CLUSTER ANALYSIS OF PC ENDOWMENT IN PRE-UNIVERSITY EDUCATION AT ROMANIAN COUNTIES LEVEL

ZAHARIA MARIAN,
PROFESSOR PHD, PETROLEUM-GAS UNIVERSITY OF PLOIESTI, ROMANIA
e-mail: *marianzaharia53@gmail.com*

PĂTRAŞCU AURELIA,
ASSOCIATE PROFESSOR PHD, PETROLEUM-GAS UNIVERSITY OF PLOIESTI, ROMANIA
e-mail: patrascuaura@yahoo.com

TĂNĂSESCU ANA,
ASSOCIATE PROFESSOR PHD, PETROLEUM-GAS UNIVERSITY OF PLOIESTI, ROMANIA
e-mail: atanasescuro@yahoo.com

Abstract

In the contemporary world, ICT has become omnipresent. Not only the younger generation, but also the other categories of the population uses it in their daily activity, on the one hand. On the other hand, however, in many counties, the education and, especially, the pre-university education remained deficient regarding the endowment of schools with personal computers (PCs) and other ICT. Based on this observation, this paper analyses the endowment with PCs of pre-university educational institutions at county level and their grouping by the average number of existing PCs. The paper emphasizes the major differences that exist between counties, in this regard.

Key words: *ICT, cluster analysis, PC, education*

JEL Classification: *A20, C10, H52*

1. Introduction

In a modern education, the use of ICT is a necessity not only for university education but also for all levels of education. However, in Romania [2], in comparison with EU states [8], the situation is far from being satisfactory.

Based on these observations and according to the organization of pre-university education in Romania and to the statistics of National Institute of Statistics from Romania [7], in this research were included pre-university educational institutions that were grouped in the following categories: primary and gymnasium education (PGE), primary and gymnasium special education (PGSE), high school (HS), vocational education (VE), post-secondary education and foreman (PSE).

The paper presents some aspects of the endowment of these educational institutions with PCs, at Romania counties level, in 2014. Of course, if we consider only one of the indicators, for instance the share of PCs in primary and gymnasium education, at county level, we see that there are counties in which that share is less than 1.5%, among them are Covasna with 1.20%, Ialomita with 1.17% and Giurgiu with 1.09%, but also counties in which the share exceeds 4%, such as: Bacau with 4.08%, Iasi with 4.66% and Bucharest with 4.85%.

The paper proposes an integrated analysis of the number of PCs on the five categories of pre-university educational institutions (PGE, PGSE, HS, VE and PSE). For this purpose, it was used the Hierarchical Clustering methodology that allows both counties classification and emphasizing the similarities and differences between them.

2. Methodology

In order to fulfil the paper main goal, the Hierarchical Clustering methodology was used. This was completed with testing of statistical hypotheses, the main tests being Student (t) test for averages and Fisher (F) test for variances. Also, for analyses and substantiation of conclusions other statistical tools and techniques were used [1,6].

A first analysis aimed to test the statistical significance of the average number of pre-university educational institutions from the 42 counties of Romania. T test was used for this purpose. The condition to reject the null hypothesis (the average value is not statistically significant), for 95% Confidence level, is:

$$Sig.(2-tailled) < \alpha = 0.05 \quad (1)$$

In the Hierarchical Clustering methodology employed to achieve the classification of the 42 counties from Romania by endowment with computers of pre-university educational institutions included in the analysis, the Euclidian distance was used [5] for generating the Proximity Matrix ($W = \|w_{ij}\|_{i=\overline{1,n},j=\overline{1,n}}$):

$$W = \|w_{ij}\|_{i=\overline{1,n},j=\overline{1,n}}, \quad w_{ij} = \sqrt{\sum_{i=1}^{n}\left(z_{ik} - z_{ij}\right)^2},\ j = \overline{1,m},\ k = \overline{1,m}\ j \neq i,\ k \neq i,\ w_{ii} = 0 \tag{2}$$

Also, Ward's method was used in order to determine the distance between clusters [3]:

$$\Delta(A,B) = \sum_{i\in A\cup B}\left\|x_i - m_{A\cup B}\right\|^2 - \sum_{i\in A}\left\|x_i - m_A\right\|^2 - \sum_{i\in B}\left\|x_i - m_{\mathbf{B}}\right\|^2 - \frac{n_{A\cap B}}{n_{A\cup B}}\left\|m_A - m_B\right\|^2 \tag{3}$$

Statistical significance testing regarding the belonging of variables to clusters was performed using ANOVA methodology if we accept the null hypothesis of Levene's Test:

$$H_{0\_1}: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \ldots = \sigma_r^2 \tag{4}$$

Data processing was performed with SPSS [4] and Excel.

3. Results and discussions

In any econometric analysis, the choice of some statistical indicators in a given context, indicators based on which analyses are done, involves, as a first step, the testing of the statistical significance of them. In order to group and characterize the Romania's counties in terms of endowment with PCs of pre-university education institutions, the statistical average was used as a first indicator. The results of testing of the statistical significance of this statistical indicator, for the five categories of educational institutions are presented in Table 1.

For Test_Value = 0, the average of PCs number in institutions of vocational education (VE), Sig.(2-tailled) = 0.091> α = 0.05, which leads to accepting the null hypothesis: the average of PCs number is not statistically significant. Consequently, the variable VE will be removed from the analysis.

Table 1 **Testing the statistical significance of the average of PCs number (Test Value = 0)**

	t	df	Sig.(2-tailled)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
PGE	16.707	41	0.000	3043.04762	2675.2066	3410.8886
PGSE	6.423	41	0.000	89.42857	61.3098	117.5473
HS	10.511	41	0.000	2759.38095	2229.1947	3289.5672
VE	1.730	41	0.091	10.64286	-1.7837	23.0694
PSE	8.065	41	0.000	89.21429	66.8744	111.5542

Source: own calculus using SPSS

For the other four variables (PGE, PGSE, HS and PES), starting from the series of data from 2014 and applying Hierarchical Clustering methodology, it was generated a classification of the counties that is shown in the dendrogram from figure 1.

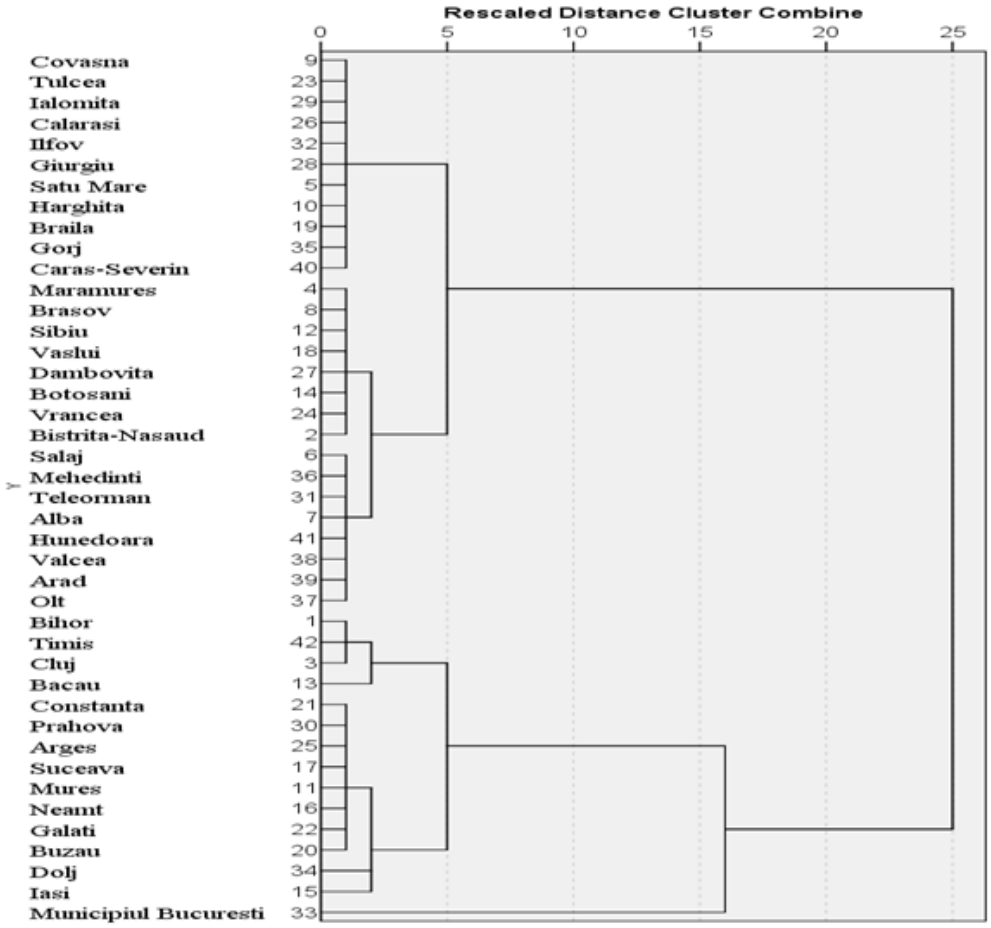

Figure 1. Dendrogram using Ward Linkage.

From the analysis of dendrogram (figure no. 1), the classification with five clusters has been chosen. Their composition is shown in Table 2.

Table 2 **Clusters composition**

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

Source: own construction

The first four clusters contain between 4 and 16 counties and the cluster C5 contains one item: Bucharest. This is due to the particular situation of Bucharest compared to the other 41 counties. As it can be seen from dendrogram (figure no. 1), the rescaled distance between this and the others is huge, so its inclusion together with C1 and C4 in a single one would affect the quality of analysis.

Analysing the main characteristics of pre-university education institutions regarding the endowment with PCs (Table 3), it results that, in this classification, Bucharest occupies a special position. Thus, in 2014, in Bucharest, in the educational institutions for primary and gymnasium education (PGE), there were 6203 PCs with 47.27% higher than the average of cluster C4 and of 3.28 times higher than the average of cluster C3. Also, in Bucharest, there were 545 PCs in primary and gymnasium special education (PGSE) that is of 15.39 times higher than the average of cluster C3, 11037 PCs in high school (HS) and 181 PCs in post-secondary education and foreman (PSE). Among the other four clusters, the highest values regarding the average number of existing PCs in primary and gymnasium education are recorded in cluster C4 (4212.80 units). Compared with this, in cluster C1 were, in average, with 17.15% less (3911.75

units) and in cluster C2 with 36.15% less. The worst situation was in the cluster C3, where the average number of PCs were of 2.23 times less than in the cluster C4.

Regarding the endowment with PCs of primary and gymnasium special education, the cluster C1 was on the first place. With an average number of 149.25 units, it exceeded with 44.9% the average number of PCs from cluster C4, being double compared with C2 and of 4.22 times higher than the average number of existing PCs, in cluster C3.

The number of existing PCs, in 2014, in high school, in the four clusters is relatively similar to the average number of PCs, from primary and gymnasium education. It should be noted, however, that in the cluster C1, the average number of PCs was with 12.93% higher than in primary and gymnasium education, while in the other three clusters the average number of PCs in high school was lower than in primary and gymnasium education with 17.19% in C2, with 13.65% in C3 and with 20.23% in C4. This situation has placed the cluster C1 on the first place regarding the average number of PCs in high school to the detriment of cluster C4 that is placed on first place regarding the average number of PCs in primary and gymnasium education.

Table 3 **Key features of PGE, PGSE, HS, PSE at cluster level**

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min	Max
						Lower Bound	Upper Bound		
PGE	C1	4	3911.75	881.74	440.87	2508.71	5314.79	3280	5214
	C2	16	2689.38	531.76	132.94	2406.02	2972.73	1851	3651
	C3	11	1890.91	445.46	134.31	1591.64	2190.17	1396	2703
	C4	10	4212.80	765.90	242.20	3664.91	4760.69	3234	5962
	C5	1	6203.00					6203	6203
PGSE	C1	4	149.25	71.40	35.70	35.63	262.87	46	210
	C2	16	74.69	42.93	10.73	51.81	97.56	14	162
	C3	11	35.36	32.41	9.77	13.59	57.14	0	116
	C4	10	103.00	48.37	15.30	68.40	137.60	28	202
	C5	1	545.00					545	545
HS	C1	4	4417.50	930.78	465.39	2936.42	5898.58	3456	5692
	C2	16	2226.56	602.11	150.53	1905.72	2547.40	1265	3228
	C3	11	1632.45	600.21	180.97	1229.23	2035.68	599	2572
	C4	10	3360.50	774.65	244.96	2806.35	3914.65	2006	4303
	C5	1	11037.00					11037	11037
PSE	C1	4	58.25	33.54	16.77	4.88	111.62	17	96
	C2	16	81.25	24.54	6.13	68.17	94.33	40	131
	C3	11	12.82	13.14	3.96	3.99	21.65	0	34
	C4	10	189.20	46.11	14.58	156.22	222.18	137	283
	C5	1	181.00					181	181

Source: own construction using SPSS

In terms of average number of existing PCs, in 2014, in post-secondary education and foreman, the cluster C4 is on first place with an average of 189.2 units, followed at a big distance by the other three clusters. Thus, in the post-secondary education and foreman, as compared with the cluster C4, the average values were of 2.33 times lower in the cluster C2, of 3.25 times lower in the cluster C1 and of 14.77 times lower in the cluster C3.

4. Conclusions

Excepting Bucharest, compared to other clusters, the better situations regarding PC endowment of pre-university education are recorded in the cluster C4 that is ranked first regarding the average number of PCs in primary and gymnasium education and in post-secondary education and foreman, as well as in the cluster C1 that is ranked first regarding the average number of PCs in primary and gymnasium special education and, also, in high school.

Definitely, the worst situation regarding the number of PCs in pre-university education is in the counties that are included in cluster C3. Thus, if in primary and gymnasium education the ratio between the highest value 4212.8 (C4) and the lowest 1890.9 (C3) was of 2.23, in the primary and gymnasium special education the ratio between the highest value 149.25 (C1) and the lowest 35.36 (C3) was of 4.22 and for the average number of PCs in high school, the higher the ratio between the highest value 4417.5 (C1) and the lowest 1632.45 (C3) was of 2.71.

5. Bibliography

[1] **Babucea, A.G.**, *Analiza datelor – Metode statistice avansate*, Editura Universitaria, Craiova, 2010;
[2] **Enăchescu, D., Zaharia, M.,** Considerations on the Impact of Information Technology Evolution in Romania in the Last Six Years**,** *Economic Insight – Trends and Challenges*, vol.II(LXV), No.2/2013, pp 82-89;
[3] **Marinoiu ,C.**, Bootstrap Stability Evaluation and Validation of Clusters Based on Agricultural Indicators of EU Countries**,** *Economic Insights – Trends and Challenges*, Vol.V(LXVIII), No. 1/2016, pp. 65 – 72;
[4] **Popa, M.,** *Statistică pentru psihologie. Teorie și aplicații SPSS*, Ed. Polirom, Iași, 2008;
[5] **Rotaru, T. (ed.), Badescu, G., Culic, I., Mezei, E., Mureșan, C.**, *Metode statistice aplicate în științele sociale*, Ed. Polirom, Iași, 2006, pp. 272-303;
[6] **Tomescu-Dumitrescu, C., Balacescu, A.,** *Statistică economică aplicată*. Editura Universitaria, Craiova, 2008;
[7] www.insse.ro - Official site of National Institute of Statistics
[8] **Zaharia, M., Enachescu, D.,** E-Commerce by Individuals - A Statistical Analysis of Evolutions of Internet Purchases by Individuals in Some Former Communist States in 2007 - 2012 Period, *Journal of Applied Computer Science & Mathematics, Stefan cel Mare University of Suceava*, Issue 17, Year 2014.