ANADOLU ÜNİVERSİTESİ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Cilt: 24, Sayı: 1, Yıl: 2024
e-ISSN: 2667-8583

AÜSBD
Anadolu Üniversitesi Sosyal Bilimler Dergisi
Anadolu University Journal of Social Sciences

Examination of Provinces in Türkiye about Sectoral Employment Share by Cluster Analysis [1]

Banu BİTGEN SUNGUR [2] - Fatma Selen MADENOĞLU [3]

Submitted by: 17.09.2023 **Accepted by:** 13.02.2024 **Article Type:** Research Article

Abstract

The significance of regional dynamics in the process of economic development and regional development has increased as a result of significant factors like competitiveness, human resource development, and observation of the global market. In this study, mathematical programming-based cluster analysis has been conducted to group the regions in Türkiye according to sectoral employment rates. A mixed integer mathematical model is presented that maximizes the smallest of the out-of-cluster distances while minimizing the largest within-cluster distance. Level 2- 26 sub-regions in Türkiye are clustered according to sectoral employment data for 2021 and 2022. As a result, two clusters were obtained for both years in our country according to employment status by gender on a sectoral basis. One of these clusters is where the employment rate of the agricultural sector is higher than other sectors, and the other is where the employment rate of the industrial and service sectors is higher. When the 2021 and 2022 clusters are compared, in total, TR22, TR32, TR33, TRC3; in men, TR21, TR22, TR32, TR52, TR81; In women, it was observed that TRC1 regions were assigned to different clusters. By implementing a successful employment policy as human resource development across the national government, it will be possible to ensure the balanced growth of provinces located in Türkiye's various geographical areas.

Keywords: *Employment Rate, Sectoral Employment Share, Clustering Analysis, Mathematical Programming, Karma Tamsayılı Doğrusal Programlama*

Citation: Bitgen Sungur, B. ve Madenoğlu, F. S. (2024). Examination of provinces in Türkiye about sectoral employment share by cluster analysis. *Anadolu Üniversitesi Sosyal Bilimler Dergisi, 24*(1), 347-366.

---

[1] This study does not require ethics committee permission.

[2] Erciyes University, Faculty of Faculty of Economics and Administrative Sciences, Department of Business Administration, **bitgenb@erciyes.edu.tr**, ORCID: 0000-0002-0233-4317

[3] Abdullah Gül University, Faculty of Managerial Sciences, Department of Business Administration, **selen.madenoglu@agu.edu.tr**, ORCID: 0000-0002-5577-4471

This work is licensed under Creative Commons Attribution-NonCommercial 4.0 International License.

ANADOLU ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Cilt: 24, Sayı: 1, Yıl: 2024
e-ISSN: 2667-8583

AÜSBD
Anadolu Üniversitesi Sosyal Bilimler Dergisi
Anadolu University Journal of Social Sciences

Türkiye'de İllerin Sektörel İstihdam Payının Kümeleme Analizi ile İncelenmesi

Banu BİTGEN SUNGUR [4] - Fatma Selen MADENOĞLU [5]

Başvuru Tarihi: 17.09.2023 **Kabul Tarihi:** 13.02.2024 **Makale Türü:** Araştırma Makalesi

Öz

Ekonomik kalkınma ve bölgesel kalkınma sürecinde bölgesel dinamiklerin önemi; rekabet gücü, insan kaynağı gelişimi ve küresel pazarın gözlemlenmesi gibi önemli faktörlerin bir sonucu olarak artmıştır. Bu çalışmada, sektörel istihdam oranlarına göre Türkiye'deki bölgeleri gruplayabilmek için matematiksel programlama tabanlı kümele analizi yapılmıştır. Küme içi uzaklığın en büyüğünü minimize ederken küme dışı uzaklıkların en küçüğünü maksimize eden karma tamsayılı bir matematiksel model sunulmuştur. Türkiye'deki 26 Düzey 2 bölgesi, 2021 ve 2022 yılları sektörel istihdam verilerine göre kümelenmiştir. Sonuç olarak ülkemizde sektörel bazda cinsiyete göre istihdam durumuna göre her iki yıl için de iki küme elde edilmiştir. Bunlardan biri tarım sektörünün istihdam oranının diğer sektörlere göre daha yüksek olduğu diğeri ise sanayi ve hizmet sektörleri istihdam oranının daha yüksek olduğu kümelerdir. 2021 yılı ve 2022 yılı kümeleri karşılaştırıldığında toplamda, TR22, TR32, TR33, TRC3; erkeklerde, TR21, TR22, TR32, TR52, TR81; kadınlarda ise TRC1 bölgelerinin farklı kümelere atandığı görülmüştür. Ulusal hükümet genelinde başarılı bir istihdam politikasının, insan kaynaklarının geliştirilmesinin uygulanmasıyla, Türkiye'nin çeşitli coğrafi bölgelerinde yer alan illerin dengeli büyümesinin sağlanması mümkün olacaktır.

Anahtar Kelimeler: *İstihdam Oranı, Sektörel İstihdam Payı, Kümeleme Analizi, Matematiksel Programlama, Mixed Integer Linear Programming*

[4] Erciyes Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, **bitgenb@erciyes.edu.tr**, ORCID: 0000-0002-0233-4317
[5] Abdullah Gül Üniversitesi, Yönetim Bilimleri Fakültesi, İşletme Bölümü, **selen.madenoglu@agu.edu.tr**, ORCID: 0000-0002-5577-4471

Bu eser Creative Commons Atıf-Gayri Ticari 4.0 Uluslararası Lisansı ile lisanslanmıştır.

Introduction

The workforce, or the human factor, must be involved in employment for technological and economic development to carry out. A country's human resource is comprised of the quantity and quality of its labor force. The dynamics of the labor market are determined by labor demand, which is a function of both economic conditions and the overall labor supply. The labor market executes economic operations and puts policies into place.

Among labor market indicators, employment and unemployment has been held a significant position. These indicators inform decision-makers on the market's operation and what must be done. Employment, a key labor market indicator, refers to both the inclusion of the labor factor in the production process and, more broadly, the participation of all production factors in the production process. One of the most crucial indicators of a nation's economic structure and degree of development is generally agreed to be how jobs are distributed among its various industries (Sema, 2012). For this reason, in this study, sectoral employment data, which is a basic labor force indicator, has been used while performing the cluster analysis of the cities.

Two primary types of clustering algorithms are partition-based and hierarchical. Under bottom-up hierarchical techniques, members initially form independent clusters, then merge according to a specified closeness metric to generate new clusters. The construction process ends when every member is in a single cluster. These new clusters also merge in an iterative manner that moves vertically from one step to the next. In contrast, partition-based methods use a pre-determined number of clusters and work horizontally to assign and reallocate individuals to groups. A wide range of clustering algorithms are based on partition-based and hierarchical clustering techniques. Although these algorithms are usually presented from a computer science or statistics perspective, it can also be useful to use the discrete optimization perspective to build models that enhance and supplement the results of other methods. Clusters are typically generated by integer programming formulations that either maximize the similarities among objects within a cluster or minimize the similarities between clusters. Compact or homogeneous clusters are generated by maximizing item similarities within clusters, whereas well-separated clusters are generated by decreasing similarities between distinct clusters (Pirim et al., 2018). Subject to certain unique clustering problem limitations, integer programming methods offer flexibility in expressing objectives. For this reason, Turkey's employment cluster analysis has been conducted with the help of a mathematical model, taking into account Turkey's sectoral employment situation and sectoral employment rates.

The rest of this paper is organized as follows. The literature review is explained in Section 1. The data and methodology are presented in Section 2 and Section 3 respectively. The empirical results are discussed in Section 4. Findings and concluding remarks are discussed in the last section.

Literature Review

There have been papers on employment analysis and clustering problem in the literature. To better understand population health in a global context, Muntaner et al. (2012) examined into the rules governing the labor market in low- and middle-income nations and suggest a taxonomy for the labor market. A hierarchical cluster analysis was used to provide a taxonomy of the labor market. McDermott et al. (2013) examined into how psychological contracts, organizational commitment, and employment traits relate to those of paid workers in a non-profit company. Their analysis contained three employee clusters based on employees' opinions of their psychological contracts using fuzzy-c-means clustering. Pașnicu et al. (2014) analyzed recent changes in labor market participation, employment, unemployment, and educational attainment with a focus on data that is subdivided by age and sex. To accomplish the goals of cohesion policies at the EU level, the paper intends to categorize the EU nations according to important indicators of the labor market in order to identify gaps and similarities. Hierarchical cluster analysis was employed to achieve this proposal. Florez and Cortissoz (2017) suggested a mathematical framework that makes advantage of a crew's members' interpersonal compatibility in order to more accurately assess production. To assess the compatibility of the team members, a personality measurement and quantification tool was suggested. To develop a probability density function that will experimentally establish the average productivity of a given cluster, cluster analysis concepts were used to group crews that have comparable compatibility and productivity scores. Seki and Arslan (2018) used the three-star method to reveal the clustering potential of the TRC2 region in the manufacturing industry and to identify sectors that can lead in regional development. Yamaç (2019) conducted the clustering analysis using the concentration coefficient in the Turkish textile industry to understand the cluster structure in Türkiye Statistical Regional Units Classification Level 2 (26 sub-regions) regions. Çamlica and Şenkayas (2020) determined the sectoral concentrations considering employment data based on manufacturing sectors in the TR32 region using location quotient (Regional Concentration) and used shift share analyzes to determine the sectors with high competitiveness. To determine the sectors with clustering potential in TRA1 provinces (Erzurum-Erzincan-Bayburt) and to comparatively reveal the competitive advantages of the companies in these sectors. Yaprakli and Aslan (2023) showed that the possible clusters in all three provinces are generally concentrated in the service sector with the results of three-star analysis. Gürler (2023) presented the clustering analysis of countries similar to each other in the travel and tourism sector using the k-means method.

Solving clustering problems using a mathematical model are discussed in the literature. Vinod (1969) introduced an integer programming for grouping problem. The results of the experimental studies showed that the proposed method is more flexible and efficient method. The proposed method converted the quadratic objective function to linear effectively. Rao (1971) presented two clustering models based on integer programming representations for the distance-based clustering problems include an m-median problem. Kusiak (1984) investigated five different mathematical models for three clustering problem which are traveling salesman problem and m-median problems. Klein and Aronson (1991) suggested a new mathematical model for clustering problem considering the total interaction of all pairs of items in a common grouping. Hansen and Jaumard (1997) presented a survey from mathematical programming viewpoint for clustering problems. Mehrotra and Trick (1998) used mathematical models for solving clique and clustering problems. Sağlam et al. (2006) presented a mixed integer programming-based clustering method to segment the customer considering demographic and transactional features. The proposed method minimized the maximum cluster diameter among all clusters. Xu et al. (2007) formulated a mixed integer quadratic problem for maximization of the network modularity. Agarwal and Kempe (2008) investigated the linear programming approach to maximize the modularity for measuring the quality of a network partitioning into communities. Pirim et al. (2012) presented a review of clustering algorithms to analyze gene expression data. Cafieri and Hansen (2014) used mathematical programming to refine heuristic solution for network clustering considering modularity maximization. Benati and García (2014) proposed a mixed integer non-linear optimization model for clustering and variable selection. Pirim et al. (2018) presented a novel mixed integer linear programming model for clustering relational networks. Benati et al. (2018) proposed a mixed integer linear model to find optimal feature selection for clustering. Wang et al. (2019) formulated the bus route clustering problem concerns the assignment of bus routes to different boarding as a mixed-integer second-order cone program. Puerto et al. (2020) formulated a mixed-integer linear programming model for clustering and asset selection problem in a unified framework. Mokhtarzadeh et al. (2021) developed a multi-objective mixed-integer non-linear programming model for solving p-mobile hob location and allocation problem with the depreciation cost of hub facilities. Caramia and Pizzari (2022) reformulated of the latter in terms of mixed-integer linear programming for clustering, location, and allocation in two stage supply chain for waste management. Doan et al. (2023) proposed two new mixed integer linear programming for clustered traveling salesman problem with relaxed priority rule. Ágoston and E.-Nagy (2023) formulated a mixed integer linear programming formulation for the minimum sum of squares clustering problem. It is seen that there is no study in the literature that deals with sectoral employment clustering for Türkiye with a mathematical model.

Data

In this study, sectoral-based employment rates data for 2021and 2022 are used. TurkStat has published this data as level 2, sectoral distribution of employment data for 26 sub-regions. In addition, the distribution of employment rates in agriculture, industry, and service sectors disaggregated by sex is presented. The descriptive statistics for sectorial distribution of employment data for 2021 and 2022 are compiled in Table 1 and Table 2, respectively.

Table 1
Descriptive Statistics for 2021

			Employment rate			
	N	**Sectors**	**Min.**	**Max.**	**Mean**	**Standard deviation**
Total	26	Agriculture	0.5	52.1	24.0	11.9
		Industry	11.4	40.4	25.0	7.4
		Service	35.1	71.7	51.0	8.8
Men	26	Agriculture	0.5	45.2	20.2	9.7
		Industry	14.7	45.0	29.6	7.8
		Service	37.4	66.7	50.2	7.6
Women	26	Agriculture	0.4	66.4	32.4	16.7
		Industry	3.2	30.5	14.4	7.2
		Service	30.4	82.7	53.2	12.5

Table 2
Descriptive Statistics for 2022

			Employment rate			
	N	**Sectors**	**Min.**	**Max.**	**Mean**	**Standard deviation**
Total	26	Agriculture	0.6	48.1	22.6	11.5
		Industry	12.3	40.9	25.3	7.4
		Service	39.4	71.4	52.1	8.2
Men	26	Agriculture	0.5	40.9	19.1	9.4
		Industry	15.9	45.8	30.0	7.7
		Service	40.3	65.9	50.9	6.9
Women	26	Agriculture	0.7	63.2	30.1	15.9
		Industry	2.9	33.5	15.1	7.3
		Service	33.8	83.1	54.8	12.1

The data shows that the employment rate in the service sector is over fifty percent of total employment for total, men and women. The lowest employment rate for total and men is in the agricultural sector and for women in the industrial sector. According to the average values of the sectors over the years, it is seen that there is a decrease in the total, male and women values in the agricultural sector in 2022 compared to the previous year. It is also seen that the employment rates of the industrial and agricultural sectors are very close to each other for total. This situation reveals that our country has development potential in the agricultural sector. When the data are analyzed in further depth, it is possible that the ratio of this service sector to total employment is higher because all other sectors use unmanned production and agricultural techniques more frequently than the service sector. The average distribution of sectoral employment for men is similar to the total. According to the sectoral employment distribution of women, women are mostly employed in the service sector, followed by the agriculture and industry sectors. It is determined that a broad policy and incentive system should be put in place to encourage our women to participate more in the industry sector. Additionally, it can be observed from the standard deviation numbers that the sectoral for women's data is typically higher than the average. In this case, it is acknowledged that there are positive actions for women's participation in the workforce in some locations, this is not the case in other regions, and the women employment rate is low. In other words, women's employment is unevenly distributed across the nation, with high women employment rates in certain areas and low in others.

The sector-based employment distribution data by province are given in Table 3. Sector-based employment distribution data by provinces reveals that there is a large difference between sectors in some of our provinces. In this case, it gives the result that the provinces are dependent on other provinces for some sectors. The fact that Istanbul, a province with a large population, has the lowest level of agricultural employment is noteworthy in this case. In addition to raising prices, this circumstance creates issues with transporting the products coming from these other provinces and servicing the city's agricultural demands from outside the province. Ankara has the highest employment rate among the three groups in the service sector. However, in terms of agricultural employment, this city has very low employment rate. The employment gap between sectors is quite high for these provinces. In the provinces of Ağrı, Kars, Iğdır and Ardahan, while the agricultural employment is the highest throughout the country, the employment rate in the industrial sector is quite low. From here, it is seen that it is necessary to develop projects and implement incentive programs to develop the industry in these provinces.

Table 3
The Sector-Based Employment Distribution Data by Province for 2021-2022

Sectors	Gender	Year	The highest employment rate	The lowest employment rate
Agriculture	Total-Men-Women	2021	TRA2 (Ağrı, Kars, Iğdır, Ardahan)	TR10 (Istanbul)
		2022	TRA2 (Ağrı, Kars, Iğdır, Ardahan)	TR10 (Istanbul)
Industry	Total- Men	2021	TR41 (Bursa, Eskişehir, Bilecik)	TRA1 (Erzurum, Erzincan, Bayburt)
		2022	TR41 (Bursa, Eskişehir, Bilecik)	TRA1 (Erzurum, Erzincan, Bayburt)
Industry	Women	2021	TR41 (Bursa, Eskişehir, Bilecik)	TRA2 (Ağrı, Kars, Iğdır, Ardahan)
		2022	TR21 (Tekirdağ, Edirne, Kırklareli)	TRA1 (Erzurum, Erzincan, Bayburt)
Service	Total-Men-Women	2021	TR51 (Ankara)	TRA2 (Ağrı, Kars, Iğdır, Ardahan)
Service	Total	2022	TR51 (Ankara)	TRA2 (Ağrı, Kars, Iğdır, Ardahan)
Service	Men	2022	TR51 (Ankara) TR61 (Antalya, Isparta, Burdur)	TR33 (Manisa, Afyonkarahisar, Kütahya, Uşak)
Service	Women	2022	TR51 (Ankara)	TRA2 (Ağrı, Kars, Iğdır, Ardahan)

Method

In this study, we present a mathematical model for the clustering problem with the objective of minimizing the difference between the largest in-cluster distance and the smallest out-of-cluster distance. Using the employment data by sectors for 2021 and 2022, Türkiye's Level 2 regions have been clustered with this mathematical model. Percentages of employment in agriculture, industry and service sectors, which are also disaggregated by gender, have been included in the analysis. Thus, sectoral employment differences by gender has been also revealed.

Cluster Analysis

Cluster analysis is described as the process of grouping objects into natural groups based on their similarities (Ferreira et.al., 2009).

Cluster analysis is an unsupervised machine learning technique that categorizes the objects into a predetermined number of groups based on similarities or differences between the objects. According to this theory, objects are grouped together into clusters based on their similarities, and each cluster is created to diverge from the others as little as possible. Most cluster analysis techniques rely on distinctions between objects. The following is a general description of clustering. There are defined the n objects $T = \{T_1, T_2, T_3, \ldots., T_n\}$ and the observed m features of the objects. The $nx\, m$ data matrix is generated. The similarities between objects is calculated by using $nx\, n$ matrix. The distance values between the T_i and T_l objects are obtained into the similarity matrix $d\left(T_{i,}T_l\right) = d_{il}$. Euclidean distance is the most popular measurement technique for determining the distance between data objects. Equation (1) must be used to determine the distance between objects (Ferreira et.al., 2009). The results of the mathematical model-based clustering method are analyzed.

$$D_{il} = \sqrt{\sum(T_i, T_l)^2} \qquad (1)$$

Cluster Validation

Cluster validation is an essential component of cluster analysis. It assesses the cluster quality and establishes the optimal number of clusters using the provided data. There are numerous cluster validation techniques to choose from. The most common index is the Silhoutte index (Rousseeuw, 1987). Equation (2) states that this indicator contrasts the proximity within a cluster with the distance between clusters. The variable $a(i)$ represents the average dissimilarity of i to all other objects in the same cluster. The variable $b(i)$ represents the lowest average distance of the ith object to other clusters. The range of Silhoutte values is -1 to 1. If it is greater than zero, the objects are well-designed within their cluster; if it is less than zero, the objects are spread among clusters. The ideal coefficient is one that is near to one.

$$Silhoutte\ (i) = \frac{b(i) - a(i)}{\max\{\, a(i), b(i)]} \qquad (2)$$

Finally, we compute the overall average silhouette index, which is the average of s(i) for all objects i in the entire dataset.

Clustering Methodology

There is a mathematical model presented for the clustering problem. The notation that will be used is listed below:

i, l	objects $(i, l \in T)$
j	clusters $(j \in S)$
n	number of objects
h	upper bound on the number of clusters
T	set of objects
S	set of clusters
d_{il}	the positive distance between object i and object l $(i \neq l)$
x_{ij}	1, if object i is assigned to cluster j; 0, otherwise $(\forall i \in T, \forall j \in S)$
z_{ilj}	1, if object i and object l are assigned to cluster$j$; 0, otherwise $(\forall i, l \in T \mid i \neq l, \forall j \in S)$
y_{il}	1, if object i and object l are in the same cluster; 0, otherwise $(\forall i, l \in T \mid i \neq l)$
p_i	total distance of object i to the other objects in the cluster it belongs to $(\forall i \in T)$
C	maximum total distance of the objects
r_i	total distance of object i to the other objects in the other clusters $(\forall i \in T)$
CC	minimum total distance of the objects

The clustering problem can be stated as follows. There are n objects, each of which can be assigned to only one cluster. The sets of $T = \{1, \cdots, n\}$ and $S = \{1, \cdots, h\}$ denote the objects and clusters respectively. The distances between the each i, l objects $(i, l \in T)$ indicated by d_{il} is calculated. For a given clusters upper bound h, the clustering problem consists in finding the best combination of objects for each cluster j, such that maximum distance of the object to the other objects in the same cluster is minimized and meanwhile minimum distance of the object to the other objects out of clusters is maximized. The mixed integer linear programming model (MILP) for the clustering problem can be formulated as follows:

$$\min C - CC \quad (3)$$

$subject\ to$:

$$\sum_j x_{ij} = 1, \forall i \in T \quad (4)$$
$$x_{ij} + x_{lj} \leq z_{ilj} + 1, \forall (i, l) \in T \mid i \neq l, \forall j \in S \quad (5)$$
$$x_{ij} + x_{lj} \geq 2z_{ilj}, \forall (i, l) \in T \mid i \neq l, \forall j \in S \quad (6)$$
$$\sum_j z_{ilj} = y_{il}, \forall (i, l) \in T \mid i \neq l \quad (7)$$
$$\sum_l d_{il} y_{il} = p_i, i \in T \quad (8)$$
$$C \geq p_i, i \in T \quad (9)$$
$$\sum_l d_{il}(1 - y_{il}) = r_i, i \in T \quad (10)$$
$$CC \leq r_i, i \in T \quad (11)$$
$$x_{ij} \in \{0,1\}, \forall i \in T, \forall j \in S \quad (12)$$
$$z_{ilj} \in \{0,1\}, \forall (i, l) \in T \mid i \neq l, \forall j \in S \quad (13)$$
$$y_{il} \in \{0,1\}, \forall (i, l) \in T \mid i \neq l \quad (14)$$
$$p_i \geq 0, \forall i \in T \quad (15)$$
$$r_i \geq 0, \forall i \in T \quad (16)$$
$$C \geq 0 \quad (17)$$
$$CC \geq 0 \quad (18)$$

The objective function (3) while minimizing the maximum distance of an object to the objects in the same cluster, the smallest distance to other objects outside the cluster is maximized. The constraints are as follows: (4) guarantees that each object i is assigned to exactly one cluster j. (5) and (6) defines the relation between x_{ij} and z_{ilj} variables: If objects i and l are assigned to the cluster j, i.e. $x_{ij} = 1$ and $x_{lj} = 1$, then $z_{ilj} = 1$. If objects i and l are not assigned to the cluster j, together, then the values of all related z_{ilj} variables are forced to be zero. Constraint (7) defines the relation between z_{ilj} and y_{il} variables: If objects i and l are assigned to the cluster j, then $y_{il} = 1$. Constraint (8) determines the summation of the distance of object i to the other objects in the cluster it belongs to. Constraint (9) determines the maximum distance of an object in the cluster to other objects in the same cluster. Constraint (10) determines the total distance of object i to the other objects in the other clusters. Constraint (11) determines the maximum distance of an object to the objects in the other cluster. (12)-(18) define the binary and integer restrictions on the variables.

Results and Discussion

Data Analysis

The data's applicability for cluster analysis has been determined. The Shapiro-Wilk test results indicate that the data are distributed normally (p≤0.05). Table 4 presents the results of the normality test. The results demonstrated that the data are normally distributed. Since all of the variables have a normal distribution, a second analysis called outliers identifying has been performed. The results indicate that there are no outliers in the data set. This detailed statistical analysis shows that the variables can be employed in clustering analysis.

Table 4
Test of Normality

	2021 Total			**2021 Men**			**2021 Women**		
Variables	**Statistic**	**df**	**Sig.**	**Statistic**	**df**	**Sig.**	**Statistic**	**df**	**Sig.**
Agriculture	0.989	26	.989	0.983	26	.927	0.985	26	.953
Industry	0.980	26	.864	0.983	26	.925	0.934	26	.095
Service	0.977	26	.793	0.952	26	.264	0.982	26	.911
	2022 Total			**2022 Men**			**2022 Women**		
Variables	**Statistic**	**df**	**Sig.**	**Statistic**	**df**	**Sig.**	**Statistic**	**df**	**Sig.**
Agriculture	0.985	26	.961	0.986	26	.973	0.982	26	.916
Industry	0.975	26	.766	0.986	26	.969	0.934	26	.098
Service	0.950	26	.235	0.933	26	.092	0.974	26	.721

Cluster analysis distance metrics are used to calculate the distance between the objects. The metric that is most frequently employed in the literature to gauge the separation between variables is the Euclidean distance. The recommended mathematical method includes grouping the regions into clusters while minimizing the maximum distances between objects and those in the same cluster and maximizing the minimum distances between objects which are in the other clusters.

Experimental results

The mixed integer programming model has been coded in AIMMS 4.78.2.4 optimization software and the solver CPLEX 20.1 has been used to solve the clustering problem described above. The models have been run on a PC with IntelI CoreI I5-1035G1 CPU @ 1.00GHz 1.19 GHz processor and 8.00 GB RAM. The running time has been limited to 3600 seconds. For cluster sizes ranging from two to ten, an experimental analysis has been done. We have presented detailed computational results in the Table 5 and Table 6.

Table 5
The Obtained Results for 2021

	Total			Men			Women		
Cluster	Objective Function Value	CPU Time	Gap %	Objective Function Value	CPU Time	Gap %	Objective Function Value	CPU Time	Gap %
2	110.683	3	0	94.858	2	0	99.520	2	0
3	-158.425	10	0	-119.441	14	0	-236.496	9	0
4	-238.473	1554	0	-203.562	2929	0	-325.263	1539	0
5	-279.149	3600	8.7	-242.620	3600	4.1	-387.405	71	0
6	-307.839	3600	8.1	-268.308	3600	8.9	-416.343	3600	1.9
7	-329.411	3600	6.0	-292.275	3600	4.6	-443.739	1519	0
8	-338.236	1. 3 600	5.3	-302.905	3600	1.7	-452.869	3600	3,8
9	-345.731	3600	4.4	-307.808	3600	2.3	-463.464	3600	2.9
10	-354.010	3600	3.0	-314.773	117	0	-471.738	50	0

Table 6
The Obtained Results for 2022

	Total			Men			Women		
Cluster	Objective Function Value	CPU Time	Gap %	Objective Function Value	CPU Time	Gap %	Objective Function Value	CPU Time	Gap %
2	80.324	2	0	85.5	2	0	102.3	3	0
3	-144.313	12	0	-126.9	13.8	0	-217	21	0
4	-225	2294	0	-200.6	2225	0	-308.7	2724	0
5	-282.5	17	0	-240.9	49.3	0	-362.9	3600	2.6
6	-296.4	3600	3.81	-261.7	3600	6.7	-394.1	693	0
7	-315.6	655	0	-274.2	3600	8	-423.5	2110	0
8	-329.1	3480	0	-288	3600	4	-433.8	3600	3.2
9	-331.4	3600	3.78	-294.5	3600	3.9	-442.7	3602	2.2
10	-336.9	3600	1.52	-307	3600	1.9	-449.8	2430	0

Cluster Validation

The silhoutte value for each cluster is calculated. The ideal cluster number is chosen based on the Silhoutte index. The Silhoutte index rates the quality of clusters for cluster analysis based on Euclidean distance. The silhouette coefficcient value for clusters is shown in Figure 1. The optimal number of clusters is two for total, men, and women, which corresponds to the greatest silhoutte coefficient for 2021 and 2022.

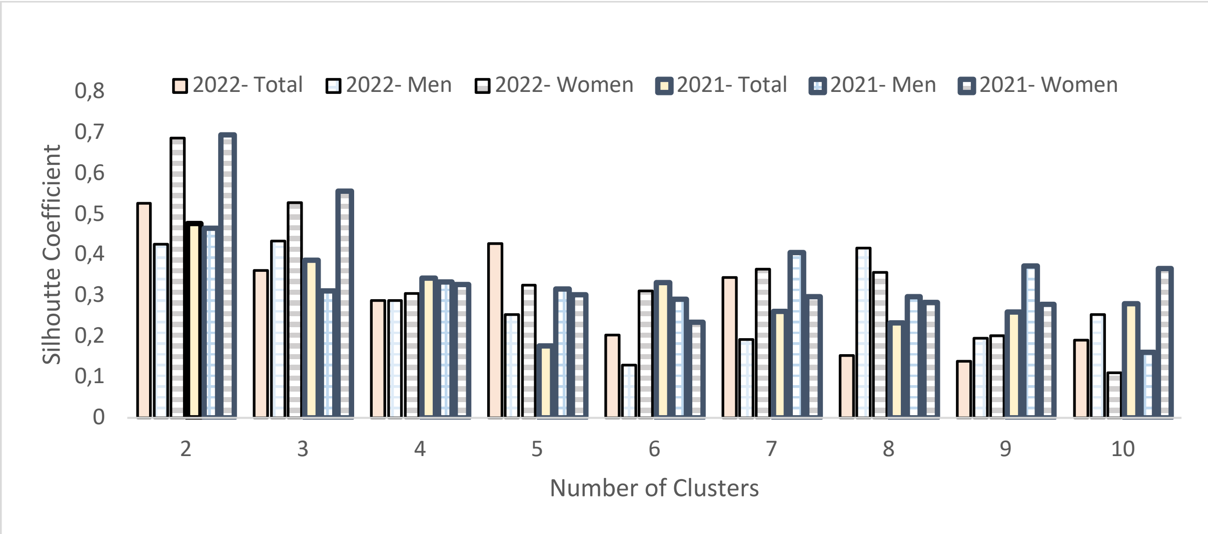

Figure 1. Silhoutte coefficient of clusters for 2021-2022

Analysis of the Experimental Results

The list of the clusters according to the total employment rate are given in Table 7. The results of 2021 have been examined. When the total and male clustering results are compared for the optimal solution; TR33, TR52 and TR81 are in opposite clusters in two clusters. When these three regions are examined, it has been seen that women have much higher employment rates than men in the agricultural sector. The high women agricultural employment rate also increased the total employment rate. For this reason, they were assigned to reverse clusters in total and male clustering analyses. Other 23 provincial groups are assigned to the same cluster in both groups. When the total and women clustering results are compared; only TR22, TR32 and TR33 are assigned with different regions. It has been observed that the agricultural employment rates of the TR33 region are much higher for women. It has been determined that women employment rates in the service sector are very high in the TR22 and TR32 regions. These three regions, which differ in women's employment rates according to their total employment rates by sector, were assigned to different clusters for both analyses. When the men and women clustering results are compared; TR52, TRC1, TR22 and TR32 are assigned with different regions. When the employment rates in TR52 and TRC1 regions are examined, it is seen that the rates in the service sector are not very different by gender, but there are significant differences in the agriculture and industrial sectors. It has been determined that in these regions, men are employed in the industrial sector, while women are employed in the agricultural sector. It is seen that TR22 and TR32 regions employ women at higher rates in the service sector.

When the total and male clusters are compared for 2022, it is seen that all regions, except TR21 and TRC3 regions, are in the same cluster. The employment rate of men in the service sector is lower in TR21 region and higher in TRC3 region than the employment rate in the same sector in other regions. When the total and women clusters were compared, only the TRC3 region was assigned to a different cluster. TRC3 region has a relatively lower female employment rate in the agricultural sector compared to other regions. When the male and women clusters are examined, it is seen that the TR21 region is assigned to different clusters. Compared to other regions, TR21 region has a relatively low rate of male employment in the service sector and a relatively high rate of female employment in the industrial sector.

When comparing 2021 and 2022; total, TR22, TR32, TR33, TRC3; for male, TR21, TR22, TR32, TR52, TR81; for women, TRC1 regions are assigned to different clusters. In TRC1 region, from 2021 to 2022, female employment rates decreased in the service sector and increased in the industry and service sectors.

Table 7
The Obtained Clusters for 2021 and 2022

Cluster			Regions
2021	Total	Cluster 1	TR10, TR21, TR31, TR33, TR41, TR42, TR51, TR61, TR62, TR63, TR72, TRC1, TRC3
		Cluster 2	TR22, TR32, TR52, TR71, TR81, TR82, TR83, TR90, TRA1, TRA2, TRB1, TRB2, TRC2
	Men	Cluster 1	TR10, TR21, TR31, TR41, TR42, TR51, TR52, TR61, TR62, TR63, TR72, TR81, TRC1, TRC3
		Cluster 2	TR22, TR32, TR33, TR71, TR82, TR83, TR90, TRA1, TRA2, TRB1, TRB2, TRC2
	Women	Cluster 1	TR33, TR52, TR71, TR81, TRC1, TR82, TR83, TR90, TRA1, TRA2, TRB1, TRB2, TRC2
		Cluster 2	TR10, TR21, TR22, TR31, TR32, TR41, TR42, TR51, TR61, TR62, TR63, TR72, TRC3
2022	Total	Cluster 1	TR10, TR21, TR22, TR31, TR32, TR41, TR42, TR51, TR61, TR62, TR63, TR72, TRC1
		Cluster 2	TR33, TR52, TR71, TR81, TR82, TR83, TR90, TRA1, TRA2, TRB1, TRB2, TRC2, TRC3
	Men	Cluster 1	TR21, TR33, TR52, TR71, TR81, TR82, TR83, TR90, TRA1, TRA2, TRB1, TRB2, TRC2
		Cluster 2	TR10, TR22, TR31, TR32, TR41, TR42, TR51, TR61, TR62, TR63, TR72, TRC1, TRC3
	Women	Cluster 1	TR10, TR21, TR22, TR31, TR32, TR41, TR42, TR51, TR61, TR62, TR63, TR72, TRC1, TRC3
		Cluster 2	TR33, TR52, TR71, TR81, TR82, TR83, TR90, TRA1, TRA2, TRB1, TRB2, TRC2

The mean values of sectors in the groups are summarized in Table 8. The table has been examined for 2021. While the total and men average values for the cluster 1 in industry and service sectors are the greatest, the average value for agriculture is lower. The opposite happens for women. In other words, whereas the industry and service sectors' average values are highest in cluster 2 for women, the agricultural sector's average value is highest in cluster 1. The table has been examined for 2022. While the total and women average values for the cluster 1 in industry and service sectors are the greatest, the average value for agriculture is lower. For men, the average values of the industry and service sectors are highest in cluster 2, and the average value of the agricultural sector is the highest in cluster 1.

Table 8
Mean Values of Sectors in Clusters

		2021		2022	
	Sectors	Cluster 1	Cluster 2	Cluster 1	Cluster 2
Total	Agriculture	15.1	32.8	13,6	31,7
	Industry	29,4	20.6	29.1	21,4
	Service	55.6	46.6	57,3	46,9
Men	Agriculture	13.3	28.3	26,1	12,1
	Industry	34.0	24.5	27,1	32,8
	Service	52.8	47.2	46,7	55,2
Women	Agriculture	46.3	18.4	17,9	44,4
	Industry	10.6	18.2	18,4	11,2
	Service	43.1	63.3	63,7	44,4

The clustered distribution of sectors according to total and gender for year 2021 is shown in Figure 2 graphically.

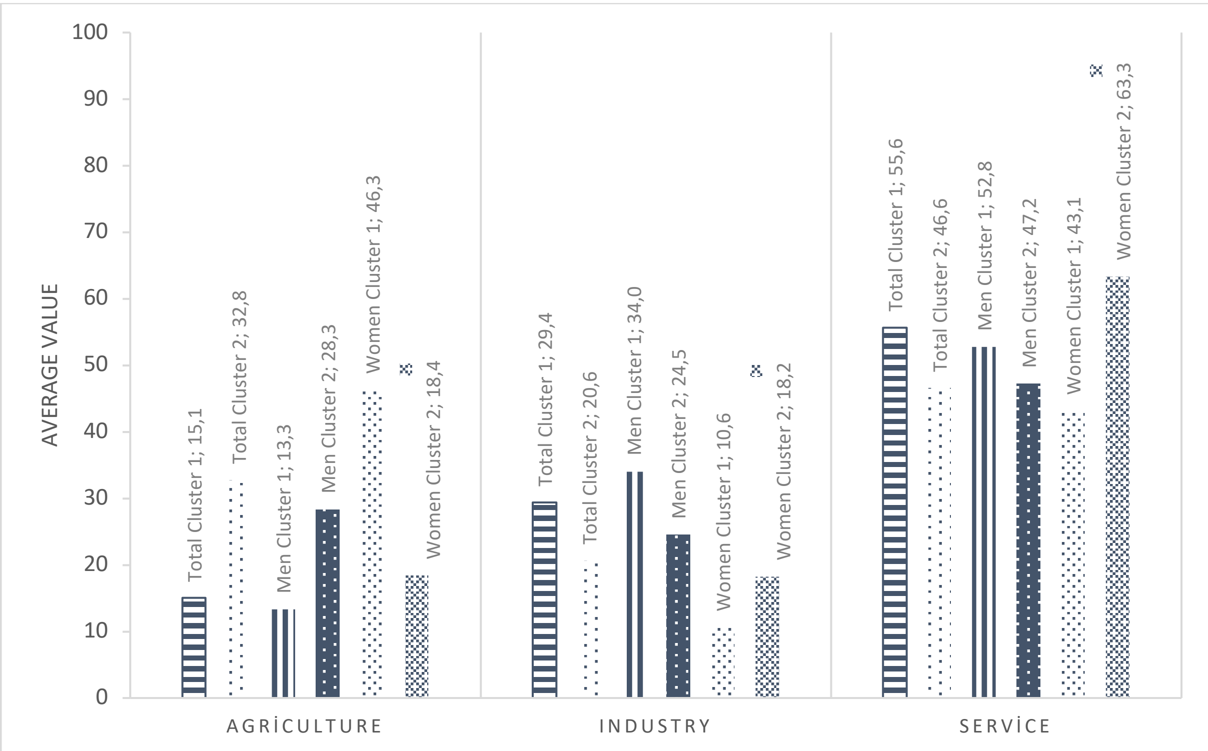

Figure 2. Average Value of Sectors in Clusters for 2021

The clustered distribution of sectors according to total and gender for year 2022 is shown in Figure 3 graphically.

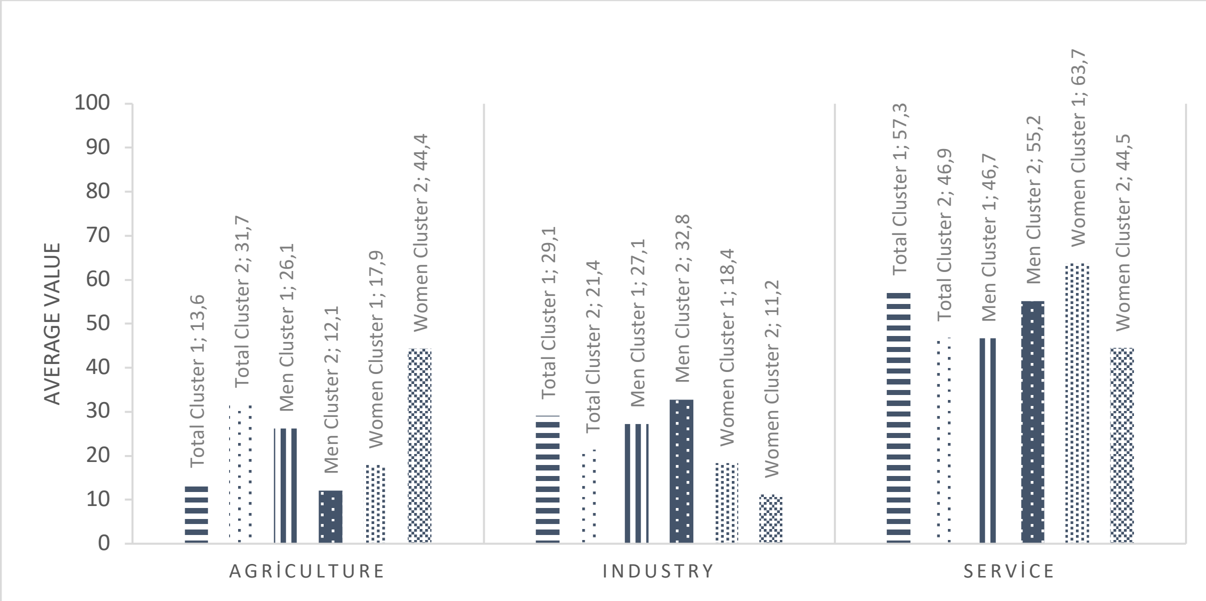

Figure 3. Average Value of Sectors in Clusters for 2022

Conclusion

With the use of employment rates, this paper aims to categorize the Türkiye provinces. TUIK publishes the sector-specific employment rate statistics that was used in the study. The analysis was conducted using 2021 and 2022 data. For women, men, and the entire sample, data were grouped into subcategories. When these data are evaluated, it becomes clear that men's sectoral employment distribution and overall employment distribution are comparable. There is a difference from the overall when the sectoral employment data of women is studied. It is clear that regulations and incentives should be put in place in this area because the employment of women in the industry sector is below average. The employment rates of women are higher than average in other sectors, excluding industry, which is good news for our nation as we strive to accomplish the 2030 Sustainable Development Goals. The incentives and norms applicable to women's participation in the labor market have a tangible impact on labor demand, female labor force participation and therefore macroeconomic outcomes. Economic opportunity policies to be encouraged for women will have a positive correlation with both women's success and economic growth (Karlilar and Kiral, 2019).

The data were used in the study's clustering analysis of the provinces. The mathematical model often saves time and can be corrected and resolved immediately if necessary; it can enable rapid response to situational changes, crises and sudden situations. While other clustering algorithms focus on a specific goal, results can be obtained by taking different objective functions into account with the mathematical model. For this reason, cluster analysis has been handled with a mathematical model in the study. In the generated model, it is intended to reduce similarity all over the clusters while enhancing similarity inside the cluster. The developed mathematical model minimizes the distance to other objects in the cluster in order to increase the similarity of the objects in the cluster and maximizes the smallest distance to the other objects outside the cluster in order to reduce the similarity of the objects to the objects in other clusters outside the cluster. The employment situation of the two clusters as determined by the study should be taken into consideration when establishing a more detailed employment policy for each cluster separately. For women, men, and the whole population, there are two clusters for 2021 and 2022.

In the clusters obtained for 2021 considering total population, cluster 1 includes the following regions: TR10, TR21, TR31, TR33, TR41, TR42, TR51, TR61, TR62, TR63, TR72, TRC1, TRC3 and Cluster 2 covers the following region: TR22, TR32, TR52, TR71, TR81, TR82, TR83, TR90, TRA1, TRA2, TRB1, TRB2, TRC2. In the clusters obtained for 2022 considering total population, cluster 1 includes the following regions: TR10, TR21, TR22, TR31, TR32, TR41, TR42, TR51, TR61, TR62, TR63, TR72, TRC1and Cluster 2 covers the following region: TR33, TR52, TR71, TR81, TR82, TR83, TR90, TRA1, TRA2, TRB1, TRB2, TRC2, TRC3. The comparison of these results show that TR22, TRC3 TR32, TR33 are assigned to different clusters.

When comparing 2021 and 2022; total, TR22, TR32, TR33, TRC3; for male, TR21, TR22, TR32, TR52, TR81; for women, TRC1 regions are assigned to different clusters.

In Turkey, employment in the agricultural sector has decreased over the years, while employment in the services sector has increased. In the TR22 and TR32 regions, employment rates decreased in the agricultural sector and increased in the services sector. In TR33 region, while the employment rate remained the same in agriculture, it decreased in industry and increased in services. In TRC3 region, contrary to Turkey as a whole, employment rate increased in the agricultural sector and decreased in the service sector.

In this paper, cluster analysis was conducted using sectoral employment rates in 2021 and 2022. It can be planned to conduct a study taking into account the data of the following years. In this way, it will be possible to compare the contributions of sectors to employment over the years.

References

Agarwal, G. and Kempe, D. (2008). Modularity-maximizing graph communities via mathematical programming. *The European Physical Journal B*, 66, 409-418. https://doi.org/10.1140/epjb/e2008-00425-1

Ágoston, K. C. and Nagy, M. E. (2023). Mixed integer linear programming formulation for K-means clustering problem. *Central European Journal of Operations Research*, 32, 1-17. https://doi.org/10.1007/s10100-023-00881-1

Benati, S. and García, S. (2014). A mixed integer linear model for clustering with variable selection. *Computers & Operations Research*, 43, 280-285. http://dx.doi.org/10.1016/j.cor.2013.10.005

Benati, S., García, S. and Puerto, J. (2018). Mixed integer linear programming and heuristic methods for feature selection in clustering. *Journal of the Operational Research Society*, *69*(9), 1379-1395. https://doi.org/10.1080/01605682.2017.1398206

Cafieri, S. and Hansen, P. (2014). *In models, algorithms and technologies for network analysis: From the third international conference on network analysis*. Switzerland: Springer International Publishing.

Caramia, M. and Pizzari, E. (2022). Clustering, location, and allocation in two stage supply chain for waste management: A fractional programming approach. *Computers & Industrial Engineering*, 169, 108297. https://doi.org/10.1016/j.cie.2022.108297

Çamlica, Z. and Şenkayas, H. (2020). Kümelenme potansiyelinin belirlenmesi: TR32 bölgesi imalat sektörlerinde bir uygulama. *Yaşar Üniversitesi E-Dergisi*, *15*, 88-105. https://doi.org/10.19168/jyasar.657876

Ferreira, L. and Hitchcock, D. B. (2009). A comparison of hierarchical methods for clustering functional data. *Communications in Statistics-Simulation and Computation*, *38*(9), 1925-1949. https://doi.org/10.1080/03610910903168603

Florez, L. and Cortissoz, J. C. (2017). Using workers compatibility to predict labor productivity through cluster analysis. *Procedia Engineering*, 196, 359-365. doi: 10.1016/j.proeng.2017.07.211

Gürler, C. (2023). Ülkelerin turizm potansiyellerine göre K-ortalamalar yöntemi kullanılarak kümelenmesi. *MANAS Sosyal Araştırmalar Dergisi*, *12*(1), 353-363. https://doi.org/10.33206/mjss.1082471

Hansen, P. and Jaumard, B. (1997). Cluster analysis and mathematical programming. *Mathematical programming*, *79*(1-3), 191-215. http:// doi.org/10.1007/BF02614317

Karlilar, S. and Kiral, G. (2019). Kadın işgücüne katılımı ve ekonomik büyüme arasındaki ilişki: Ülke grupları için panel veri analizi. *Third Sector Social Economic Review*, *54*(2), 935-948. http:doi.org/10.15659/3.sektor-sosyal-ekonomi.19.04.1099

Klein, G. and Aronson, J. E. (1991). Optimal clustering: A model and method. *Naval Research Logistics (NRL)*, *38*(3), 447-461. https://doi.org/10.1002/1520-6750(199106)38:3<447::AID-NAV3220380312>3.0.CO;2-0

Kusiak, A. (1984). Analysis of integer programming formulations of clustering problems. *Image and Vision Computing*, *2*(1), 35-40. https://doi.org/10.1016/0262-8856(84)90042-8

McDermott, A. M., Heffernan, M. and Beynon, M. J. (2013). When the nature of employment matters in the employment relationship: A cluster analysis of psychological contracts and organizational commitment in the non-profit sector. *The International Journal of Human Resource Management*, *24*(7), 1490-1518. https://doi.org/10.1080/09585192.2012.723635

Mehrotra, A. and Trick, M. A. (1998). Cliques and clustering: A combinatorial approach. *Operations Research Letters*, *22*(1), 1-12. https://doi.org/10.1016/S0167-6377(98)00006-6

Mokhtarzadeh, M., Tavakkoli-Moghaddam, R., Triki, C. and Rahimi, Y. (2021). A hybrid of clustering and meta-heuristic algorithms to solve a p-mobile hub location–allocation problem with the depreciation cost of hub facilities. *Engineering Applications of Artificial Intelligence*, *98,* 104-121. https://doi.org/10.1016/j.engappai.2020.104121

Muntaner, C., Chung, H., Benach, J. and Ng, E. (2012). Hierarchical cluster analysis of labour market regulations and population health: a taxonomy of low-and middle-income countries. *BMC Public Health*, *12*(1), 1-15. Erişim adresi: http://www.biomedcentral.com/1471-2458/12/286

Paşnicu, D., Pîrciog, S., Ghinăraru, C. and Tudose, G. (2014). Cluster analysis of the EU countries in terms of labor market indicators. *International Conferans on Procedia of Economics and Business Administration*, *0*(0), 222-229. https://icesba.eu/RePEc/icb/wpaper/ICESBA2014_27PASNICU_P222-229.pdf

Pirim, H., Ekşioğlu, B., Perkins, A. D. and Yüceer, Ç. (2012). Clustering of high throughput gene expression data. Computers & Operations Research, *39*(12), 3046-3061. doi: 10.1016/j.cor.2012.03.008

Pirim, H., Eksioglu, B. and Glover, F. W. (2018). A novel mixed integer linear programming model for clustering relational networks. *Journal of Optimization Theory and Applications*, *176,* 492-508. https://doi.org/10.1007/s10957-017-1213-1

Puerto, J., Rodríguez-Madrena, M. and Scozzari, A. (2020). Clustering and portfolio selection problems: A unified framework. *Computers* & *Operations Research*, *117,* 104891. https://doi.org/10.1016/j.cor.2020.104891

Rao, M. R. (1971). Cluster analysis and mathematical programming. *Journal of the American Statistical Association*, *66*(335), 622-626. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/01621459.1971.10482319

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster Analysis. *Journal of Computational and Applied Mathematics*, *20,* 53-65. https://doi.org/10.1016/0377-0427(87)90125-7

Sağlam, B., Salman, F. S., Sayın, S. and Türkay, M. (2006). A mixed-integer programming approach to the clustering problem with an application in customer segmentation. *European Journal of Operational Research*, *173*(3), 866-879. https://doi.org/10.1016/j.ejor.2005.04.048

Seki, İ. and Arslan, M. (2018). TRC2 (Diyarbakır–Şanlıurfa) Bölgesi kümelenme potansiyeli analizi. *Al Farabi Uluslararası Sosyal Bilimler Dergisi*, *2*(1), 32-53. Retrieved from https://dergipark.org.tr/en/download/article-file/431072

Sema, A. Y. (2012). Türkiye'de işsizliğin nedenleri: istihdam politikaları üzerine bir değerlendirme. *Yönetim ve Ekonomi Dergisi*, *19*(2), 321-341. Retrieved from https://dergipark.org.tr/en/download/article-file/146108

Vinod, H. D. (1969). Integer programming and the theory of grouping. *Journal of the American Statistical association*, 64(326), 506-519. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/01621459.1969.10500990

Yamaç, B. (2019). Tekstil sektöründe kümelenme Türkiye tekstil sektörü örneği. *Kahramanmaraş Sütçü İmam Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *9*(2), 215-232. Retrieved from http://iibfdergisi.ksu.edu.tr/en/download/article-file/906847

Yaprakli, S. and Aslan, Ö. F. (2023). Tra1 bölgesi illerinde (Erzurum, Erzincan, Bayburt) kümelenme potansiyeli ve yerel rekabet gücü: Üç yıldız analizine dayalı bir saha araştırması. *Anadolu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 24*(2), 429-458. https://doi.org/10.53443/anadoluibfd.1115658

Wang, S., Zhang, W., Bie, Y., Wang, K. and Diabat, A. (2019). Mixed-integer second-order cone programming model for bus route clustering problem. *Transportation Research Part C: Emerging Technologies*, 102, 351-369. https://doi.org/10.1016/j.trc.2019.03.019

Xu, G., Tsoka, S. and Papageorgiou, L. G. (2007). Finding community structures in complex networks using mixed integer optimisation. *The European Physical Journal B*, *60*, 231-239. Retrieved from https://link.springer.com/article/10.1140/epjb/e2007-00331-0

Genişletilmiş Özet

Amaç

Ülkelerin ekonomik olarak kalkınmalarının, bölgesel farklılıklarını gözeterek geliştirdikleri kalkınma ve istihdam planlarıyla ilişkili olduğu aşikardır. Farklı coğrafi bölgelerin yapısına uygun olarak ortaya konulan politikalar insan kaynaklarının geliştirilmesine ve bölgesel dengeli büyümeye katkı sağlamaktadır. Bu çalışmanın amacı, bölgesel sektörel istihdam verilerini kullanarak ülkemiz şehirlerinin benzerliklerine ve farklılarına göre istihdam kümelerinin oluşturulmasıdır.

Tasarım ve Yöntem

Çalışmanın literatür araştırması kısmında istihdam konusundaki çalışmaların taranması sonucunda ülkemiz özelinde sektörel bazlı istihdam durumu değerlendirmesinin kümeleme analizi ile yapılmadığı ortaya konmuştur. Bu çalışmada sektörel bazlı istihdam oranları kullanılarak kümeleme analizi yapılmıştır. Ülkemizde, istihdamın sektörel dağılımının şehirler bazında incelenmesi için şehirlerin kümeleme analizi Düzey 2 bölgeleri çerçevesinde 26 bölge için gerçekleştirilmiştir. Çalışmada veri olarak, TÜİK'in 2021ve 2022 yılları için yayınlamış olduğu tarım, sanayi ve hizmet sektörlerine özel toplam, erkek ve kadın istihdam oranı istatistikleri kullanılmıştır. Kümeleme analizi, çok değişkenli istatistiksel tekniklerdendir. Kümeleme analizi, birimlerin değişkenlere göre benzerlikleri bakımından ayrık kümelerde toplanmasını sağlayan bir yöntemdir. Kümeleme analizi yönteminin amacı birimleri, sahip oldukları karakteristik özelliklerine göre, gruplandırmaktır. Aynı kümedeki birimlerin benzerliklerinin yüksek olması istenirken bunun neticesi olarak farklı kümelerdeki nesnelerin de benzerliklerinin az olması beklenmektedir. Çalışmada, kümeleme analizi karma tamsayılı doğrusal programlama modeli ile yapılmıştır. Modelin amaç fonksiyonu küme içi nesnelerin uzaklıklarını minimize ederken farklı kümelerdeki nesneler arasındaki uzaklığı da maksimum edecek şekilde formüle edilmiştir. Aynı kümedeki birimlerin her biri için küme içindeki diğer birimlere olan uzaklıkları hesaplanmış ve küme içi uzaklığı en yüksek olan birimin uzaklık değeri belirlenmiştir. Her bir nesne için farklı kümelerdeki nesnelere olan toplam uzaklıklar hesaplanmış ve küme dışı uzaklığı en düşük olan birimin uzaklık değeri belirlenmiştir. Küme içi en yüksek uzaklık değerinin küme dışı en düşük uzaklık değerinden farkı maksimize edilmiştir.

Bulgular

Matematiksel model 3600 saniye üst sınır süresince çalıştırılmıştır. 2021 yılı verileriyle toplam, erkek ve kadın için yapılan üç analizde de iki, üç ve dört küme için optimum çözümler bulunmuştur. Kadın sektörel istihdam analizi için beş ve on küme, erkek istihdam analizi için de on kümede optimum çözümler elde edilmiştir. 2022 yılı verileriyle toplam için yapılan sektörel istihdam analizi için iki, üç, dört, beş yedi ve sekiz küme, erkek istihdam analizi için iki, üç, dört ve beş kümede, kadın için yapılan analizlerde de iki, üç, dört, altı, yedi ve on kümeler için optimum çözümler elde edilmiştir. Sonuçların doğrulamasını yapmak için iki ila on küme arası Silhoutt endeksi hesaplanmıştır. Doğrulama analizi sonuçları iki küme için en iyi endeks değerini vermiştir. 2021 yılı için toplam ve erkek istihdam kümeleme sonuçları karşılaştırıldığında; TR33, TR52 ve TR81 bölgelerinin farklılık gösterdiği diğer 23 bölgenin aynı kümelerde yer aldığı görülmektedir. Toplam ve kadın kümeleme sonuçları karşılaştırıldığında; TR22, TR32, TR33 bölgelerinin farklılık gösterdiği tespit edilmiştir. Erkek ve kadın kümeleme sonuçları karşılaştırıldığında; TR52, TRC1, TR22 ve TR32 bölgelerinin sektörel bazda cinsiyete göre değişiklik gösteren istihdam oranlarından dolayı farklı kümelere atandığı belirlenmiştir. 2022 yılı için toplam ve erkek kümeleri karşılaştırıldığında TR21 ve TRC3 bölgeleri dışında diğer bölgelerin aynı kümede olduğu görülmektedir. Erkeklerin hizmet sektöründeki istihdam oranı diğer bölgelerin aynı sektördeki istihdam oranından TR21 bölgesinde daha düşük, TRC3 bölgesinde ise daha yüksektir. Toplam ve kadın kümeleri karşılaştırıldığında sadece TRC3 bölgesi farklı kümeye atanmıştır. TRC3 bölgesinde kadın istihdam oranının diğer bölgelere göre nispeten daha düşük tarım sektörü istihdam oranına sahip olduğu görülmektedir. Erkek ve kadın kümeleri incelendiğinde de TR21 bölgesinin farklı kümeye atandığı görülmektedir. Diğer bölgelerle karşılaştırıldığında TR21 bölgesi erkek istihdamında hizmet sektöründe nispeten düşük, kadın istihdamında ise sanayi sektöründe nispeten yüksek istihdam oranına sahiptir.

2021 yılı ve 2022 yılı karşılaştırıldığınde ise; toplamda, TR22, TR32, TR33, TRC3; erkek için, TR21, TR22, TR32, TR52, TR81; kadın için, TRC1 bölgelerinin farklı kümelere atandığı görülmektedir. TRC1 bölgesinde, 2021 yılından 2022 yılına geçildiğinde kadın istihdamının tarımda 0,08 düştüğü, sanayide 0,02 arttığı ve hizmet sektöründe 0,06 arttığı görülmektedir.

Sınırlılıklar

Çalışmanın uygulaması Türkiye için yapılmıştır. Bu nedenle elde edilen sonuçlar ve yapılan değerlendirmeler ülkemizle sınırlı kalmaktadır. Türkiye'de Düzey 2 olarak 26 bölge tanımlanmıştır. Bu 26 bölgenin tarım, sanayi ve hizmet sektörleri için 2021 ve 2022 yılları istihdam oranları ile kümeleme analizi yapılmıştır.

Öneriler

Toplam nüfus, erkekler ve kadınlar için sektörel bazda istihdam oranlarının dikkate alınarak yapıldığı kümeleme analizi neticesinde elde edilen iki küme incelenmiştir. 2021 yılında, toplam ve erkek için bu iki kümenin en düşük ortalama değerleri birinde tarım sektörüde diğerinde sanayi sektöründedir. Kadın için yapılan analiz sonucu elde edilen her iki kümenin de en düşük değeri sanayi sektöründedir. 2022 yılında, toplam için bu iki kümenin en düşük değerleri bir kümede tarım sektöründe diğerinde sanayi sektöründe, erkek için iki kümede de tarım sektöründedir. Kadın için yapılan analiz sonucunda her iki kümenin de en düşük değerinin yine sanayi sektöründe olduğu sonucuna ulaşılmıştır. Ülkemizde kadınlar tarım ve hizmet sektöründe istihdam edilmektedir. Sanayi sektöründe kadınların desteklenmesi gerekmektedir. Her iki yıl cinsiyet bazında yapılan analiz neticesinde elde edilin iki küme içinde hizmet sektörü istihdam oranlarının yüksek olması, ülkemiz istihdamının hizmet sektöründe yoğunlaştığını göstermektedir.

Özgün Değer

Literatürde yaptığımız araştırmaya göre bu çalışma ülkemizdeki sektörel istihdam durumunu matematiksel model ile kümeleyen ilk çalışmadır. Matematiksel modelin özgünlüğü ise içerisindeki mesafeyi en aza indirirken aynı zamanda küme dışındaki mesafelerin de en büyük olmasını amaçlıyor olmasıdır. Buna ek olarak da ülkemizdeki bölgelerin istihdam oranlarına göre kümelemesi çalışmanın diğer bir özgün değeridir.

Araştırmacı Katkısı: Banu BİTGEN SUNGUR (%50), Fatma Selen MADENOĞLU (%50).