Theoretical and Applied Economics
Volume XXII (2015), No. 1(602), pp. 215-226

Socioeconomic development index ranking calculations of cities with fuzzy clustering method: case of Turkey

Necati Alp ERİLLİ
Cumhuriyet University, Sivas, Turkey
aerilli@cumhuriyet.edu.tr

Abstract. *In economy policies, socioeconomic indicators have an important place in determining the development levels. The determination and classification of the current social and economic structure of cities is considerably important in analyzing the development of cities and their probable development tendencies and in developing the regional development policies in parallel with this. In this study, cities are classified by fuzzy clustering c-means method according to their development levels with the help of socio-economic indicators. The aim is to determine the homogeneous city groups having the same characteristics based on these indicators. The data gathered from the study have been compared with some other classification studies in the literature.*
In this paper, development levels of cities were classified with socioeconomic indicators using FCM. Development Index values used for cities classification upon socioeconomic indicators. Results were also compared with DPT and Is-Bank development index values and interpreted.

Keywords: Fuzzy Clustering Analysis, FCM, Socioeconomic Indicators, Development Levels.

JEL Classification: C38, F63.

1. Introduction

The primary targets of national economies should be a high rate of growth and ever increasing levels of development. The upward movement of these indicators also means higher living standards and high levels of welfare. The concept of development encompasses the progress in the societal, political, cultural and similar institutions throughout the country and forms a whole. Socio-economic development can be summarized as the increase in the per capita income. In addition to this, structural and humanitarian developments can also be included in this notion. Socio-economic data, the comparisons of which give results close to accurate, are used to equilibrate among cities and areas of a country. Socio-economic data can be grouped in three main groups as social, economic and geographical indicators. Social indicators are demographic, educational, employment and social security indicators; economic indicators are monetary and financial, manufacturing industry, agriculture, international trade, energy, housing and substructure; geographical indicators are altitude, presence or absence of a shore, climate and structure of land.

Economic development is a structural change, that is, it is the transition from one structure to another. However, such a transition can be only one indicator of development. For development, there should also be increases in investment, rise in level of education as well as changes in socioeconomic structures since a developed geographical unit's meeting its every need depends on its ways and possibilities of production and its social institutions. With changes in structure, there will be a transition from traditional way of production to modern production and a new understanding of life. An important point here is that problems which are difficult to solve arise in development, transition from one structure to another. That is, when there are changes in socioeconomic structure which aim to reach stability in production, the resulting stability is a relative stability since relations of production will not stop and instability will continue. Instability is an absolute perspective while stability is a relative perspective (Özgüven, 1988).

These researches which aim to present the socioeconomic development levels of towns, cities and regions in line with the primary objectives stated in development plans enable monitoring and comparing towns, cities and regions in terms of economic and social sectors and also serves as a basic foundation in determining policies on deciding which regions are priority regions for development, allocation of public resources and directing the investments in private sector (Dinçer et al., 2003).

In the 7th five-year development plan (1996-2000), there is such a statement: "the fact that every region in the country has different means, characteristics and problems make it imperative to have a new planning approach that addresses sectoral preferences and spatial analysis together." Along with these measures, there is also the statement that "a settlement centers planning will be made in line with the decisions of the plan in order to reach a balanced settlement order throughout the country" (Dinçer et al., 1996).

In the 8th five-year development plan prepared for 2001-2005, one of the primary objectives stated was the development of national resources in a way to provide the highest economic and social benefits and the minimization of inter regional instability. In the plan, it was also stated that in the current new world order, the significance of national dynamics in the process of economic development and regional development increased and thus, studies on determining the characteristics, differences, development levels and main problems and potentials of regions in order to minimize the developmental differences among countries and to find out the means for development will continue (Dinçer and Özaslan, 2004).

The comparison of cities or regions in a country has an important place in determining the feasibility of existing policies and continuing necessary works by producing new policies. The aim here is to bring levels of inter-city and inter-regional development differences to acceptable levels and to help underdeveloped cities and regions to develop.

Intense migration caused by differences in socioeconomic development creates a big population pressure and cause urbanization problems that become an important problem within the country. In the cities crowded with migration, deficiencies arise in educational and health services and deficiencies in municipal services such as housing, water, electricity and infrastructure increase. In addition to all these, problems such as traffic, noise, crowded population and environmental pollution increase public investment needs and bring additional burden on public finance. Thus, inter-city and inter-regional socioeconomic differences become an important problem that needs to be fixed not only for underdeveloped but also for developed regions (Albayrak, 2005).

The main problem in inter-regional development differences that can be considered as acceptable within limits is about the size of instability. If developmental differences are in a level that enables long term development, it can be seen as necessary for the effective use of resources. However, if these differences cause an intense migration, settlement problems and eventually societal unrest throughout the country, it becomes one of the important instability elements that suspend the long term development of the country (Dinçer et al., 2003).

There are a great number of studies in literature conducted with socioeconomic data. These studies are on the classification of cities or regions. In other studies conducted with socioeconomic data, cities have been classified or clustered. Different statistical methods have been applied on socioeconomic data and interpreted in these studies. Some of these are: Erilli et al. (2009) used fuzzy clustering analysis; Kaygısız et al. (2005) used both path analysis and clustering analysis; Özdemir et al. (2005) used factor analysis; Eraydın et al. (2012), Vyas and Kumaranayake (2006) and Albayrak (2005) used principle components analysis; Karabulut et al. (2004) used clustering analysis; Kılıç et al.(2011) used multi-dimensional scaling analysis, Varfis and Versino (1992) used ANN, Wagner et al. (2007) used panel data and Ganzeboom et al. (1992) and McLaughlin et al. (2002) used multivariate statistical techniques. Whether they used the same statistical techniques or different ones, all of these studies gave different results. Even if their numbers of clusters were the same, their cluster members were different. The most important reasons for different results can be listed as; different number of variables used, years for collecting the variables and different statistical package programs used for analysis. Nevertheless, these differences are the basis of researches on development assessed by socioeconomic data and they are helpful for the future.

2. Material and Method

2.1. Fuzzy Clustering Analysis

This approach comes into the picture as an appropriate method when the clusters cannot be separated from each other distinctly or when some units are uncertain about membership. Fuzzy clusters are functions modifying each unit between 0 and 1 which is defined as the membership of the unit in the cluster. The units which are very similar to each other hold their places in the same cluster according to their membership degree.

Similar to other clustering methods, fuzzy clustering is based on distance measurements as well. The structure of the cluster and the algorithm used to specify which of these distance criteria will be used. Some of the convenient characteristics of fuzzy clustering can be given as follows (Naes and Mevik, 1999):
i. It provides membership values which are convenient to comment on.
ii. It is flexible on the usage of distance.
iii. When some of the membership values are known, they can be combined with numeric optimization.

The advantage of fuzzy clustering over classical clustering methods is that it provides more detailed information on the data. On the other hand, it has disadvantages as well. Since there will be too much output when there are too many individuals and clusters, it is difficult to summarize and classify the data. Moreover, fuzzy clustering algorithms, which are used when there is uncertainty, are generally complicated (Oliveira and Pedrycz, 2007).

2.2. Fuzzy C-Means (FCM) Algorithm

Fuzzy C-Means algorithm forms the basis of all clustering techniques that depend on objective function. It was developed by Bezdek (1974a and 1974b). When the FCM algorithm comes to a conclusion, the dots in the p dimension space become a sphere-shaped figure. It is assumed that these clusters are approximately the same size. Cluster centers represent each cluster and they are called prototypes. Euclidean distance d_{ik} between the data and the cluster center is used as the distance measurement and can be calculated by formula given;

$$d_{ik} = d(x_i, v_k) = \left[\sum_{j=1}^{p} \left(x_{ji} - v_{jk} \right)^2 \right]^{\frac{1}{2}}$$

where x_i represents the position observation value in the coordinated system, and v_k represents the cluster center.

It is necessary to know the number of clusters and the membership degrees of the individuals beforehand to be able to put this technique into practice. Since it is difficult to know these parameters before the application, it is possible to find these values through the method of trial and error or through some techniques developed.

The objective function used for this clustering method is as follows:

$$J(u,v) = \sum_{j=1}^{n} \sum_{k=1}^{c} u_{jk}^{m} \left\| x_{ji} - v_{jk} \right\|^2$$

This function is the weighted least square function. *n* parameter represents the number of observations, and *c* represents the number of clusters. u_{jk}^{m} is the membership of x_j in *k*-th cluster, $J(u,v)$ value is a measure of the total of all weighted error sum of squares.

If the $J(u,v)$ function is minimized for each value of c, in other words if it is derived from the 1st degree according to v_i 's and made equal to 0, the prototype of FCM algorithm can be given as follows:

$$v_{jk} = \frac{\sum_{j=1}^{n} u_{jk}^{m} x_{ik}}{\sum_{j=1}^{n} u_{jk}^{m}}$$

The required steps for the FCM algorithm are as follows:

Step 1: Set the initial values: the number of values *c*, fuzzy index *m*, end process criterion *ε* and membership degrees matrix *U* or *V* cluster prototypes are generated randomly.

Step 2: Considering that the *U* cluster prototypes are randomly generated, the membership degrees matrix is calculated by using u_{ik} values which is given below.

$$u_{ik} = \left[\sum_{j=1}^{c} \left(\frac{d_{ji}}{d_{jk}} \right)^{\frac{2}{m-1}} \right]^{-1}$$

Step 3: *U* cluster prototypes are updated following the Step 2 equation.

Step 4: In case of $\left\| U^{(t)} - U^{(t-1)} \right\| \leq \varepsilon$ stop, otherwise go back to Step 2.

After the FCM algorithm is applied, membership degrees are used to decide upon which individual will enter in which cluster. The individuals are included in the clusters by considering in which of these clusters they have the biggest membership. However, each individual can enter the other clusters with a certain degree of membership.

The results of the FCM algorithm depend on the randomly generated values to some extent. Thus, various algorithms have been developed and still being developed to solve the problems resulting from randomness (Halkidi et al., 2001).

FCM updates the cluster centers and the membership degrees for each data point through iteration and carries the clusters centers to where they should be inside the data set.

Since the initial places (values) of the cluster centers are generated by using the U matrix, the value of which is given randomly in the first place, FCM will not be able to guarantee getting close to the optimal result (Sintas et al., 1999).

Performance of cluster centers depends on the starting points of the centers (Dave, 1996). The following are two ways defined for a stronger approach:

i. Using an algorithm to define all the centers.

ii. Restarting FCM with different starting centers.

2.3. Fuzzy Clustering Validity Index

Clustering analysis aims similar objects into same groups. In many of the clustering algorithm, the number of cluster must be known beforehand. In studies based on real data, if the researchers do not have preliminary information about the number of cluster, it cannot be known whether the number of cluster which calculated is more or less than the real number of cluster. Determination processes of the optimal number of clusters are generally called as Cluster Validity. So, after clustering processes are carried out the validity of the number of cluster which calculated can be determined (Erilli, 2009: 47).

Many fuzzy clustering analysis validity indexes are used in literature (Bezdek, 1974 and 1981; Rezaee et al., 1998; Kwon, 1998; Xie and Beni, 1991). Convenient clustering validity analyses are used depending on data structure and the number of variables. In this study, Artificial Neural Networks Based Cluster Validity Index was used.

2.4. Artificial Neural Networks Based Cluster Validity Index

This method was proposed by Erilli et al. (2011). In this method at first the lowest and the highest number of cluster which are convenient to data are decided. The most convenient determined number of cluster will be in this interval. Lets the optimal number of cluster is c_{opt} , maximum number of the cluster is $c_{\max}$ and minimum number of the cluster is $c_{\min}$, are defined. The relation between them will be like that; $c_{\min} \le c_{opt} \le c_{\max}$. Then, feed-forward artificial neural networks are implemented for each possible numbers of clusters in the manner that its output will be data matrix and its target value will be the number of cluster to which each data is appointed as a result of fuzzy clustering. The median of RMSE (root-mean-square error) value which is obtained through artificial neural networks according to several hidden layer unit number are calculated for each number of clusters. The graph or obtained median values of each number of clusters or classification error is drawn and the first jumping (where median value of RMSE overgrows for the first time) is observed. Then pre-jumping value is determined as the most convenient number of cluster (Erilli et al., 2011).

3. Application

In this study, the socioeconomic development index of the cities of Turkey was found through the help of fuzzy clustering. The cities were ranked based on the index calculations obtained and they were compared with the ranking studies by State Planning Organization (SPO) and İş Bankası and they were interpreted.

The data used in the “Socioeconomic Development Ranking of Cities and Regions” published by SPO in 2003 and 2011 were used in this study. 58 variables of 81 cities used in 2003 work and 61 variables of 81 cities used in 2011 work in Turkey were analyzed by SPO. So common variables which is 58 used in this study by fuzzy clustering for comperasions. These 58 variables were presented in 10 different headings. For each heading, index rankings were calculated separately.

The clusters that were formed as a result of fuzzy clustering were separated from each other with cluster membership degrees. The rankings of these memberships from the biggest to the smallest were used in the related variable groups in the calculations of index rankings. For index calculations, 81 cities were first clustered with the fuzzy c-means clustering method based on 58 variables. The results of the clustering are presented in Table 1 (Erilli et al., 2009).

Table.1. *Distribution of cities in clusters based on the fuzzy clustering analysis*

1st degree developed cities	2nd degree developed cities	3rd degree developed cities	4th degree developed cities	5th degree developed cities
İSTANBUL	ADANA	ADIYAMAN	BATMAN	ARDAHAN
	ANKARA	AFYON	BİLECİK	ARTVİN
	BURSA	AĞRI	BİTLİS	BARTIN
	ESKİŞEHİR	AKSARAY	DİYARBAKIR	BAYBURT
	GAZİANTEP	AMASYA	EDİRNE	DÜZCE
	İZMİR	ANTALYA	ELAZIĞ	GÜMÜŞHANE
	KAYSERİ	AYDIN	ERZURUM	KARS
	KIRIKKALE	BALIKESİR	HAKKARİ	MUĞLA
		BİNGÖL	ISPARTA	MUŞ
		BOLU	KARABÜK	NEVŞEHİR
		BURDUR	KARAMAN	NİĞDE
		ÇANAKKALE	KIRKLARELİ	SİNOP
		ÇANKIRI	KIRŞEHİR	ZONGULDAK
		ÇORUM	KİLİS	
		DENİZLİ	KOCAELİ	
		ERZİNCAN	KONYA	
		GİRESUN	MALATYA	
		HATAY	MANİSA	
		IĞDIR	MARDİN	
		K.MARAŞ	MERSİN	
		KASTAMONU	OSMANİYE	
		KÜTAHYA	RİZE	
		ORDU	SAKARYA	
		SAMSUN	SİİRT	
		TOKAT	SİVAS	
		TRABZON	ŞANLIURFA	
		VAN	ŞIRNAK	
		YOZGAT	TEKİRDAĞ	
			TUNCELİ	
			UŞAK	
			YALOVA	

In this study, artificial neural networks based cluster validity ındex has been used for determine the number of cluster. According to the fuzzy clustering results obtained, the cities in Turkey were grouped into five clusters. Each city can be a member of the cluster it is in with membership degrees of a certain extent. Membership degrees in each cluster were ranked from the greatest to the smallest and thus membership sizes were ranked and socioeconomic index ranking was determined.

The same operations were calculated for the data of the related 10 different categories and index rankings of the variables were made. In its 2003 and 2011 study, SPO calculated the index rankings for socioeconomic, educational, health and production data. Socioeconomic development index 2011 is presented in Table 2.

When Table 2 is analyzed, we can see the biggest cities of Turkey in the first five. The most important factor for this can be the calculation of the important variables in the principal components analysis of the study with specific coefficients. In the fuzzy clustering analysis, all the variables were included in the calculations with equal coefficients.

Tablo.2. *SPO socioeconomic development rankings of cities (2011)*

SOCIOECONOMIC DEVELOPMENT RANKING OF CITIES (2011)								
1	İSTANBUL	4,5154	**28**	KARABÜK	0,2916	**55**	AKSARAY	-0,3671
2	ANKARA	2,8384	**29**	ZONGULDAK	0,2758	**56**	NİĞDE	-0,3761
3	İZMİR	1,9715	**30**	GAZİANTEP	0,2678	**57**	TOKAT	-0,3821
4	KOCAELİ	1,6592	**31**	TRABZON	0,2218	**58**	TUNCELİ	-0,3892
5	ANTALYA	1,5026	**32**	KARAMAN	0,1864	**59**	ERZURUM	-0,4327
6	BURSA	1,3740	**33**	SAMSUN	0,1579	**60**	K.MARAŞ	-0,4677
7	ESKİŞEHİR	1,1671	**34**	RİZE	0,1550	**61**	ORDU	-0,4810
8	MUĞLA	1,0493	**35**	DÜZCE	0,1056	**62**	GÜMÜŞHANE	-0,4814
9	TEKİRDAĞ	0,9154	**36**	NEVŞEHİR	0,1029	**63**	KİLİS	-0,5733
10	DENİZLİ	0,9122	**37**	AMASYA	0,0510	**64**	BAYBURT	-0,5946
11	BOLU	0,6394	**38**	KÜTAHYA	0,0198	**65**	YOZGAT	-0,6079
12	EDİRNE	0,6383	**39**	ELAZIĞ	-0,0103	**66**	ADIYAMAN	-0,9602
13	YALOVA	0,6263	**40**	KIRŞEHİR	-0,0211	**67**	DİYARBAKIR	-1,0014
14	ÇANAKKALE	0,5999	**41**	KIRIKKALE	-0,0687	**68**	KARS	-1,0923
15	KIRKLARELİ	0,5923	**42**	MALATYA	-0,0785	**69**	IĞDIR	-1,1184
16	ADANA	0,5666	**43**	AFYON	-0,0797	**70**	BATMAN	-1,1203
17	KAYSERİ	0,5650	**44**	ARTVİN	-0,1046	**71**	ARDAHAN	-1,1384
18	SAKARYA	0,5641	**45**	ERZİNCAN	-0,1056	**72**	BİNGÖL	-1,1920
19	AYDIN	0,5597	**46**	HATAY	-0,1302	**73**	ŞANLIURFA	-1,2801
20	KONYA	0,5308	**47**	KASTAMONU	-0,1471	**74**	MARDİN	-1,3591
21	ISPARTA	0,5272	**48**	BARTIN	-0,1976	**75**	VAN	-1,3783
22	BALIKESİR	0,4764	**49**	SİVAS	-0,2208	**76**	BİTLİS	-1,4003
23	MANİSA	0,4711	**50**	ÇORUM	-0,2405	**77**	SİİRT	-1,4166
24	MERSİN	0,4636	**51**	SİNOP	-0,2479	**78**	ŞIRNAK	-1,4605
25	UŞAK	0,3737	**52**	GİRESUN	-0,2564	**79**	AĞRI	-1,6366
26	BURDUR	0,3684	**53**	OSMANİYE	-0,2892	**80**	HAKKARİ	-1,6961
27	BİLECİK	0,3634	**54**	ÇANKIRI	-0,3312	**81**	MUŞ	-1,7329

Index rankings obtained from the fuzzy clustering analysis are presented in Table 3.

Table.3. *Development index rankings based on fuzzy clustering analysis*

	Demo-graphic	Emplo-yment	Educa-tion	Health	Indus-try	Agri-culture	Construc-tion	Finan-cial	Infra-structure	Othe r	FCM Mean.
İSTANBUL	1	3	23	26	4	74	3	1	36	13	5
ADANA	2	11	61	63	15	11	36	16	29	7	17
BURSA	3	1	6	13	38	2	41	10	2	10	2
GAZİANTEP	4	10	36	34	1	22	27	40	35	27	8
ESKİŞEHİR	5	9	20	22	33	36	13	13	22	5	4
İZMİR	6	14	10	28	3	6	1	15	11	12	1
KIRIKKALE	7	20	25	3	41	71	44	12	30	42	24
ANKARA	8	15	22	31	6	4	2	18	18	3	3
KAYSERİ	9	8	26	61	14	17	31	28	16	31	13
VAN	10	78	73	67	74	26	70	62	62	60	68
KÜTAHYA	11	34	19	55	79	24	6	31	24	20	26
DENİZLİ	12	12	13	14	75	14	38	24	34	4	12
AKSARAY	13	80	48	47	78	45	78	47	3	39	52
TOKAT	14	67	43	51	7	13	51	75	48	38	40
TRABZON	15	41	30	11	81	75	58	80	52	37	53
ÇORUM	16	52	37	80	9	21	49	27	50	53	37
AYDIN	17	51	66	12	10	10	20	38	26	9	19
AĞRI	18	68	74	71	50	77	76	70	74	72	77
BİNGÖL	19	71	80	65	44	51	64	64	61	67	69
ÇANKIRI	20	48	29	24	53	76	56	66	60	77	57
ORDU	21	61	41	30	22	20	72	48	64	51	46
IĞDIR	22	72	81	76	43	64	81	58	78	59	75

	Demographic	Employment	Education	Health	Industry	Agriculture	Construction	Financial	Infrastructure	Other	FCM Mean.
SAMSUN	23	43	53	60	32	9	21	26	66	29	32
BOLU	24	18	7	4	34	81	14	9	45	1	9
K.MARAŞ	25	32	39	29	24	18	46	81	57	46	38
YOZGAT	26	69	52	52	13	15	43	51	27	69	42
BALIKESİR	27	17	27	45	21	5	32	32	28	8	14
ÇANAKKALE	28	16	5	8	77	43	11	7	12	32	11
HATAY	29	30	40	59	17	27	39	23	10	30	27
KASTAMONU	30	81	57	33	25	47	47	22	68	26	47
AFYON	31	56	31	35	37	25	53	72	1	41	36
AMASYA	32	37	64	46	80	31	17	43	25	28	39
GİRESUN	33	76	46	58	72	58	60	77	76	44	71
ADIYAMAN	34	63	62	53	39	67	45	52	81	76	67
ERZİNCAN	35	44	59	25	51	79	30	56	40	36	50
BURDUR	36	24	12	5	23	35	66	25	13	11	16
ANTALYA	37	36	9	10	29	7	35	42	58	15	22
ISPARTA	38	19	8	15	28	41	9	41	6	16	6
ŞIRNAK	39	58	79	78	66	69	79	69	65	58	78
SAKARYA	40	29	1	43	71	48	8	35	4	17	25
MALATYA	41	40	34	1	8	19	7	76	53	40	29
HAKKARİ	42	55	71	69	57	59	73	63	49	75	73
ERZURUM	43	47	38	75	26	12	42	53	41	74	49
SİİRT	44	53	72	70	55	53	62	46	72	64	70
KIRŞEHİR	45	42	67	6	12	44	48	79	9	56	41
KIRKLARELİ	46	6	18	9	69	34	24	4	15	23	15
EDİRNE	47	25	24	23	36	33	19	6	17	25	18
YALOVA	48	13	17	49	65	63	29	5	37	54	35
KARAMAN	49	28	3	48	76	46	54	29	5	34	34
TUNCELİ	50	59	56	7	60	80	34	45	73	63	58
TEKİRDAĞ	51	2	21	18	11	39	15	11	21	48	10
ELAZIĞ	52	45	54	17	27	30	12	30	54	35	31
KOCAELİ	53	7	11	56	31	73	33	19	8	24	28
BİTLİS	54	66	78	74	48	49	74	68	63	68	76
BİLECİK	55	5	4	38	46	38	16	8	7	18	7
UŞAK	56	33	65	39	70	40	25	44	31	33	48
DİYARBAKIR	57	50	70	66	20	23	50	73	80	62	64
RİZE	58	31	63	21	40	55	67	21	67	43	51
KİLİS	59	26	58	50	63	42	65	33	32	80	56
SİVAS	60	49	33	81	16	16	5	74	42	52	45
ŞANLIURFA	61	60	76	62	18	29	52	57	46	81	62
BATMAN	62	38	77	79	49	50	55	55	77	65	72
MANİSA	63	23	51	54	5	1	40	17	23	14	23
MERSİN	64	22	14	36	19	3	37	14	33	22	20
MARDİN	65	64	75	57	52	65	71	59	51	71	74
OSMANİYE	66	46	45	27	62	72	23	61	59	79	61
KONYA	67	27	2	16	2	8	69	39	14	21	21
KARABÜK	68	4	60	19	47	61	10	36	70	47	43
ZONGULDAK	69	35	68	42	30	68	4	3	38	6	33
DÜZCE	70	21	16	64	59	52	26	67	55	61	55
NİĞDE	71	70	44	44	73	37	61	20	19	45	54
GÜMÜŞHANE	72	65	47	20	42	56	77	54	47	78	65
MUĞLA	73	54	15	2	61	28	57	2	39	2	30
MUŞ	74	79	69	73	54	70	80	71	56	73	81
BAYBURT	75	62	49	40	67	62	28	50	43	57	59
ARDAHAN	76	73	32	77	56	78	75	65	71	66	80
KARS	77	75	50	72	58	57	63	60	79	70	79
ARTVİN	78	39	55	41	68	54	59	34	69	50	63
NEVŞEHİR	79	74	28	32	35	32	68	37	20	19	44
BARTIN	80	57	35	37	64	66	22	49	75	55	60
SİNOP	81	77	42	68	45	60	18	78	44	49	66

According to the results of the fuzzy clustering analysis, the rankings of 58 variables and rankings of demographic indicators were the same. Because of this, only the rankings of demographic variables were presented in the table. According to these results, the most developed city was İstanbul, followed by Adana, Bursa, Gaziantep and Eskişehir. The most interesting result here is Kırıkkale ranking 7th in FCM analysis while it ranked 41th in SPO ranking. Only 5 of the cities that were in the first ten of SPO results were in the first ten of FCM analysis.

While there was a significant association between FCM index ranking and SPO index ranking ($p = 0{,}016$) the correlation value between them was found to be $r = 0{,}27$. We can say that there is a quite low association between these two calculations.

In their 2011 study, SPO calculated development index ranking according to education and health data as well as all the variables. Education and health association between FCM index ranking and SPO index ranking was compared: there is a significant association in terms of education data ($p = 0{,}00$) and the correlation value is $r = 0{,}67$. There is also a significant association in terms of health data and the correlation value is $r = 0{,}68$.

In the last column of Table 3, the means of FCM ranking values of 10 data groups were taken and they were ranked again, mean index ranking obtained was presented. While there was a significant association between the FCM index ranking of the whole data and the index value of FCM means ($p = 0{,}00$), the correlation value between these was $r = 0{,}44$. In addition, while there was a significant association between the index value taken from FCM means and the SPO index value ($p = 0{,}00$), the correlation value between these was $r = 0{,}91$. We can say that there is a very high association between them.

Table 4 presents the socioeconomic development ranking of cities by SPO in 2003 and 2011 (Dinçer et al., 2003; Dinçer and Özaslan, 2004), 2012-development levels of cities in Turkey prepared by Türkiye İş Bankası (Work Bank) (Eraydın et al., 2012) and the correlation matrix of fuzzy clustering analysis results. Since the calculations of İş Bankası and SPO were made through principal components analysis, there was a high correlation between them. FCM mean index values were found to be high with these three studies. This result gives an idea about the success of the study.

Table.4. *Correlation matrix of the specified studies*

	SPO-2011	FCM	FCM-Mean	İş	SPO-2003
SPO-2011	1,000	0,249	0,887	0,942	0,75
FCM	0,249	1,000	0,442	0,259	0,267
FCM-Mean	0,887	0,442	1,000	0,868	0,906
İş	0,942	0,259	0,868	1,000	0,956
SPO-2003	0,75	0,267	0,906	0,956	1,000

4. Discussion and results

Recently, socioeconomic development ranking is considered to be the most important indicator of development. These data which are also an indicator of development have an important place for State Planning Organization. With these data, cities are grouped into levels and the state and private investments to be made and the shares for these are prepared based on these data.

The distribution of settlement in Turkey centers around the cities. The content of this distribution and its effect on the neighboring cities is directly proportional to the intensity of financial and social activities in the centers. Spatial development tendencies will form an infrastructure to regional and integral developments. However, no matter how positive studies are, it is not possible for economic and social developments to be balanced throughout the country. Studies on how to minimize the differences in regional development can help to decrease these differences. The correct use of investment resources and the scientific studies conducted on these will yield desired results.

This study calculated development index values with fuzzy clustering method by using 58 demographic data of 81 cities. The purpose of clustering is to distribute into groups that are homogenous in itself. As a result of the study, the cities were grouped in five clusters. The validity index values of these cities that were calculated by the fuzzy clustering analysis membership degrees gave the development rankings of the cities.

The results obtained were compared with the development index rankings published in the report of State Planning Organization and they were interpreted. High correlation results between the index values from the means of FCM index values and index values of SPO and İş Bankası have shown that the study will be an alternative method.

Fuzzy clustering and BCO algorithm that have increased in popularity recently give better results when the number of data and the number of variables increase. Clustering analysis has been shown to give effective results when we have difficulty in deciding while classifying the data and when there are obvious differences between classes. While classifying, it can produce more clear results with complicated data structures when compared with other clustering methods. With this study, it has been presented that fuzzy clustering analysis can be successfully used for index calculation or ranking measures.

References

Albayrak, A.S. (2005). "Türkiye'deki illerin Sosyoekonomik Gelişmişlik Düzeylerinin Çok Değişkenli İstatistik Yöntemlerle İncelenmesi". *ZKÜ Sosyal Bilimler Dergisi*, Cilt:1, Sayı:1, Zonguldak

Bezdek, J.C. (1974a). "Cluster validity with fuzzy sets". *J. Cybernetics* 3, 58-73

Bezdek, J.C. (1974b). "Numerical Taxonomy with Fuzzy Sets", *J.Math Biol*. 1, 57-71

Bezdek, J.C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*, NY Plenum Press

Dave R.N. (1996). "Validating fuzzy partition obtained through c-shells clustering", *Pattern Recognition Lett*., pp. 613-623

Dinçer, B., Özaslan, M. (2004). *İlçelerin sosyoekonomik gelişmişlik sıralaması araştırması*. Bölgesel gelişme ve yapısal uyum genel müdürlüğü, DPT, Ankara

Dinçer, B., Özaslan M., Satılmış, E. (1996). *İllerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması,* DPT, Ankara

Dinçer, B., Özaslan, M., Kavasoğlu, E. (2003). *İllerin ve Bölgelerin sosyoekonomik gelişmişlik sıralaması araştırması*. Bölgesel gelişme ve yapısal uyum genel müdürlüğü, DPT, Ankara

Eraydın, K., Gül, E., Çevik, B., Demir, E., (2012). *Türkiye'de illerin gelişmişlik düzeyi araştırması*. Türkiye iş bankası, iktisadi araştırmalar bölümü, İstanbul

Erilli, N.A. (2009). *Kümeleme Analizine Bulanık Yaklaşım Algoritmaları ve Uygulamaları*, 19 Mayıs Üniv., Fen Bilimleri Enst., Yayınlanmamış Yüksek Lisans Tezi, Samsun

Erilli, N.A., Tunç, T., Öner, Y., Yolcu, U., (2009). "İllerin sosyoekonomik verilere dayanarak bulanık kümeleme analizi ile sınıflandırılması". *e-Journal of New World Sciences Academy*. Vol. 4, No. 1, 3A0001

Erilli, N.A., Yolcu, U., Eğrioğlu, E., Aladağ, Ç.H. ve Öner, Y. (2011) "Determining the Most Proper Number of Cluster in Fuzzy Clustering by Artificial Neural Networks", *Expert Systems with Applications*, 38, pp. 2248-2252

Ganzeboom, H.B., De Graaf, P.M., Treiman, D.J. (1992). "A Standard International Socio–Economic Index of Occupational Status", *Social Science Research,* 21, pp. 1-56

Halkidi M., Batistakis Y., Vazirgiannis M. (2001). "On Clustering Validation Techniques", *Journal of Intelligent Systems*, K.A. Publishes, Holland, pp. 107-145

Kalaycı, Ş. (2005). *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri*, Asil Yayıncılık, Ankara

Karabulut, M., Gürbüz, M., Sandal, E.K. (2004). "Hiyerarşik kluster (küme) tekniği kullanılarak Türkiye'de illerin sosyo-ekonomik benzerliklerinin Analizi". *Coğrafi Bilimler Dergisi*, Cilt:2, Sayı:2

Kaygısız, Z., Saraçlı, S., Dokuzlar, K., (2005). "İllerin Gelişmişlik Düzeyini Etkileyen Faktörlerin Path Analizi ve Kümeleme Analizi İle İncelenmesi", VII. Uluslararası Ekonometri ve İstatistik Sempozyumu, İstanbul Üniversitesi İktisat Fakültesi, Ekonometri Bölümü- İstanbul, 26-27 Mayıs

Kılıç, İ., Saraçlı, S., Kolukısaoğlu, S. (2011). "Sosyo-ekonomik göstergeler bakımından illerin bölgesel bazda benzerliklerinin çok değişkenli analizler ile incelenmesi". *İstatistikçiler Dergisi*, sayı:4, pp. 57-68

Kwon, S.H. (1998) "Cluster Validity Index for Fuzzy Clustering" *Elec. Lett,* 34(22), pp. 2176-2178

McLaughlin, S., McKenna, J., Cooper, J.A.G. (2002). "Problems encountered in the incorporation of socio-economic data into a coastal vulnerability indices: constraints and opportunities", *Journal of Coastal Research,* 36, pp. 487-497 Special Issue

Naes, T., Mevik, T.H. (1999). "The Flexibility of Fuzzy Clustering Illustred by Examples", *Journal Of Chemo Metrics*

Oliveira, J.V., Pedrycz, W. (2007). *Advances In Fuzzy Clustering And Its Applications*, John Wiley &Sons Inc. Pub.,West Sussex, England

Özdemir, A.İ., Altıparmak, A. (2005). "Sosyo-ekonomik göstergeler açısından illerin gelişmişlik düzeyinin karşılaştırmalı analizi. Erciyes Üniversitesi", *İktisadi ve İdari Bilimler Fakültesi Dergisi*, Sayı: 24, Ocak - Haziran, pp. 97-110

Özgüven, A. (1988). *İktisadi Büyüme*, *İktisadi Kalkınma, Sosyal Kalkınma ve Japon Kalkınması*, İstanbul

Rezaee, M.R., Lelieveldt, B.P.F. ve Reiber, J.H.C. (1998). "A New Cluster Validity Index for the FCM" *Pattern Recognition Lett.,* 19, pp. 237-246

Sintas, A.F., Cadenas, J.M., Martin, F. (1999). "Membership functions in the Fuzzy c-Means Algorithm", *Fuzzy Sets and Systems*, 101

Tatlıdil, H. (2002). *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Akademi Matbaası, Ankara

Varfis, A., Versino, C. (1992). "Clustering of socio-economic data with Kohonen maps", *Neural Network World,* Vol. 2, No. 6, pp. 813-834

Vyas, S., Kumaranayake, L. (2006). "Constructing Socio-economic Status Indices: How to Use Principal Components Analysis", *Health Policy and Planning,* 21, pp. 459-68

Wagner, G.G., Frick, J.R., Schupp, J. (2007). "The German Socio-Economic Panel Study- Evolution, Scope and Enhancements". *Schmollers Jahrbuch,* Vol. 127, pp. 139-69

Xie, L. ve Beni, G. (1991) "A Validity Measure for Fuzzy Clustering", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 13(4), pp. 841-846