PARADOKS Ekonomi, Sosyoloji ve Politika Dergisi
PARADOKS Economics, Sociology and Policy Journal

Categorisation of Turkish Cities as Per Their Education and Health Indicators

Türkiye'deki İllerin Eğitim ve Sağlık Göstergelerine Göre Sınıflandırılması

Dr. Burcu AVCI ÖZTÜRK
Uludag University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Quantitative Methods

Doç.Dr. Sevda GÜRSAKAL
Uludag University, Faculty of Economics and Administrative Sciences, Department of Econometrics, Statistics

Temmuz/July 2015, Cilt/Vol: 11, Sayı/Num: 2, Page: 83-101
ISSN: 1305-7979

PARADOKS Ekonomi, Sosyoloji ve Politika Dergisi
PARADOKS Economics, Sociology and Policy Journal
Temmuz/July 2015, Cilt/Vol: 11, Sayı/Num: 2
ISSN: 1305-7979

PARADOKS Ekonomi, Sosyoloji ve Politika Dergisi
PARADOKS Economics, Sociology and Policy Journal
Temmuz/July 2015, Cilt/Vol: 11, Sayı/Num: 2
ISSN: 1305-7979

PARADOKS Ekonomi, Sosyoloji ve Politika Dergisi
PARADOKS Economics, Sociology and Policy Journal

Temmuz/July 2015 - Cilt/Vol: 11 - Sayı/Num: 02
Sayfa/Page: 83-.101

CATEGORISATION OF TURKISH CITIES AS PER THEIR EDUCATION AND HEALTH INDICATORS

TÜRKİYE'DEKİ İLLERİN EĞİTİM VE SAĞLIK GÖSTERGELERİNE GÖRE SINIFLANDIRILMASI

Dr. Burcu AVCI ÖZTÜRK
Uludag University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Quantitative Methods

Doç.Dr. Sevda GÜRSAKAL
Uludag University, Faculty of Economics and Administrative Sciences, Department of Econometrics, Statistics

Abstract:

In this study, 81 Turkish cities were categorized as per their education and health indicators. First, the variables, which are education and health indicators, were categorized using factor analysis. Next, multi-dimensional scaling analysis was carried out in order to determine the position of 81 cities in the multi-dimensional space via the factor loading obtained through the factor analysis. MDS analysis revealed that the positions of the cities in the multi-dimensional space were not that different from their positions in Turkey's map, and the difference was observed in only five big cities.

Key Words: *Socio-economic development, MDS Analysis, Factor Analysis*

Özet:

Bu çalışmada Türkiye'deki 81 il eğitim ve sağlık göstergeleri dikkate alınarak sınıflandırmaya tabi tutulmuştur. Çalışmada ilk olarak eğitim ve sağlık göstergesi olarak kullanılan değişkenler faktör analizi kullanılarak sınıflandırılmış, ardından da elde edilen faktör yükleri kullanılarak 81 ilin çok boyutlu uzaydaki konumlarını ortaya koymak amacıyla Çok Boyutlu Ölçekleme Analizi yapılmıştır. Yapılan ÇBÖ analizi sonucunda ise illerin çok boyutlu uzaydaki konumları Türkiye haritasından çok da farklı çıkmamış, farklılık sadece beş büyük il bazında olmuştur.

Anahtar Kelimeler: *Sosyo-ekonomik Gelişmişlik, ÇBÖ Analizi, Faktör Analizi*

1. INTRODUCTION

Any kind of inequality observed among countries and among regions and cities of a given country is qualified as interregional imbalance. The development levels of different regions in a country are determined by economic, social, historical, geographical, climatic and strategic factors (Şahin, 2002: 539). The most important reason for interregional imbalance is that economic resources are not evenly distributed and that there is no balance in equality of opportunities and income distribution. Interregional imbalance gives rise to inequality of social opportunities leading to the difference in socio-economic development (Gürbüz, Karabulut and Sandal, 2005: 159).

Interregional differences in socio-economic development have been observed throughout the history of the world, but such differences began to increase with the Industrial Revolution in the 18th century, and the gap widened after the World War Two. In today's world, the differences in socio-economic development can be observed in almost all countries and regions, yet such differences are much more apparent in developing countries than in developed countries (Dinler, 2002: 120). Therefore, it can be said that the development of a country's cities and regions also reflects the development of that country as a whole.

The development level of a country is closely related to the development of its cities, which are its important settlement areas. In Turkey, as in every country, there are differences among cities in terms of their socio-economic development levels. Such differences can be traced back to the Ottoman Empire. The ports of the Empire (İzmir, Mersin, Trabzon, etc.), which connected the country to European countries, came to the forefront as a result of the commercial privileges granted to Western countries while the cities in Central Anatolia and in the East and South-East of Anatolia gradually lost their economic importance (Dinler, 2002: 174).

Cities are not only administrative but also socio-economic units forming the sub-dwelling units in a country. Thus, they have to be the starting point of any planned development in a country. There are a great many factors that affect and determine the socio-economic structure of cities. These factors have been categorized in three main groups, which are social, economic and geographical. Demographic indicators and indicators related to education, health, employment and social security are grouped under social indicators. Economic indicators cover fiscal and financial variables and variables related to manufacturing industry, agriculture, foreign trade, energy and infrastructure. Geographical indicators include the variables such as altitude, proximity to the sea, and climate (Albayrak, 2005: 153).

This study first offers a literature review summarizing the related studies carried out in Turkey. Then, it provides theoretical information on factor analysis and multi-dimensional scaling analysis, which constitute the statistical analysis part of this study. The last part of this study covers the findings and the conclusions.

2. LITERATURE REVIEW

State Planning Organization (SPO) conducted a survey titled "Socio-Economic Development Ranking of Cities and Regions". In this survey, which ranked cities in terms of their socio-economic development, principal component analysis was applied and 81 provinces were first examined using 58 socio-economic variables. In this ranking, the loading obtained from the initial principal component analysis was used. Next, these 81 cities were categorized in five groups based on the ranking (SPO, 2003). The Ministry of Development updated this survey and analyzed the 81 cities using 61 socio-economic variables in addition to the principal component analysis. The cities were again ranked using the loads from the principal component analysis. After that, 81 cities were categorized under five groups based on this ranking (SDI, 2011). The study Socio-Economic Developmet Index (SDI) 2011

measured the socio-economic development level of cities and regions, which had already been measured in 2003, with a new set of data and using a more advanced method. Recently, the concept of regional development has been handled from a multi-dimensional perspective and in addition to economic and social development, concepts like technological development and quality of life have gained importance. As a result, SDI (2011) has had to use a different set of data. In SPO (2003) study, 58 variables from ten different fields had been used. SDI (2011), however, some of these 58 variables were omitted as they were no longer up-to-date and their production had ceased. Using a new set of data, the ultimate SDI ranking provided ranks on the basis of cities and Level-2.

There are numerous studies on the socio-economic indicators of cities and various sets of data have been used to carry out such studies. For example, Göçer and Çıracı (2003) carried out a study titled "The Relation Between Social and Economic Indicators of Turkish Cities", and in this study the cities were ranked using the factor loading obtained from the factor analysis. In a study by Özdemir and Altıparmak (2005), variables factor analysis was used and the variables were categorized in three groups, which are health, education and ratio of schooling at primary/secondary level. 81 cities were ranked based on the factor loading obtained out of these factors. In Albayrak's (2005) study, however, the variables were categorized in 8 groups and the cities were ranked based on the factor loading of these groups. Next, the cities were divided into three groups: developed, developing and under-developed. After that, discriminant analysis was carried out in line with this grouping and the classification percentages were checked. In another study by Gürbüz, Karabulut and Sandal (2005), 58 socio-economic variables were used and the cities that were identical in terms of socio-economic variables were categorized using hierarchical clustering analysis. Kaygısız, Saraçlı and Dokuzlar (2005) conducted a study with the aim of determining the factors that affect the development level of cities and they used Path analysis and clustering analysis. Şen, Çemrek and Özaydın (2006), applied

factor analysis and principal component analysis in order to rank the 81 cities as per their development level and they used 28 socio-economic variables. They concluded that the most developed city is İstanbul and the least developed one is Bilecik. Kavasoğlu (2007) tried to determine the socio-economic development of cities on the basis of selected socio-economic variables and aimed to create development ranks as per geographical regions considering the cities that had similar features. Kılıç, Saraçlı and Kolukısaoğlu (2011) examined the regional similarities among Turkish cities in terms of socio-economic indicators using clustering and multi-dimensional scaling analysis. Their research revealed that the results of clustering analysis and the results of multi-dimensional scaling analysis were similar. Yıldız, Sivri and Berber (2012) identified the socio-economic development ranking of cities using 2010 data. They compared their results with the study conducted by State Planning Organization in 2003. Albayrak and Karamustafa (2013) found out the most important determiners of socio-economic development level using factor analysis and they examined the 2012 socio-economic development level of Turkish cities using principal component analysis. Çelik (2013) categorized cities as per their health indicators using clustering analysis. Erilli (2014) used fuzzy clustering method and categorized districts as per their development level with the help of socio-economic indicators. He divided the cities in TR72 zone into four groups.

Different from other cities, in this study multi-dimensional scaling analysis was used and the position of 81 cities in the multi-dimensional space was determined based on their education and health indicators. The shape obtained at the end of the study was the map of Turkey.

3. STATISTICAL METHODS USED IN THE RESEARCH

3.1 Factor Analysis

Factor analysis is a multivariate statistical analysis that brings together interrelated variables in a p-dimensioned space in order to find out fewer, new and

unrelated variables (Tatlıdil, 2002: 167). Basically, the variables which are grouped under one factor have a high correlation among themselves and their correlation with the variables in the other group is relatively lower (Johnson and Wichern, 1998: 514-515).

The targeted formula in a factor analysis can be defined as Z = AF + BU in matrix form. In this formula, Z refers to pxn dimensioned standardized datum matrix. A signifies the pxm dimensioned loading matrix. F stands for mxn dimensioned factor matrix. B signifies pxp dimensioned diagonal coefficient matrix and U refers to pxn dimensioned special factor matrix.

To determine the appropriate number of factors, some methods like Kaiser criterion, slope test, disclosed variance test, Joliffe criterion and comprehensibility are used. In this study, Kaiser and disclosed variance criteria were taken into consideration. The appropriate number of factors was determined considering the condition that the number should be equal to the number of roots that are bigger than one in the correlation matrix or that the cumulative variance disclosed by latent values should be minimum 67%.

Besides, factor rotation is used to ensure conceptual meaningfulness. In this study, Verimax factor rotation was used to reach conceptual meaningfulness.

Finally, it should be noted that KMO sampling adequacy test was applied to check the appropriateness of the data -provided at the beginning of the analysis- for the factor analysis.

3.2. Multi-dimensional Scaling Analysis

Multi-dimensional scaling analysis is very frequently used in the analysis of behavioral data such as preferences, attitudes, inclinations and expectations. The objective of this analysis is to draw maps that show the relations between objects by looking at the distances between them. Another objective is to produce shapes that are as close as possible to the original shapes by using minimum number of

dimensions (Tatlıdil, 2002: 353). Although this method is among Q analysis methods, it is also regarded as an R type analysis technique when metric sampling methods are applied. In fact, if the matrix of differences has been obtained by using the Euclidian distances from the data matrix, metric scaling values correspond to the score values obtained from the data matrix in the principal components analysis (Tatlıdil, 2002: 367).

In multi-dimensional scaling analysis, stress value stands an important measure in deciding on the appropriateness of the number of dimensions. This scale is also used as the goodness of fit scale. Stress value is calculated by taking the square roots of the normalized residual roots. It is said that the closer the stress value to zero is, the more similar the obtained shape to the original shape is. To be more specific;

Sk ≥ 0,20 means weak

Sk ≤ 0,10 means mediocre

Sk ≤ 0,05 means good

Sk = 0,00 means perfect fit.

In multi-dimensional scaling, different methods are applied depending on which scale is used to measure the data. The method which uses the serial numbers of the distances between objects is called as non-metric scaling method whereas the method that shows the position of a point by using distance values is called as metric scaling method. In addition to these, there is also a semi-metric method in which the data are obtained through sequencer scale and which assumes that metric outputs will be obtained at the end. Briefly, if the data have been measured using a categorizing or sequencing method, non-metric system is applicable. However, if the data have been measured with uniform or proportional scale, metric method is applicable.

In this study, the data used in multi-dimensional scale are factor scores. Therefore, considering the proportionally scaled data group, metric method was

applied. Besides, the data were converted by using the Euclidian distances as there was no direct distance matrix.

Since the objective was to map the cities as per their education and health indicators, the compatibility between the final map drawn as a result of the analysis and the original distances in the set of data was assessed using stress values. Finally, as the main objective of MDS analysis is to measure the similarity between the estimated distances based on the matrix of differences and the matrix of distances obtained from direct data, the similarity between these two distance matrixes was checked with scatter diagram modeled according to Euclidian distances. That this diagram shows a linear relation means that the model has the ideal goodness of fit (Kalaycı, 2005: 396).

4. FINDINGS

In this study, the 2013 regional statistics concerning 81 cities and provided by Turkish Statistics Institute were used and a total of 9 variables obtained out of these statistics were used. These variables can be seen in the table below.

At first, the multi-dimensional scaling analysis did not help to obtain meaningful conclusions when all of the 9 variables mentioned in Table 1 were used. As a result, the number of variables was decreased through dimension reduction and reduction of the number of variables, which are the features offered in a factor analysis.

Table 1: Variables Used in the Research

Number of people per doctor	Schooling ratio in pre-school education
Number of people per nurse	Net schooling ratio at primary school level
Number of people per pharmacist	Net schooling ratio at secondary school level
Number of people per dentist	Net schooling ratio at vocational high school level
Number of people per midwife	

First of all, Bartlett test was applied to understand whether the variables were appropriate for factor analysis. It was concluded that these variables were suitable for factor analysis.

Table 2: KMO and Bartlett Tests

KMO	0.865
Bartlett	3821.345(p = 0.000)

KMO statistics was higher than 0.50, which shows that the sampling was adequate. As a result of the Bartlett test, the zero hypothesis that the sets of data on which factor analysis was to be applied did not have correlation was rejected (p = 0.0000), which showed that factor analysis was an appropriate method.

Before the number of factors was decided, the number of latent values bigger than one and the ratio of explained variances were considered. As the table below shows, there are two latent values bigger than one and the percentage of explained values of two factors is 73%. Therefore, the number of factors was fixed as two.

Table:3 Latent Values of the Factors and Ratio of Explained Variances

Factor	**Latent Value**	**Ratio of Explained Variances (%)**
1	11.534	67.844
2	3.627	21.338
Total	15.161	89.182

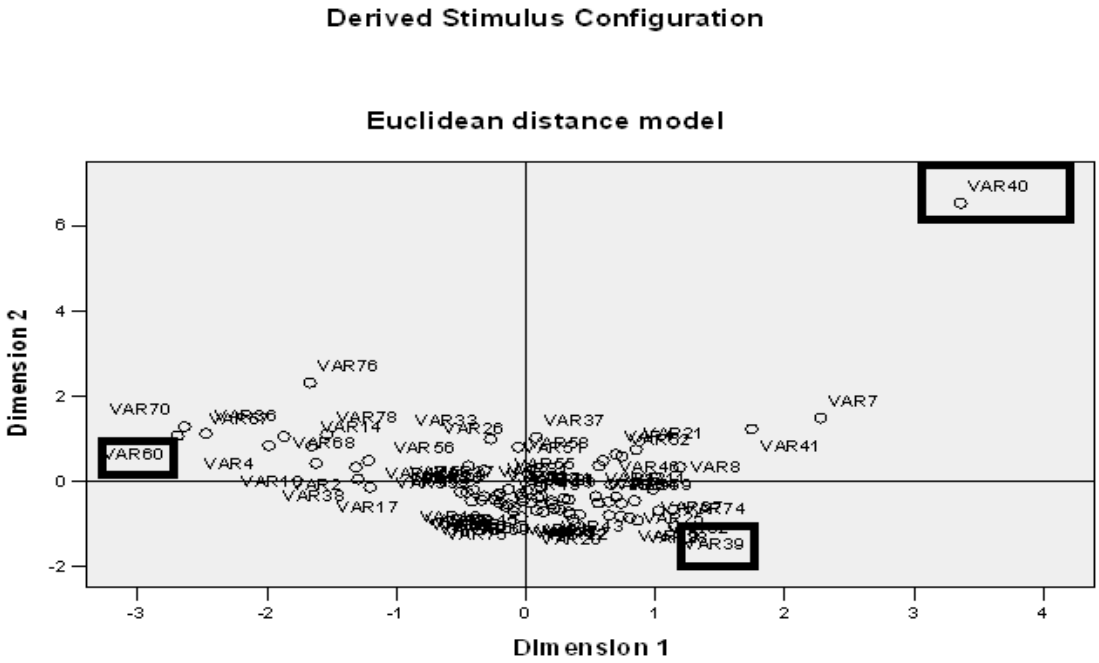
Among factor rotation methods, which aim at conceptual meaningfulness, Varimax was used and as the table below shows, the first five variables were grouped under the 1st factor while the remaining 4 variables were grouped under the 2nd factor.

Table 4: Rotated Factor Loads

	Component	
	1	**2**
Number of people per dentist	0.891	
Number of people per pharmacist	0.882	
Number of people per doctor	0.810	
Number of people per nurse	0.808	
Number of people per midwife	0.747	
Net schooling ratio at primary school level		0.986
Net schooling ratio at secondary school level		0.977
Net schooling ratio at pre-school level		0.957
Net schooling ratio at vocational high school level		0.975

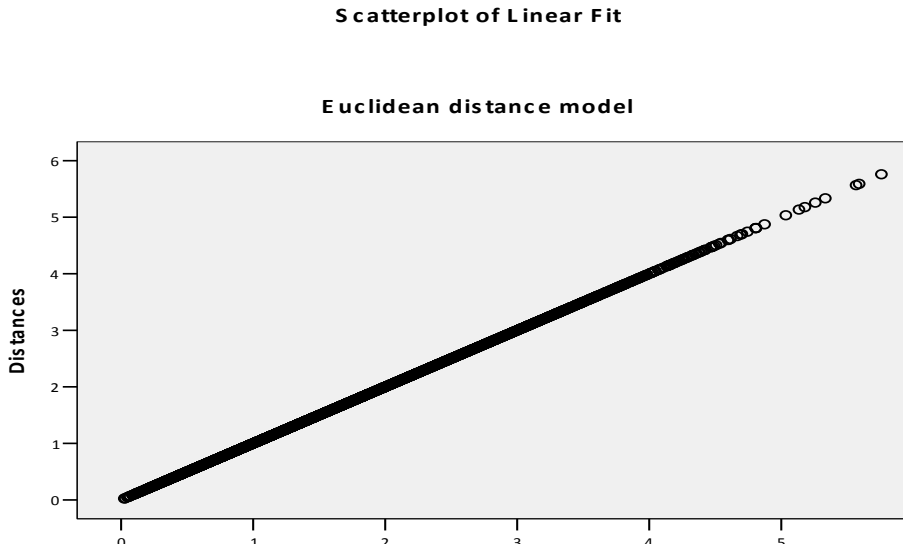
Following the factor rotation, two factors were obtained considering the variables they covered and the first factor was called "Health Indicators" while the second one was called "Education Indicators".

After the number of variables was decreased to two through factor analysis, MDS analysis was applied using the factor loading of the two factors obtained and the cities were shown on the two-dimensional space based on these two factors. The results are as follows:



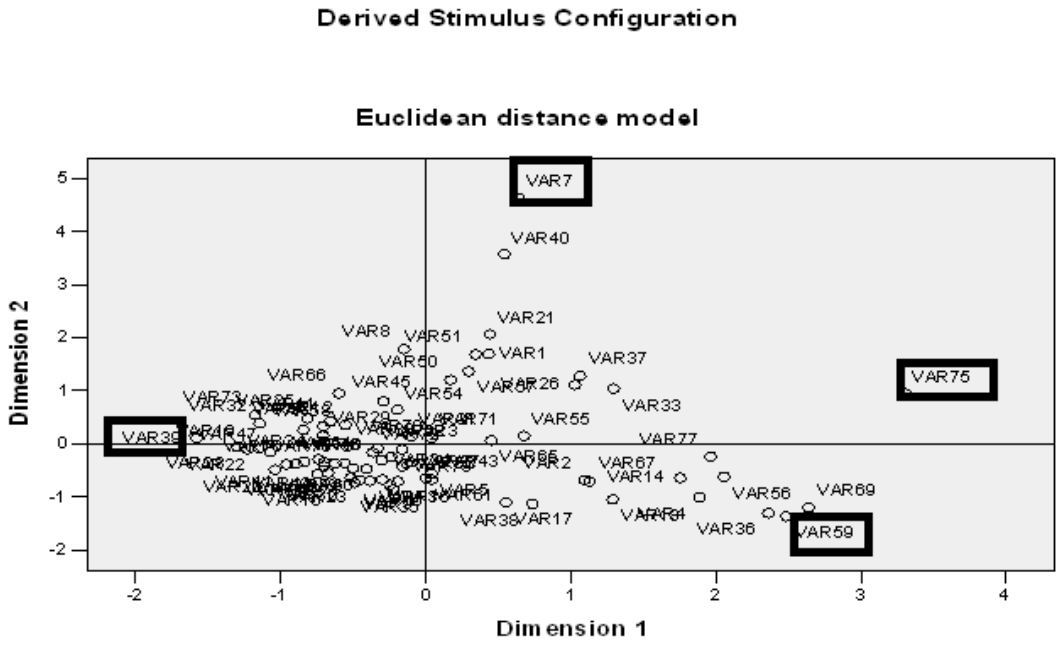
Graph 1: Euclidian Distances Model

As Graph 2 shows, the differences between units are in a linear compatibility compared to distances. Graph 1 called "Euclidian Distances Model" shows that the distance between İstanbul and Muş is greater than all other distances in the first dimension while in the second dimension the distance between İstanbul and Isparta is greater than all the other ones. The difference between İstanbul and Muş results from the education dimension while the difference between İstanbul and Isparta can be explained by the health dimension. As a matter of fact, both the research conducted by State Planning Organization in 2003 and the research into liveability of cities conducted by CNBC-E have revealed that Isparta ranks number 1 in the list of most liveable cities (in terms of health). The same study also revealed that Muş ranked near the bottom of the list in education ranking. Thus, it can be stated that the findings of our study are in harmony with the results of the mentioned research. Regarding İstanbul, it can be said that unlike other cities, it will inevitably come to the forefront as it is the most crowded city in Turkey considering its population and surface area. Therefore, İstanbul, qualified as a megacity, was excluded from our analyses in the second phase of our study and the cities were repositioned.



Graph 2: Scatter Diagram with Euclidian Distance Model

The multi-dimensional scaling analysis excluding İstanbul showed that the stress value is again zero (s = 0,000), which proves that the analysis is perfectly compatible with the data. The map below shows that the distance between Isparta and Urfa is greater than the distances between all other cities in the health dimension and the distance between Ankara and Muş is the greatest in the education dimension.



Graph 3: Euclidian Distances Model (İstanbul excluded)

5.CONCLUSION

The objective of this study is to position the 81 Turkish cities based on their education and health indicators in the multi-dimensional space. To this end, 9 variables, 6 of which were about health and 3 of which were about education, were used in this study. Since the multi-dimensional scaling analysis which covered all of the variables did not adequately show the position of the cities in the two-dimensional space, factor analysis was applied in order to decrease the number of

variables. As a result of the factor analysis, the 9 variables were categorized in two groups considering the 73% ratio, which was the percentage of the variances explained. Given the variables they covered, the factors were called education and health indicators.

After the number of variables was decreased to two following the factor analysis, the factor loads of these two factors were used and MDS analysis was applied. Next, the cities were positioned in the two-dimensional space considering their education and health indicators. According to this positioning, the greatest distance in the first (education) dimension is between Ankara and Muş whereas the greatest distance in the second (health) dimension is between Isparta and Urfa It is not unexpected that Ankara stood as the most developed city in the education dimension. In the health dimension, however, Isparta stood as the most developed city. It is thought that this ranking could be related to the Faculty of Medicine under Süleyman Demirel University. In fact, Isparta had come into prominence also in the ranking of SPO titled "Socio-Economic Development Ranking of Cities and Regions". Besides, Isparta ranked the third in the list of most liveable Turkish cities drawn by CNBC-E Business magazine. It has been revealed that Isparta came to the forefront in this list mainly due to its health indicators. In the same study, Ankara ranked the first in the education indicators.

The shape obtained at the end of the factor analysis and multi-dimensional scaling analysis was roughly similar to Turkey's map and it is clearly seen that there is a difference in the education and health indicators between the east and west of our country as there is a difference in economic development between two parts. The stress value also shows that the shape obtained in this study is quite compatible with the original shape.

Given that the development of cities and regions in a country is an indicator of the development of the country itself, it is of great importance to remove the obvious development differences between the east and west of Turkey.

REFERENCES

ALBAYRAK, A. Sait (2005), "Türkiye'de İllerin Sosyo-ekonomik Gelişmişlik Düzeylerinin Çok Değişkenli İstatistik Yöntemlerle İncelenmesi", **ZKÜ Sosyal Bilimler Dergisi**, 1(1): 153–177.

ALBAYRAK, A. Sait and KARAMUSTAFA, Osman (2013), " Türkiye'de İllerin Sosyo-ekonomik Gelişmişlik Düzeylerinin En Önemli Belirleyicileri ve İllerin 2012 Yılı Sosyo-ekonomik Gelişmişlik Sıralaması", **Recep Tayyip Erdoğan Üniversitesi Araştırma Projesi**.

ÇELİK, Şenol (2013), "Kümeleme Analizi ile Sağlık Göstergelerine Göre Türkiye'deki İllerin Sınıflandırılması", **Doğuş Üniversitesi Dergisi**, 14(2): 175–194.

DİNLER, Zeynel (2002), Bölgesel İktisat, Bursa: Ekin Kitabevi.

ERİLLİ, N. Alp (2014), "TR72 Bölgesi İlçelerinin Sosyo-Ekonomik Verilere Göre Bulanık Kümeleme Analizi ile Sınıflandırılması", **Ekonomik ve Sosyal Araştırmalar Dergisi**, 10(2): 33–45.

GÖÇER, Kenan and ÇIRACI, Hale (2003), "Türkiye'de Kentlerin Sosyal ve Ekonomik Göstergeleri Arasındaki İlişki", **İTÜ Dergisi/a, Mimarlık, Planlama ve Tasarım**, 2(1): 3–14.

GÜRBÜZ, Mehmet, Karabulut, Murat and Sandal, E. Kaya (2005), "Hiyerarşik Kümeleme Analizi Tekniği Kullanılarak Türkiye'nin Sosyo-ekonomik Bölgelerinin Belirlenmesi" **14. İstatistik Araştırma Sempozyumu (5-6 Mayıs) Bildiri Kitabı**: 381–397.

JOHNSON, R. Arnold and WICHERN, Dean W. (1998), **Applied Multivariate Statistical Analysis**, London: Prentice-Hall.

KALAYCI, Şeref (2005), **SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri**, Ankara: Asil Yayınevi.

KAVASOĞLU, Taner (2007), " Sosyo-Ekonomik Gelişmişlik Araştırması", **2. Bölgesel Kalkınma ve Yönetişim Sempozyumu (25-26 Ekim).**

KAYGISIZ, Zeliha, SARAÇLI, Sinan and DOKUZLAR, K. Ulaş (2005), "İllerin Gelişmişlik Düzeyini Etkileyen Faktörlerin Path Analizi ve Kümeleme Analizi ile İncelenmesi", **VII. Ulusal Ekonometri ve İstatistik Sempozyumu Bildiler Kitabı**: 1–33.

KILIÇ, İbrahim, SARAÇLI, Sinan and KOLUKISAOĞLU, Sıdıka (2011), "Sosyo-Ekonomik Göstergeler Bakımından İllerin Bölgesel Bazda Benzerliklerinin Çok Değişkenli Analizler ile İncelenmesi", **İstatistikçiler Dergisi**, 4: 57–68.

ÖZDEMİR, A. İhsan and ALTIPARMAK, Aytekin (2005), "Sosyoekonomik Göstergeler Açısından İllerin Gelişmişlik Düzeyinin Karşılaştırmalı Analizi", **Erciyes Üniversitesi İ.İ.B.F. Dergisi**, 24: 97–110.

SOCIO-ECONOMIC DEVELOPMENT INDEX (SDI) (2011), İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması SEGE-2011, **Ankara: T.C. Kalkınma Bakanlığı Yayını.**

STATE PLANNING ORGANIZATION (SPO,DPT) (2003), http://bssupgrade.oceaninfo.ru/

library/files/39490.pdf (Accsess Date: 02.02.2014).

ŞAHİN, Hüseyin (2002), **Türkiye Ekonomisi**, Bursa: Ezgi Kitabevi.

ŞEN, Hülya, ÇEMREK, Fatih and ÖZAYDIN, Özer (2006), "Türkiye'deki İllerin Sosyo-Ekonomik Gelişmişlik Düzeylerinin Belirlenmesi", **Selçuk Üniversitesi İ.İ.B.F. Sosyal ve Ekonomik Araştırmalar Dergisi**, 11: 155–171.

TATLIDİL, Hüseyin (2002), **Uygulamalı Çok Değişkenli İstatistiksel Analiz**, Ankara: Ziraat Matbaacılık.

YILDIZ, Ezgi, SİVRİ, Uğur and BERBER, Metin (2012), "Türkiye'de İllerin Sosyo-Ekonomik Gelişmişlik Sıralaması (2010)", **Erciyes Üniversitesi İ.İ.B.F. Dergisi**, 39: 147–167.