

ANALYSIS OF THE PRE-UNIVERSITY SYSTEM IN ROMANIA USING THE CLUSTERING TECHNIQUE

*Ioana Gabriela Grigorescu**

Abstract

A country is all the more competitive as the pre-university education system is better developed. In this article, a study was made of the pre-university education system in Romania in terms of infrastructure indicators, human resources involved in it and the number of students enrolled in various forms of education. After registering the values of the indicators from the Tempo database of the National Institute of Statistics (INS), with the help of the clustering technique and using the SAS Enterprise Guide program, the counties were divided into three distinct, homogeneous classes. Further on, the results found were analyzed in dynamics, referring to the previous year, 2019, highlighting the differences and finding solutions to improve the existing situation.

Keywords: Romania, education, clustering, strategy, classification, SAS Enterprise Guide

JEL Classification: I20, I21, I25

1. Introduction

The present time brings to the fore a rather controversial education system, given the changes that have taken place. Starting from the unanimously valid principle that the formation of a child starts from the embryonic stage, here the role of the mother being an essential one, several stages of this process are identified: the pre-school stage (nursery), the preschool stage (kindergarten), the primary education stage (preparatory class and grades I-IV), gymnasium stage (grades V-VIII) and high school stage (IX-XII). This extensive process of education of an individual has its origins in the family of origin, but is continued by the public environment of the state: nurseries, kindergartens, schools and high schools.

The training course of a child in Romania starts early. In this sense, from the age of one he can attend pre-school education through the regulated institutions of the state. This stage was implemented in the system to help the child to develop emotionally, cognitively and sensorial.

Continuing education with the second stage, preschool, when the child begins to be aware of his abilities, he realizes that he needs to be part of a group that will help him in his own development. Other factors play an essential role in the harmonious development of a child: the size of the group, the ability of the educator to properly manage this relationship between the child and the teacher. That is why the groups should be of

* Grigorescu Ioana Gabriela is Assistant Professor PhD. at the Romanian-American University, Bucharest. E-mail address: grigorescu.ioana.gabriela@profesor.rau.ro

maximum 25 children, and the teacher must be specialized in this respect. Unfortunately, the system does not always offer what it should and then either the number of children in a group far exceeds the required maximum number, or the educator is overwhelmed by other shortcomings of the system and then acts as he sees fit.

The third stage is that of the primary cycle, when the child walks shyly towards the new world of the challenges that will follow. This stage also comes with a recalculation of the plans, considering that, from the school year 2012-2013, a preparatory class was introduced at the school, appearing a new step and thus giving a new valence to the parents, who due to age fears were later school start due to difficulties that may have existed. At this stage the child enters a new stage of its existence and development where the emphasis is on the connection between the student and the teacher. The teaching-learning methods used by the educator, but also the pedagogical ability represent defining elements of the teaching profession. This stage is a defining one in the mental and educational evolution of the future student. Thus, teachers have an important role to play in this stage because they must combine training in various fields with elucidating behavioral typologies and forming the characters of future adults.

Another defining stage in this educational path is the middle school stage, when students have to make the transition from teacher to more teachers with multiple requirements and different grading methods.

The last stage according to the classification made above is the high school one. Unfortunately, although it should be an important period from the perspective of a child's evolution to adulthood, it seems that some of the students drop out. Either they come from disadvantaged families where education is not a priority and then they have to work to survive, or the lack of cooperation between students, teachers and family pushes students to make certain momentary decisions, which will mark their whole life.

2. Defining the problem, data and coding of the indicators

The problem I set out to address in this research paper is the analysis of each county in Romania in terms of infrastructure, staff and students in the pre-university education system. Thus, I propose to make a grouping of counties according to specific indicators.

The indicators chosen further in the analysis are presented in table 1.

Table 1. Coding of indicators

Crt. no.	Code of indicator	Description of each indicator chosen in the analysis
1	I1	Number of educational units in each county
2	I2	Number of school laboratories in each county
3	I3	Number of classrooms (school offices / amphitheatres) in each county
4	I4	The teaching staff from each county
5	I5	The number of students enrolled in pre-school education
6	I6	Number of students enrolled in secondary education
7	I7	Number of students enrolled in high school education

I1, the number of educational units in each county, quantifies the totality of the administrative education units in a county with legal personality, in which one or more levels of teaching operate, having a single leadership. In Romania, the following are considered educational units: nurseries, kindergartens, primary and secondary school units, special primary and secondary education units, high schools and others.

Indicator I2, the number of school laboratories in each county, quantifies the total number of rooms in all educational units in a county endowed with facilities, equipment, instruments, substances, teaching materials in which teaching experiences and practical works are carried out.

I3, the number of classrooms (school offices / amphitheatres) in each county, has in its composition the total number of rooms within all the educational units in a county destined for the process of education and school training. In the teaching process the same classroom can be used by two or at most three classes of students in a row, during a day.

I4, the teaching staff from each county, counts all the people in a county who are employed in the education system and teach in the educational and training process, having full or part time. Each teacher registers only once and only at the educational unit where he / she has the basic function or teaches the largest number of teaching hours.

I5, the number of students enrolled in pre-school education, quantifies the number of minor children enrolled in the preschool education system, at the level of each county.

The I6 indicator, the number of students enrolled in secondary education, is represented by the number of minors in the secondary education system, in each county.

I7, the number of students enrolled in high school education, represents the totality of students from each county who attend the form of high school education.

The indicators described above were registered for 2020 from the Tempo database, the National Institute of Statistics (INS), their values being presented in the table 2.

Table 2. The values of the indicators registered for each county in Romania

County	I1	I2	I3	I4	I5	I6	I7
Alba	140	338	2733	3914	8332	25574	11195
Arad	154	618	3465	4850	10577	32801	12968
Arges	215	731	4354	6669	15345	46513	20915
Bacau	187	594	4631	6723	15580	50471	18389
Bihor	204	796	5044	7970	16398	48511	17386
Bistrita-Nasaud	116	256	2588	3417	8607	25702	8864
Botosani	132	278	3120	4723	10408	34588	13264
Braila	104	241	2072	2986	6919	22460	8856
Brasov	186	792	3815	6725	16075	48703	14951
Buzau	150	375	2843	4357	9937	33993	12698
Calarasi	96	190	1763	2609	6566	24187	6455
Caras-Severin	120	286	2294	3218	6234	17972	8895
Cluj	271	2286	6263	11667	20338	53830	18973
Constanta	228	871	4201	8030	17945	59571	22831
Covasna	79	183	1863	2715	6083	18804	5105
Dambovita	153	422	3201	4888	11306	37544	13999

Dolj	214	1129	4431	7981	15464	46668	20465
Galati	188	853	3091	5884	11783	40211	17719
Giurgiu	87	170	1434	2353	5385	21255	5568
Gorj	116	406	2460	3809	7619	24092	12642
Harghita	130	255	3028	4259	10176	27840	8741
Hunedoara	111	459	2983	3804	8039	26540	12595
Ialomita	91	208	1629	2309	5860	20587	7299
Iasi	260	2158	6396	11630	21357	72607	24554
Ilfov	132	212	2011	3243	11904	33979	7096
Maramures	182	398	3747	5357	12818	35480	17083
Mehedinti	106	229	1822	2732	5460	16995	9476
Bucuresti	597	3974	12482	26675	44103	151097	72820
Mures	187	586	4686	7355	15984	46932	15300
Neamt	140	457	3530	4640	11026	37493	15094
Olt	161	277	2610	4043	9348	28359	13878
Prahova	236	713	4236	6960	15751	56638	23114
Salaj	111	308	2086	2607	7346	19443	6922
Satu Mare	131	276	2776	3806	9959	28052	9301
Sibiu	176	677	3510	5407	12547	35523	11294
Suceava	224	721	5357	8019	20315	61670	23925
Teleorman	140	224	1987	3253	6945	22118	9339
Timis	238	1644	5715	9871	18036	52213	20519
Tulcea	80	256	1293	2063	4745	16180	5779
Valcea	129	329	2556	4016	7954	25336	11308
Vaslui	148	278	2981	4494	10181	33879	12412

2. The research methodology

The research methodology has the following stages:

- identifying the indicators and registering their values for each county in Romania from the Tempo, INS database;
- classification of counties according to all registered indicators using specific data mining algorithms (clustering);
- grouping the counties by homogeneous classes and their characterization according to the aggregate indicators calculated at the level of each class: minimum, maximum and average;
- dynamic analysis of aggregate indicators according to 2019;
- identification of the main differences existing at the level of the classes determined at the previous step, in 2020, compared to 2019;
- establishing the measures that are required at the level of the entire Romanian economy in order to improve the situation in the pre-university system.

3. Results and discussions

The next step in the proposed research methodology is to conduct cluster analysis. This involves first of all obtaining information about the distribution of data series: whether or not the series are close to the normal distribution, whether they have larger or smaller queues in a certain directive, etc. (Table 3).

Table 3. Descriptive statistics for the indicators according to which the companies will be classified

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Bimodality
I1	166.0	84.7939	3.3394	15.9369	0.6339
I2	637.3	713.1	3.1841	11.8640	0.7378
I3	3459.4	1925.3	2.7018	10.9713	0.5843
I4	5608.6	4111.5	3.5375	16.6307	0.6803
I5	12028.1	6800.7	2.7047	11.2112	0.5757
I6	37843.6	22649.3	3.2283	14.8056	0.6332
I7	14776.8	10651.6	4.0946	21.8304	0.7089

The purpose of the cluster analysis is to group the sample of counties according to the choice of the classification algorithm and well-established criteria. Thus, Ward's classification algorithm was chosen, which involves calculating the distance between an object and the centroid of each of the classes. The object will be assigned to the class where this distance will be minimal. The same will be done with the other objects in the county sample. In the end, there will be an aggregation graph (figure 1) and, depending on the assumed error, the classes will be delimited.

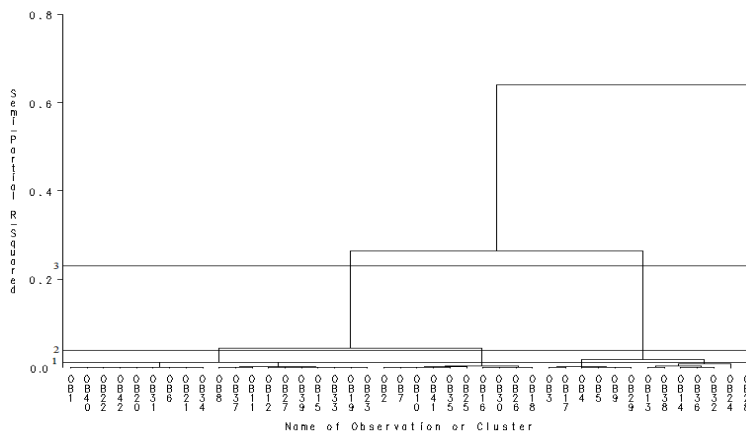


Figure 1. Dendrogram resulting from the analysis

In order to classify the counties, the classes must be chosen according to an assumed error. Thus, if the dendrogram is sectioned at the level of line 1 (figure 1) then the error will tend to zero, there will be five classes and one outlier (Bucharest). The classes, in this case, will be very close to each other, the objects can be easily classified into several classes. If the cut is made at the level of line 1 in the graph, then there will be two large classes and an outlier (Bucharest), the objects in each class will be quite different, and the error will be well above the allowed limit of 5%. If, instead, the cut will be made at the level of line 2, the error will be within the allowed limit and there will be 3 homogeneous classes and an outlier (Bucharest). Furthermore, the county of Bucharest will be eliminated from the analysis, it can be analyzed separately.

The classification of the counties according to all the indicators is the following:

- class 1: Alba, Bistrita-Nasaud, Braila, Calarasi, Caras-Severin, Covasna, Giurgiu, Gorj, Harghita, Hunedoara, Ialomita, Mehedinti, Olt, Salaj, Satu Mare, Teleorman, Tulcea, Valcea, Vrancea;
- class 2: Arad, Botosani, Buzau, Dambovita, Galati, Ilfov, Maramures, Neamt, Sibiu, Vaslui;
- class 3: Arges, Bacau, Bihor, Brasov, Cluj, Constanta, Dolj, Iasi, Mures, Prahova, Suceava, Timis.

Class 1 is represented by counties with rather low values of the indicators chosen for the study. Thus, at the class level, on average, each county has 114 educational units equipped with 274 school laboratories and 2220 classrooms. The teaching staff is also reduced compared to the other classes of counties, so on average, at the level of each county in this group its number rises to 3234. In general, in the pre-university education system, the infrastructure is built according to the number of potential students. For this group of counties, the indicators that quantify the number of students have low values compared to the other groups. Thus, on average, I5 is 7368, I6 has the value 23080 and I7 is 9098 (table 4).

Class 2 is composed of counties with a higher number of children who are enrolled in the pre-university education system (on average, I5 = 11249, I6 = 35549 and I7 = 13363). Consequently, the number of teachers is higher, I4 = 4784, but the infrastructure indicators are not with double average values compared to class 1, as it was expected to be, I1 = 156, I2 = 457, I3 = 3150. In this group, an almost double number of children compared to class 1 is enrolled with about 40 additional educational units and with only half of the necessary teachers (I4 in this class is higher than I4 in class 1 with 1550, on average) (Table 4).

Class 3 contains, on average, counties with a large school age population, so the indicators regarding infrastructure and the number of teachers will follow this trend (table 4).

**Table 4. Aggregate indicators calculated at the class level
in the years 2019 and 2020**

		2019						
		I1	I2	I3	I4	I5	I6	I7
Class 1	minim	80	168	1297	2037	5012	16730	4991
	maxim	161	457	2988	4223	10485	29542	14412
	average	114.05	280.74	2223	3229.11	7648.42	23749.26	9180.32
Class 2	minim	132	199	1888	3111	10457	33433	6881
	maxim	187	864	3755	5940	13102	41655	17571
	average	156.5	463.9	3151.9	4767.9	11629.1	36440.7	13340.5
Class 3	minim	186	598	3777	6658	16163	47598	14727
	maxim	272	2201	6480	11573	22546	73626	24283
	average	221.08	1085	4930.17	8268.17	18082.25	54508.92	20027.33

		2020						
		I1	I2	I3	I4	I5	I6	I7
Class 1	minim	79	170	1293	2063	4745	16180	5105
	maxim	161	459	3028	4259	10176	28359	13878
	average	114.11	273.89	2220.16	3233.95	7368.47	23079.84	9097.68
Class 2	minim	132	212	2011	3243	9937	32801	7096
	maxim	188	853	3747	5884	12818	40211	17719
	average	155.5	456.8	3149.9	4784.3	11248.7	35549.1	13362.7
Class 3	minim	186	586	3815	6669	15345	46513	14951
	maxim	271	2286	6396	11667	21357	72607	24554
	average	220.83	1085.08	4927.42	8300	17382.33	53693.92	20110.17

In order to highlight the impact of the Covid-19 pandemic on the pre-university system in Romania, the classification algorithm for the data registered in 2019 must be applied, having the same sample of counties and the values for the same indicators. Following its application, the structure of the groups has not changed, the only changes are on the aggregate indicators calculated at the level of the classes. Thus, some indicators increased compared to the previous year, but most of them have lower values (table 5).

Table 5. Absolute and relative changes of the aggregate indicators at the level of each class

		2020 - 2019						
		I1	I2	I3	I4	I5	I6	I7
Class 1	minim	-1	2	-4	26	-267	-550	114
	maxim	0	2	40	36	-309	-1183	-534
	average	0.06	-6.85	-2.84	4.84	-279.95	-669.42	-82.64
Class 2	minim	0	13	123	132	-520	-632	215
	maxim	1	-11	-8	-56	-284	-1444	148
	average	-1	-7.1	-2	16.4	-380.4	-891.6	22.2
Class 3	minim	0	-12	38	11	-818	-1085	224
	maxim	-1	85	-84	94	-1189	-1019	271
	average	-0.25	0.08	-2.75	31.83	-699.92	-815	82.84
		2020 / 2019						
		i1	i2	i3	i4	i5	i6	i7
Class 1	minim	-1.25	1.19	-0.31	1.28	-5.33	-3.29	2.28
	maxim	0	0.44	1.34	0.85	-2.95	-4	-3.71
	average	0.05	-2.44	-0.13	0.15	-3.66	-2.82	-0.9
Class 2	minim	0	6.53	6.51	4.24	-4.97	-1.89	3.12
	maxim	0.53	-1.27	-0.21	-0.94	-2.17	-3.47	0.84
	average	-0.64	-1.53	-0.06	0.34	-3.27	-2.45	0.17
Class 3	minim	0	-2.01	1.01	0.17	-5.06	-2.28	1.52
	maxim	-0.37	3.86	-1.3	0.81	-5.27	-1.38	1.12
	average	-0.11	0.01	-0.06	0.38	-3.87	-1.5	0.41

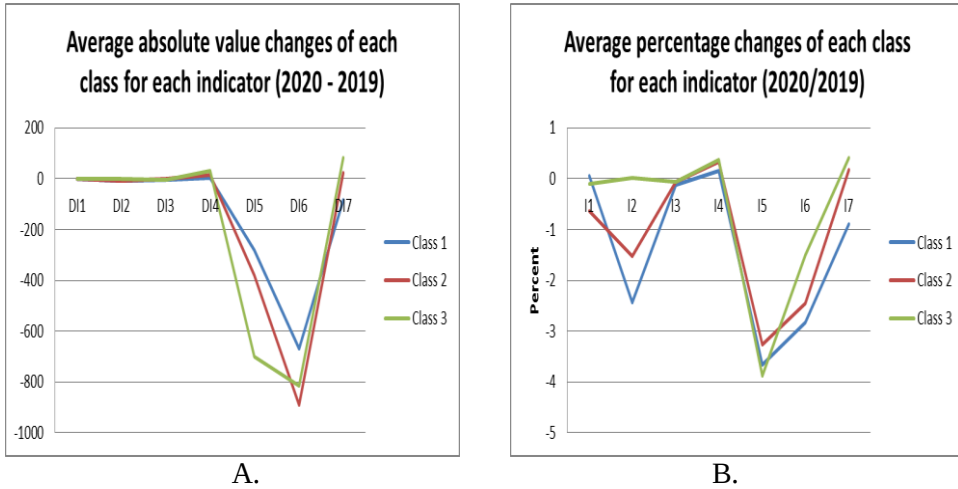


Figure 2. Average absolute value changes (A) and average percentage changes (B) for each indicator from each class, in the period 2020 - 2019.

According to table 5, respectively to graphs A and B from figure 2, it is observed that in 2020, compared to 2019, at the level of each class found, the indicators in absolute value tend to remain constant or to register slight decreases. And from a percentage point of view, the trend is maintained, proving once again that the infrastructure (educational units, classrooms, laboratories, etc.) and human resources (teachers) are developing or contracting depending on demand (number of children enrolled in the pre-university education system).

Some solutions can be highlighted to improve the activity of the pre-university education system as follows:

- ✓ Organizing specialized monthly trainings attended by all teachers from the entire pre-university system, regardless of background;
- ✓ Development of new learning methods based on innovation and personal development of the student;
- ✓ The use of different new teaching means, which combine traditional methods with elements of originality and thus highlight the abilities of each student;
- ✓ Increasing teachers' salaries, settlement of transportation and housing provided to those who choose to go to disadvantaged environments to practice;
- ✓ Finding methods to stimulate the increase of nature: vouchers, facilities offered to parents, increase of allowances for raising children and other allowances;
- ✓ Encouraging teachers with good results by offering certain merit scholarships for a certain limited period of time;

4. Conclusion

A country is more developed the more functional and competitive the education system is. For an individual to go through the whole pre-university educational process, from kindergarten to high school, it requires quite a lot of resources. Human, financial and infrastructure resources must be adapted to the number of children, and a possible

imbalance caused by more or less predictable causes can be overcome without greatly disrupting the educational process. Looking at the two years analyzed as a whole, it can be concluded that although the pandemic had a major impact on all economic and social activities around the globe, in the end people have adapted their actions so that they can lead a life as close as possible to normalcy. From the point of view of future research, it is recommended to minimize the error assumed. This can be done by applying other classification methods, such as discriminated analysis. Attempts can also be made to improve the results by using special artificial intelligence algorithms, such as training neural networks for classification or genetic algorithms.

Looking at the two years as a whole, it can be concluded that although the pandemic had a major impact on all economic and social activities around the world, in the end people have adapted their actions so that they can lead a life as close as possible to normality.

References

1. Mahmudovna Nigmatova Mavjuda (2021), *Improving the preschool education system from the perspective of the age values of children*, Asian Journal of Multidimensional Research, Vol. 10, No. 1, p. 182-185.
2. Bound, John, Breno Braga, Gaurav Khanna, and Sarah Turner (2021), *The Globalization of Postsecondary Education: The Role of International Students in the US Higher Education System*, Journal of Economic Perspectives, 35 (1): 163-84.
3. Rona Tamir, Dafna Regev (2021), *Characteristics of parent-child art psychotherapy in the education system*, The Arts in Psychotherapy, Vol. 72, 101725.
4. Młodak, A. (2020), *k-Means, Ward and Probabilistic Distance-Based Clustering Methods with Contiguity Constraint*, Journal of Classification, <https://doi.org/10.1007/s00357-020-09370-5>
5. Iordache Ana Maria Mihaela (2011), *Using Sas Enterprise Guide Software In Classification*, Journal of Doctoral Research in Economics, The Bucharest University of Economic Studies, Vol. 3, No. 1, p. 3-12.
6. Tudorache Ana Maria Mihaela (2008), *Data mining and the process of taking decisions in e-business*, Romanian Economic and Business Review, Vol. 3, No. 4, p. 111-116.
7. Iordache Ana Maria Mihaela, Codruța Cornelia Dura, Cristina Coculescu, Claudia Isac, Ana Preda (2021), *Using Neural Networks in Order to Analyze Telework Adaptability across the European Union Countries: A Case Study of the Most Relevant Scenarios to Occur in Romania*, International Journal of Environmental Research and Public Health, Vol. 18, No. 20, id: 10586.
8. Koti Neha, M Yogi Reddy (2020), *A Study On Applications Of Data Mining*, INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 02, FEBRUARY 2020 ISSN 2277-8616.
9. Irit Belity (2021), *Working with Parents within the Education System*, Routledge, ISBN: 9781003008323.
10. Joseph Dunne (2005), *What's The Good of Education?*, Routledge, ISBN: 9781003209317.
11. Innovative Cluster Model For Improving The Quality Of Education (2021), Link: <https://cyberleninka.ru/article/n/innovative-cluster-model-for-improving-the-quality-of-education>

12. Institutul National de Statistica (INS), link: <http://statistici.insse.ro:8077/tempo-online/#/pages/tables/insse-table>
13. <http://www.sar.org.ro/wp-content/uploads/2012/12/Restructurarea-sistemului-de-educatie-din-Romania-din-perspectiva-evolutiilor-pe-piata-interna-si-impactul-asupra-progresului-cercetarii.pdf>
14. http://www.store.ectap.ro/articole/703_ro.pdf
15. https://ibn.idsi.md/sites/default/files/imag_file/114-118_32.pdf
16. https://www.xn--libriecolar-1nbi872b.ro/image/data/rasfoire/Sistemul_de_invatamant_2.pdf