SANITAS MAGISTERIUM

Received: 01.06.2022

Published: 14.01.2024

https://dergipark.org.tr/ijhadec

January 2024 •

Classification of OECD Countries According to Health Data With Clustering Analysis

Mustafa Filiz[1], Olkan Budak [2]

Abstract

This study was carried out to determine how OECD countries are clustered according to the determined health data, which ones are similar, and which countries are better. 36 countries were included in the study and 10 variables, which are among the health indicators of the countries, were used. Centroid tree graph and k-means clustering analysis, one of the non-hierarchical clustering analysis methods, were used to analyze the data. With the ANOVA test, the differences in the variables according to the clusters were determined. It was observed that seven clusters were formed in the centroid method. As a result of the K-mean clustering analysis, it was seen that the distance from the selected countries was the USA the least and Turkey the most. It has been seen that among the variables selected in the clustering of OECD countries under seven clusters, variables such as life expectancy at birth, infant mortality rate, per capita health expenditure, Gini coefficient, crude death rate, the share of health in GDP, and the number of nurses/midwives play an important role. It was concluded that the countries in the 1st cluster had the best values in terms of health indicators of 36 countries, and the countries in the 5th cluster had the worst values. In addition, as a result of the ANOVA test, it was decided that other health indicators other than maternal mortality rate, number of patient beds, and number of physicians play an important role in clustering OECD countries under seven clusters.

Keywords: Health, Health indicators, K-mean, OECD, Tree graph.

Citation: : Mustafa Filiz, Olkan Budak (2024) Classification of OECD Countries According to Health Data With Clustering Analysis, International Journal of Health Administration and Education (Sanitas Magisterium), 10(1), 1-15.

[1] Dr. Öğretim Üyesi, Artvin Çoruh Üniversitesi İşletme Fakültesi Sağlık Yönetimi, mustafafiliz1109@gmail.com. ORCID: 0000-0002-7445-5361.

[2] Dr. Öğretim Üyesi, İstanbul Atlas Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, olkan.budak@atlas.edu.tr ORCID: 0000-0002-2276-2300.

Introduction

Various data are used when evaluating the development levels of countries. Among these data, health data is also important. Health societies formed by healthy individuals are among the development indicators of countries (Akyürek, 2012: 125). For this reason, the production and delivery of health services are of great importance not only at the individual level but also at the social level (Altay, 2007).

To create a healthy society, it is necessary to focus on the health status of individuals. It is of great importance for all people to be in equal health conditions to create a flourishing egalitarian society. A healthy life should be ensured for individuals of all ages and classes, health services should be improved and the needs of the population should be met (Costa et al., 2019: 2). Health services must be produced in sufficient quantity and quality and must be accessible. Increasing demand for health, increasing costs, technological developments, aging of the population, decreasing quality, medical errors, inequalities, and uncertainties have increased the interest in comparative studies in health services (Kelley & Hurst, 2006: 9). To increase the number of healthy individuals, it is necessary to understand why some communities are healthier than others (Costa et al., 2019: 2).

One of the most important strategies of international organizations such as the European Union, the World Health Organization, and the Organization for Economic Co-operation and Development (OECD) in the field of health is to determine the health indicators and to collect them accurately and reliably (Çelebi & Cura, 2013: 54). OECD is an organization that makes up for the deficiency in the collection and storage of health data of member countries and with its comparative statistical studies every year. OECD has become a reliable institution in the determination of health indicators with all these studies (Demir & Bakırcı, 2014: 116).

The OECD shares with the public the data it collects from many countries other than its member and partner countries, with the statistics it regularly publishes every year. The organization reports these statistics within the framework of the health indicators it has determined, both on its official website and in its book, Health at a Glance, which it publishes regularly every year. Within the framework of the health indicators determined in this book, both the health status of the society and the health service performances of countries that are members of the OECD, partners, and other strong economies are compared (OECD, 2015: 9).

This study, it was aimed to classify OECD member countries according to the determined health indicators and cluster analysis methods were used for classification. More than one variable is effective in determining the health status or health performance of countries. Handling and analyzing more than one variable one by one may not be possible and may lead to wrong evaluations. Because evaluating each variable by analyzing it alone means ignoring its relations with other variables. For this reason, it was found correct to use cluster analysis, one of the multivariate statistical methods, in the study. This study aims to evaluate the health system of OECD member countries with the determined variables and to classify the countries. The purpose of the research is to evaluate which OECD countries are similar to each other with cluster analysis and the differences between clusters according to the data obtained. It is anticipated that the study will add new information to the literature since it is included in the new OECD countries in 2020 and 2021.

Cluster Analysis

In every period of history, people have tended to divide the objects and assets around them into groups. In primitive times, plants were divided into various groups. Over time, the scope and purpose of this grouping process have changed. Bringing similar ones together is the most important detail in dividing them into groups (Everitt & Dunn, 2010). People have had difficulty classifying due to the increase in the details they use over time. To resolve this situation, they used various methods of classifying. One of these methods is cluster analysis. Cluster analysis was developed by Linnaeus in 1753 to classify plants and animals. The data used in cluster analysis are classified according to the similarity status. It provides summary data to help researchers interpret the issue (Blashfield & Aldenferder, 1978).

According to the similarity of the data used; The process of separating into groups is expressed as clustering analysis. These groups that are separated from each other are also called clusters (Dinçer, 2006). In the analysis, it is expected that the similarity between the clusters is low and the similarity between the cluster elements is high (Berkhin, 2002). Although cluster analysis has been known for a long time, its use has become widespread recently. Hierarchical and non-hierarchical clustering methods are used in cluster analysis. According to the subject of the studies to be done, one of these two methods is generally preferred (Girginer, 2013). The most used method in Cluster Analysis is the hierarchical cluster analysis method (Kalaycı, 2016: 358). In the hierarchical clustering analysis method, if there is no idea about the number of clusters, the units and variables are clustered according to different similarity criteria with each other (Koyuncugil & Özgülbaş, 2009). There are two hierarchical methods, grouping and divisive (Hubert, 1974). In the grouping hierarchical method, each unit or each observation is initially considered as a cluster, then the two closest clusters (or observations) are combined into a new cluster. With this process, the number of clusters is reduced by one in each step. In this method, graphs called tree graphs or icicles are used for easy understanding of stages and clusters. In the divisive hierarchical method, unlike the hierarchical grouping method, the analysis starts with a large set of all observations and the process continues until each observation forms a cluster alone (Everitt et al., 2001; Çelik, 2013: 181). Single linkage method, full linkage method, average linkage method, central method (Centroid), and ward's method are the most widely used grouper hierarchical clustering analysis methods (Uçar, 2014: 359).

In the non-hierarchical method, the entire distribution is evaluated and divided into certain clusters (Girginer, 2013). Graphs expressed as tree diagrams (dendrograms) were used to reveal the cluster analysis process more clearly. The clusters formed at the beginning of the process show the tree branches and the ones formed at the end show the trunk of the tree (Tatlıdil, 2002). The non-hierarchical clustering method is preferred when the researcher has prior knowledge about the number of clusters. After determining the number of clusters in which the units can be separated, it is decided which clusters the units will enter according to the cluster determination criteria and assignment procedures are performed (Çelik, 2013: 182). The most widely used method among non-hierarchical clustering methods is the k-means clustering method (Özdamar, 2010: 311). In the k-means clustering method, a typical observation is selected from each determined cluster and similar observations are clustered around this typical observation (Kalaycı, 2016: 360).

Cluster analysis has four basic stages. These; are the selection of the variables to be used and the creation of the data matrix, the selection of the distance measure to determine the distance or closeness of the observations, and the determination of the clustering method and type to be used in the analysis, and finally the interpretation of the data divided into clusters (Karagöz, 2019).

In the analysis, it is desired that the similarity between the clusters is low and the similarity between the cluster elements is high (Berkhin, 2002). To follow the process of cluster analysis more easily, graphs called dendrograms are used. The clusters formed at the beginning of the process represent the tree branches and the ones formed at the end represent the trunk of the tree. In practice, it is expected that the sum of the squares within the cluster or the group will be minimized (Sharma, 1996).

Literature Review

There are many studies in the literature on the health indicators of countries. In this section, studies in the literature will be given in chronological order.

Sığırlı et al., (2006) conducted a study on health level measures for 25 countries, and as a result of applying multidimensional scaling analysis, it was seen that countries formed three different groups in two-dimensional space according to the variables of interest. He found that it differs from other countries except for Turkey, Slovakia, Hungary, and the Czech Republic in terms of basic health indicators and especially in terms of health expenditures and the share allocated to health from national income.

Ersöz (2009) classified the 2004 health data of 30 OECD countries using clustering and separation methods. The ratio of health expenditures to gross domestic product (GDP), per capita health expenditure, life expectancy at birth, and infant mortality per 1000 births were used. According to the K-Means analysis result; It has been observed that 4 selected health indicators in the clustering of 30 countries are significantly effective.

Nirel et al., (2015) evaluated the relationship between nurses' supply and demand projections in Israel. According to the study, considering the aging of the Israeli population and the increasing demand, it is predicted that the nursing supply will increase. In the study, it is predicted that the demand for services will increase seriously and a significant nursing deficit will increase by 2030.

Alptekin and Yeşilaydın (2015), in their study, classified OECD countries according to certain health indicators using the fuzzy clustering technique. As a result, it was seen that the number of clusters was five. It has been observed that Turkey is in the same group as Chile, Mexico, Estonia, Poland, and Hungary.

Köksal et al., (2016) found significant differences in health indicators between Turkey and European Union countries. It has been found that Turkey can approach the average health level indicators of European Union countries with the development of maternal and child health services, increasing the share of national income for the health, and improving manpower planning in the field of health.

Method

1.The population of the Study: The population of the study consists of OECD countries. These countries are Germany, USA, Australia, Austria, Belgium, United Kingdom, Czech Republic, Denmark, Estonia, Finland, France, Netherlands, Ireland, Spain, Israel, Sweden, Switzerland, Italy, Iceland, Japan, Canada, Colombia, Korea, Costa Rica, Latvia, Lithuania, Luxembourg, Hungary, Mexico, Norway, Poland, Portugal, Slovakia, Slovenia, Chile, Turkey, New Zealand, and Greece. Since the data of Colombia and New Zealand countries could not be partially accessed in the research, they were excluded from the study sample. Therefore, while the universe of the research is 38 OECD countries, 36 OECD countries are the sample.

The reason for choosing OECD countries in the research is that the data of the said countries are collected according to the same criteria by the OECD and they enable comparative research (Afonso & Aubyn, 2005: 228).

2.Variables Used in the Research: The selection of health indicators used in the research is important to give more objective and accurate results in the evaluation of countries according to health indicators (Asanduluia et al., 2014: 265). In determining the variables to be used in the research, a literature review was conducted for previous studies in the field. In line with the data obtained from the literature, care has been taken to select the variables that will best show the health system of the countries. Accordingly, health indicators and definitions used in the clustering of OECD countries are given in Table 1.

Table 1. Health Indicators and Definitions Used in the Study

No	Health Indicators	Definitions
1	Number of physicians (per 1,000 people)	The number of physicians per 1,000 people in the total country population (DB, 2022).
2	Number of patient beds (per 1,000 people)	The number of total beds per 1,000 people remaining, excluding long-term care beds, in the total country population (OECD, 2021).
3	Share of GDP allocated to health (%)	It is the share of the monetary value of all final goods and services produced within the borders of a country in a given period (OECD, 2021).
4	Gini coefficient	It is a coefficient that measures the equality in the distribution of national income in a country. It ranges from 0 to 1. As the value approaches zero, inequality decreases, while it increases as it approaches one (OECD, 2021).
5	Life expectancy at birth (years)	It is the average number of years a newborn is expected to live if she/he does not die in an accident or similar special situation during her/his life (OECD, 2021).
6	Infant mortality rate (per 1,000 live births)	It is found per 1,000 live births, dividing the number of babies who die before one year of age in a year by the number of babies born alive in that year (OECD, 2021).
7	Number of nurses and midwives (per 1000 people)	The number of nurses/midwives per 1,000 people in the total country population (OECD, 2021).
8	Health expenditure per capita ($)	It is found by dividing the total health expenditure within the country's borders by the total country population (OECD, 2021).
9	Maternal mortality rate (100,000)	It is found by dividing the number of women who die during pregnancy, childbirth, and within 42 days of birth in a given period in a society by the number of live births in the same society in the same period (OECD, 2021).
10	Crude Death rate (per 1,000 people)	It is the number of deaths per 1000 people in a society (OECD, 2021).

Data Collection and Analysis

Data on health indicators of 36 countries were used in the study. Research data were obtained from statistics regularly published by the OECD (http://stats.oecd.org/). Only the data for the variable "number of physicians per 1000 people" were obtained from the World Bank's website (http://data.worldbank.org/indicator). In the statistics in the OECD database on the number of physicians, the total number of physicians is not found, instead; It is seen that there are data according to different types of physicians (practitioners, specialists, assistants, lecturers, physicians who do not practice their profession, cannot find a job or are retired) (Alptekin & Yeşilaydın, 2015: 145). Due to the use of such detailed data and the difficulty of access, the data on the number of physicians were obtained from the World Bank.

In the study, data for 2019 were used because many countries were not yet entered into the variables for the year 2020-2021. The deficiencies in the data for 2019 were used for the most recent year. Although this situation is undesirable, it is a method mostly used in studies on the country comparison (Moran & Jacops, 2013).

The classification of 36 countries according to their health indicators was carried out using the cluster analysis method. In the research, various analyzes such as grouping hierarchical methods such as single linkage, full linkage, average linkage, and centralized method were carried out. As a result of the data obtained, tree graphs (Dendrograms) were evaluated and it was decided to prefer the Centroid method. The mathematical usability and superiority of the Centroid method make it preferred in many disciplines (Cohen & Shannon, 1981).

According to Kalaycı (2016), measurement tools change as the variety of variables in the analysis increases. Variables need to be standardized to obtain accurate results. In other words, the variables must be set to the same state. In this study, to eliminate the scale difference, the variables were transformed into values, which are also expressed as "Z scores", which are generally preferred in standardization, before the clustering analysis is performed.

After the variables were standardized, the number of clusters was determined as a result of the Centroid analysis. After these stages, k-means clustering analysis, one of the non-hierarchical clustering analysis methods, was used. While non-hierarchical methods are used, hierarchical methods are primarily used in the study to guide the determination of the number of clusters.

After the number of clusters was determined, as a result of the k-mean cluster analysis, the clusters of the countries, the centers of each cluster according to the variables, and the distance between the clusters were created in the tables. With the ANOVA test, the differences in the variables according to the clusters were revealed. Analyzes were carried out using Microsoft Excel and SPSS 25 package programs.

Ethical Aspect of Research

Since the data obtained in the study are ready data published by the OECD and the World Bank, no permission was required. Therefore, no ethical committee, informed consent, or legal permission was obtained to conduct the study.

Results

Table 2: Descriptive Information on Variables

Health Indicators Used in Analysis	Min.	Max.	Mean	S.D.
Years of Life Expected at Birth	75,1	84,4	81,02	2,523
Infant Mortality Rate (Per 1,000 People)	1,100	13,100	3,782	2,284
Health Expenditure Per Capita ($)	1133	10.948	4.161,9	1,999
Maternal Mortality Rate (100,000)	1,100	37,600	9,400	9,052
Gini Coefficient	0,222	0,497	0,316	0,059
Crude Death rate (Per 1,000 People)	0,97	4,05	1,997	0,801
Share of GDP Allocated to Health (%)	4,344	16,767	8,826	2,340
Number of Patient Beds (Per 1.000 Person)	1,000	12,800	4,519	2,632
Number of Physicians (Per 1,000 People)	1,810	8,010	4,093	1,382
Number of Nurses and Midwives (Per 1,000 People)	2,400	18,000	9,008	4,162

According to the data obtained in Table 2, it was seen that the average life year at birth was 81.02, and the infant mortality rate was 3.782. On the other hand, while the minimum health expenditure per capita was $1,133, it was found to be $10,948 at the maximum. The maternal mortality rate was found to vary between 1,100 and 37,600. In addition, the Gini coefficient was found to have a mean of 0.316 and a standard deviation of 0.059.

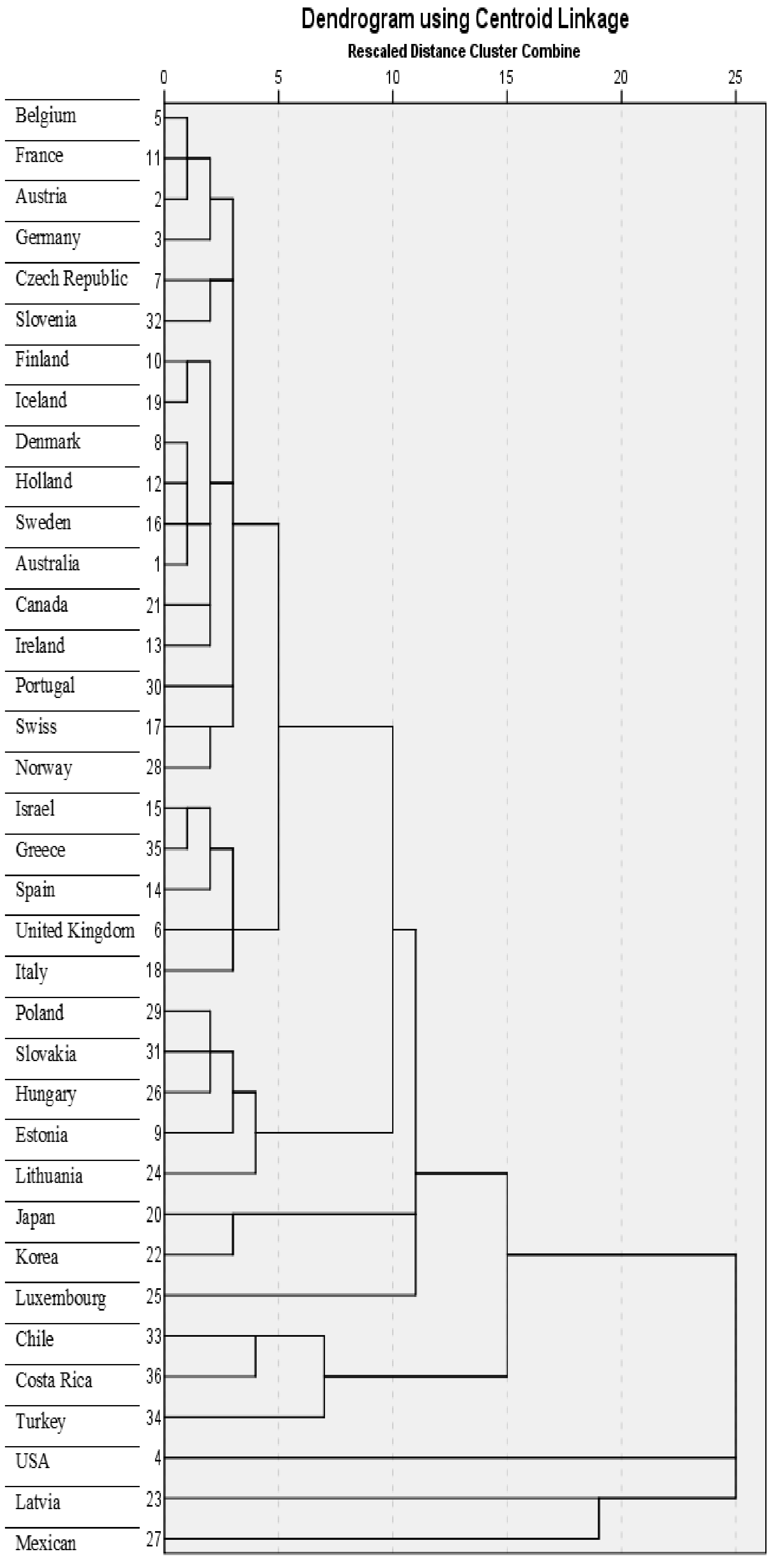
Dendrogram using Centroid Linkage
Rescaled Distance Cluster Combine
0
5
10
15
20
25
Belgium 5
France 11
Austria 2
Germany 3
Czech Republic 7
Slovenia 32
Finland 10
Iceland 19
Denmark 8
Holland 12
Sweden 16
Australia 1
Canada 21
Ireland 13
Portugal 30
Swiss 17
Norway 28
Israel 15
Greece 35
Spain 14
United Kingdom 6
Italy 18
Poland 29
Slovakia 31
Hungary 26
Estonia 9
Lithuania 24
Japan 20
Korea 22
Luxembourg 25
Chile 33
Costa Rica 36
Turkey 34
USA 4
Latvia 23
Mexican 27

Figure 1:Tree Chart Created Using the Centroid Method

In Figure 1, the tree graph obtained by the Centroid method is given. In the tree graph, scaled from left to right by five units, the countries that are most similar to each other come together at the closest distance. As the distance increases, new countries are included in the first cluster, depending on the similarity. When the distance reaches 25 units, a single cluster is formed and all countries are included in this cluster (Gan et al., 2007). The tree graph in Figure 1 was examined and it was seen that the most appropriate number of clusters was 7. The 7-cluster result obtained using the centroid method is given in Table 3.

Table 3. Clustering Results Obtained Using the Centroid Method Method

1. Cluster (17)		**2. Cluster (5)**	**3. Cluster (5)**	**4. Cluster (2)**	**5. Cluster (1)**	**6. Cluster (3)**	**7. Cluster (3)**
Belgium	Iceland	Israel	Poland	Japan	Luxembourg	Turkey	Latvia
France	Denmark	Greece	Slovakia	Korea		Costa Rica	USA
Austria	Holland	United Kingdom	Hungary			Chile	Mexican
Germany	Sweden	Spain	Estonia				
Czech Republic	Swiss	Italy	Lithuania				
Slovenia	Australia						
Finland	Canada						
Portugal	Ireland						
Norway							

In table 3, it was decided to divide the OECD countries into seven clusters according to the determined health data. According to this, it is seen that there are 17 countries in the 1st cluster, 5 countries each in the 2nd and 3rd clusters, 2 countries in the 4th cluster, 1 country in the 5th cluster, and 3 countries each in the 6th and 7th clusters. After the cluster numbers were determined, the k-means clustering analysis method, which is one of the non-hierarchical clustering analysis methods, was applied. The results of the analysis are given in table 4.

Table 4. Cluster Memberships and Distances Based on K-means Cluster Analysis

OECD Countries	**Cluster**	**Distance**	**OECD Countries**	**Cluster**	**Distance**
Australia	6	0,276	Canada	2	0,082
Austria	2	0,252	Korea	7	0,031
Germany	1	0,282	Latvia	3	0,248
USA	4	**0,000**	Lithuania	3	0,406
Belgium	2	0,071	Luxembourg	2	0,045
United Kingdom	6	0,142	Hungary	3	0,150
Czech Republic	7	0,041	Mexican	5	0,200
Denmark	2	0,026	Norway	1	0,055
Estonia	3	0,186	Poland	3	0,033
Finland	6	0,081	Portugal	7	0,029
France	2	0,178	Slovakia	3	0,132
Holland	2	0,286	Slovenia	7	0,072
Ireland	2	0,369	Chile	3	0,030
Spain	7	0,224	Turkey	5	**0,670**
Israel	7	0,472	Greece	3	0,088
Sweden	2	0,099	Costa Rica	5	0,266
Swiss	1	0,337			
Italy	7	0,277			
Iceland	6	0,101			
Japan	6	0,049			

In table 4, the cluster memberships and distances resulting from the K-means cluster analysis are given. Accordingly, the countries included in the analysis were found to be the USA at least (0.000) and Turkey (0.670) at the most. The clusters formed by the countries as a result of the clustering obtained by the K-mean clustering analysis method are given in table 5.

Table 5. Clustering Results Obtained by K-means Cluster Analysis Method

7 Cluster Result						
1. Cluster (3)	**2. Cluster (9)**	**3. Cluster (8)**	**4. Cluster (1)**	**5. Cluster (3)**	**6. Cluster (5)**	**7. Cluster (7)**
Germany	Austria	Estonia	USA	Mexican	Australia	Czech Republic
Swiss	Belgium	Letonya		Costa Rica	United Kingdom	Spain
Norway	Denmark	Latvia		Turkey	Finland	Israel
	France	Hungary			Iceland	Italy
	Holland	Poland			Japan	Korea
	Ireland	Slovakia				Portugal
	Sweden	Chile				Slovenia
	Canada	Greece				
	Luxembourg					

In table 5, the cluster memberships formed as a result of the K-means cluster analysis are given. According to this, while the 2nd Cluster has the most with 9 countries, the 4th Cluster has the least number of countries with 1 country. The averages of health indicators used in clustering by clusters are given in table 6.

Health Indicators	**Clusters**						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Years of Life Expected at Birth	82,8	82,5	78,15	78,9	78,0	82,2	82,3
Infant Mortality Rate	2,97	3,41	3,79	5,70	10,03	2,42	2,61
Health Expenditure Per Capita	6800	5453	2321	10948	1333	4642	3376
Maternal Mortality Rate	3,93	7,91	11,79	17,40	22,47	7,54	5,51
Gini Coefficient	0,287	0,285	0,320	0,395	0,437	0,309	0,307
Crude Death Rate	1,47	1,51	2,98	2,65	2,64	1,64	1,62
Share of GDP Allocated to Health	11,170	9,501	7,157	16,767	5,684	9,667	8,474
Number of Patient Beds	5,33	4,01	5,18	2,80	1,67	5,06	5,16
Number of Physicians	4,50	4,30	4,06	2,60	3,18	4,17	4,23
Number of Nurses and Midwives	16,60	10,98	5,25	12,00	2,90	12,4	7,29
: The best health indicators : Worst health indicators							

Table 6. Last Cluster Centers

According to table 6, it was seen that the cluster with the highest life years at birth was the 1st cluster (82.8), while the lowest cluster was the 5th cluster (78,0). It was observed that the cluster with the lowest infant mortality rate was the 1st cluster (2.97), while the highest cluster was the 5th cluster (10.03). The crude death rate was found to be at least in the 1st cluster (1.47) and at the highest in the 3rd cluster (2.98).

When table 6 is evaluated in general, it is seen that the 1st cluster has the best health data and the 5th cluster has the worst health data.

The values of the distances between the last cluster centers are given in table 7 below.

Table 7. Distances Between Last Cluster Centers

Clusters	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1							
2	1,347						
3	4,479	3,131					
4	4,147	5,495	8,627				
5	5,467	4,119	0,987	**9,614**			
6	2,157	**0,810**	2,321	6,305	3,309		
7	3,424	2,076	1,054	7,572	2,042	1,266	

When table 7 is examined, it is seen that the 2nd and 6th clusters are the closest (0.810) clusters to each other, while the 4th and 5th clusters are the farthest (9.614) clusters from each other regarding the distance values between the last cluster centers. In cluster analysis, an ANOVA test was applied to find out the difference in health data clusters. The results of the ANOVA test for the clusters formed as a result of the K-mean clustering analysis are given in table 8.

Table 8. K-means Cluster Analysis ANOVA Results

Health Indicators	**Cluster Mean Squares**	**df**	**MSE**	**df**	**F**	**p**
Years of Life Expected at Birth	24,204	6	2,678	29	9,037	**0,000**
Infant Mortality Rate	23,828	6	1,364	29	17,468	**0,000**
Health Expenditure Per Capita	23,088,482	6	49,088	29	470,343	**0,000**
Maternal Mortality Rate	142,401	6	69,438	29	2,051	0,091
Gini Coefficient	0,010	6	0,002	29	4,780	**0,002**
Crude Death Rate	2,333	6	0,291	29	8,022	**0,000**
Share of GDP Allocated to Health	23,325	6	1,782	29	13,088	**0,000**
Number of Patient Beds	6,572	6	7,004	29	0,938	0,483
Number of Physicians	0,965	6	2,106	29	0,458	0,833
Number of Nurses and Midwives	86,617	6	2,988	29	28,990	**0,000**

According to Table 8, when the ANOVA results are examined, among the variables selected in the clustering of OECD countries under seven clusters, life expectancy at birth ($p:0.000<0.05$), infant mortality rate ($0.000<0.05$), health expenditure per capita ($0.000<0$) .05), Gini coefficient ($0.002<0.05$), the crude death rate ($0.000<0.05$), the share of health from GDP (%) ($0.000<0.05$) and the number of nurses/midwives play an important role. seen playing.

When table 8 is examined, it is seen that among the variables selected in the clustering of OECD countries under seven clusters, such as maternal mortality rate ($p:0.091>0.05$), the number of patient beds ($0.483>0.05$), and the number of physicians ($p:0.833>0.05$), there are no variables. did not appear to play a role.

Conclusion

In this study, a cluster analysis was conducted by using 10 health data in total to compare the health indicators of 36 OECD countries. In the study, important data were obtained in the comparison of countries.

According to the data obtained, it is seen the average life expectancy at birth in 36 OECD countries is 81.02 years. Considering that the average worldwide human lifespan is 72.74 years (World Bank, 2022), it can be said that it is at a very good level. It has been determined that the infant mortality rate of 36 OECD countries is 3,782. Considering that the worldwide infant mortality rate is 49.4 (World Bank, 2022), this average is at a very good level. The maternal mortality rate of 36 OECD countries has been calculated to be 9,400. Considering that the average worldwide maternal mortality rate is 216 (World Bank, 2022), it can be argued that it is at a very good level.

In the study, it was seen that 7 clusters were formed by using the centroid method. It has been observed that there are 17 countries in the 1st cluster, 5 countries in the 2nd and 3rd clusters, 2 countries in the 4th cluster, 1 country in the 5th cluster, and 3 countries in the 6th and 7th clusters. It has been observed that the countries are at least the USA (0.000) and the maximum is Turkey (0.670). In other words, it has been determined that the USA has the best data among the selected health indicators and Turkey has the lowest data. In terms of clusters, the cluster with the best health indicators is the cluster with Germany, Switzerland, and Norway, while the cluster with Mexico, Costa Rica, and Turkey has the lowest health data. It has been seen that the cluster formed by the USA and the cluster formed by Mexico, Costa Rica, and Turkey are the clusters with the most distant and greatest differences in terms of health indicators. It is argued that the cluster consisting of countries such as Austria, Belgium, Denmark, France, Netherlands, Ireland, Sweden, Canada, and Luxembourg, and the cluster formed by Australia, United Kingdom, Finland, Iceland, and Japan are the countries with the least difference and the most similar to each other in terms of health indicators. drivable. In the study, it was seen that the infant mortality rate, per capita health expenditure, Gini coefficient, crude death rate, the share of health in GDP, and the number of nurses/midwives used in the study were effective in clustering countries.

Considering that the countries in the 1st cluster have the best values in terms of health indicators examined in the research, and the countries in the 5th cluster have the worst values, they are in the 5th cluster in terms of other health indicators other than maternal mortality rate, the number of patient beds and number of physicians. It can be concluded that developing countries are areas of development. It may be beneficial for the countries in this cluster to be in strategies and practices to improve their health indicators.

Conflict of Interest Disclosure: There are no potential conflicts of interest regarding the research, authorship, and publication of this article.

Support/Financing Information: No financial support has been received for the research, authorship, and publication of this article.

Ethics Committee Decision: Since the data obtained in the study are ready data published by the OECD and the World Bank, no permission was required. Therefore, no ethical committee, informed consent, or legal permission was obtained to conduct the study.

References

Afonso, A. & Aubyn, M.S. (2005). Non-Parametric approaches to education and health efficiency in OECD countries. *Journal of Applied Economics, 8(*2). 227-246.

Akyürek, Ç.E. (2012). Sağlıkta bir geri ödeme yöntemi olarak global bütçe ve Türkiye. *Sosyal Güvenlik Dergisi, 2*(2), 124-153.

Alptekin, N. & Yeşilaydın, G. (2015). OECD ülkelerinin sağlık göstergelerine göre bulanık kümeleme analizi ile sınıflandırılması. *Journal of Business Research-Türk, 7*(4), 137-155.

Altay, A. (2007). Sağlık hizmetlerinin sunumunda yeni açılımlar ve Türkiye açısından değerlendirilmesi. *Sayıştay Dergisi.* 64, 12-33.

Asandului, L., Roman, M. & Fatulescu, P. (2014). The efficiency of healthcare systems in Europe: a data envelopment analysis approach. *Procedia Economics and Finance*, 10, 261-268.

Berkhin, P. (2002). Survey of Clustering Data MiningTechniques, San Jose, California, USA, Accrue Software Inc, s,2.

Blashfıeld, R.K. & Aldenferder, M.S. (1978). The literature on cluster analysis. *Multivariate Behavioral Research,*13, 271-295.

Cohen, G.L., & Shannon, A.G. (1981). John ward's method for the calculation of pi. *Historia Mathematica,* 8, 133-144.

Costa, C., Freitas, Â., Stefanik, I., Krafft, T., Pilot, E., Morrison, J. & Santana, P. (2019). Evaluation of data availability on population health ındicators at the regional level across the european union. *Population Health Metrics, 17*(11), 1-15.

Çelebi, A.K., & Cura, S. (2013). Etkinlik Göstergeleri Açısından Sağlık Sistemleri: Karşılaştırmalı Bir Analiz. *Maliye Dergisi, 164*(6), 47-67.

Çelik, Ş. (2013). Kümeleme analizi ile sağlık göstergelerine göre Türkiye'deki illerin sınıflandırılması. *Doğuş Üniversitesi Dergisi. 14*(2), 175-194.

Demir, A. & Bakırcı, F. (2014). OECD üyesi ülkelerin ekonomik etkinliklerinin veri zarflama analiziyle ölçümü. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi. 28*(2), 109-132.

Dinçer, E. (2006). Veri madenciliğinde k-means algoritması ve tıp alanında uygulanması, yüksek lisans tezi, Kocaeli Üniversitesi.

World Bank (2022). Access address: [http://data.worldbank.org/indicator]. Access Date: 21.05.2022.

Ersöz, F. (2009). Türkiye ile OECD'ye üye ülkelerin seçilmiş sağlık göstergelerinin kümeleme ve ayırma analizi ile karşılaştırılması. *Türkiye Klinikleri Tıp Bilimleri Dergisi, 29*(6), 1650-1659.

Everıtt, B., Landau, S. & Leese, M. (2001). Cluster analysis. London: Oxford University Press.

Everitt, E. & Dunn, G. (2010). Applied multivariate data analysis. Wiley, New York.

Gan, G., Ma, C. & Wu, J. (2007) Data clustering theory, algorithms and applications (asa-sıam series on statistics and applied probability), Canada: SIAM society for ındustrial and applied mathematics publishing.

Girginer, N. (2013). A comparison of the healthcare ındicators of Turkey and the European Union members countries using multidimensional scaling analysis and cluster analysis. *İktisat, İşletme ve Finans*, 28, 323-362.

Http://Data.Worldbank.Org/İndicator (Access Date: 28.05.2022).

Http://Stats.Oecd.Org/ (Access Date: 28.05.2022)

Hubert, L. (1974). Approximate evaluation techniques for the single-link and complete- link hierarcihal clustering procedures. *Journal of the American Statistical Association*, 69, 698-704.

Kalaycı, Ş. (2016). SPSS uygulamalı çok değişkenli istatistik teknikleri. Asil Yayınevi. Ankara.

Karagöz, Y. (2019). Spss Amos Meta Uygulamalı İstatistiksel Analizler. Nobel Yayıncılık. Ankara.

Kelley, E. & Hurst, J. (2006). Health Care Qualıty Indıcators Project Conceptual Framework Paper. Oecd Health Workıng Papers No: 23. Access address: [https://www.oecd.org/els/healthsystems/36262363.pdf]. Access Date: 21.05.2022.

Koyuncugil, A., S. & Özgülbaş, N. (2009). Veri madenciliği: tıp ve sağlık hizmetlerinde kullanımı ve uygulamaları. *Bilişim Teknolojileri Dergisi, 2*(2), 21-32.

Köksal, S.S., Sipahioğlu, N.T., Yurtsever, E., & Vehid, S. (2016). Temel sağlık düzeyi göstergeleri açısından Türkiye ve Avrupa Birliği ülkeleri. *Turkish Journal of Family Medicine and Primary Care, 10*(4),205-212.

Moran, V. & Jacobs, R. (2013). An ınternational comparison of efficiency of ınpatient mental health care systems. *Health Policy,* 112, 88– 99.

Nırel, N., Grınstıen Cohen, O., Eyal, Y., Samuel, H. & Ben Shoham, A. (2015). ‘‘Models for projecting supply and demand for nurses in Israel. *Israel Journal of Health Policy Research, 4*(46), 1-12.

OECD (2015). Health at a Glance 2015: OECD Indicators. OECD Publishing. Paris.

OECD (2021). Health at a Glance 2021: OECD Indicators. OECD Publishing. Paris. https://www.oecd.org/health/health-at-a-glance/. Access Date: 18.04.2022.

Özdamar, K. (2010). Paket programlar ile istatistiksel veri analizi- 2 (çok değişkenli analizler). Kaan Kitabevi. Eskişehir.

Sharma, S. (1996). Applied MultivariateTechniques, John Wiley and Sons.

Sığırlı, D., Ediz, B., Cangür, Ş., Ercan, İ. & Kan, İ. (2006). Türkiye ve Avrupa Birliği'ne üye ülkelerin sağlık düzeyi ölçülerinin çok boyutlu ölçekleme analizi ile incelenmesi. *İnönü Üniversitesi Tıp Fakültesi Dergisi, 13(*2), 81-85.

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

Uçar, N. (2014). Kümeleme analizi. içinde: SPSS uygulamalı çok değişkenli istatistik teknikleri Ed: Kalaycı, Ş. Asil Yayıncılık. Ankara.