

**Maria Parlińska<sup>1</sup>, Piotr Pomichowski<sup>2</sup>**  
Warsaw University of Life Sciences - SGGW

## Evaluation of Human Capital in Selected EU Countries Using Cluster Analysis

**Abstract.** The study evaluates the quality of human capital on the basis of the Eurostat data. There was also prepared a short bibliography review of the previous research which indicates various definitions of the term human capital. To evaluate its quality in selected EU countries. There was used cluster analysis – the k-medoid method. The analysis was performed for three periods: 2008, 2011 and 2014. The selected states were divided into 4 clusters, and Norway constitute a single-element group. In the other cases, it may be stated that the quality of capital may be determined through the joint economic past or historical traditions.

**Keywords:** human capital, cluster analysis, k-medoid method

### Introduction

One of the most important tasks of the European Economic Community is to continuously improve the quality of human capital, in compliance with the requirements of competitive knowledge-based economy. The respective EU countries conduct activities in that regard in compliance with their individual development strategies. The improvement in the quality of human capital, understood as the resources in the form of knowledge, skills, health, strength, as well as vital energy, guarantees long-lasting economic growth which is conducive to the improvement of the well-being of the society (Becker, 1975). Many discussions are conducted regarding human resources and their role in the development of the economic development of different countries. As a result, there exist multiple methods for comparing them among different states, and to group them by degree of probability.

This article assesses the quality of human capital in selected EU countries, with the use of cluster analysis. Cluster analysis has the main aim of classifying and dividing the collection of elements on the basis of the similarities among the selected properties, into the specified number of groups (Parlińska, Wasilewska, 2016). The result of the analysis will be separation of the groups that are most homogenous in terms of the selected characteristics of the countries according the quality of human capital.

### Description of data and applied grouping method

The subject literature contains numerous definitions of human capital and its numerous interpretations, resulting from the application of different research assumptions by the

---

<sup>1</sup> Doctor habilitated, Professor of the WULS-SGGW, Department of Agricultural Economics and International Economic Relations, Faculty of Economic Sciences WULS-SGGW, ul. Nowoursynowska 166, 02-787 Warszawa, e-mail: maria\_parlinska@sggw.pl

<sup>2</sup> PhD student at Faculty of Economic Sciences, e-mail: piotr\_pomichowski@sggw.pl

persons handling that subject. The definition of that term often depends on the level of research conducted on a macro scale or at the level of an organization.

An example of such deliberations is the study by Stanisław R. Domański entitled “Kapitał ludzki i wzrost gospodarczy” (“Human capital and economic growth”). He characterized that term as “the resources in the form of knowledge and skills, health, vital energy, included in the society or nation as a whole”. He also believed that human capital guarantees the ability to work, satisfaction and sufficient salaries.

In turn, Nobel Prize winners, such as G.S. Becker (1975) and T.W. Schultz (1981) mentioned the significance of that term as an investment into people for the benefit of development of a country. Schultz emphasized that the well-being of a society depends on the quality of human resources, such as education, experience and health, and the factor that is decisive for economic growth, is an acting person, through such investments as gaining education and experience, developing its skills and looking after its health. Additionally, G. Becker indicated the need to educate the unemployed, thus introducing, in economy, the notion of investing in people.

After a brief analysis of the above studies, there may be differentiated 4 groups of indices that are decisive for the quality of human capital which will be used in the cluster analysis below. These are: labour market, education, health and investments in people.

The source of data for the analysis was the Statistical Office of the European Union – Eurostat. The following 17 states were selected on the basis of the criteria of availability of data of European countries: Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Ireland, Spain, France, Latvia, Lithuania, Hungary, Poland, Portugal, Romania, Slovenia and Norway. For comparison purposes, the analysis was performed for the years 2008, 2011 and 2014. The year 2015 was omitted due to the large amount of missing data in the form of the variables being the indices of quality of human capital for the respective states. In the selection of variables representing the above-mentioned groups of indices, there was applied the amount of information provided by the given variable, as well as the access to its levels in the respective countries in the researched years.

Table 1 contains all the variables used for characterizing the quality of human capital in the selected states. The following variables characterizing the labour market were selected: GDP adjusted for purchasing power per one citizen and unemployment rate calculated as the proportion of the unemployed persons in comparison with the professionally active population. Those measures reflect quite well the labour market and the quality of the human capital in that market, and they are easily accessible. Education was characterized with the use of the variables depicting the level of education in the society, such as the percentage of people leaving the education system at the age of 18-24, the percentage of people aged 15-64 with primary or lower secondary education, and the percentage of people aged 15-64 with higher secondary education but without higher education. Those characteristics are to depict the resources of knowledge represented by the given society. Another factor determining the quality of human capital is healthcare. It was characterized on the basis of the life expectancy of women and men, as well as the number of physicians per 100,000 citizens. These characteristics may have different impact on the quality of human capital, because it might result from higher awareness of healthy lifestyle or from the higher availability of qualified labour in hospitals. The last group of indices of quality of human capital, called investment in people, includes the amounts of money spent by the state on development per one citizen, the number of inventions reported in the patent

office per one million citizens and the percentage of people aged 15-24 and 65-74 employed in the areas of science and technology.

All of the above variables help to characterize the human capital of the given society and are available for most states in the respective years. In the case of lack of the value of any variable for the given country, it is filled in with the value that is the arithmetic mean of the two adjacent values (one year before and later) for the given country. In some situations, this required filling in with the use of the value from several years (like in Latvia, where the variable “number of physicians” for 2008 was filled in with the value from 2011).

Table 1. Selected diagnostic variables for analysing the human capital in selected EU countries.

Variable symbol	Diagnostic variables
X <sub>1</sub>	GDP <i>per capita</i> by purchasing power parity
X <sub>2</sub>	Percentage of unemployed person in the professionally active population
X <sub>3</sub>	Percentage of the population aged 18-24 who left the education system early
X <sub>4</sub>	Percentage of the people aged 15-64 with primary or lower secondary education
X <sub>5</sub>	Percentage of the people aged 15-64 with lower or higher secondary education, but without higher education
X <sub>6</sub>	Life expectancy for women
X <sub>7</sub>	Life expectancy for men
X <sub>8</sub>	Number of physicians per 100,000 citizens
X <sub>9</sub>	Expenditure on research and development <i>per capita</i>
X <sub>10</sub>	Number of patents submitted in the EPO per one million of citizens
X <sub>11</sub>	Percentage of the population aged 15-24 and 64-74 employed in areas of science and technology

Source: own study.

The research into human capital may be conducted with the use of various methods, but one can also come across the studies where that subject was presented with the method of grouping.

A good example is the joint study by Małgorzata Stec, Agata Janas and Artur Kuliński entitled “Grupowania państw Unii Europejskiej ze względu na zasoby kapitału ludzkiego i intelektualnego” (“Grouping of European Union states based on their human and intellectual capital resources”), in which the condition of human capital was presented using Ward’s method (Stec, Janas, Kuliński, 2005).

Another study worth mentioning is one by Gabriela Wronowska, entitled: “Kapitał ludzki w krajach Unii Europejskiej – analiza porównawcza” (“Human capital in the states of the European Union – a comparative analysis”), in which the author also grouped the states with the use of the same method as used by the above-mentioned authors. Additionally, she classified human capital in the respective EU countries on the basis of the control indices, using the data for 2006. In her work she demonstrated that Denmark was the leader, while Poland had the lowest level of development of human capital at that time (Wronowska, 2009).

The study presented in our article for verifying the quality of human capital in the respective states uses cluster analysis. Cluster analysis is a method of classifying, aimed at

dividing the collection of elements on the basis of the similarities among the selected properties, into the specified number of groups. The aim of that method is to separate the groups that are most homogenous in terms of the selected characteristics of the elements included in them. The popular algorithms for the above task include the k-means algorithm and k-medoid algorithm. The k-means method is used most frequently due to its simplicity and low computational complexity, but it is encumbered with numerous flaws. These include the need to declare the number of clusters and low resistance to noisy data. One of the alternative methods not resistant to diverging observations, applied in cluster analysis, is the k-medoid method. The difference between the operation of those algorithms consists that in the case of k-means, the centre of a cluster is the “centre of gravity of the group”, the so-called centroid, while in the other method, it is the so-called medoid, i.e. a representative element of the group that lies closest to its centre. Therefore, in the former method the central point of the group is determined with the use of an algebraic method, and it does not constitute an element, unlike in the latter method, where medoid is an element belonging to the group. That is why the k-medoid method was applied in this study.

The general steps in the grouping algorithm may be defined as:

1. Draw the initial k number of medoids as the example central points of clusters
2. Assign an element to the group in which the distance between the element and the centre of the group is smallest
3. Select new central points of clusters among the elements not being medoids
4. Go back to point 2 and perform steps 2-4 until a new grouping of the elements no longer improves the value of the criterion function.

As regards the listed methods, the notion of criterial function which constitutes the criterion of stopping the grouping algorithms, should be explained. That function determines whether the next grouping of elements is better than the previous grouping. An example of such a function may be the average distance of all the elements in the group from the centre of the group, or the sum of squares of deviations<sup>3</sup>. Attention should also be paid to selection of the optimum number of clusters. One of the criteria found in literature is the Caliński-Harabasz index. A detailed mathematical description of that index is presented in the article issued by Caliński and Harabasz in 1974. For an optimum number of clusters, that index assumes the maximum value.

For the purposes of this article, the calculations make use of the R software as well as the packages used for cluster analysis: “cluster” and “ClusterSim” (Nowak-Brzezińska, 2012). The next chapter presents the results of the research.

## **Presentation and analysis of results**

In the cluster analysis method, the most important aim is to determine the optimum number of clusters. In order to do that, there was applied the Caliński-Harabasz index which assumes the maximum value for the optimum number of clusters for the given group and its characteristics. In order to compare the migration of states between the clusters more easily, the decision was made to assume the same number of elements for all the periods studied. The figure demonstrates that the Caliński-Harabasz index achieves the local maximum for the number of clusters of 4, except for the data for 2008.

---

<sup>3</sup> Function used in Ward’s method.

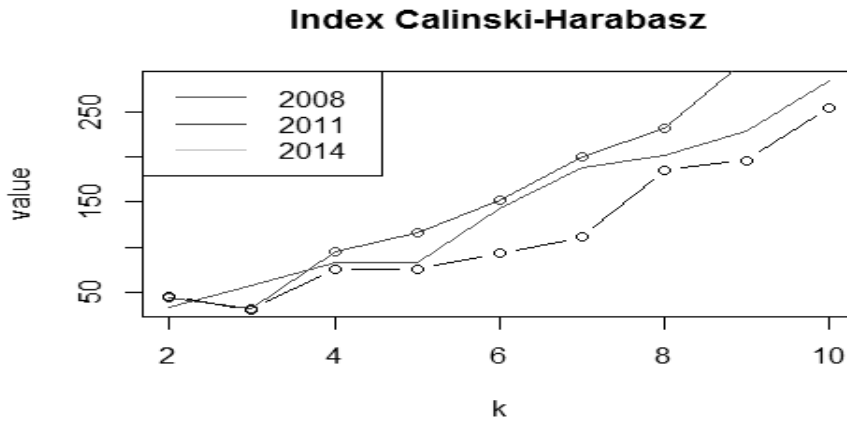


Fig. 1. The Caliński-Harabasz index

Source: own study on the basis of the *Eurostat* statistical data in the R software.

Despite that, it was assumed that the optimum number of clusters was 4, and the states were divided based on the variables reflecting the quality of the human capital of their residents. Below please find the results obtained using the k-medoid method.

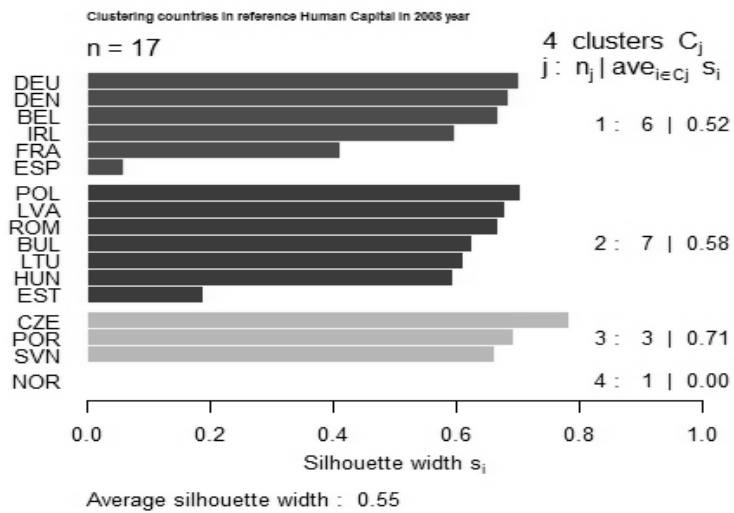


Fig. 2. Cluster analysis for 2008

Source: own study on the basis of the *Eurostat* statistical data in the R software.

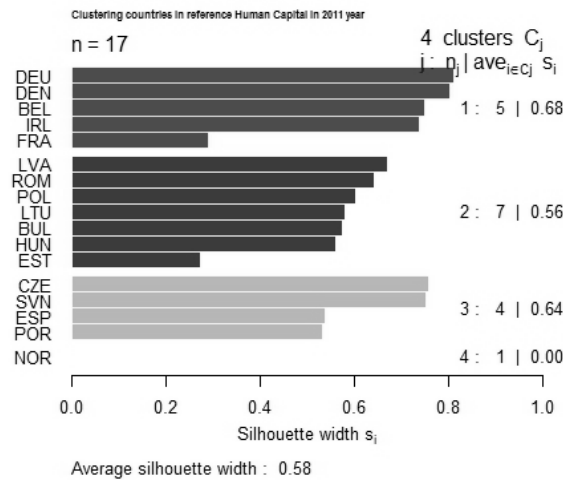


Fig. 3. Cluster analysis for 2011

Source: own study on the basis of the *Eurostat* statistical data in the R software.

It follows from the figure analysis that the first cluster comprises the states of the “old” EU, such as: Germany, Denmark, Belgium, France, Ireland and Spain only for the year 2008. In 2011, Spain found itself in the third group, together with such states as: Czech Republic, Slovakia, Portugal. The second, largest cluster comprised the Central and Eastern European countries (except for the Czech Republic): Poland, Latvia, Romania, Bulgaria, Lithuania, Hungary and Estonia. In both cases Norway constituted a separate cluster.

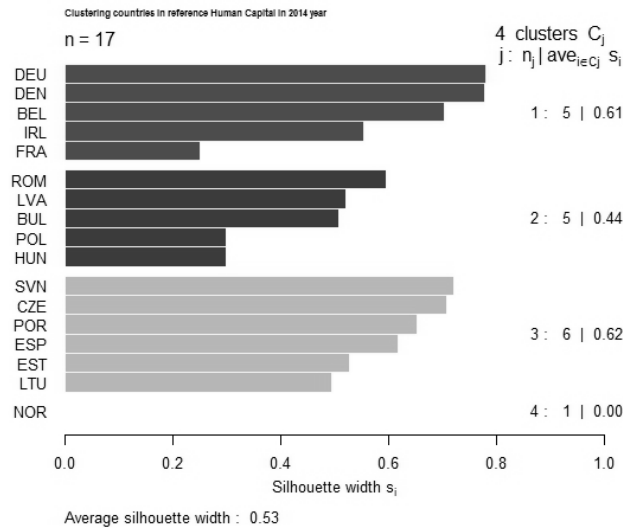


Fig. 4. Cluster analysis for 2014

Source: own study on the basis of the *Eurostat* statistical data in the R software.

For 2014, the first and last cluster did not change in comparison with 2011. The changes resulting from the division based on quality of capital were recorded in the second and third groups. As a result of grouping, Estonia and Lithuania moved to the third group, with states such as Spain, Slovenia, Portugal and Czech Republic.

## Conclusions

After conducting the above analyses for the three periods of 2008, 2011 and 2014, it may be stated that the states being the first members of the EU demonstrate high similarity of capital, and the lowest fluctuations in recent years. It should be noted that these countries present high quality of capital due to the presence of Germany in that group, which country is exceptional in terms of economic growth and development in comparison with other states. The core of the second concentration between 2008 and 2011 comprised the states with common past, as well as historical and economic tradition. These are the countries of the former Soviet Union, the legacy of which is visible in certain areas of life until this day. The Czech Republic is the only country from Central and Eastern Europe that did not belong to that group. In 2014, Estonia and Lithuania migrated from the second to the third cluster. It may result from the change in quality of capital, which may, in turn, stem from the beneficial development of those states, as a result of which they adopted the joint currency – EURO. In all the periods, Norway constitute a separate group in terms of quality of capital. That country is not a member of the EU, but has agreements with the community. Norway as a state is characterized by highly developed economy which may result from the high quality of the possessed human capital, and it is not subject to the EU policies.

The above analysis demonstrates that the joint past and geographical location may determine the quality of capital in the selected countries.

## Literature

- Becker, G.S. (1975). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 2d ed., New York: Columbia University Press for NBER.
- Caliński, T., Harabasz, J. (1974). A dendrite method for cluster analysis, *Communications in Statistics*, vol. 3, 1-27.
- Nowak-Brzezińska, A. (2012). Konspekt do zajęć: Statystyczne metody analizy danych. [Available at:] [http://zsi.tech.us.edu.pl/~nowak/odzw/smad\\_lab11.pdf](http://zsi.tech.us.edu.pl/~nowak/odzw/smad_lab11.pdf). [Access: September 2016].
- Parlińska, M., Wasilewska, E. (2016). Analiza skupień jako przykład zastosowania w ocenie rozwoju regionów. *Zeszyty Naukowe WSE-S w Ostrołęce*.
- Schulz, T.W. (1981). *Investing in People. The Economics of Population Quality*.
- Stec, M., Janas, A., Kuliński, A. (2005). Grupowanie państw Unii Europejskiej ze względu na zasoby kapitału ludzkiego i intelektualnego, [in:] *Nierówności społeczne a wzrost gospodarczy*, Uniwersytet Rzeszowski, no 6, 135-146.
- Wronowska, G. (2009). Kapitał ludzki w krajach Unii Europejskiej – analiza porównawcza, [in:] Kopycińska D. (red.) *Kapitał ludzki jako czynnik przewagi konkurencyjnej*, Katedra Mikroekonomii Uniwersytetu Szczecińskiego, Szczecin.