

İstanbul Üniversitesi  
İktisat Fakültesi  
Maliye Araştırma Merkezi Konferansları  
47. Seri / Yıl 2005  
Prof. Dr. Türkan Öncel'e Armağan

**TÜRKİYE'DE İLLERİN GSYİH'İNİN  
SEKTÖR PAYLARINA GÖRE  
GRUPLANDIRILMASINA  
İSTATİSTİKSEL BİR YAKLAŞIM**

**Yrd. Doç. Dr. Mürüvvet PAMUK**

İstanbul Üniversitesi  
İktisat Fakültesi  
Ekonometri Bölümü

## ÖZET

Bu çalışmanın amacı, Türkiye’de iller arasındaki farklılığı GSYİH (Gayri Safi Yurt İçi Hasıla)’nin sektör paylarına göre incelemek ve birbiriyle ilişkili olan sektörleri bulmaktır.

Türkiye’de iller arasında sektörler açısından farklılıklar olduğu bilinmektedir. Bu farklılığı ortaya koymak için Türkiye’nin 81 ilinde 2000 yılına ait GSYİH’nin 10 farklı sektördeki payları alınmış ve birbiriyle ilişkili sektörler, faktör analizi ve kümeleme analizi teknikleri kullanılarak gruplandırılmıştır.

## ABSTRACT

The aim of this study is to work out the differences between provinces of Turkey according to GDP (Gross Domestic Product) sector shares in 10 different sectors and to find out the sectors that are related with each other.

It is already known that these sectors are different in their financial situations. In order to show these differences, GDP (Gross Domestic Product) sector share values in 10 different sectors in the year 2000 from 81 provinces of Turkey are taken; then the sectors related with each other are grouped by using factor and cluster analysis techniques.

## 1. GİRİŞ

Bölgeler arası farklılıklar birçok gelişmekte olan ülkelerde, hatta bazı gelişmiş ülkelerde olduğu gibi, ülkemiz sınırları içinde de yoğun olarak görülmektedir. Diğer ülkelerde olduğu gibi ülkemizde de 10 farklı sektörde görülen bu farklılık, iller arasındaki coğrafi yapıdan, iklim özelliklerinden, ekonomik ve sosyal yapıdan, bölgeler arası ticaret, teknoloji ve bilgi yayılımından kaynaklanmaktadır. (Exp. Spa. Data Analy., 1980-1985, s. 2)

Çalışmada, sektörler itibariyle bölgesel farklılıkların ortaya çıkarılmasına, benzer özellikleri taşıyan sektörleri bir araya getirerek, illerin GSYİH bakımından bu sektörlerle yaptığı katkıyı belirlemeye çalışılmıştır. Üretilen tüm mal ve hizmetlerin toplam değeri olan GSYİH, ekonomik gelişmişlik düzeylerinin karşılaştırılmasında kullanılan önemli değişkenlerden biridir (Dincer, B., Özaslan, M., Kavasoğlu, T., 2003, s. 36). Bölgesel dengesizlikler ulusal bir sorun olarak toplumların ekonomik ve siyasal gündeminde yer almaktadır. Bu bakımdan illere yapılan kamu kaynaklarının tahsisi ve özel sektör yatırımlarının yönlendirilmesi konuları önemli olmaktadır. Bu çerçevede “İllerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması” karar organlarına önemli bilgiler sunmaktadır. (Dincer, B., Özaslan, M., Kavasoğlu, T., 2003).

Çalışmada, Devlet İstatistik Enstitüsü’nün (DİE) Türkiye’de 81 ilde 10 sektöre ait 2000 yılının GSYİH sektör payları kullanılmıştır. (İllere göre Gayri Safi Yurtiçi

Hasıla, Türkiye, DİE, 2000) Bu sektörler; 1. Tarım, 2. Madencilik, 3. İmalat, 4. Elektrik, gaz, su, 5. İnşaat, 6. Ticaret, 7. Ulaştırma ve haberleşme, 8. Mali kuruluşlar, 9. Konut sahipliği, 10. Serbest meslek ve hizmetler. İmalat ve Elektrik, gaz, su sanayi sektörünün alt başlıkları olarak yer almaktadır.

Ekonomik ve sosyal araştırmalarda, araştırmacılar çok sayıda değişkenle uğraşmak zorunda olduklarından, çok değişkenli istatistik yöntemler kolaylık sağlamaktadırlar. Bu çalışmada yapılan istatistiksel analizler, Çok Değişkenli Analiz Tekniklerinden olan Faktör Analizi ve Kümeleme Analizi Teknikleri kullanılarak SPSS for Windows 11.0 ve S-Plus 2000 istatistiksel paket programlarında çözümlenmiştir.

Kullanılan yöntemler ikinci bölümde kısaca anlatılmakta olup, üçüncü bölümde uygulama yer almaktadır.

## 2. YÖNTEM

Bu bölümde çalışmada kullanılan Faktör Analizi ve Kümeleme Analizi Teknikleri'nden genel olarak bahsedilecektir.

### 2.1. FAKTÖR ANALİZİ

Bilindiği gibi Faktör Analizi birbiriyle ilişkili çok sayıdaki değişkenleri temsil eden daha az sayıdaki hipotetik faktörleri belirlemek amacıyla kullanılır. Temel bileşenler analizinde olduğu gibi bir boyut indirgeme ve bağımlılık yapısını yok etme yöntemidir. Veri matrisi olarak, n tane bağımsız birime (gözleme) ilişkin p tane değişkenin olduğu  $X_{p \times n}$  ham veri matrisi veya standartlaştırılmış  $Z_{p \times n}$  matrisi kullanılır.

Faktör modelinde gözlenebilir değişkenler  $x_j$  ( $j=1, \dots, p$ ) gözlenemeyen ancak istatistik yöntemlerle tahmin edilen faktörlerin  $f_k$  ( $k=1, \dots, m < p$ ) doğrusal fonksiyonu olduğu kabul edilmektedir. Böylece,  $x_j$  değişkeni

$$x_j = \sum_{k=1}^m a_{jk} f_k + e_j, \quad (j=1, \dots, p), \quad m < p \quad (1)$$

olarak ifade edilebilir. Burada,  $a_{jk}$ : j. değişkendeki k. faktörün önem derecesini belirten faktör yükü ve  $e_j$ : j. artık veya kalıntı değeridir. Modelde yapılan varsayımlar için bakınız. (Pamuk, 1997, s. 5)

Faktör analizinin uygulanışını 4 ana adımda toplayabiliriz: (1) verilerin toplanması ve ilgili kovaryans matrisinin hazırlanması, (2) ilk faktörlerin bulunması, (3) döndürme ve (4) faktör skorların bulunmasıdır. (Kim, and Mueller, 1978, s. 46)

#### 1) Verilerin Toplanması ve Kovaryans Matrisinin Hazırlanması

Faktör analizinde ilk adım verilerin toplanması ve ilgili kovaryans veya korelasyon matrisinin hazırlanmasıdır. Faktör analizinin araştırmaya ait sorgulayıcı olarak kullanımında korelasyon matrisinin seçilmesi uygun olur.

#### 2) İlk Faktörlerin Çıkarılması

Bu aşamada gözlenen değişkenler arasındaki korelasyonları yeterli derecede açıklayan ilk faktörler elde edilir. Bunun için çeşitli yöntemler kullanılabilir. En fazla kullanılan yöntem temel (asal) faktör yöntemidir. Bu yöntem farklı durumlarda farklı korelasyon matrislerine uygulanabilir. Eğer köşegeninde birim elemanları olan korelasyon matrisine uygulanırsa, temel (asal) bileşenler; eğer köşegeninde ortak varyans tahminleri yer alıyorsa, ortak faktörler elde edilir. (Kim, and Mueller, 1978, s. 21) Temel faktör yönteminin özelliği her bir faktörün yapısında yer alan değişkenlerin varyansının mümkün olan en fazla miktarını hesaba katmasıdır.

### 3) Döndürme Yöntemleri

Faktör analizinde elde edilen ilk faktörlerin döndürülmesindeki asıl amaç anlamlı faktörler ve basit faktör yapısının bulunmasıdır. Değişkenlerin kümesi döndürmeden sonra daha açık olarak belli olur. Dik ve eğik döndürme olmak üzere çeşitli döndürme yöntemleri bulunmaktadır. Eğer faktörler arasındaki korelasyonlar ihmal edilebilirse, dik döndürme yöntemi ; eğer korelasyonlar önemli görülürse eğik döndürme yöntemi seçilir. (Pamuk, 1997, s. 57) Uygulamada en fazla kullanılan yöntem dik döndürme yöntemlerinden varimax yöntemidir. Bu yöntemde sütunların veya faktörlerin basitleştirilmesi amaçlanmaktadır. Böyle bir basitleştirme, her bir sütundaki ve her bir faktörün kareli yüklerinin varyansını maksimize etmeye eşittir. Varyansın en büyük yapılması için 1 ve 0'a yakın yüklerin sayısının artırılması gerekir. Böylece varimax çözüm sayesinde, daha az sayıda değişken üzerinde fazla yükleri olan faktörler üretilir.

### 4) Faktör Skorların Bulunması

Kullanıcıların isteğine bağlı olarak faktör skorları da bulunabilir. Bunlar başka çalışmalarda değişken olarak kullanılabilir. Döndürmeden sonra elde edilen faktör yükleri matrisi G ise, her bir gözlem sayısı için her bir faktörün skoru

$$F=(G^1G)^{-1} G^1X \quad (2)$$

eşitliğinden bulunur. (Bryan, 1986, s. 78)

## 2. 2. KÜMELEME ANALİZİ

Kümeleme analizi, p sayıdaki değişkeni n sayıda birimde saptanan özelliklerine göre ortak özellikleri açıkladığı varsayılan alt kümelere ayırmak amacıyla kullanılan bir yöntemdir (Özdamar, 2002, s. 280). Elde edilen kümeler olabildiğince kendi içlerinde türdeş, kendi aralarında da farklı gruplardır.

Kümeleme analizinde kullanılan yöntemler genel olarak aşamalı (hiyerarşik) ve aşamalı olmayan (hiyerarşik olmayan) olarak ikiye ayrılmaktadır.

### 1) Aşamalı Kümeleme Yöntemi

Bu kümeleme yönteminde küme sayısı hakkında ön bilgi gerekmez (Bharati, 2003, s. 4). Kümeleme sürecinin başlangıcında her birim (gözlem) bir kümedir. Süreç sonunda ise tüm birimleri bir kümede toplamak mümkündür. Birimlerin p değişkene göre birbirleri arasındaki uzaklıkları hesaplamak için Öklid uzaklığı, karesel Öklid uzaklığı ve Mahalonobis uzaklığı gibi uzaklık ölçü birimleri kullanılır. Öklid uzaklığı en çok kullanılan ölçü birimlerinden biridir. Basit olarak, çok boyutlu uzayda geometrik

uzaklığı belirtir. Öklid uzaklığı genellikle ham veriden hesaplanır. Bu yöntemin avantajı, iki birim arasındaki uzaklık, analize yeni bir bireyin ilavesiyle etkilenmez. Fakat, değişkenlerin birimlerinde farklılık olduğunda Öklid ve karesel Öklid uzaklıkları önemli derecede etkilenir ve kümeleme analizi sonuçları çok farklı çıkabilir. Bu durumda aynı ölçü birimlerine çevirmek veya standardize veriyle çalışmak uygun olur (Cluster Analysis, 1985-2004). Karesel Öklid uzaklığı ise Öklid uzaklığında karekök alınmaksızın elde edilen uzaklıktır. Mahalonobis uzaklığı ise, değişkenler arasındaki korelasyonlar dikkate alındığından tercih edilen bir uzaklık ölçüsüdür. Mahalonobis uzaklığında, korelasyonlar sıfır olduğunda, standardize edilmiş değişkenlerle hesaplanmış Öklid uzaklığına eşittir (Öğüş, 1985, s. 16). Aşamalı kümeleme yönteminde en çok kullanılan tek bağlantı tekniği (minimum uzaklık veya en yakın komşuluk), tam bağlantı tekniği (maksimum uzaklık veya en uzak komşuluk) ve son zamanlarda sıkça kullanılan ortalama bağlantı tekniğidir (ortalama uzaklık). Aşamalı kümeleme yönteminde, bir önceki aşamada yanlış gruplandırılan bireylerin tekrar atanmasına izin verilmez (Johnson and Wichern, 1992, s. 596). Aşamalı kümeleme yöntemlerinin sonuçları ağaç diyagramı (dendogram) olarak iki boyutlu diyagram şeklinde gösterilebilirler. Ağaç diyagramları uygun küme sayısı hakkında bilgi vermede yararlı olmaktadır ve yeniden ölçeklendirilen uzaklıkları gösterirler. Daha büyük uzaklıklar kümelerin elemanlarının benzer olmadığını; daha küçük uzaklıklar küme elemanlarının oldukça türdeş (homojen) olduğunu gösterir. Kümeler arasındaki bu uzaklıklar küme sayısına karar vermede yardımcı olurlar (Bharati, 2003, s. 5).

## 2) Aşamalı Olmayan Kümeleme Yöntemi

Aşamalı kümelemede hem birimler hem de değişkenler birbiriyle değişik benzerlik düzeylerinde kümeler oluştururken, aşamalı olmayan yöntemlerde birimlerin uygun oldukları kümelerde toplanmaları ve n birimin k sayıda kümeye parçalanması hedeflenir (Özdamar, 2002, s. 325). Küme sayısı konusunda ön bilgi var ise yada araştırmacı anlamlı olacak küme sayısına karar vermiş ise çok uzun zaman alan aşamalı yöntemler yerine aşamalı olmayan kümeleme yöntemleri tercih edilmelidir. (Tatlidil, 1992, s. 258) Bu yöntemler benzerlik yada uzaklık matrisi yerine X veri matrisini yada standartlaştırılmış Z veri matrisini kullanır. Bu yöntemde k-ortalamlar tekniği, yığıma kümeleme tekniği ve en çok olabilirlik tekniği yer almaktadır. En çok kullanılan k-ortalamlar tekniğidir. Bu teknikte ilk olarak k sayıda çekirdek nokta oluşturulur. Eğer işleme çekirdek noktalarla başlanmışsa, her birim (gözlem) en yakın çekirdek noktaya atanır, çekirdek noktalar algoritma boyunca değişmez. Eğer işleme çekirdek kümelerle başlanmışsa, kümelerin ortalama vektörü alınarak en yakın uzaklık ölçüsüne göre birimler en yakın kümeye atanır. Sonra yeni küme ortalamaları bulunur. Küme içi varyansın minimum, kümeler arası varyansın maksimum olduğu kümeleme yapısına ulaşıncaya kadar tüm birimler k-sayıda kümeye atanır.

### Küme Sayısının Belirlenmesi:

Kümeleme analizinde önemli bir sorun küme sayısının belirlenmesidir. Kümeleme algoritması uygulandıktan sonra kullanıcı veri setinde kaç tane küme olduğuna karar vermelidir. Küme sayısının belirlenmesi için çeşitli yöntemler önerilmiştir. “Bunlardan birisi, Marriot (1971)’un önerdiği, W küme içi kareler toplamı matrisi olmak üzere,  $k^2W$  değerini minimum yapan k değerini bulmaktır. Fakat, Everitt (1974) k’nın doğru değerini bulmada  $k^2W$  kriterinin başarısız olduğunu söyler. Diğer bir yöntem, Calinski ve Harabasz (1974)’ın önerdiği  $C = [\bar{Iz}(B)/(k-1)]/[\bar{Iz}(W)/(n-k)]$  kriteridir. Burada, B ve W sırasıyla kümeler arası ve küme içi kareler toplamı matrisleridir. k ile birlikte monoton olarak artan bir C değeri küme yapısının olmadığını belirtir, bunun yanı sıra, k ile birlikte monoton olarak azalan bir C değeri aşamalı bir yapı olduğunu belirtir. Bununla beraber, C’yi maksimum yapan k değeri küme sayısını belirtir.” (Seber, 1984, s. 388-389). “Basit bir kural olarak, aşamalı yöntemlerde, ağaç diyagramındaki (dendogram) büyük bir değişiklik, doğru küme sayısının belirleyicisi olarak kabul önerilir. Böyle bir yaklaşım gerekli, fakat yeterli değildir” (Seber, 1984, s. 388). “Küme ağaç diyagramı belli bir yükseklikten kesilmelidir. Küme sayısı ağaç diyagramının nereden kesileceğine bağlıdır. Mojena , aşamalı kümeleme yöntemleri için küme sayısını belirlemede, ağaç diyagramının nereden kesileceği konusunda “durma kuralı” veya “kesme yüksekliği” önermiştir. Bu kural, küme ağacının k-1 yüksekliğine dayalı tek-tarafli güven aralığına benzer. Mojena’nın “durma kuralı”  $\bar{h} + \alpha s_h$  olarak tanımlamıştır. Burada  $\bar{h}$  , tüm k-1 kümeler için ortalama yükseklik ve  $s_h$  , yüksekliklerin yansız standard sapmasıdır. Mojena, başlangıçta  $\alpha$ ’nın 2, 75 - 3, 50 sınırları içinde belirlenmesini önermiştir. Fakat, Milligan ve Cooper (1985) çalışmalarında, Mojena’nın “durma kuralı”nın en iyi performansı  $\alpha$ ’nın 1,25 değerinde ulaştığı sonucuna varmışlardır” (Sebert, 1998, s. 464-465). Araştırmacı veri setinin yapısına göre ağaç diyagramını inceleyerek  $\alpha$  değerini belirlemelidir.

Aşamalı olmayan kümeleme analizinde küme sayısına karar vermede ise, önce aşamalı kümeleme analizi yapılarak, ağaç diyagramından anlamlı kaç küme oluşacağını belirlemek ve bu sayıyı diğer yönetime uygulamaktır.

## 3. UYGULAMA

Analizde kullanılan çok değişkenli analiz yöntemlerinden faktör analizi ve kümeleme analizi sonuçları bu bölümde verilmiştir.

### 3. 1. FAKTÖR ANALİZİ SONUÇLARI

Çalışmada kullanılan veriler olarak 81 ile ait 10 sektörün GSYİH payları alınmıştır. Dolayısıyla, veri matrisi 81 ile ait 10 sektörü kapsamakta olup,  $X_{10 \times 81}$  boyutlu matristir.  $X_{i \times j}$  değerlerinin birimleri aynı olmasına rağmen, daha fazla yada daha az ağırlıklandırmaları önlemek için veri matrisi standartlaştırılarak  $Z_{10 \times 81}$  veri matrisi haline dönüştürülmüştür. Böylece girdi matrisi olarak standartlaştırılmış  $Z_{10 \times 81}$  matrisi

kullanılmıştır. Buradan korelasyon matrisi hesaplanarak, analizin ilk aşamasına ait bilgiler Tablo:1’de Açıklanan Toplam Varyans Tablosunda verilmiştir.

**Tablo 1: Açıklanan Toplam Varyans**

	Başlangıç Özdeğerleri			Döndürülmüş Kareli Yüklerin Toplamı		
Bileşenler	Toplam	Varyans Yüzdeleri	Kümülatif Yüzde	Toplam	Varyans Yüzdeleri	Kümülatif Yüzde
1	2, 730	27, 300	27, 300	2, 554	25, 543	25, 543
2	1, 675	16, 753	44, 054	1, 559	15, 593	41, 136
3	1, 329	13, 289	57, 343	1, 408	14, 079	55, 214
4	1, 069	10, 693	68, 035	1, 282	12, 821	68, 035
5	, 910	9, 095	77, 131			
6	, 798	7, 985	85, 115			
7	, 561	5, 612	90, 728			
8	, 435	4, 355	95, 083			
9	, 351	3, 513	98, 596			
10	, 140	1, 404	100, 000			

Başlangıç Metodu: Temel Bileşenler Analizi.

Özdeğeri 1’den büyük olan temel bileşenler dikkate alındığında, 4 temel bileşen elde edilmiştir. Bu bileşenler toplam değişkenliğin %68,035’ini açıklamaktadır. Temel bileşenler yöntemiyle elde edilen ilk çözüme dik döndürme yöntemlerinden Varimax yöntemi uygulanmıştır. Varimax yöntemiyle elde edilen Döndürülmüş Faktör Bileşenleri Matrisi Tablo:2’de verilmiştir.

**Tablo 2: Döndürülmüş Faktör Matrisi\***

	Faktörler			
	1	2	3	4
X10-S. Mes. ve Hiz.	<b>, 756</b>	-, 296	-, 152	-, 240
X9-Konut Sahipliği	<b>, 754</b>	, 240	-5, 237E-02	-6, 810E-02
X6-Ticaret	<b>, 751</b>	1, 808E-02	-5, 354E-02	1, 722E-02
X1-Tarım	<b>-, 669</b>	, 381	-, 472	3, 816E-02
X8-Mali Kuruluş.	<b>, 463</b>	9, 927E-02	-8, 229E-02	, 381
X5-İnşaat	, 232	<b>, 782</b>	5, 306E-02	-8, 589E-02
X3-İmalat	, 320	<b>, 782</b>	, 104	-, 351
X2-Madencilik	-, 135	-, 120	<b>, 808</b>	, 161
X4-Elektrik, gaz, su	-3, 134E-02	, 242	<b>, 695</b>	-, 423
X7-Ulaştır. ve Hab.	-, 107	7, 060E-02	1, 731E-02	<b>, 859</b>

\*Varimax Yöntemi ile elde edilmiştir.

Döndürmeden sonra elde edilen faktör matrisinin formatı temel bileşenler matrisinininkinden farklıdır. Burada da 4 faktör elde edilmiştir ve bu 4 faktörün toplam varyansı açıklama oranı %68'dir. Fakat burada, Tablo:2'den görüldüğü gibi daha açık, net ve basit faktör yapısı elde edilmiştir. Faktörler ise, sırasıyla toplam varyansın %25,5, %15,5, %14 ve %12,8'ini açıklamaktadırlar. Bu değerler Tablo:1'de yer almaktadır. Tablo: 2'den görüldüğü gibi toplam varyansın % 25,5'ini açıklayan Faktör 1'in X10 (serbest meslek ve hizmetler), X9 (konut sahipliği), X6 (ticaret), X8 (mali kuruluşlar) için yüksek pozitif yükleri ve X1 (tarım) için negatif yükleri vardır. Tarım değişkeni ile zıt bir ilişki olduğu söylenebilir. Böylece bu faktör tarımdan ziyade serbest meslek, konut sahipliği ve ticaret'te GSYİH değerlerini ölçer ve **“konut ve serbest ticaretin önemi”** faktörü olarak isimlendirilebilir. Faktör 2'nin X5 (inşaat) ve X3 (imalat) değişkenleri için pozitif ve yüksek oranda yükleri vardır. Şehirlerde imalat sektörü önemli bir alanı oluşturduğunda, buna paralel olarak inşaat sektörü de gelişme gösterecektir. Bu nedenle bu değişkenler bu sektörde endüstrinin önemini belirtmektedirler. Dolayısıyla bu faktör **“endüstrileşmenin önemi”** olarak isimlendirilebilir. Faktör 3'ün X2 (madencilik) ve X4 (elektrik, gaz, su) değişkenleri için pozitif yükleri vardır ve **“madencilik ve enerjinin önemi”** olarak isimlendirilebilir. Son olarak Faktör 4 için X7 (ulaştırma ve haberleşme) değişkeninin pozitif yükü vardır ve **“haberleşmenin önemi”** olarak isimlendirilebilir.

Tablo: 2'de verilen varimax yöntemi sonucunda elde edilen döndürülmüş faktör bileşenleri matrisini aşağıdaki şekilde yazmak mümkündür.

$$\begin{aligned}
 X_1 &= -0,669F_1 + 0,381F_2 - 0,472F_3 + 0,038F_4 \\
 X_2 &= -0,135F_1 - 0,120F_2 + 0,808F_3 + 0,161F_4 \\
 X_3 &= 0,320F_1 - 0,782F_2 + 0,104F_3 - 0,350F_4 \\
 X_4 &= -0,031F_1 + 0,242F_2 + 0,695F_3 - 0,423F_4 \\
 X_5 &= 0,232F_1 + 0,782F_2 + 0,053F_3 - 0,085F_4 \\
 X_6 &= 0,751F_1 + 0,018F_2 - 0,053F_3 + 0,017F_4 \\
 X_7 &= -0,107F_1 + 0,070F_2 + 0,017F_3 + 0,859F_4 \\
 X_8 &= 0,463F_1 + 0,099F_2 - 0,082F_3 + 0,381F_4 \\
 X_9 &= 0,754F_1 + 0,240F_2 - 0,052F_3 - 0,068F_4 \\
 X_{10} &= 0,756F_1 - 0,296F_2 - 0,152F_3 - 0,240F_4
 \end{aligned}$$

Yukarıdaki döndürülmüş faktörler (F'ler), G matrisi kullanılarak,

$$F = (G^1 G)^{-1} G^1 X$$

eşitliğinden (Bryan, 1986, s. 78)  $X_i$  değişkenlerinin doğrusal kombinasyonu olarak ifade edilebilir. Burada,

$F^1 = (F_1, F_2, F_3, F_4)$ ,  $X^1 = (X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10})$  ve G matrisi de  $10 \times 4$  boyutlu döndürülmüş faktör yükleri matrisidir. Bu eşitlikler,



$$F_1 = -0, 255X_1 - 0, 034X_2 + 0, 053X_3 - 0, 020X_4 + 0, 140X_5 + 0, 309X_6 + 0, 036X_7 + 0, 228X_8 + 0, 318X_9 + 0, 266X_{10}$$

$$F_2 = 0, 188X_1 - 0, 072X_2 - 0, 445X_3 + 0, 236X_4 + 0, 554X_5 + 0, 058X_6 - 0, 050X_7 + 0, 053X_8 + 0, 216X_9 - 0, 131X_{10}$$

$$F_3 = -0, 336X_1 + 0, 599X_2 + 0, 011X_3 + 0, 471X_4 + 0, 067X_5 - 0, 022X_6 + 0, 108X_7 - 0, 005X_8 - 0, 021X_9 - 0, 136X_{10}$$

$$F_4 = -0, 119X_1 + 0, 225X_2 - 0, 175X_3 - 0, 306X_4 - 0, 129X_5 + 0, 072X_6 + 0, 705X_7 + 0, 340X_8 - 0, 023X_9 - 0, 121X_{10}$$

Veri değerleri kullanılarak, yukarıdaki eşitliklerden elde edilen 81 il için faktör skorları Tablo 3’de verilmiştir. (Veri değerleri,  $X_i$  ‘ler sıfır ortalaması ve 1 standard sapma değerleri olarak standardize edilmiştir).

**Tablo 3: Faktör Skorları**

	F1	f2	F3	f4
Adana	0, 78175	-1, 12794	0, 12096	0, 57517
Adıyaman	-0, 6956	-0, 62764	2, 50346	0, 71072
Afyon	-0, 25185	0, 17712	-0, 40833	0, 62666
Ağrı	-1, 25133	0, 88212	-1, 01694	-0, 6018
Amasya	-0, 05583	0, 02809	-0, 42933	0, 90915
Ankara	3, 48956	2, 58559	0, 40019	0, 85122
Antalya	1, 51248	0, 84175	-0, 15173	0, 46124
Artvin	-0, 67174	-0, 70995	0, 83264	1, 83108
Aydın	0, 87558	0, 62375	-0, 83168	-0, 63671
Bahkesir	0, 07554	-0, 35322	-0, 43643	0, 27566
Bilecik	-0, 53759	-1, 07144	0, 00072	-0, 30765
Bingöl	-1, 26455	0, 48623	-0, 3337	-1, 13797
Bitlis	-1, 25218	0, 01132	-0, 95692	-0, 49374
Bolu	-0, 99753	-0, 45575	-0, 68885	-0, 10821
Burdur	0, 09704	-0, 41602	-0, 2751	1, 85234
Bursa	0, 68179	-0, 86447	0, 14919	-0, 56004
Çanakkale	-0, 11841	-0, 53574	-0, 25409	0, 12077
Çankırı	0, 54947	0, 652	-0, 42911	0, 43582
Çorum	1, 19701	0, 44378	-0, 38265	0, 32372
Denizli	0, 76392	0, 10155	-0, 16112	0, 21329
Diyarbakır	-0, 15341	0, 74825	1, 36177	-0, 3051
Edirne	0, 23546	-0, 0571	-0, 82406	0, 41098
Elazığ	-0, 11238	1, 13866	1, 04103	-1, 24166
Erzincan	-0, 92372	-0, 4602	-0, 58305	0, 45773
Erzurum	0, 0653	0, 79494	-0, 37443	-0, 40958
Eskişehir	0, 69488	-0, 49454	1, 82985	0, 6442
Gaziantep	2, 21846	0, 697	-0, 10375	-0, 6114
Giresun	1, 54652	0, 67291	-0, 1909	3, 05298
Gümüşhane	-1, 13488	0, 93709	-0, 55784	1, 3919
Hakkari	-1, 18989	1, 39604	-0, 3911	-0, 99312
Hatay	0, 25545	-0, 75344	-0, 57364	0, 43932
Isparta	0, 04588	0, 3983	-0, 56124	-0, 21786
İçel	0, 70873	-1, 03304	-0, 61244	-0, 0445
İstanbul	3, 34262	-0, 77483	-0, 03547	0, 6447
İzmir	1, 58324	-0, 78794	-0, 0808	-0, 31571
Kars	-0, 64545	0, 63345	-0, 62235	-0, 35422
Kastamonu	-0, 73778	0, 14897	-0, 0616	1, 07161

Kayseri	0, 99268	0, 64335	0, 11313	-0, 39468
Kırklareli	-0, 49496	-0, 19861	2, 64743	-2, 39614
Kırşehir	0, 02201	0, 18276	-0, 07977	0, 92032
Kocaeli	-0, 23552	-2, 5908	0, 24548	-0, 95131
Konya	0, 42329	0, 56258	0, 04936	0, 45419
Kütahya	-0, 53226	0, 25159	4, 58907	0, 32538
Malatya	0, 19549	-0, 7151	-0, 26522	-0, 42691
Manisa	-0, 08165	-1, 07824	0, 87046	-0, 87222
K. Maraş	0, 31574	2, 97581	1, 86007	-1, 47349
Mardin	-0, 68918	-0, 47356	-0, 08779	1, 70318
Muğla	0, 05607	0, 94556	1, 2544	-0, 74348
Muş	-1, 49505	0, 85605	-1, 03789	-0, 98551
Neveşehir	-0, 62884	-0, 17211	-0, 68841	0, 77419
Niğde	-1, 30482	0, 64917	-1, 06135	-0, 41216
Ordu	-0, 19709	0, 49113	-0, 71751	0, 20546
Rize	-0, 22386	-0, 76293	0, 84194	1, 85696
Sakarya	0, 03426	-0, 24851	-0, 54865	-0, 65527
Samsun	0, 24825	0, 59939	0, 26221	-0, 94728
Siirt	-0, 90401	-0, 59949	0, 2768	-0, 11022
Sinop	-0, 20685	0, 01152	-0, 43284	1, 39793
Sivas	-0, 12576	0, 96518	1, 55602	-0, 11802
Tekirdağ	0, 00085	0, 68963	1, 43666	-1, 72349
Tokat	-0, 35795	-1, 40531	-0, 33915	-0, 59075
Trabzon	0, 49707	0, 20945	-0, 20059	0, 27026
Tunceli	-1, 07695	0, 60467	-0, 324	-0, 25257
Şanlıurfa	-0, 83864	1, 58249	0, 09295	-1, 28222
Uşak	0, 71234	0, 35651	-0, 53982	-0, 1152
Van	-0, 7066	-0, 31476	-0, 25507	0, 55428
Yozgat	-0, 78138	-0, 32674	-0, 53146	0, 7201
Zonguldak	-0, 38212	-2, 49342	1, 29364	0, 31912
Aksaray	-0, 53552	0, 66891	-0, 74225	0, 8305
Bayburt	-0, 50794	0, 3872	-0, 4744	1, 23769
Karaman	-1, 31834	0, 26964	-1, 02199	-0, 98928
Kırıkkale	0, 98484	-2, 93809	-0, 61192	-2, 08051
Batman	-0, 94348	-1, 26564	2, 40244	0, 70773
Şirnak	-1, 23741	-0, 45244	0, 07549	1, 20834
Bartın	-0, 27011	0, 2231	-0, 00922	1, 77742
Ardahan	-1, 27118	0, 37649	-1, 21029	-0, 5552
Iğdır	-0, 65896	1, 72137	-0, 6254	0, 03009
Yalova	1, 72921	-1, 80122	-0, 60895	-1, 39516
Karabük	1, 41309	-0, 63073	-0, 2816	-1, 33991
Kilis	0, 96119	-0, 38561	-1, 25725	-1, 06906
Osmaniye	1, 34295	0, 33396	-1, 03828	-1, 34726
Düzce	-0, 64947	-1, 58171	-0, 78764	-1, 02681

Tablo 3'den görüldüğü gibi Faktör 1 Ankara, İstanbul ve Gaziantep'de konut ve serbest ticaretin önemini işaret eder. Faktör 2 değerleri K. Maraş ve Ankara gibi illerde endüstrileşme alanında GSYİH'nın sektör paylarının daha fazla etkili olduğunu ve Kırıkkale, Kocaeli ve Zonguldak gibi illerde ise ters yönde olduğunu göstermektedir. Faktör 3 Kütahya, Kırklareli ve Batman illerinde madencilik ve enerji sektöründe GSYİH'nın sektör paylarının daha fazla olduğunu göstermektedir. Son olarak Faktör 4 Giresun ilinde haberleşme alanında GSYİH'nın sektör paylarının daha etkili olduğunu göstermektedir.

### 3. 2. KÜMELEME ANALİZİ SONUÇLARI

Çalışmada ele alınan verilere kümeleme ikinci aşamada kümeleme analizi uygulanmıştır. Önce birimlere (gözlem) göre kümeleme analizi yapılmıştır. Aşamalı kümeleme yöntemlerinden ortalama bağlantı tekniği seçilmiş ve küme uzaklığı olarak Öklid uzaklığı kullanılmıştır. Sonuçları ağaç diyagramı olarak Şekil:1’de verilmiştir. Küme yükseklikleri ise Tablo: 4’de yer almaktadır. Ağaç diyagramının nerede kesileceğine karar vermek için bölüm: 2.2’de verilen Mojena’nın “durma kriteri” kullanılmıştır.  $\alpha=2$  alınarak ağaç diyagramı 26, 02 noktasından kesilmesine karar verilmiştir. Böylece 4 tane küme oluşmaktadır. Daha sonra aşamalı olmayan kümeleme yöntemlerinden k-ortalamar tekniği uygulanmıştır. Küme sayısı ortalama bağlantı tekniği sonucuna göre 4 olarak belirlenmiştir. Son küme merkezleri arasındaki Öklid uzaklıkları Tablo:5’de verilmiştir. Tablo: 5’den görüldüğü gibi 1 ve 2. kümelerin birbirinden uzaklığı 23, 083; 1 ve 3. kümelerin 16,651; 1 ve 4. kümelerin 30, 543; 2 ve 3. kümelerin 29, 440; 2 ve 4. kümelerin 39, 936; 3 ve 4. kümelerin ise 16, 830 ‘dur. Bu değerler her küme için değişkenlerin ortalama değerlerini göstermektedir.

Tablo: 6’da varyans analizi yer almaktadır. Küme profili bu tablodan incelenebilir. F-değerleri küme ortalama kareler ile hata ortalama karelerin oranıdır. Buradan görüldüğü gibi bu dört kümenin oluşumunda en etkili değişkenin X3-İmalat olduğu görülmektedir, bu değişkene ait F-değeri 78, 368’dir. Buna göre, farklı kümeler içinde iller, imalat değişkenine göre geniş şekilde ayrılmıştır, fakat aynı küme içinde iller küçük bir değişime gösterirler. Diğer etkili değişkenlerin X1-Tarım, X6-Ticaret, X9-Konut Sahipliği ve X10-Serbest Meslek ve Ticaret olduğunu söyleyebiliriz. Tablo:7’de ise hangi illerin hangi kümelerde yer aldığı görülmektedir.

**Tablo 4: Küme Yükseklikleri**

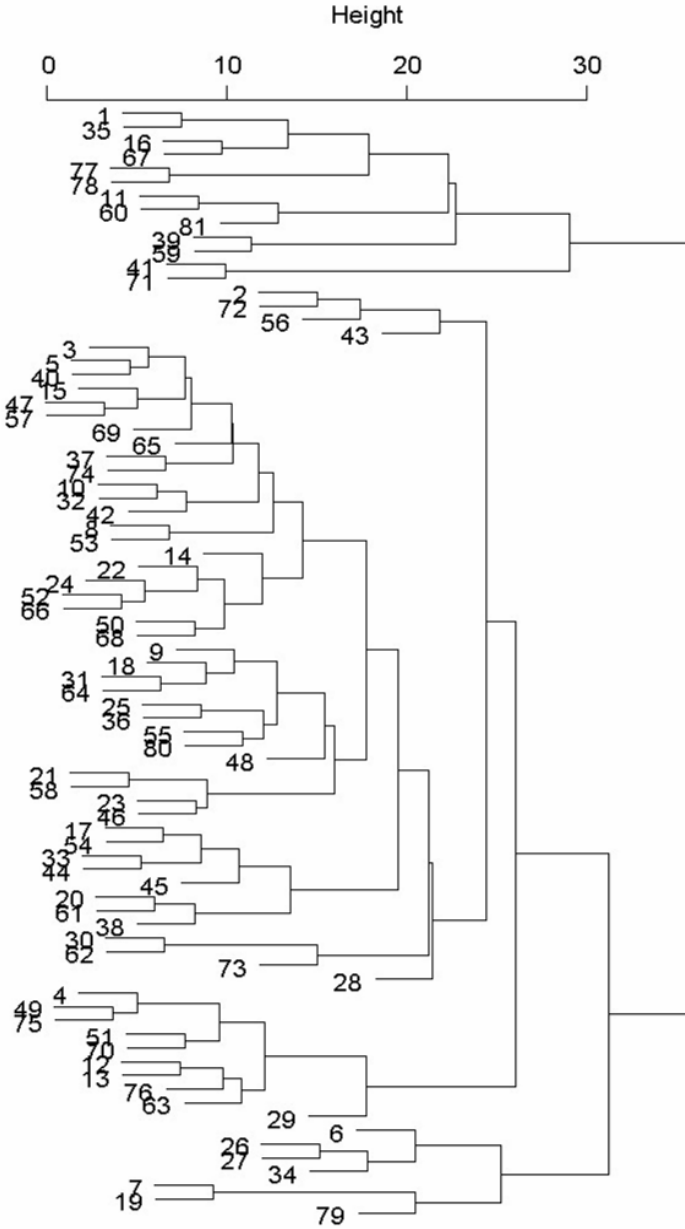
[1]	7. 457211	13. 382510	9. 698969	17. 878297	6. 789698	22. 341084	8. 417838
[8]	12. 858694	22. 772452	11. 377170	29. 083464	9. 932271	35. 781706	15. 015658
[15]	17. 418630	21. 859483	24. 454590	5. 605349	4. 596738	7. 641551	5. 022186
[22]	3. 178050	8. 037423	10. 293677	10. 352318	6. 607571	11. 761519	6. 078651
[29]	7. 714425	12. 569881	6. 766092	14. 211766	11. 965443	8. 318794	5. 426580
[36]	4. 101219	9. 823196	8. 199390	17. 790376	10. 434819	8. 831500	6. 315853
[43]	12. 807646	8. 549269	12. 034392	10. 864621	15. 415990	16. 014277	4. 536518
[50]	8. 930122	8. 305420	19. 535010	6. 471476	8. 582627	5. 192302	10. 665697
[57]	13. 538842	5. 955670	8. 193616	21. 270555	6. 500769	15. 022491	21. 472130
[64]	26. 045912	5. 001493	3. 631804	9. 548834	7. 651144	12. 144302	7. 407429
[71]	9. 817564	10. 830690	17. 765844	31. 254568	20. 463699	15. 141995	17. 824000
[78]	25. 251692	9. 234717	20. 514237				

**Tablo 5: Kümeler Arası Uzaklık**

Kümeler	1	2	3	4
1		23, 083	16, 651	30, 543
2	23, 083		29, 440	39, 936
3	16, 651	29, 440		16, 830
4	30, 543	39, 936	16, 830	

**Tablo 6: Varyans Analizi**

	Küme		Hata		F-Değeri	Olasılık
	Ortalama Kareler	Ser. Derecesi	Ortalama Kareler	Ser. Derecesi		
X1-Tarım	2353, 427	3	37, 607	77	62, 579	, 000
X2-Madencilik	41, 804	3	14, 277	77	2, 928	, 039
X3-İmalat	2737, 385	3	34, 930	77	78, 368	, 000
X4-Elekt. , gaz, su	36, 200	3	15, 284	77	2, 368	, 077
X5-İnşaat	5, 258	3	3, 255	77	1, 615	, 193
X6-Ticaret	813, 782	3	38, 739	77	21, 007	, 000
X7-Ulaş. ve Hab.	128, 389	3	25, 692	77	4, 997	, 003
X8-Mali Kuruluş.	4, 271	3	5, 192	77	, 823	, 485
X9-Konut Sahip.	13, 061	3	1, 915	77	6, 819	, 000
X10-S. Mes. ve Hiz.	16, 392	3	1, 701	77	9, 636	, 000

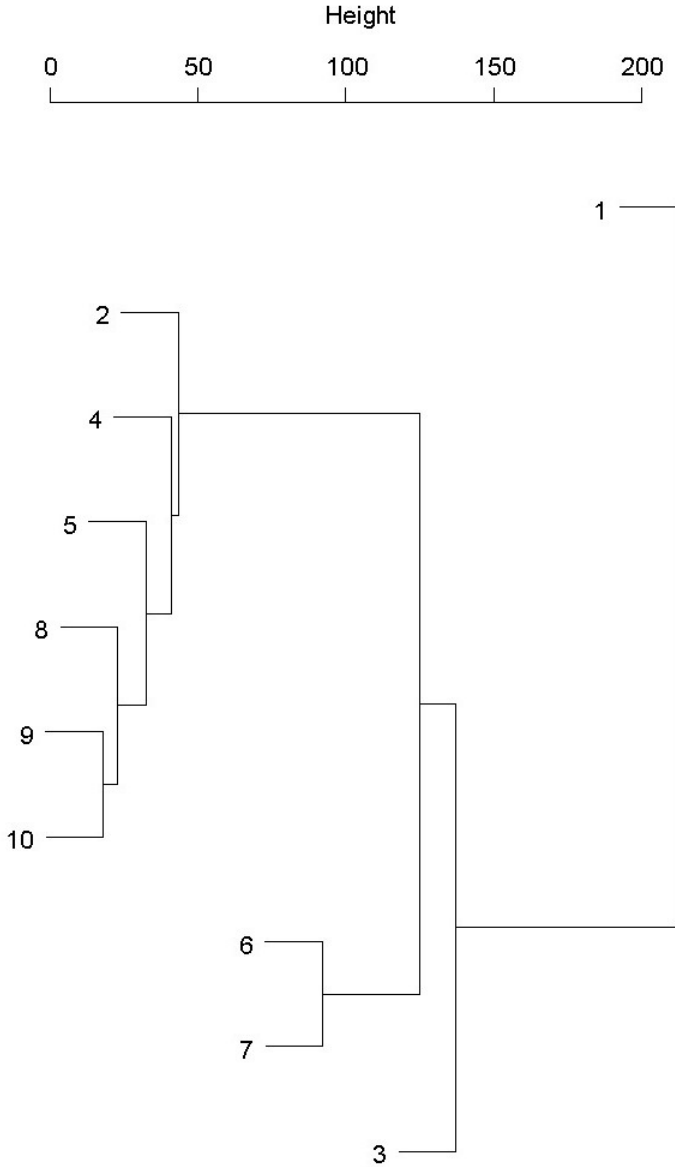


Şekil 1: İllere (birimlere) Göre Ağaç Diyagramı

**Tablo 7: Küme Elemanları**

<b>Küme No.</b>	<b>İller</b>
1	Adana, Ankara, Antalya, Çorum, Denizli, Eskişehir, Gaziantep, Hatay, İçel, İstanbul, Kayseri, Malatya, Sakarya, Trabzon, Kilis
2	Bilecik, Bursa, İzmir, Kırklareli, Kocaeli, Manisa, Tekirdağ, Tokat, Zonguldak, Kırıkkale, Yalova, Karabük, Düzce
3	Adıyaman, Afyon, Amasya, Artvin, Aydın, Balıkesir, Burdur, Çanakkale, Çankırı, Diyarbakır, Elazığ, Erzurum, Giresun, Isparta, Kastamonu, Kırşehir, Konya, Kütahya, K. Maraş, Mardin, Muğla, Rize, Samsun, Siirt, Sinop, Sivas, Tunceli, Uşak, Van, Batman, Şırnak, Bartın, Osmaniye
4	Ağrı, Bingöl, Bitlis, Bolu, Edirne, Erzincan, Gümüşhane, Hakkari, Kars, Muş, Nevşehir, Niğde, Ordu, Şanlıurfa, Yozgat, Aksaray, Bayburt, Karaman, Ardahan, Iğdır

Daha sonra kümeleme analizi değişkenlere göre tekrarlanmıştır. Burada da, aşamalı kümeleme yöntemlerinden ortalama bağlantı tekniği kullanılmıştır. Bu analizin sonuçları Şekil:2’de ağaç diyagramı olarak verilmiştir. Buradan görüldüğü gibi, faktör analizi sonuçlarından farklı olarak, X1-tarım ve X3-imalat değişkenleri ayrı birer küme oluşturmaktadır. Bu değişkenlerin tek başına önemli olduğu söylenebilir. X6-ticaret değişkeni ile X7-ulaş. ve haberleşme değişkenleri de bir arada bir küme oluşturmuştur. Geriye kalan diğer değişkenlerde ayrı bir küme teşkil etmektedir.



**Şekil 2: Sektörlere (değişkenlere) Göre Ağaç Diyagramı**

#### 4. SONUÇ

Bu çalışmada, Türkiye’de illerin GSYİH bakımından, ele alınan 10 sektöre yaptığı katkı incelenmiş, faktör analizi ve kümeleme analizi yöntemleri uygulanarak,

sektörler ve iller gruplandırılmaya çalışılmıştır. İlk uygulama olan faktör analizi sonuçlarında, toplam varyansın dörtte birini açıklayan faktör 1'in içinde yer alan tarım değişkeni negatif bir etki yapmaktadır; buna karşılık ikinci uygulama olan kümeleme analizi sonucunda tarım değişkeninin tek başına bir grup oluşturduğu görülmektedir. Aynı şekilde imalat değişkeni kümeleme analizi sonucunda tek olarak bir grup oluşturmaktadır. Diğer taraftan kümeleme analizinde, küme oluşumunda etkili olan ticaret, konut sahip. ve, s. meslek. ve hiz. 'ler değişkenlerinin faktör analizinde faktör 1'de yer aldığı görülmektedir. Tablo 3'de faktör skorları tablosunda, faktör 1'in (konut ve serbest ticaretin önemi) Ankara, İstanbul ve Gaziantep'te önemli olduğu, faktör 2'nin (endüstrileşmenin önemi) Kırıkkale, Kocaeli ve Zonguldak'ta ters yönde etkili olduğu görülmekte; kümeleme analizinde de bu illerin aynı şekilde gruplandığı görülmektedir.

Bu bulgular, Türkiye'nin ulusal teknoloji politikalarının oluşturulmasında iller arasındaki farklılığı göz önüne almak açısından önemli olmaktadır. Diğer taraftan kamu ve özel sektör yatırımlarının nasıl dağılacığı ve teşvik indirimlerinin uygulanması konularında yardımcı olacaktır.

#### KAYNAKÇA

- Bharati, P. (2003). **Critique of the 'Digital World' Myth: An Empirical Investigation**. <http://www.mngt.waikato.ac.nz/research/ejrot/cmsconference/2003/proceedings/technology/Bharati.pdf>.
- Bryan F. J. Manly. (1986). **Multivariate Statistical Methods. A Primer**. Chapman and Hall.
- Cluster Analysis**. (1984-2004). <http://www.statsoft.com/textbook/stcluan.html>.
- Dincer, B., M. Özasan, T. Kavasoglu. **İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması (2003)**. Yayın No: DPT2671. <http://ekutup.dpt.gov.tr/bolgesel/gosterge/2003-05.pdf>.
- Exploratory Spatial Data Analysis of the distribution of regional per capita GDP in Europe**. 1980-1985. [http://ungaro.u-bourgogne.fr/ertur/e2000-09.\(27/08/2001\)](http://ungaro.u-bourgogne.fr/ertur/e2000-09.(27/08/2001)).
- Johnson, R. A. and D. W. Wichern. (1992). **Applied Multivariate Statistical Analysis**. Prentice-Hall International Inc.
- Kim, J. O. and C. W. Mueller. (1978). **Introduction to Factor Analysis**. Sage Publications. Beverly-Hill-London.
- Öğüş, E. (1985). **Çeşitli Kümeleme Yöntemlerinin Karşılaştırılması**. Doktora Tezi. A. Ü. Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı. Ankara.
- Özdamar, K., (2002). **Paket Programlar ile İstatistiksel Veri Analizi-2**. Kaan Kitabevi. Eskişehir.
- Pamuk, M. (1997). **Faktör Analizi Yöntemi ve Ekonomik Verilere Uygulanışı**. Doktora Tezi. İ. Ü. İktisat Fakültesi. İstanbul.
- Seber, G. A. F. , (1984). **Multivariate Observations**. John Wiley & Sons.
- Sebert, D. M., D. C. Montgomery, and D. A. Rollier. (1998). "A Clustering Algorithm for Identifying Multiple Outliers in Linear Regression". **Computational Statistics & Data Analysis**. 27. s. 461-484.
- T. C. Başbakanlık Devlet İstatistik Enstitüsü, **İllere göre Gayri Safi Yurtiçi Hasıla (2000)**.
- Tatlıdil, H. (1992). **Uygulamalı Çok Değişkenli İstatistiksel Analiz**. Ankara.