Akademik Araştırmalar ve Çalışmalar Dergisi 2024, 16(31), 331-344
Journal of Academic Researches and Studies 2024, 16(31), 331-344
https://doi.org/10.20990/kilisiibfakademik.1507942

Makale Türü: Araştırma Makalesi
Paper Type: Research Paper

Geliş Tarihi/Received Date: 01.07.2024
Kabul Tarihi/Accepted Date: 16.09.2024

Clustering Türkiye's Provinces According to Education Level Indicators with Two-Stage Cluster Analysis

Eğitim Düzeyi Göstergelerine Göre Türkiye İllerinin İki Aşamalı Kümeleme Analizi ile Kümelenmesi

Selim TÜZÜNTÜRK[1]

Abstract

Purpose: This study aims to cluster similar provinces of Türkiye according to (i) below-high school and (ii) high school and above-high school education level indicators using cluster analysis methods.

Design/Methodology: The annual most recent data for 81 provinces were taken from the education statistics for the year 2022 published on the official website of the Turkish Statistical Institute. A two-stage cluster analysis method was used to identify similar provinces.

Findings: According to indicators of education level below the high school level, the first cluster consists of 17 provinces (Adıyaman, Ağrı, Bingöl, Bitlis, Diyarbakır, Gaziantep, Hakkâri, Kars, Mardin, Muş, Siirt, Şanlıurfa, Van, Batman, Şırnak, Iğdır and Kilis) and the second cluster consists of 64 provinces (other remaining provinces). The 17 provinces in the first cluster are eastern provinces where education needs to be increased. According to the education level indicators at high school and high school level, the first cluster consists of 6 provinces (Ankara, Çanakkale, Eskisehir, Isparta, Istanbul, and Izmir), and the second cluster consists of 75 provinces (the remaining other provinces). The six provinces in the first cluster are western provinces where education is advanced.

Limitations: The limitation of the study is that cluster analyses were conducted for Türkiye's 81 provinces and 2022 cross-sectional data.

Originality/Value: The use of the two-stage cluster analysis method, which allows the number of clusters to be automatically determined according to AIC and BIC criteria, unlike the Ward and K-means methods used in cluster analysis, contributes to the original value of the research. Considering that the increase in education benefits people and society, it points out the importance of the study in providing meaningful and valuable information by creating awareness among individuals and policymakers in the relevant provinces by identifying similar provinces with clustering methods.

Keywords: Education, Türkiye, Two-Step Cluster Analysis, Multivariate Statistics.

Öz

Amaç: Bu çalışmanın amacı, (i) lise altı ve (ii) lise ve lise üstü eğitim düzeyi göstergelerine göre Türkiye'nin birbirine benzer illerinin kümeleme analiz yöntemleri ile kümelenmesidir.

Tasarım/Yöntem: 81 ilin yıllık en güncel verileri Türkiye İstatistik Kurumu'nun resmi internet sitesinde yayınlanan 2022 yılı eğitim istatistiklerinden alınmıştır. Benzer illerin belirlenmesi için iki aşamalı kümeleme analizi yöntemi kullanıldı.

Bulgular: Lise düzeyinin altındaki eğitim düzeyi göstergelerine göre 17 ilden oluşan birinci küme (Adıyaman, Ağrı, Bingöl, Bitlis, Diyarbakır, Gaziantep, Hakkâri, Kars, Mardin, Muş, Siirt, Şanlıurfa, Van, Batman, Şırnak, Iğdır ve Kilis) ve ikinci küme 64 ilden (diğer kalan iller) oluşmaktadır. Birinci kümedeki 17 il eğitimin arttırılması gereken doğu illeridir. Lise ve lise düzeyindeki eğitim düzeyi göstergelerine göre birinci küme 6 ilden (Ankara, Çanakkale, Eskisehir, Isparta, İstanbul ve İzmir) ve ikinci küme ise 75 ilden (geri kalan diğer iller) oluşmaktadır. Birinci kümedeki 6 il eğitimin ileri düzeyde olduğu batı illeridir.

Sınırlılıklar: Kümeleme analizlerinin Türkiye'nin 81 ili ve 2022 yılı kesit verileri için yapılması çalışmanın sınırlılıklarıdır.

Özgünlük/Değer: Kümeleme analizlerinde kullanılan Ward ve k-ortalama yöntemlerinde farklı olarak küme sayısının AIC ve BIC kriterlerine göre otomatik olarak belirlenmesine imkân tanıyan iki aşamalı kümeleme analiz yönteminin kullanılması araştırmanın özgün değerine katkı sağlamaktadır. Eğitim artışının insanlar ve toplum üzerinde bir fayda sağladığı düşünüldüğünde, benzer illerin kümeleme yöntemleri ile belirlenmesiyle hem ilgili illerdeki bireylere hem de politika yapıcılara farkındalık yaratılarak anlamlı ve faydalı bilgilerin sağlanması çalışmanın önemine işaret etmektedir.

Anahtar Kelimeler: Eğitim, Türkiye, İki Aşamalı Kümeleme Analizi, Çok Değişkenli İstatistik.

[1] Doç. Dr., Bursa Uludağ Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri, selimtuzunturk@uludag.edu.tr, ORCID: 0000-0002-8987-2280

1. INTRODUCTION

Education is the teaching and learning process, usually at school, college, or university (Longman Dictionary, 2003: 501). It is the process of providing, training, and developing knowledge and skills on a particular subject in a specific branch of science (Püsküllüoğlu, 2004: 321). Education means studying to acquire more profound knowledge, skills, and understanding of various subjects that can be applied to daily life. Education is one of the essential parts of people's lives. Therefore, it becomes crucial to understand some facets of the benefits of education. In this context, education's individual and social benefits can be mentioned (Johnston, 2004: 5). Education policies impact earnings, employment, development and economic growth, poverty, hunger, health, crime, and many other life outcomes. Higher education can enhance the quality of life for individuals and countries (Chan, 2016: 7).

Education is essential for individuals to have a job, a higher probability of finding a better job, and earn more money and higher income. More education means increased opportunities (Baum et al., 2013: 10). Well-educated people are less likely to be unemployed than poorly educated and more likely to work full-time (Ross & Wu, 1995: 721). A considerable number of studies (e.g. Blaug, 1972; McMahon, 2002; Johnston, 2004; Murray, 2009; Faridi et al., 2010; Baum et al., 2013; Hermannsson, et al., 2017; Rizvanovic & Efendic, 2021; Ma & Pender, 2023) in the literature is seen concentrated on the direct economic effects of education. These studies generally focus on the correlation between education and earnings and the causal effects of education on earnings. Studies have reported that higher wages are achieved individually as education increases, and higher employment, lower unemployment, and higher national income are achieved socially. Moreover, Lee et al. (2024) reported that parents' education positively affects children's earnings.

Human capital (knowledge, skills, and experience) determines economic growth (Barro, 2013: 9). Therefore, education is the fundamental basis for development and growth in individuals and society. Education helps to enhance an individual's career and personal development. Greater productivity and creativity are achieved with educated individuals. Education encourages entrepreneurship, technological advances, and economic growth, increasing a country's GDP. The literature review shows that education's impact on economic growth was studied for different economies. The relationship was explored by Lin (2004) for Taiwan, by Brempong et al. (2006) for African countries, by Aghion et al. (2009) for the United States., by Huang et al. (2009) for China, by Odit et al. (2010) for Mauritius, by Ochilov (2012) for Uzbekistan, by Reza and Widodo (2013) for Indonesia, by Mercan and Sezer (2014) for Türkiye, Lv et al. (2017) for the regions of China, Dudzeviciute and Simelyte (2018) for the selected European Union Countries, Irughe and Edafe (2020) for Nigeria, Aizizi (2021) for China and by Bah (2023) for cross countries. Studies have reported that higher education positively and significantly affects economic development for related data sets of different economies.

Education helps eliminate poverty and hunger, giving people a chance at a better quality of life. It empowers individuals with the knowledge and skills to achieve better job opportunities and higher incomes. When the literature is examined, it is seen that research generally focuses on the effect of education on reducing poverty in different cultures (e.g., Njong (2010) for Cameroon, Julius and Bawane (2011) for Kenya, Gounder and Xing (2012) for Fiji, Thapa (2013) for Nepal, Shimeles and Chouchane (2016) for Sudan, Arsani et al. (2020) for Indonesia, Hofmarcher, (2021) for European countries). Education and poverty are inversely related (Awan et al., 2011: 1): The higher the level of education of the population, the lesser the number of low-income persons because education imparts knowledge and skills, which is supportive of higher wages. As more people access quality education, the poverty rate decreases, living standards are improved, and income inequality is reduced. At the macro level, it is generally examined that low-income countries have low levels of education. At the micro level, children of low-income families do not attend schools (Awan et al., 2011: 2).

High levels of labor force participation, employment, and earnings increase the material well-being of individuals and the wealth of society and also carry psychological benefits (Baum et al., 2013: 10). Since well-educated people have higher incomes than uneducated people, they experience less economic distress, and thus the working and economic conditions of well-educated people can

better protect their health (Ross & Wu, 1995: 719). Education contributes to human capital by developing personal control, and personal control can impact individuals' attitudes and behaviors, including health behaviors (Zimmerman & Woolf, 2014: 4). People who do not have control mechanisms possibly exhibit uncontrolled behaviors such as substance use, alcohol, and drug addictions. Moreover, education paves the way for the adoption of healthier lifestyles and the prevention of diseases. Educated people are likelier to live healthy lifestyles such as smoking less, exercising more, and loving obesity rates. More highly educated people live longer and more nutritious (Baker et al. 2011: 307). Low educational attainment is associated with high rates of infectious disease, many chronic noninfectious diseases, self-reported poor health, shorter survival when sick, and shorter life expectancy (Ross & Wu, 1995: 719). In this context, it is possible to find several studies in the literature that have shown a positive correlation between education and health outcomes and also that have demonstrated associations between higher education and better health (e.g. Johnston, 2004; Silles, 2009; Baker et al., 2011; Grounder & Xing, 2012; Raghupathi & Raghupathi, 2020; Arsani et al., 2020).

The review of the literature shows that education has an impact on the reduction of crime rates (e.g., Johnston, 2004; Buonanno & Leonida, 2006; Groot & Van Den Brink, 2010; Lochner, 2011; Bell et al., 2022; Meghir et al., 2023). Studies state that policies that increase education help to reduce crime rates. Besides the main benefits of education explained above, education has advantages in different facets, such as human rights, social peace, democracy, gender equality, environmental awareness, national security, etc.

The vital importance of education in human life within the framework of the above explanations is undeniable, and conducting educational research is very important. In this context, this study focuses on clustering Türkiye's provinces according to education level indicators. Considering that the increase in education benefits people and society, identifying similar provinces through clustering methods and providing meaningful and valuable information by raising awareness to both individuals and policymakers in the relevant provinces points to the importance of the current study. When the literature was examined, it was seen that two main groups of clustering studies were conducted using education-level indicators. One is for the countries, and the other is for the provinces. Güler and Veysikarani (2022) classified OECD countries by cluster analysis according to education level, using data on 2018 PISA scores. Astuti and Rezania (2022) grouped the sub-districts of Sidoarjo Regency in Indonesia based on the education indicators using cluster analysis. Bhalla and Meher (2022) clustered major states of India by cluster analysis according to the education indicator variables. De Souza and Taceneli (2022) used educational indicators to cluster municipalities of Paraná, Brazil, by using cluster analysis. Based on the education indicators using cluster analysis, Kaynak and Rashid (2020) clustered Economic Alptekin (2015) clustered EU countries and Türkiye. Cooperation Organization member countries depending on the education indicators using cluster analysis. Akın and Eren (2012) clustered OECD countries using cluster analysis methods based on the education indicators. Öz et al. (2009) clustered EU countries and Türkiye based on the education indicators using cluster analysis techniques. On the other hand, as a result of the literature review, a small number of studies were found on the clustering of provinces in Türkiye according to education indicators, which is also the subject of this study. Özaydın et al. (2022) used 2020-2021 education statistics (10 variables) to cluster 81 provinces of Türkiye by using the k-means cluster technique. Öztürk and Gürsakal (2015) used 2013 education statistics (4 variables) to cluster 81 provinces of Türkiye by using multidimensional scaling analysis. Using the k-means cluster technique, Uzgören et al. (2013) used basic education indicators (30 variables) to cluster 81 provinces of Türkiye. Cengiz and Öztürk (2012) used 2010 education statistics (8 variables) to cluster 81 provinces of Türkiye by using the k-means cluster technique. This research is completely different from the studies of Özaydın et al. (2022), Öztürk and Gürsakal (2015), and Uzgören et al. (2013) in terms of the variables used, period, purposes of the studies, and the method used. This research is also different from Cengiz and Öztürk (2012) in terms of the variables' definitions, the period, the purposes of the study, and the method used.

The aims of this study are (i) the determination of similar provinces of Türkiye according to indicators of education level lower than high school level and (ii) the determination of similar

provinces of Türkiye according to indicators of education level of high school and higher than high school level. The findings will be interpreted and turned into helpful information for people living in the related provinces and for policymakers in the Turkish education system. Clustering technique selection and variable selection are essential for obtaining meaningful clusters in cluster analysis. A two-stage clustering analysis was used in this study to cluster similar provinces. This technique has significant advantages, such as being applied to large data sets containing quantitative and qualitative variables and automatically determining the most appropriate number of clusters with AIC and BIC. To obtain different meaningful clusters based on education level, education indicator variables were divided into two poles: (i) less than high school level and (ii) high school level and above. The rest of the paper is organized as follows: A two-stage cluster analysis is mentioned in the study's second part. The third section includes data, analysis, and findings. The last section consists of the conclusion.

2. TWO-STEP CLUSTER ANALYSIS

Analysis of the variables simultaneously that are more than two($p > 2$; such as $X_1, X_2, \ldots. X_p$) is considered as multivariate analysis (Hair et al., 2013: 3). The essence of multivariate analysis is understanding the relationships of many variables with each other (Sarma & Vardhan, 2019: 4). All variables are examined simultaneously to uncover the patterns and critical features in the data (Everitt, 2007: 1).

Multivariate data is characterized by a mean vector (**x**), data matrix (X), covariance matrix (S), and correlation matrix (R) and displayed with profiles, stars, glyphs, faces, and boxes (Alpar, 2011: 43-88; Rencher & Christensen, 2012: 57-71). Cluster analysis is one of the multivariate analysis methods that covers a collection of methods developed to cluster (divide into homogeneous subgroups) the observations (sometimes variables) in the **X** data matrix within the framework of their characteristics (Alpar, 2011: 309). Because the similarities are determined using a set of variables selected, variable selection is vital in clustering (Morrison, 1967: 779; Li, 2006: 457).

There are two main clustering, including hierarchical and non-hierarchical methods (Tatlıdil, 2002: 334; Turanlı et al., 2006: 99). When the number of clusters is unknown, hierarchical clustering is used (Bülbül & Camkıran, 2018: 371). The number of clusters is discovered by using a diagram called a dendrogram. A researcher with a naked eye selects the number of clusters (k) from a dendrogram by cutting across the branches at a given level of distance measured by one of the axes (Rencher & Christensen, 2012: 544). In non-hierarchical methods, the number of clusters is predetermined. These methods begin with a predetermined classification and, through various iterative processes, try to find a revised classification that will optimize a measure of homogeneity of the cluster (Blashfield & Aldenderfer, 1978: 272). A part from the two approaches above, the two-step clustering method is another approach (Alpar, 2011: 314). The most important features of this method are (Alpar, 2011: 339): (i) it can be applied to large data sets, (ii) it can be applied to both qualitative and quantitative variables, and (iii) the most appropriate number of clusters is automatically determined according to AIC and BIC criteria. In the first step, data is assigned to small clusters to decrease the distance matrix's dimension; in the second step, hierarchical clustering is applied to pre-clusters.

In addition, different clustering techniques will often generate distinct solutions to the same data set (Morey et al., 1983: 309). For this reason, a reasonable method should be chosen. In the validation of the cluster solutions, analysis of variance (ANOVA) tests (F tests) are used (Morey et al., 1983: 317).

3. DATA, ANALYSIS AND FINDINGS

The annual most recent data for 81 provinces were taken from the education statistics for the year 2022 published on the official website (https://www.tuik.gov.tr/) of the Turkish Statistical Institute. The original data consists of numerical values (number of people). The variables' provincial values(number of people) representing education levels were proportioned to the total number of people, and these proportional data were used in the analysis.

The data includes eight variables (illiterate, literate without a diploma, primary school, middle school, high school, universities, and other higher educational institutions, master (including five or

six years' faculties), and doctorate) representing the distribution of the education level of the population aged six and over by province. These eight variables were divided into two by taking the high school education level as a threshold. The first group of indicators is called indicators of education level lower than high school level. The first group of indicators is called indicators of education level high school and higher than high school level.

Table 1: Indicators of Education Level Lower Than High School Level

Symbols	Names
L1	Illiterate
L2	Literate without a diploma
L3	Primary school
L4	Middle school

Table 2: Indicators of Education Level High School and Higher Than High School Level

Symbols	Names
U1	High school
U2	Universities and other higher educational institutions
U3	Master (Including five or six years' faculties)
U4	Doctorate

The research questions are as follows: (i) Are there provinces of Türkiye that are similar in terms of the indicators of education levels lower than high school? and (ii) Are there provinces of Türkiye that are similar in terms of the indicators of education levels of high school and higher high school level?

Two-Step Cluster Analysis was performed with SPSS 23 software. As mentioned, the two-step cluster analysis automatically discovers the number of clusters. If continuous variables are to be analyzed, the Euclidean distance is chosen; if both categorical and continuous, the log-likelihood is selected in a two-step cluster analysis (Tkaczynski, 2017: 110). So, in this study, Euclidean distance was used as a distance measure. Schwarz's Bayesian Criterion (BIC) was used as a clustering criterion. The two-step cluster analysis procedures evaluated the importance of cluster quality and predictor (X variables). Finally, F tests were used to determine the significant and insignificant variables in forming clusters.

3.1. Analyses of Lower than High School Level

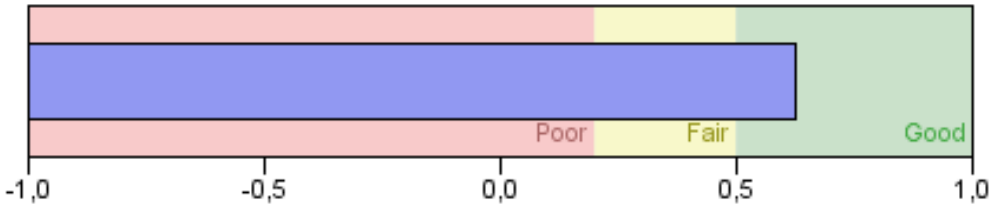
Two cluster solutions were found in the two-step cluster analysis with four input variables (L1, L2, L3, and L4). The cluster memberships are found as seen in the following table below:

Table 3: Two-Step Cluster Analysis Cluster Memberships

Clusters	Number and Percentage of Provinces in Each Cluster	Province
Cluster 1	17 (21 %)	Adıyaman Ağrı Bingöl Bitlis Diyarbakır Gaziantep Hakkâri Kars Mardin Muş Siirt Şanlıurfa Van Batman Şırnak Iğdır Kilis
Cluster 2	64 (79 %)	Other remaining provinces

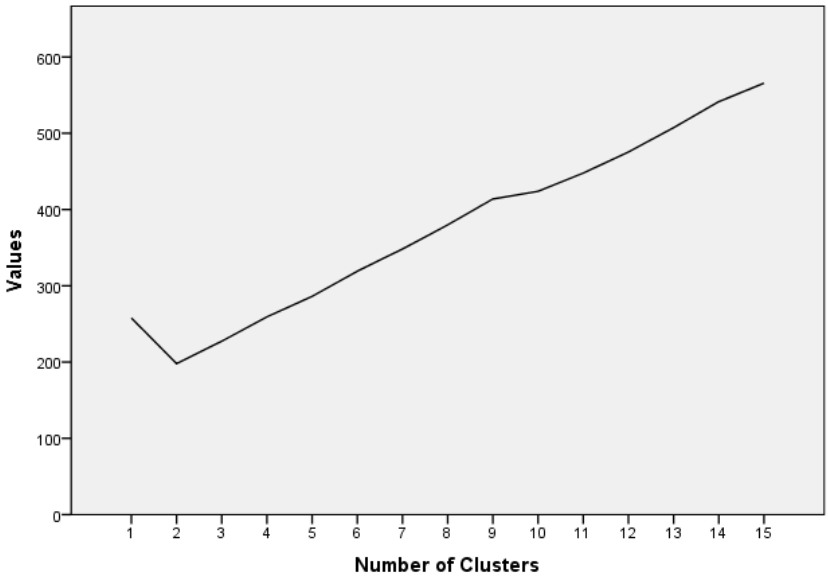
Cluster 1 is composed of 17 provinces (Adıyaman, Ağrı, Bingöl, Bitlis, Diyarbakır, Gaziantep, Hakkâri, Kars, Mardin, Muş, Siirt, Şanlıurfa, Van, Batman, Şırnak, Iğdır and Kilis) and Cluster 2 composed of 64 provinces (other remaining provinces). Figure 1 shows the Silhouette measure of cohesion and separation as is seen below:

Figure 1: Cluster Quality: The Silhouette Measure of Cohesion and Separation



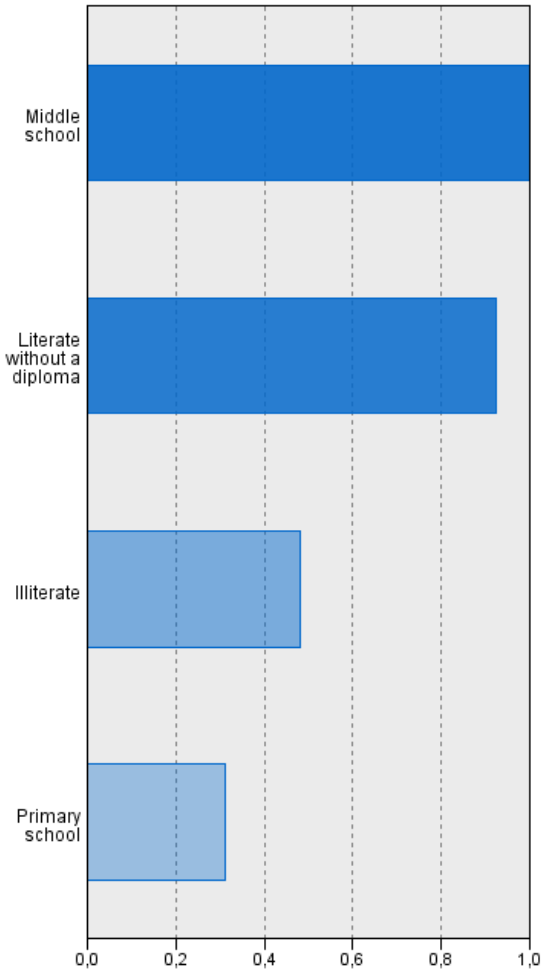
The silhouette measure of cohesion and separation measures how similar a province is to its cluster compared to others. The measure of cohesion and separation measure ranges between −1 to +1. A high value (close to 1) indicates that the province is well suited to cluster. In this Figure, the Silhouette measure of cohesion and separation shows that the cluster quality is good. The BIC criterion is used in a two-step clustering analysis to validate that the two-cluster solution is optimal. Figure 2 shows the line graph of BIC values for different cluster numbers.

Figure 2: Auto Clustering Schwarz's Bayesian Criterion (BIC)



The minimum BIC value in the graph points out the two-cluster solution, which means that the two-cluster solution seems ideal. Figure 3 shows the predictors' importance.

Figure 3: Predictors' Importance



The order of significance levels of the predictor variables that are important in the formation of clusters are as follows: Middle school (L4) 1.00, literate without diploma (L2) 0.93, illiterate (L1) 0.48, and primary school (L3) 0.31. Whether each of these variables affects dividing the provinces into two clusters can be investigated with the F test. The following table shows the F test results:

Table 4: Analysis of Variance Tests (F Tests) Results

	Cluster		Error			
	Mean Square	df	Mean Square	df	F	Sig.
Illiterate (L1)	0.007	1	0.000	79	58.486	0.000
Literate without a diploma (L2)	0.040	1	0.000	79	162.750	0.000
Primary school (L3)	0.026	1	0.001	79	35.612	0.000
Middle school (L4)	0.052	1	0.000	79	185.616	0.000

Here, the null hypothesis is the variable L_i, which does not affect the division of provinces into two clusters, and the alternative hypothesis is the variable L_i, which impacts the division of provinces into two clusters. The null hypothesis is rejected (Sig. < α=0.05) for the variables L1, L2, L3, and L4, which means that each variable affects the division of provinces into two clusters.

3.2. Analyses of High School Level and Higher Than High School Level

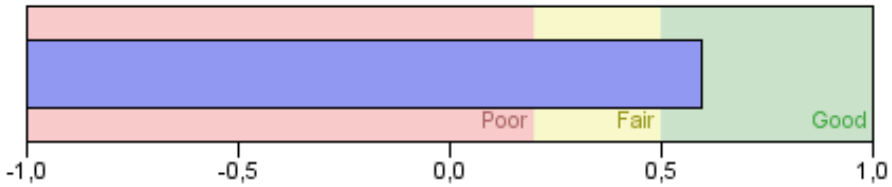
Two cluster solutions were found in the two-step cluster analysis with four input variables (U1, U2, U3, and U4). The cluster memberships are found as seen in the following table below:

Table 5: Two-Step Cluster Analysis Cluster Memberships

Clusters	Number and Percentage of Provinces in Each Cluster	Province
Cluster 1	6 (7.4 %)	Ankara Çanakkale Eskişehir Isparta İstanbul İzmir
Cluster 2	75 (92.6 %)	Other remaining provinces

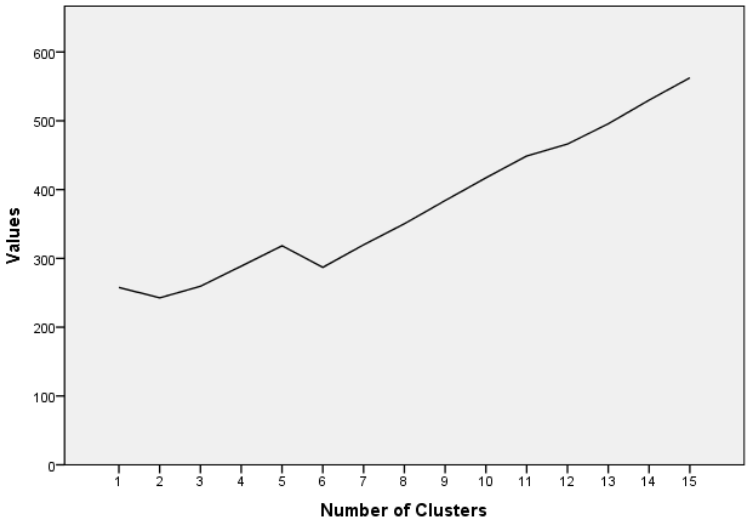
Cluster 1 is composed of 6 provinces (Ankara, Çanakkale, Eskişehir, Isparta, İstanbul and İzmir), and Cluster 2 is composed of 75 provinces (other remaining provinces). Figure 4 shows the Silhouette measure of cohesion and separation as seen below:

Figure 4: Cluster Quality: The Silhouette Measure of Cohesion and Separation



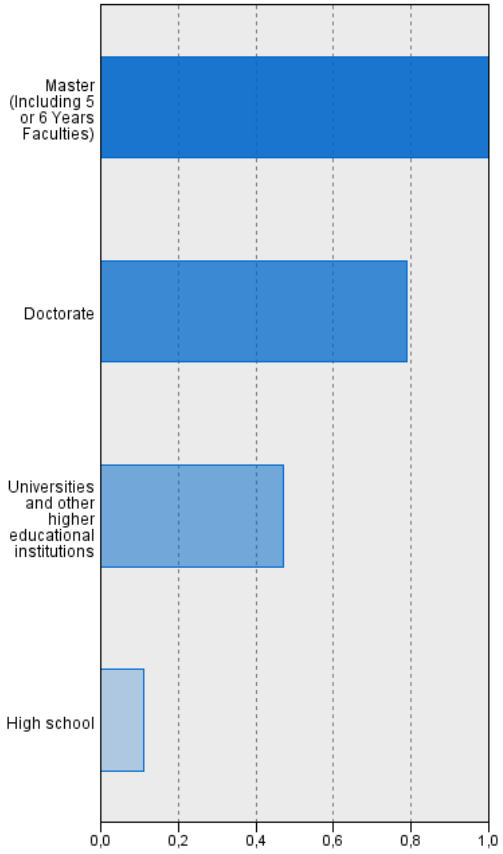
In this Figure, the Silhouette measure of cohesion and separation shows that cluster quality is good. The BIC criterion is used in a two-step clustering analysis to validate that the two-cluster solution is optimal. Figure 5 shows the line graph of BIC values for different cluster numbers.

Figure 5: Auto Clustering Schwarz's Bayesian Criterion (BIC)



The minimum BIC value in the graph points out the two-cluster solution, which means that the two-cluster solution seems ideal. Figure 6 shows the predictors' importance.

Figure 6: Predictors' Importance



The order of significance levels of the predictor variables that are important in the formation of clusters is as follows: Master (U3) 1.00, doctorate (U4) 0.79, universities and other higher educational institutions (U2) 0.47, and high school (U1) 0.11.

Whether each of these variables affects dividing the provinces into two clusters can be investigated with the F test. The following table shows the F test results:

Table 6: Analysis of Variance Tests (F Tests) Results

	Cluster		Error			
	Mean Square	df	Mean Square	df	F	Sig.
High school (U1)	0.045	1	0.000	79	136.939	0.000
Universities and other inst. (U2)	0.036	1	0.000	79	100.508	0.000
Master (U3)	0.000	1	0.000	79	31.690	0.000
Doctorate (U4)	0.000	1	0.000	79	22.263	0.000

Here, the null hypothesis is that the variable U_i does not affect the division of provinces into two clusters, and the alternative hypothesis is that the variable U_i affects the division of provinces into two clusters. The null hypothesis is rejected (Sig. < α=0.05) for the variables U1, U2, U3, and U4, which means that each variable affects the division of provinces into two clusters.

4. CONCLUSION

The aims of this study were: (i) determination of similar provinces of Türkiye according to indicators of education level lower than high school level and (ii) determination of similar provinces of Türkiye according to indicators of education level of high school and higher than high school level. For these purposes, this study sought answers to the following research questions: (i) Are there provinces of Türkiye that are similar in terms of the indicators of education levels lower than high school? and (ii) Are there provinces of Türkiye that are similar in terms of the indicators of education levels of high school and higher high school level? As a result of the empirical studies, the answer to

both research questions was obtained as yes. Two province groups were found for each one of the cases. In addition, all variables used for both research questions were statistically significant, indicating that the variables effectively divided the provinces into two clusters.

According to the indicators of education level lower than high school level, Cluster 1 is composed of 17 provinces (Adıyaman, Ağrı, Bingöl, Bitlis, Diyarbakır, Gaziantep, Hakkâri, Kars, Mardin, Muş, Siirt, Şanlıurfa, Van, Batman, Şırnak, Iğdır and Kilis) and Cluster 2 composed of 64 provinces (other remaining provinces). These are provinces where individuals should give importance to education for a better life, and policymakers should give importance to education, develop and encourage education for the development of individuals, society, and, therefore, the country. Education and the economy have a reciprocal relationship (Ilgar, 2023: 24). Regional income differences result in educational differences. The eastern provinces that make up the 1st Cluster above indicate provinces that are both underdeveloped in terms of economic development and have low education levels. Implementing economic policies by policymakers to improve and develop the regional economy will be an important initiative to eliminate the educational differences between the region's people and other regions. In regional educational differences, social and cultural values and economic situations effectively access education (Ilgar, 2023: 24). The reflection of the family's low level of education on the child or the inability to study or girls not being sent to school due to the necessity of working in agricultural activities are examples of socio-cultural reasons. In other words, Cluster 1 provinces are the eastern provinces where education should be increased, and inequalities of opportunity in education should be eliminated. In this regard, many responsibilities fall on society and policymakers, starting with families. Bilgin and Erbuğ (2021) stated the economic, social, regional, biological, and political factors that cause inequality in education opportunities for people. Everyone needs to do their best to improve these factors. According to the indicators of education level of high school and higher high school level, Cluster 1 is composed of 6 provinces (Ankara, Çanakkale, Eskişehir, Isparta, İstanbul and İzmir), and Cluster 2 is composed of 75 provinces (other remaining provinces). Cluster 1 provinces are western provinces where education is at an advanced level. These provinces are provinces that individuals and policymakers should take as examples for improving education. Ankara, İstanbul, and İzmir are the biggest provinces that receive intense immigration and have many education opportunities.

Policies in the field of education are primarily effective in building the qualifications and equipment of the individuals who make up society and, therefore, play a vital role in social transformation and development (Yıldız & Karakaş, 2019: 272). Policies that increase education have long-lasting impacts on several important life outcomes, as mentioned, such as earnings, employment, development and economic growth, poverty, hunger, health, and crime (Meghir, Palme & Schnabel, 2023: 1). Thus, both individuals and society as a whole can benefit from increased levels of education in terms these outcomes and also outcomes more as human rights, social peace, democracy, gender equality, environmental awareness, national security and so on. In this context, the primary responsibility falls on educational policymakers. Policymakers can consider and benefit from the above findings for 2022 when deciding whether the government should invest in education and develop education policies in the future.

In the study, cross-sectional data for units consisting of provinces for the year 2022 was used. Therefore, the findings and evaluations show the picture for 2022. In future research, different findings may be obtained in the light of positive or negative developments related to the important points mentioned regarding the focused topic. Again, in the future, different aspects of the subject can be addressed with other variables compiled from different sources.

Etik Beyan: *Bu çalışmada "Etik Kurul" izini alınmasını gerektiren bir yöntem kullanılmamıştır.*
Ethics Statement: *In this study, no method requiring the permission of the "Ethics Committee" was used.*

REFERENCES

Aghion, P., Boustan, L. Hoxby, C. & Vandenbussche, J. (2009). The causal impact of education on economic growth, evidence from the United States. In D. Romer & J. Wolfers (Eds.). *Brookings Papers on Economic Activity* (1-74).

Aizizi, D. (2021). An analysis of the relationship between education and economic growth. *International Journal of New Developments in Education,* 3(4), 65-67. https://doi.org/10.25236/IJNDE.2021.030411

Akın, H. B. & Eren, Ö. (2012). OECD Ülkelerinin eğitim göstergelerinin kümeleme analizi ve çok boyutlu ölçekleme analizi ile karşılaştırmalı analizi. *Öneri Dergisi,* 10 (37), 175-181. https://doi.org/10.14783/od.v10i37.1012000178

Alpar, R. (2011). *Uygulamalı çok değişkenli istatistiksel yöntemler.* Detay Yayıncılık.

Alptekin, N. (2015). Bulanık kümeleme analizi ile Türkiye ve AB ülkelerinin eğitim göstergeleri açısından sınıflandırılması. *Kafkas Üniversitesi Fen Bilimleri Enstitüsü Dergisi,* 8(2), 1-16.

Arsani, A. M., Ario, B. & Ramadhan, A. F. (2020). Impact of education on poverty and health: evidence from Indonesia. *Economics Development Analysis Journal,* 9(1), 87-96. https://doi.org/10.15294/edaj.v9i1.34921

Astuti, C. C. & Rezania, V. (2022). Cluster analysis for grouping districts in Sidoarjo regency based on education indicators. *ICIGR 2021 Conference Proceedings,* Volume 2022, 311-317. https://doi.org/10.18502/kss.v7i10.11233

Awan, M. S., Malik, N., Sarwar, H. & Wagas, M. (2011). *Impact of education on poverty reduction.* (MPRA Paper No. 31826).

Bah, I. A. (2023). The relationship between education and economic growth: A cross-country analysis. *Research, Society and Development,* 12(5), 1-13. https://doi.org/10.33448/rsd-v12i5.40522

Baker, D. P., Leon, J., Greenaway, E. G. S., Collins, J. & Movit, M. (2011). The education effect on population health: a reassessment. *Population and Development Review,* 37(2), 307-332. https://doi.org/10.1111/j.1728-4457.2011.00412.x

Barro, R. J. (2002). Education in the twenty-first century. In E. P. Lazear (Ed.), *Education as a determinant of economic growth* (pp. 9-24). Hoover Press.

Baum, S., Ma, J. & Payea, K. (2013). *Education pays 2013 the benefits of higher education for individuals and society.* Trends in Higher Education Series, College Board, New York.

Bell, B., Costa, R. & Machin, S. (2022). Why does education reduce crime?. *Journal of Political Economy,* 130(3), 732-765. https://doi.org/10.1086/717895

Bhalla, R. & Meher, S. (2022). Interstate disparities in education and health infrastructure: A multivariate analysis. *IASSI Quarterly: Contributions to Indian Social Science,* 41(4), 534-548.

Bilgin, R. & Erbuğ, E. (2021). Eğitimde fırsat eşitsizliği üzerine eleştirel bir değerlendirme. *International Journal of Economics Administrative and Social Sciences,* 4(2), 231-239.

Blashfield, R. K. & Aldenderfer, M. S. (1978). The literature on cluster analysis. *Multivariate Behavioral Research,* 13, 271-295. https://doi.org/10.1207/s15327906mbr1303_2

Blaug, M. (1972). The correlation between education and earnings: What does it signify? *Higher Education,* 1, 53-76. https://doi.org/10.1007/BF01956881

Brempong, K. G., Paddison, O. & Mitiku, W. (2006). Higher education and economic growth in Africa. *The Journal of Development Studies,* 42(3), 509-529. https://doi.org/10.1080/00220380600576490

Buonanno, P. & Leonida, L. (2006). Education and crime: evidence from Italian regions. *Applied Economics Letters,* 13(11), 709-713. https://doi.org/10.1080/13504850500407376

Bülbül, Ş. & Camkıran, C. (2018). Bankaların klasik ve bulanık yaklaşımlarla sınıflandırılması. *Trakya Üniversitesi Sosyal Bilimler Dergisi,* 20(2), 367-385. https://doi.org/10.26468/trakyasobed.464442

Cengiz, D. & Öztürk, F. (2012). Türkiye'de illerin eğitim düzeylerine göre kümeleme analizi ile incelenmesi. *Trakya Üniversitesi Sosyal Bilimler Dergisi,* 14(1), 69-84.

De Souza, D. C. & Taconeli, C. A. (2022). Spatial and non-spatial clustering algorithms in the analysis of Brazilian educational data. *Communications in Statistics: Case Studies, Data Analysis and Applications,* 8(4), 588-606. https://doi.org/10.1080/23737484.2022.2117744

Dudzeviciute, G. & Simelyte, A. (2018). Education and economic development in the selected European Union countries. *European Journal of Sustainable Development,* 7(2), 14-28. https://doi.org/10.14207/ejsd.2018.v7n2p14

Everitt, B. (2007). *An R and S-PLUS companion to multivariate analysis*. Springer.

Faridi, M. Z., Hussain, S. & Bashir, F. (2010). Impact of education on student's earnings: a case of public sector universities in Pakistan. *International Research Journal of Finance and Economics,* 42, 171-178. https://doi.org/10.2139/ssrn.1662910

Gounder, R. & Xing, Z. (2012). Impact of education and health on poverty reduction: monetary and non-monetary evidence from Fiji. *Economic Modeling,* 29, 787-794. https://doi.org/10.1016/j.econmod.2012.01.018

Groot, W. & Van Den Brink, H. M. (2010). The effects of education on crime. *Applied Economics,* 42, 279-289. https://doi.org/10.1080/00036840701604412

Güler, E. Ö. & Veysikarani, D. (2022). Sosyo-Ekonomik göstergeler ve PISA skorlarına göre ülkelerin sınıflandırılması: OECD örneği. *Ömer Halisdemir Üniversitesi İktisadi ve İdari Bilimler Fakültesi,* 15(3), 506-522. https://doi.org/10.24287. 958903

Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R. E. (2013). *Multivariate data analysis*. Pearson.

Hermannsson, K., Lisenkova, K., Lecca, P., McGregor, G. & Swales, J. K. (2017). The external benefits of higher education. *Regional Studies,* 51(7), 1077-1088. https://doi.org/10.1080/00343404.2016.1172062

Hofmarcher, T. (2021). The effects of education on poverty: A European perspective. *Economics of Education Review,* 83, 1-16. https://doi.org/10.1016/j.econedurev.2021.102124

Huang, F., Jin, L. & Sun, X. (2009). Relationship between scale of higher education and economic growth in China. *Asian Social Science,* 5(11), 55-60. https://doi.org/10.5539/ass.v5n11p55

Ilgar, R. (2023). Türkiye'de eğitimde yaşanan bölgesel farklılıklar. *Uluslararası Ssosyal Bilimler Eğitimi Dergisi,* 9(2), 20-44. https://doi.org/10.47615/issej.1329419

Irughe, R. I. & Edafe, J. (2020). *Education and economic growth: Empirical evidence from Nigeria.* (AERC Research Paper No. 407).

Johnston, G. (2004). *Healthy, wealthy, and wise? A review of the wider benefits of education.* (New Zealand Treasury Working Paper No. 04/04). https://hdl.handle.net/10419/205544

Julius, M. K. & Bawane, J. (2011). Education and poverty, relationship and concerns. a case for Kenya. *Problems of Education in the 21st Century,* 32(1), 72-85. https://doi.org/10.33225/pec/11.32.72

Kaynak, S. & Rashid, Y. (2020). Sosyo-Ekonomik göstergelerine göre ECO üyesi ülkelerin hiyerarşik kümeleme metoduyla kümelenmesi. *Hitit Üniversitesi Sosyal Bilimler Enstitüsü Dergisi,* 13(1), 69-81. https://doi.org/10.17218.hititsosbil.713776

Lee, S. Y., Roys, N. A. & Seshadri, A. (2024, March). *The causal effects of parents' education on children's earnings.* (National Bureau of Economic Research Working Paper No. 32223). https://doi.org/10.3386/w32223

Li, B. (2006). A new approach to cluster analysis: the clustering-function-based method. *Journal of the Royal Statistical Society. Series B: Statistical Methodology,* 68(3), 457-476. https://doi.org/10.1111/j.1467-9868.2006.00549.x

Lin, T. C. (2004). The role of higher education in economic development: an empirical study of Taiwan case. *Journal of Asian Economies,* 15, 355-371. https://doi.org/10.1016/j.asieco.2004.02.006

Lochner, L. (2011). Education policy and crime. In P. J. Cook, J. Ludwig & J. MacCrary (Eds.), *Controlling crime: strategies and tradeoffs* (pp. 465-515). University of Chicago Press.

Longman dictionary of contemporary English. (2003). Pearson Longman.

Lv, K., Yu, A., Gong, S., Wu, M. & Xu, X. (2017). Impacts of educational factors on economic growth in regions of China: A spatial econometric approach. *Technological and Economic Development of Economy,* 23(6), 827-847. https://doi.org/10.3846/20294913.2015.1071296

Ma, J. & Pender, M. (2023). *Education pays 2023 the benefits of higher education for individuals and society.* Trends in Higher Education Series, College Board, New York.

MacMahon, W. W. (2002). *Education and development measuring social benefits*. Oxford University Press.

Meghir, C., Palme, M. & Schnabel, M. (2023, February). *The effect of educational policy on crime: an intergenerational perspective.* (National Bureau of Economics Research Working Paper No. 18145). https://doi.org/10.1920/WP.IFS.2011.1111

Mercan, M. & Sezer, S. (2014). The effect of education expenditure on economic growth: the case of Turkey. *Procedia-Social and Behavioral Sciences,* 109, 925-930. https://doi.org/10.1016/j.sbspro.2013.12.565

Morey, L. C., Blashfield, R. K. & Skinner, H. A. (1983). A Comparison of cluster analysis techniques within a sequential validation framework. *Multivariate Behavioral Research,* 18, 309-329. https://doi.org/10.1207/s15327906mbr1803_4

Morrison, D. G. (1967). Measurement problems in cluster analysis. *Management Science,* 13(12), 775-780. https://doi.org/10.1287/mnsc.13.12.B775

Murray, J. (2009). The wider social benefits of higher education: What do we know about them? *Australian Journal of Education,* 53(3), 230-244. https://doi.org/10.1177/000494410905300303

Njong, A. M. (2010). The effects of educational attainment on poverty reduction in Cameroon. *Journal of Educational Administration and Policy Studies,* 2(1), 1-8. https://doi.org/10.5897/IJEAPS10.058

Ochilov, A. (2012). Education and economic growth in Uzbekistan. *Perspectives of Innovations, Economics & Business,* 12(3), 21-33. http://dx.doi.org/10.15208/pieb.2012.14

Odit, M. P., Dookhan, K. & Fauzel, S. (2010). The impact of education on economic growth: the case of Mauritius. *International Business & Economics Research Journal,* 9(8), 141-152. https://doi.org/10.19030/IBER.V9I8.620

Öz, B., Taban, S. & Kar, M. (2009). Kümeleme analizi ile Türkiye ve AB ülkelerinin beşeri sermaye göstergeleri açısından karşılaştırılması. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi,* 10(1), 1-29.

Özaydın, Ö., Çemrek, F. & Balbağ, M. Z. (2022). Statistical analysis of the indicators for the education of provinces in Turkey. *Osmangazi Journal of Educational Research,* 9(2), 80-100.

Öztürk, B. & Gürsakal, S. (2015). Categorization of Turkish cities as per their education and health indicators. *Paradoks Economics, Sociology and Policy Journal,* 11(2), 83-101.

Püsküllüoğlu, A. (2004). *Arkadaş Türkçe sözlük.* Arkadaş Yayınevi.

Raghupathi, V. & Raghupathi, W. (2020). The influence of education on health: an empirical assessment of OECD countries for the period 1995-2015. *Achieves of Public Health,* 78(20), 1-18. https://doi.org/10.1186/s13690-020-00402-5

Rencher, A. C. & Christensen, W. F. (2012). *Methods of multivariate analysis.* John Wiley & Sons.

Reza, F. & Widodo, T. (2013). The impact of education on economic growth in Indonesia. *Journal of Indonesian Economy and Business,* 28(1), 23-44. https://doi.org/10.22146JIEB.6228

Rizvanovic, A. & Efendic, A. (2021). The effect of education on earnings in a post-conflict economy: evidence from Bosnia and Herzegovina. *Financial Internet Quarterly,* 17(2), 1-15. https://doi.org/10.2478/fiqf-2021-0008

Ross, C. E. & Wu, C. (1995). The links between education and health. *American Sociological Review,* 60(5), 719-745. https://doi.org/10.2307/2096319

Sarma, K. V. S. & Vardhan, R. V. (2019). *Multivariate statistics made simple: a practical approach*. CRP Press.

Shimeles, A. & Chuchane, A. V. (2016). The key role of education in reducing poverty in South Sudan. *African Development Review,* 28(2), 162-176. https://doi.org//10.1111/1467-8268.12199

Silles, M. A. (2009). The causal effect of education on health: evidence from the United Kingdom. *Economics of Education Review,* 28, 122-128. https://doi.org/10.1016/j.econedurev.2008.02.003

Tatlıdil, H. (2002). *Uygulamalı çok değişkenli istatistiksel analiz.* Akademi Matbaası.

Thapa, S. B. (2013). Relationship between education and poverty in Nepal. *Economic Journal of Development Issues,* 15(1-2), 148-161. https://doi.org/10.3126/ejdi.v15i1-2.11873

Tkaczynski, A. (2017). Segmentation in social marketing: Process, methods, and applications. In T. Dietrich, S. Rundle-Thiele & K. Kubacki (Eds.), *Segmentation using two-step cluster analysis* (pp.109-125). Springer.

Turanlı, M., Özden, Ü. H. & Türedi, S. (2006). Avrupa birliği'ne aday ve üye ülkelerin ekonomik benzerliklerinin kümeleme analiziyle incelenmesi. *İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi*, 9, 95-108.

Uzgören, N., Keçek, G. & Uzgören, E. (2013). Türkiye'de illerin beşeri sermayenin unsuru olan temel eğitim göstergeleri bakımından sınıflandırılması. *TİSK AKADEMİ,* 11, 119-133.

Yıldız, S. & Karakaş, D. G. (2019). Türkiye'de eğitim eşitsziliğinin farklı yüzleri. *Anadolu Üniversitesi Sosyal Bilimler Dergisi,* 19(2), 271-292. https://doi.org/10.18037/ausbd.566795

Zimmerman, E. & Woolf, S. H. (2014). *Understanding the relationship between education and health.* (Institute of Medicine, Discussion Paper). https://doi.org/10.31478/201406A