A CLUSTER ANALYSIS OF THE ROMANIAN COUNTIES IN TERMS OF SCHOOL POPULATION, DIDACTIC STAFF AND NUMBER OF CLASSES

Cristina OPREA[1]
Delia Mioara POPESCU[2]
Ramona-Elena GABREA[3]
Alexandru STOICA[4]

ABSTRACT
It is well known that after 1990 the number of school population has decreased. Also, because of multiple curriculum modifications, among other reasons, the education quality, especially pre-university one, has recorded a decrease that is not yet stable. Of all complex aspects that concern education, this paper analyzes the transformations produced in Romania's counties from the school population number, teaching staff number and number of classrooms point of view, between 2000 - 2015. The results obtained by cluster analyses for 2000 and 2015 show a tendency of convergence and uniformisation for more than a half of Romania's counties, although emphasizing differences in counties such as Cluj, Iași, Constanța, Brașov or Bucharest.

KEYWORDS: *cluster analysis, education, hierarchical cluster, school population.*

JEL CLASSIFICATION: *A20, C10, H52*

1. INTRODUCTION

A significant number of research and statistiques have demonstrated that differences between regions, under different aspects: geographical, etnic, cultural, socio-economic development etc, have a great impact of education. Reorientation of educational sistems towards promoting etiquette is a necessary remedy for scholar success (Faubert, 2012; Field, Kuczera & Pont, 2007; Heckman, 2011).

Jigău and Farturnic (2012) group schools by the socio-economical development level of the region where they are. These emphasise the fact that, in disadvantaged communities from the socio-economic point of view, the drop out of school and early abandoning of school has an increased intensity.

Bădescu and Petre (2013) make a classification of scholar population in urban environment and rural environment based on the risk of dropout of school. The study shows that the risk of drop out is bigger in the rural areas.

In one study, researchers examined using the clustering method whether the performance of students at a university was influenced by the demographic context. (Yorke et al., 2005)

Thus the current paper aims to analyse the transformations produced in Romania's counties from the scholar population number, of professors' number and of class number between 2000 - 2015.

[1] Valahia University of Târgoviște, Târgoviște, România, oprea_cris2005@yahoo.com
[2] Valahia University of Târgoviște, Târgoviște, România, depopescu@yahoo.com
[3] Valahia University of Târgoviște, Târgoviște, România, ramona_gabrea@yahoo.com
[4] Valahia University of Târgoviște, Târgoviște, România, stoica.dumitru.alexandru@gmail.com

2. METHODOLOGY

To analyze transformations recorded in Romania's counties regarding School population, Teaching staff and Infrastructure implicated in the period 2000 - 2015, there were chosen five indicators: Total school population (TSP), Total teaching staff (TTS) and Total classrooms (TCR) School population per teaching staff (SPTS) and School population per classroom (SPCR). Using the data series there were formed doua matrix (one matrix for year 2000 and one for year 2017) as such:

$$X=(x_{ij})_{i=\overline{1,n},j=\overline{1,n}} \quad (1)$$

where n is the number of counties in Romania, and m is the number of indicators (variables) considered. For generating Romania's counties considering the values of the five indicators, it was used the hierarchical cluster methodology (Rotaru et al., 2006).
Starting from matrix X and using Euclidean squared distance it was generated the Proximity matrix and it was applied the Z-sore transformation, having the following form:

$$W=(w_{ij})_{i=\overline{1,n},j=\overline{1,n}} \quad (2)$$

For this there were used the transformations [Zaharia & Gogonea, 2016]:

$$y_{ij}=\frac{x_{ij}-x_j}{\sigma_j}, \text{ unde } x_j=\frac{\sum_{i=1}^{n}x_{ij}}{n}, \text{ iar } \sigma_j=\sqrt{\frac{\sum_{i=1}^{n}(x_{ij}-x_j)^2}{n-1}} \quad (3)$$

$$w_{ij}=\sum_{i=1}^{n}(y_{ik}-y_{ij})^2, j=\overline{1,n}, k=\overline{1,n}, j\neq i, k\neq i \quad (4)$$

To determine the distance between clusters it was used the Average linkage between groups method.
The data processing and analise were done using SPSS (Howitt & Cramer, 2005, Popa, 2008) and Excel (Oprea & Zaharia, 2011). Conffidence level 95% was used (α=0.05).

3. RESULTS AND DISCUSSIONS

Following the tests carried out for the year 2000, a structure was chosen in seven clusters (Table 1), one of these having only one element (Bucharest Municipality). This particularity is the consequence of the particular situation of the Bucharest Municipality in comparison with the other counties.
Thus, in 2000, in Bucharest there were on average 488,832 pupils, 28,515 teachers, 8,566 classrooms, 17 pupils and 57 pupils per teaching staff, 8.16 times higher more than in C 2 cluster (Tulcea, Ilfov, Braila, Calarasi, Giurgiu, Ialomita), and 1.8 times more than in the C5 (Dolj, Timis, Arges, Cluj, Iasi).
Of the other 6 clusters, the highest values on the average number of the school population and the average number of teachers are recorded in the C5 cluster (169,276.6 pupils and 1,168.2 teachers respectively).
Regarding the average number of pupils in a classroom, we can say that the worst situation, after Bucharest, is in the counties of the C3 cluster, where there are 41 pupils average per classroom.

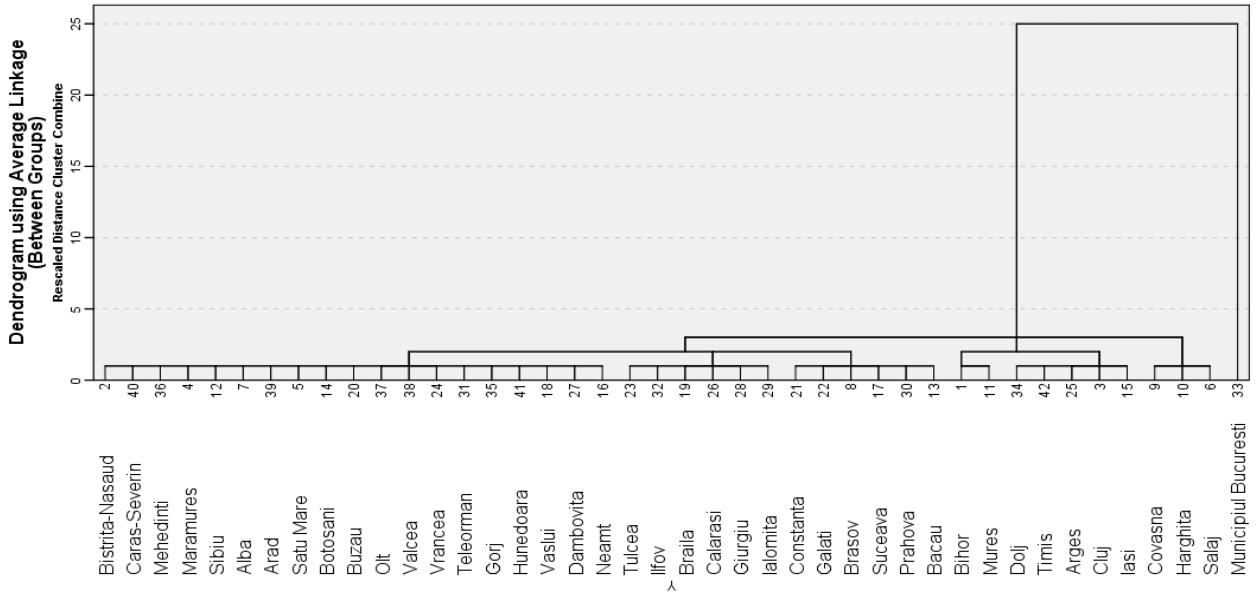

Figure 1 Dendogram of counties clusters in the year 2000

Regarding the number of pupils in the teaching staff, clusters C2 and C3 ranked second, after Bucharest, with an average of 16 pupils / teacher, exceeded by 3 pupils the average number of pupils to cluster 6.

Table 1. The structure of the clusters in the year 2000

Cluster	Counties
C1	Bistrița-Năsăud, Caras-Severin, Mehedinți, Maramureș, Sibiu, Alba Arad, Satu Mare, Botoșani, Buzău, Olt, Vâlcea, Vrancea, Teleorman, Gorj, Hunedoara, Vaslui, Dâmbovița, Neamț
C2	Tulcea, Ilfov, Brăila, Călărași, Giurgiu, Ialomița
C3	Constanta, Galați, Brașov, Suceava, Prahova, Bacău
C4	Bihor, Mureș
C5	Dolj, Timiș, Argeș, Cluj, Iași
C6	Covasna, Harghita, Sălaj
C7	Municipiul București

Source: Own construction using SPSS

The verification of the statistical significance of the membership of the five cluster variables was performed using the ANOVA methodology (Table 2). Given that for all the variables F_{stat} is significantly higher than the critical value ($F_{\alpha, m-1, n-m-1} = F_{0,05,6,35}$), it results that the variables TSP, TTS, TCR, SPTS and SPCR are significant in in terms of cluster membership.

Tabelul 2. Rezultatele testării semnificației statistice a apartenenței variabilelor TSP, TTS, TCR, SPTS and SPCR la clusteri, pentru anul 2000

	df1	df2	Fstat	$F_{0.05;6;35}$	Sig.F
TSP	6	35	136.112	2.38	0.000
TTS	6	35	105.270	2.38	0.000
TCR	6	35	63.240	2.38	0.000
SPTS	6	35	18.060	2.38	0.000
SPCR	6	35	20.850	2.38	

Source: own construction using SPSS

Taking into account that the conditions of statistical significance for all seven clusters are met, the mean values of the analyzed variables corresponding to clusters (cluster centers) are presented in Table 3.

Tabelul 3. The values of means for TSP, TTS, TCR, SPTS and SPCR în the year 2000

Cluster	Number of counties	TSP scholars	TTS teachers	TCR classrooms	SPTS scholars/teacher	SPCR scholars/classroom
C1	19	85938.3	5609.3	2628.6	15.4	32.7
C2	6	52708.3	3148.5	1420.0	16.8	37.2
C3	6	145680.5	8771.8	3583.0	16.7	41.3
C4	2	123370.0	8872.0	4062.0	13.5	30.5
C5	5	169276.6	11682.2	4502.6	14.8	37.6
C6	3	53387.7	4056.3	1998.7	13.3	26.3
C7	1	488832.0	28515.0	8566.0	17.0	57.0

Source: own construction using SPSS

In the year 2015 the dendrogram (Figure 2) highlights a significant trend of convergence of the values of the analyzed indicators in more than half of the counties of Romania. Unlike the year 2000, where the most numerous cluster included 19 counties, in 2015 the largest cluster includes 24 counties (Table 5). Thus, in the period of 2000 - 2015, cluster C1 was joined by seven counties: Bacău from cluster C3, Bihor and Mureş from C4, Argeş from C5 and Covasna, Harghita and Salaj from cluster C6. During the same period, Vrancea and Teleorman counties passed from cluster C1 to cluster C2.

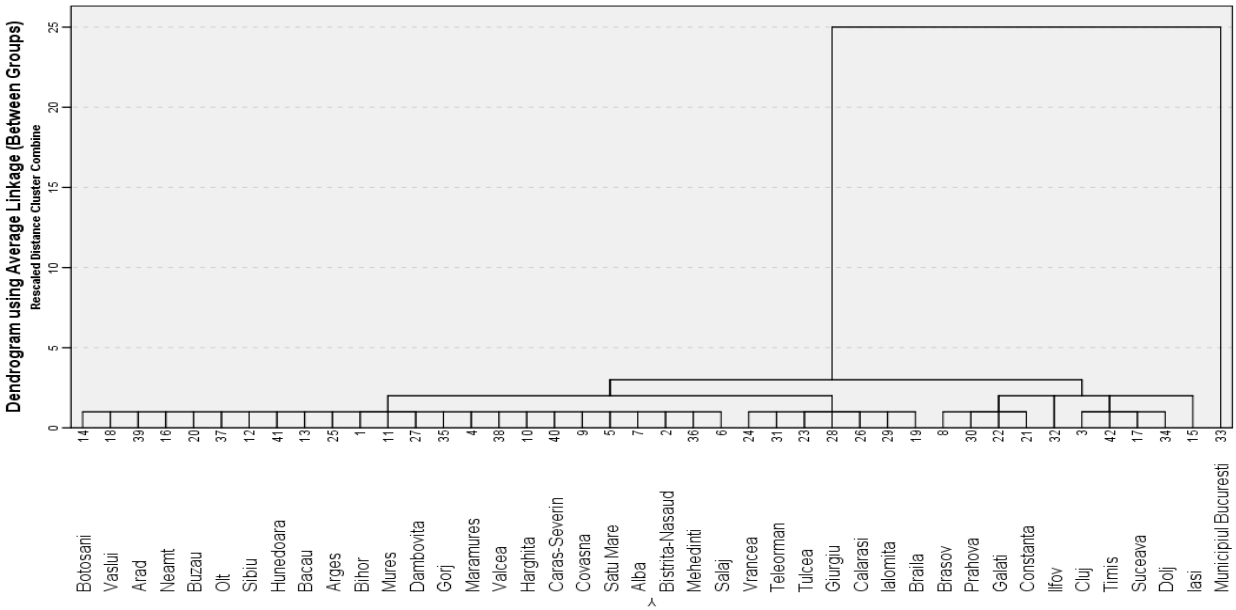

Figure 2. Dendogram of counties clusters in 2015 year

The verification results of the statistical significance of the membership of the five cluster variables, corresponding to 2015 using the ANOVA methodology are presented in Table 4. Given that in this case, for all variables, F_{stat} is significantly higher than the critical value, it results that, at the level of 2015, the variables analyzed are significant in terms of cluster membership.

Table 4. The results of testing the statistical significance of the membership of TSP, TTS, TCR, SPTS and SPCR variables on clusters for 2015

	df1	df2	Fstat	$F_{0.05;6;35}$	Sig.F
TSP	6	35	77.351	2.38	0.000
TTS	6	35	57.765	2.38	0.000
TCR	6	35	40.261	2.38	0.000
SPTS	6	35	10.342	2.38	0.000
SPCR	6	35	24.839	2.38	

Source: own construction using SPSS

Also, three other counties have changed their position in clusters: Ilfov passed from C2 to C4, Iași switched from C3 to C6, and Suceava passed from cluster C2 to cluster C5.
Taking into account these mutations as well as the evolution of the values of the five indicators in Romania's counties in the period 2000 - 2015, the mean values of the analyzed clusters (cluster centers) are presented in Table 6.

Table 5. The structure of the clusters in 2015 year

Cluster	Counties
C1	Bistrita-Nasaud, Caras-Severin, Mehedinti, Maramures, Sibiu, Alba, Arad, Satu Mare, Botosani, Buzau, Olt, Valcea, Gorj, Hunedoara, Vaslui, Dambovita, Neamt, Bacau, Bihor, Mures, Arges, Covasna, Harghita, Salaj
C2	Vrancea, Teleorman, Tulcea, Braila, Calarasi, Giurgiu, Ialomita
C3	Constanta, Galati, Brasov, Pahova
C4	Ilfov
C5	Dolj, Timis, Cluj, Suceava,
C6	Iasi
C7	Municipiul Bucuresti

Source: own construction using SPSS

The school population and the number of teaching staff decreased in 2015 compared to 2012 in almost all counties. It should be noted that in the county of Iasi, cluster C6, the average of the school population increased 3 times in the year 2015 compared to the year 2000, from 53387 pupils to 191754.
Bucharest Municipality records a drop in the school population, the teaching staff, but also the number of pupils returning to a class.

Tabelul 7. The values of means for TSP, TTS, TCR, SPTS and SPCR în the year 2015

Cluster	Number of counties	TSP scholars	TTS teachers	TCR classrooms	SPTS scholars/teacher	SPCR scholars/classroom
C1	24	68846.2	4759.8	2609.3	14.3	26.0
C2	7	43220.7	2825.3	1379.7	15.4	31.3
C3	4	115301.5	6842.8	3033.0	16.8	38.3
C4	1	48103.0	2762.0	1233.0	17.0	39.0
C5	4	137507.8	9158.0	4381.5	15.3	31.5
C6	1	191754.0	11495.0	5317.0	17.0	36.0
C7	1	436685.0	25172.0	10276.0	17.0	42.0

Source: own construction using SPSS

4. CONCLUSIONS

The results obtained on the basis of the cluster analyzes carried out for the years 2000 and 2015 show a tendency of convergence and uniformization for more than half of Romania's counties, but also emphasising differences in counties such as Cluj, Iaşi, Constanţa, Braşov or Bucharest.
The group of counties in cluster C2, Tulcea, Brăila, Călăraşi, Giurgiu, Ialomiţa, where, in 2015, School population per teaching staff are 15 pupils on average, the number of classes decreases from 1420 in 2015 to 1379 in 2010 and a class has on average 31 pupils.
With the exception of Iași, Covasna, Harghita și Sălaș Counties, in all counties in Romania, both the school population and the number of teachers are decreasing. The main factors that led to the drop in the school population are drop in birth rates, emigration and school drop-out.
Certainly, the most serious situation with regard to the school population of a teacher, as well as the school population corresponding to a class, is in the city of Bucharest.
The scientific contribution of this research is found in the group of Romanian counties according to the number of the school population, the teaching staff and the number of classes, using the cluster method. It provides groups that form spacious county units with similar demographic indicators in which appropriate urgent measures of a demographic, economic and educational policy must be implemented.

REFERENCES

Bădescu, G., & Petre, N. (2012). *The Child Wellbeing Rural Romania*, Risoprint, Cluj-Napoca, p. 29

Faubert, B., (2012), In-school policies and practices for overcoming school failure: A Literature Review, OECD Education Working Paper, OECD, Paris.

Field S., Kuczera M., & Pont, B. (2007). No More Failures: Ten Steps to Equity in Education, OECD, Paris

Heckman, J. (2011). The Economics of Inequality: The Value of Early Childhood Education. *American Educator*, 35, 31-47

Horwitt, D., & Cramer, D. (2005), *Introduction to SPSS in Psychology*, Third Edition, Pearson Education Limited.

Jigău, M., & Fartuşnic, C. (coord.) (2012). Estimarea dimensiunii fenomenului de abandon şcolar folosind metodologia analizei pe cohortă. Bucureşti: Editura Alpha MDN.

Oprea C., & Zaharia M. (2011). *Elemente de analiza datelor şi modelare utilizând Excel*, Editura Universitară, București.

Rotaru, T. (coordonator), Bădescu G., Culic I., Mezel E., & Mureșan C. (2006). *Metode statistice aplicate în științele sociale*, Ed. Polirom, Iași, p.272-303.

Popa, M., *Statistică pentru psihologie. Teorie și aplicații SPSS* (2008). Editura Polirom, Iași.

Zaharia M., & Gogonea, R.M., (2016). A Cluster Analysis of Industrial Production Indices in Some European Country and Turkey, *Economic Insight – Trends and Challenges*, Vol. VII (LXVII), no. 2/2016, p.59-71.

Yorke, M., Barnett, G., Evanson, P., Haines, C., Jenkins, D., Knight, P., & Woolf, H. (2005). Mining Institutional Datasets to Support Policy Making and Implementation. *Journal of Higher Education Policy and Management*, 27(2), 285-298

http://statistici.insse.ro/shop/index.jsp?page=tempo3&lang=ro&ind=SCL103F
http://statistici.insse.ro/shop/index.jsp?page=tempo3&lang=ro&ind=SCL104A
http://statistici.insse.ro/shop/index.jsp?page=tempo3&lang=ro&ind=SCL105C
http://www.worldvision.ro/_downloads/allgemein/
The_Child_Wellbeing_ Rural_Romania_EN.pdf