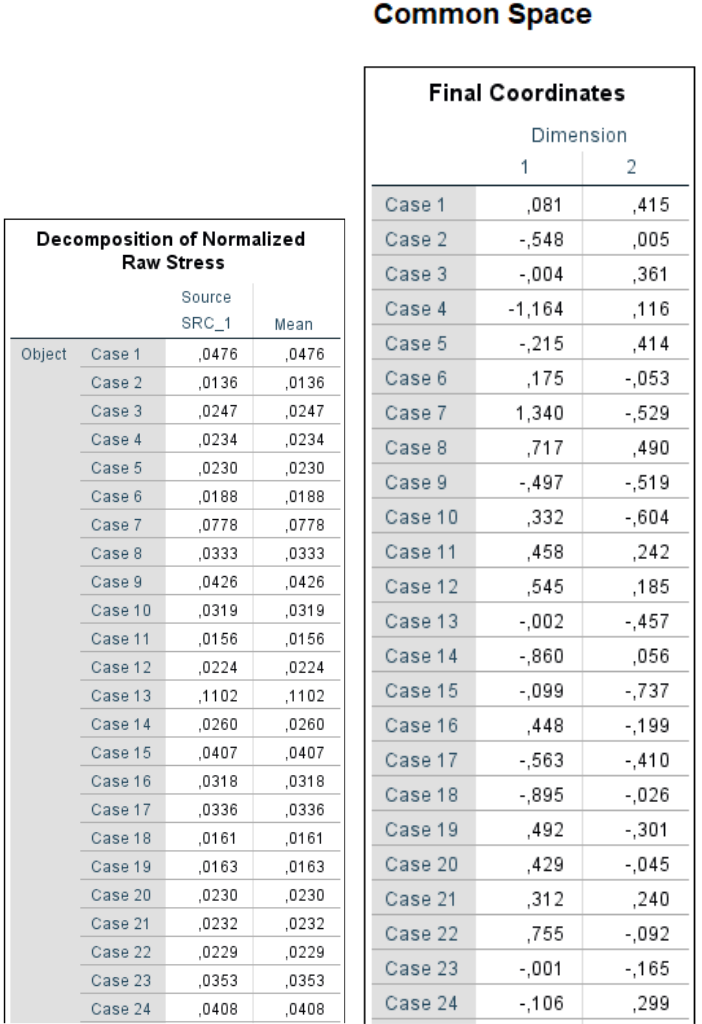


Şekil 6.1: Stress and Fit Meausers

Şekil 6.1’de bulunan Stress and Fit Measures tablosu, değişkenlerin stress değerlerini vermektedir. Doğru bir çok boyutlu ölçeklendirme analizi için bahsi geçen stress değerleri olabildiğince küçük, olabildiğince 0’a yakın olmalıdır.

44 Bölüm 6. Çok Boyutlu Ölçeklendirme Analizi

**6.1.2 Decomposition of Normalised Raw Stress**



(a) Decomposition of Nor- (b) Final Coordinates malized Raw Stress

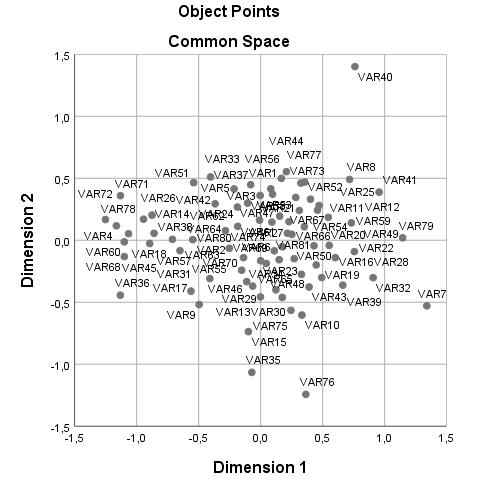
Şekil 6.2: Raw Stress ve Final Coordinates tabloları

Şekil 6.2.a’da gözüken Decompostion of Normalized Raw Stress tablosu, deney birimimleri açısından stress seviyesinin ne kadar değiştiğini gösterir. Deney birimlerinin stress seviyelerinin 0’a yakın olması beklenir.

Şekil 6.2.b’de gözüken Final Coordinates tablosunda, deney birimleri açısın-dan iki boyutlu düzlemde final koordinat değerlerini verir.

|  |  |  |
| --- | --- | --- |
| 6.1. Proximities - Proxscal | | 45 |
|  |  |  |

**6.1.3 Object Points**



Şekil 6.3: Object Points

Object Points grafiğine bakıldığı zaman, deney birimlerinin birbirine çok yakın olduğu görülmektedir, yoğunluk Y(-0.5,0.5) ve X(-1,1) aralığında dağılmıştır. Bu dağılım dar bir alanda olduğu için kullanılan deney birimleri üst üste binmiştir. Bu durum çok başarılı bir scaling yapılamadığı olgusunu kanıtlar.

**Sonuç**

Bu projede ilk olarak veri seti tanımlanarak uygun analiz yöntemleri belirlenmiştir. Çoklu doğrusal bağlantı sorunu sonucunda temel bileşen analizi uygun görülmüştür. Bu analiz yöntemi ile 15 değişkenin (şehirler eksik) arasındaki ilişkiler incelenmiştir. Temel bileşenler analizinden elde edilen sonuçlar ile faktör analizi uygulanılarak 15 değişken 4 yapay değişkene indirgenmiştir. Bu dört faktör ; yüksek okur-yazarlık, düşük okur-yazarlık, düşük modernite ve yoksulluk, göç şeklinde olmuştur.

Ayrıştırma analizinde illerdeki Göç oranı veriyi en iyi şekilde özetleyeceğinden, Göç alan ile Göç veren binary değişkenlerinin araştırma konusu olması ve değişkenlerin faktör yüklerinin fazla olması nedeniyle grup değişkeni olarak belirlenmiştir. Sınıflandırma analizleri değerlendirilmiştir. Faktör analizine göre 4 faktör birbirinden bağımsızdır. Wilk’s lambda’nın sigma değeri olabildiğince küçük olmalıdır ve bu veri seti için 0,315’lik sigma değerine sahiptir ve istatistiksel olarak anlamlıdır. Kanonik ayrıştırma fonksiyonu ise 0,211’lik kanonik korelasyona sahiptir ve istatistiksel olarak anlamlıdır.

Bu çalışma boyunca temel bileşenler analizi, faktör analizi ve ayrıştırma analizi öğrenilmiştir. Bu analizlerin amaçları, analize uygunluk koşulları , uygulama yöntemleri ve sonuçların yorumlanması gibi adımlarda gereken değerler kazanılmıştır.

**Kaynaklar**

Akkuş, Z., Sanisoğlu, S. Y., Akyol, M., & Çelik, M. Y. (2006). Değişken yapılarına göre istatistiksel yaklaşım. *Dicle Üniversitesi Tıp Fakültesi Dergisi (33)*, *2* .

Baydemir, M. B. (2020). Çok değişkenli istatistiksel yöntemlerin karşılaştırmalı analizi.

Bruin, J. (2011). Factor analysis | spss annotated output.

URL <https://stats.oarc.ucla.edu/spss/output/factor-analysis/>

Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve uygulamada eğitim yönetimi*, *32* (32), 470–483.

Cangül, O. (2006). *Diskriminant analizi ve bir uygulama denemesi*. Ph.D. thesis, Bursa Uludag University (Turkey).

Yıldız, K., Çamurcu, Y., & Doğan, B. (2010). Veri madenciliğinde temel bileşenler analizi ve negatifsiz matris çarpanlarına ayırma tekniklerinin karşılaştırmalı analizi. *Akademik Bilişim*, *10* , 248.

Çil, A. (2014). *Faktör Analizi*. [https://abdulrezzakcil.blogspot.com/2014/](https://abdulrezzakcil.blogspot.com/2014/05/faktor-analizi-ileri-istatistik.html?view=magazine) [05/faktor-analizi-ileri-istatistik.html?view=magazine](https://abdulrezzakcil.blogspot.com/2014/05/faktor-analizi-ileri-istatistik.html?view=magazine) [Accessed: Whene-ver].