Eliminating Correlations and Redundancy in Academic Evaluation Through Multivariate Analysis

Grizendi, J. C.[1] M,. Gomes, F. J[2] and Souza, R. C[3]

***Abstract.* In Brazil, according to the federal legislation, Faculties and Universities must implement institutional self-evaluations as a necessary activity inside their planning schedule. To help in these procedures, the Education Ministry issues some directives for these evaluations, which include, normally, qualitative questionnaires to be filled in by the students. These questionnaires can be helpful in the assessment of the academic quality and infrastructure conditions of the institution. The importance of multivariate techniques for structuring these evaluation questionnaires is based on the possibility of incorporating multiple variables for clarifying some relationship that, otherwise, would not be possible. Utilizing statistics tools, like variable reductions and patterns classification, an analysis of an evaluation questionnaire, utilized by a local University, shows several correlations and redundant questions. Analysing this situation, this paper shows that a simpler questionnaire, with a reduced number of questions, more specifically, two questions, based on multivariate techniques, can conduce to better results, which demand less effort, is simpler and objective and more effective.**

Index Terms – Education Assessment, Evaluation Procedures, Multivariate Techniques, Statistic tools

INTRODUCTION

It is a well known fact the problems and complexities embedded in any educational assessment procedure. In Brazil, the educational evaluation system encloses the earlier stages of apprenticeship until the Faculties and Universities, public and privacy. According to the federal legislation, the Faculties and Universities must implement institutional self-evaluations as an integral part of their planning procedures. These directives are based on the Federal legislation, issued by the Federal government, on December 15th, in the year 2003, thought the *Medida Provisória nº 147*. To help in these procedures, the Education Ministry issues some directives for these evaluations that include, normally, the application of qualitative questionnaires to be filled in by the students. The answers obtained can help in the evaluation of the academic quality and infrastructure conditions of the institution. The main objective of these questionnaires is to evaluate the student's satisfaction relative to the professor's methodology and his enrollment with the course. Based on the quantitative and qualitative analysis of the results, it is expected that they can help the institution in their planning procedures and corrective actions.

It is worth to say that these evaluation procedures can also be helpful for these institutions when they are planning their market strategies. This importance comes from the fact that there is in Brazil, actually, an accelerated growing of the educational market and a sharp concurrence among these institutions and, in such conditions, any quality differential can be very important.

This paper deals with an analysis procedure of such questionnaires, filled in by the students and utilized on the evaluation procedure at a private University located on the city of Juiz de Fora, Brazil. The questionnaire was structured with 18 questions, was filled in by 800 students and comprised 155 teachers working at the university. As the results were accessible, it was possible to analyze these results trough multivariate techniques and get the conclusions presented in this work.

Although it was not difficult to fill in the formularies and this process was also not time consuming, the general impression one could gather was that the process could be optimized and an analysis of the whole process was realized.

The analysis process was, initially, based on a factorial correspondence that tries to describe the variability of some random variables, named common factors, related with the original vector through a linear model. In this model, one can consider the vector variability derived from the common factors and from variables not included in the model, associated with random errors. Through this procedure it is possible to collect the variables into new vectors mutually not correlated.

For proceeding with the work a cluster analysis was utilized. This methodology allows separation of the sample universe into groups in such a way that homogeneous variables, considering the selected characteristics of the variables, are grouped together, but they are heterogeneous when considering their relation with other groups variables. In this procedure, the Euclidian distance was selected as the distinctive measure.

The paper is structured as follows: the section one discuss the theoretical basis for the analysis and the second presents a cluster analysis for the system. The validation results and conclusions end the paper.

[1] Grizendi, J. C. M, Estácio de Sá University, grizendi@acessa.com
[2] Gomes, F. J., Engineering Faculty, Federal University of Juiz de Fora, chicogomes@terra.com.br
[3] Souza, R. C., PUC-Rio, reinaldo@ele.puc-rio.br

THEORETICAL BASIS

The multivariate techniques [1] - [2] importance lies on the possibility of incorporating multiple variables in the analysis procedure, highlighting relations that do not appear on the singlevariate and bivariate techniques. It is necessary to consider, however, that when the number of variables selected for the analysis process increases, increases also the possibility that they are not de-correlated and have distinct meanings and concepts. For this reason, a fundamental step in the analysis procedure, for getting a better data interpretation, is to understand clearly the correlations among variables. A powerful tool that can be utilized in this step is the factorial analysis [3] [4]. It can be helpful for the definition of a structure embedded in the data matrix, making it possible to select individual dimensions, as well the correlation between each variable and its dimension. Effectively, with the factorial analysis is possible to get a data reduction without loss of information. This technique analysis deals with interrelationship techniques, which means that all the variables are simultaneously analyzed for getting their correlation with all the other variables. For this, the factor concept is utilized, allowing the maximization of the explaining strength of the whole universe of variables.

After the correlation among variables is ready, the group analysis was then utilized. His main objective is clustering together objects with similar characteristics, selecting for the same group individuals with similarities. It is possible, in this way, to maximize homogeneity of those elements belonging to one cluster and, at the same time, to maximize the difference among the different clusters [5] [6]. Some kind of similarity measures must then be utilized an, among then, the most commons are the correlated measures, distance measures and association measures. An important point is that this analysis has not an inferential characteristic; it just describes the clusters. Besides, it is influenced by the variables selected as the basis for the similarity measures.

Another step necessary in this process is the discriminante analysis [6]. In this step, a statistical variable, defined as a linear combination of several other variables and that integrate the discriminante function, must be determined. For this, it is necessary to weight properly the several statistical variables for maximizing the variance among the clusters relative to the variance inside each cluster. With this procedure, one tries to get the difference among clusters and to have a probability prediction that an individual belongs to same particular class or group.

METHODOLOGY

The questionnaires could be accessed virtually utilizing the Internet or the University intranet and was filled in on-line. As soon as one student was logged in, the system selected the disciplines he was attending and the professors responsible for them. Each student had to answer 18 questions, each question comprising 5 possible answers; of these, the student had to select just only one. The answers were qualitative ones, changing from A – Excellent to E – Not Acceptable. The system was structured in such a way that it was not possible, for the student, to select more than one option or to answer questions concerning the evaluation of a teacher whose disciplines he was not attending.
For the analysis procedure, it was necessary to change these data from qualitative to quantitative ones. So, a linear relation between each qualitative concept and a sharp value was established: A=100; B=75, C=50, D=25 e E=0. The professors analysis was based on 18 variables, displayed on Table 01 and Figure 01, where is also possible to see their statistical description. It is worth to say that there were no missing answers and at least once of all the variables got the maximum value, 100.

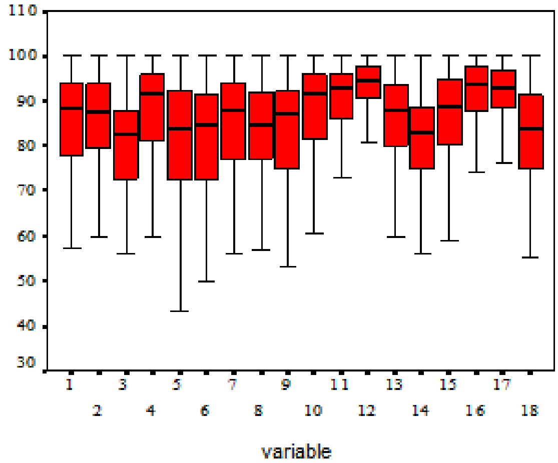

FIGURE 01
BOX PLOT FOR THE QUESTIONNAIRES

Trough the factor analysis concept it was possible to establish the correlation structure of the 18 variables and detect the strongly correlated ones. From this, it was also possible to create indices for explaining the variability among variables. The variables explaining factors are displayed on Table 02

TABLE 01
STATISTICS DESCRIPTION

QUESTIONS: THE PROFESSOR…	VARIABLE	MIN	MEAN	ST DEV
Presents a teaching schedule.	1	36.36	84.38	13.22
Explains the relevance of the discipline in the professional formation.	2	33,09	84,88	12,45
Explains the interrelationship among the disciplines of the course	3	27,21	79,24	12,67
Fulfills the program adopted for the discipline.	4	45,45	87,53	11,89
Explains the contents clearly.	5	16,91	80,56	16,00
Stimulates the participation of the students in the discussions	6	23,53	80,76	14,39
Has a good didactic and pedagogic basis	7	29,41	83,59	14,08
Always answer the students demands, even out of classroom Issues a good didactic material for the students	8	45,45	83,18	11,41

Tries to get an equilibrium situation among theory and practice	9	28,68	83,20	12,95
Has a good planning for the class	10	50,00	88,06	11,28
Follows the time schedule for the classes	11	56,25	89,47	10,03
Is always present for the classes	12	53,42	92,92	6,97
Utilizes evaluation procedures clear, objective and coherent	13	27,94	84,74	12,41
Bases on the evaluation results to reinforce the apprenticeship	14	41,18	80,57	11,20
Has a behavior that stimulate the apprenticeship	15	16,41	85,65	13,21
Knows the state-of-art of the discipline	16	42,42	90,96	9,63
Has a ethical posture	17	37,12	91,48	8,63
Utilizes the TIC as a didactic material	18	13,64	80,64	15,61

Trough the factor analysis concept it was possible to establish the correlation structure of the 18 variables and detect the strongly correlated ones. From this, it was also possible to create indices for explaining the variability among variables. The variables explaining factors are displayed on Table 02.

It is clearly seen, on Table 02, that almost 90% of the variability of the 18 variables can be explained by only 4 latent factors. And more: utilizing only two or three factors it is possible to get an explanation strength of 84,26% e 87,78%, respectively, what can be considered a very satisfactory result. Although the preliminaries results could be considered very good, the final decision about the number of latent factors will be taken considering also the eigenvalues position relative to latent factors, as displayed at Figure 02.

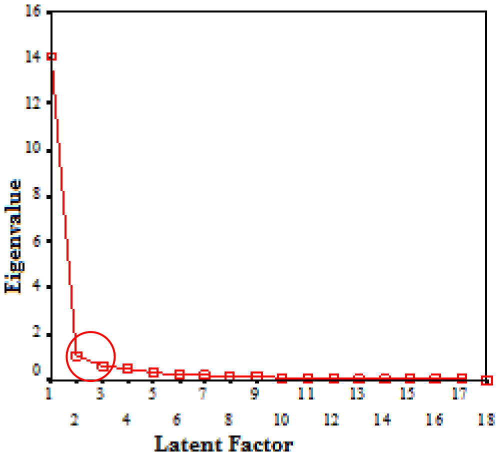

FIGURE 02
SCREEN PLOT FOR THE QUESTIONNAIRES

According to the Kaiser criterion [3] - it suggests that the eigenvalues less than one shall not be taken into consideration - it is possible to see that only two latent factor are enough and adequate for explaining the variability of the 18 variables, with an explanation strength of 84,26%.

TABLE 02
VARIANCE DESCRIPTION

Component	Eigenvalues		
	Total	% of Variance	% Cumulative
1	14,09	78,30	78,30
2	1,07	5,96	84,26
3	,63	3,52	87,78
4	,50	2,76	90,54
5	,32	1,77	92,31
6	,25	1,37	93,68
7	,21	1,18	94,86
8	,19	1,04	95,89
9	,14	,80	96,69
10	,12	,68	97,37
11	,11	,61	97,98
12	8,66E-02	,48	98,46
13	6,74E-02	,37	98,83
14	6,41E-02	,36	99,19
15	5,04E-02	,28	99,47
16	3,83E-02	,21	99,68
17	2,96E-02	,16	99,85
18	2,76E-02	,15	100,00

Method of extration: Analysis of the main components

For improving the latent factor interpretability, the component matrix was then rotated, according to the Varimax method [7] – Table 03. This make possible to see the association between the variables and the latent factors, as well the correlation among the original variables and their respective latent factor. It is clear that there is no perfect association, because one variable can be partially correlated with two factors.

TABLE 03
FACTOR MATRIX ROTATION

	Factor	
	1	2
5	.913	
9	.911	
15	.905	
6	.897	
7	.893	
2	.974	
3	.859	
13	.857	
16	.838	
14	.838	
1	.810	
8	.790	
4	.784	.517
17	.751	
10	.710	.542
11		.837
12		.808
18		.730

The variables were then plotted on a factorial plan [3], Figure 03, utilizing the factors strength as coordinates, for facilitating its interpretation. It is also expected that the factor axis be highly correlated with the different variable groups, what can be seen on Figure 04. Additionally, the factors themselves should have a low correlation, as displayed on Table 04.

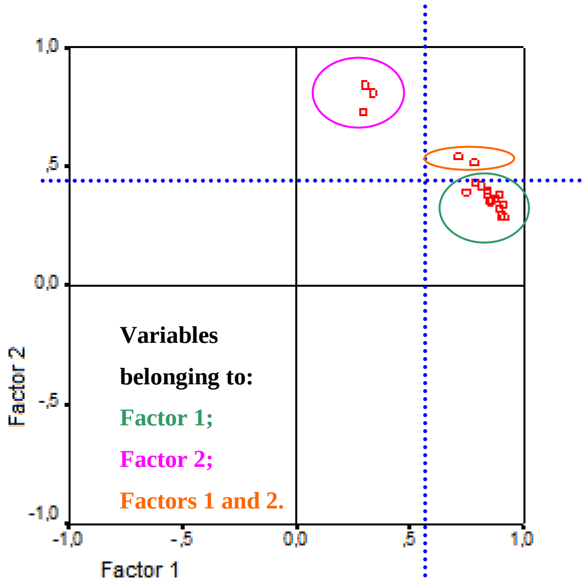

FIGURE 03
VARIABLES AND FACTORS CORRELATION

Analyzing the significance of each variable (questions related to each teacher) and their association with the latent factors, these could be interpreted as follows: the factor 1 is related with what could be called "class quality of the teacher" while the second factor could be associated with the enrollment of the professor, out of class, relative to his disciplines.

TABLE 04
FACTORS CORRELATION

		Factor 1	Factr 2
Factor 1	Pearson Correlation	1	,000
	Sig. (2-tailed)	,00	1,000
Factor 2	Pearson Correlation	,000	1
	Sig. (2-tailed)	1,000	,00

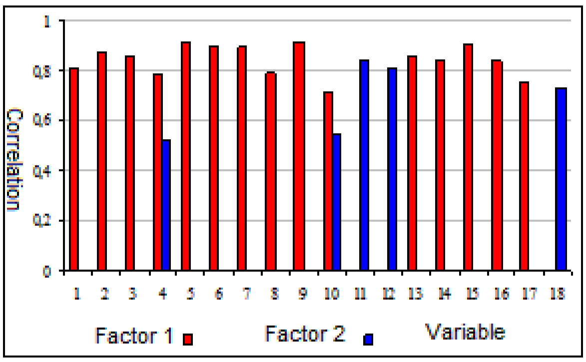

FIGURE 04
VARIABLES AND FACTORS CORRELATION

CLUSTER ANALYSIS

The cluster analysis methodology [5] was then utilized to identify cluster of professors with similar characteristics relative to their latent factors. As this technique is not a supervised one, the number of groups is not necessarily known in advance and it is necessary to utilize some techniques like the correlation between Euclidian distance and the number of clusters – Figure 05. For this analysis it was utilized the methodology "*Furthest neighbor*", since it can generate very bounded clusters, exactly in the way it is desired: maximum heterogeneity among the clusters and maximum homogeneity inside the cluster

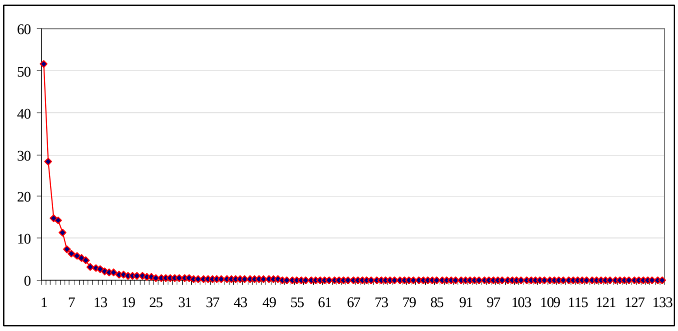

FIGURE 05
EUCLIDIAN DISTANCE X CLUSTER NUMBERS

There are two procedures for measuring the similarity among the elements. The first one works with an association measure: the greater the correlation coefficient, the greater will be the similarity. The second one utilizes an evaluation of the similarity trough the "proximity" among the elements, considering that shorts distances account for greater similarity.

Both the procedures – the number of elements in each cluster as well as the correlation between the Euclidian distance and the number of clusters – show that the professors should be divided into three groups. This is also clear on figure 05 where one can see that, when the number of groups equals three, the result is a steepest descent of the Euclidian distance. The final conclusion is that these three groups can give a good classification of the professors: the good evaluated, the medium evaluated and the bad evaluated, with their respective grades – Figure 06 and 07.

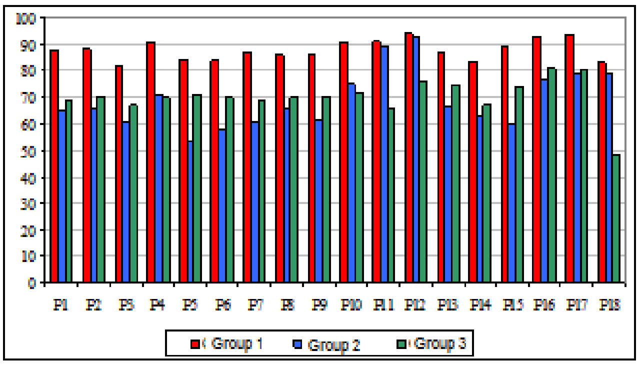

FIGURE 06
VARIABLES MEAN VALUE AND GROUP NUMBER

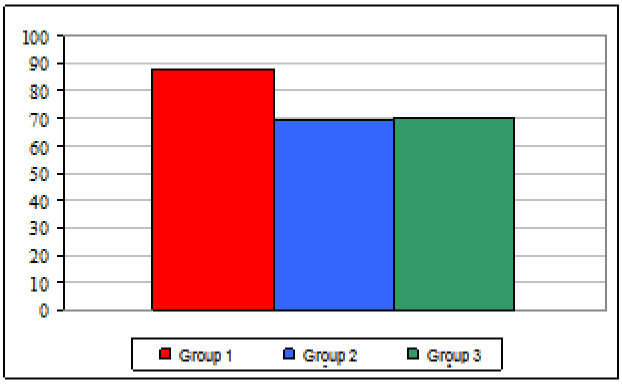

FIGURE 07
FINAL GROUP GRADES – 18 QUESTIONS

According to the Figure 05, 06 and 07, we can say that the first group comprises those professors with the greatest grades, that is, those with the best evaluation from the students. The group 2 comprises the professors whose evaluation grades were only median and, in the third group, that professors that got the worst evaluation from the students.

Another interesting aspect of this analysis comes out when the latent factors are confronted with the clusters. As displayed on Figure 04, the first group is composed by the professors with a reasonable evaluation on both factors, considering their behavior inside classroom as well their behavior out of class when dealing with the students. The second group, although showing a mean value very close to the third group, has a greater variance and includes the professors that had a bad evaluation concerning their behavior inside classroom, but had a good evaluation concerning their responsibility considering the subjects they teach. Finally, in the third group are the teachers that were bad evaluated when considering their didactics procedures, and the worst evaluation when considering the enrollment with the subject they are responsible for. The figure 08 displays, graphically, these results.

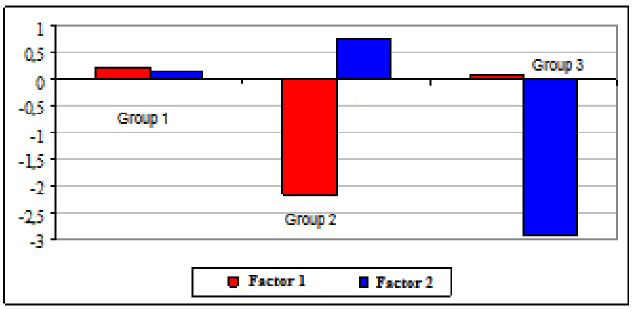

FIGURE 08
LATENT FACTOR DISTRIBUTION

RESULTS VALIDATION AND CONCLUSIONS

The procedures utilized suggested that are necessary only two questions for evaluating the professors. And more: it was also possible to conclude, according to the evaluation results that the professors can be divided into three groups. It is necessary, now, to check this procedure against another method to get a confirmation, as well as to apply the questionnaires with two questions to see the results.

The first procedure utilized was the PBM [8] method that comes from the names of the authors Pakhira, Bandyopadhyay e Maulik. The PBM index is defined as the product of three factors, whose maximization assures the best partition: a small number of compact groups, very distant one from another. Mathematically, the PBM index is defined as:

$$PBM(k) = \left(\frac{1}{k}.\frac{E_o}{E_k}.D_k\right)^2 \qquad (1)$$

where K is the number of groups, E_0 is the summation of the distance from each register until the geometric center of the data cluster w_0 defined as:

$$E_0 = \sum_{t-1..N} d(x(t), w_o) \qquad (2)$$

E_k is the summation of the intra-clusters distances for the K groups expressed by:

$$E_k = \sum_{t=1..N} \sum_{i=1..N} u_i(t) d(x(t), w_i)^2 \qquad (3)$$

The D_k value means the maximum separation among clusters:

$$D_k = \max(d(w_i, w_j)) \qquad (4)$$

The best partition will be given by the greatest value of BPM index. When this procedure was applied to the problem in focus, the best BPM index value was for a number of groups equal three, as suggested by the preceding analysis – Figure 09.

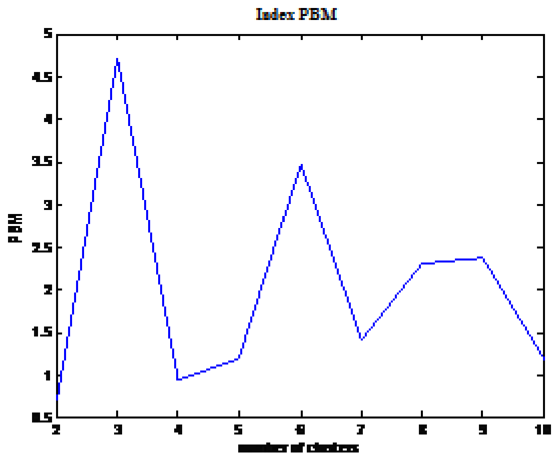

FIGURE 09
BMP OTIMIZATION PROCEDURE

One semester after the questionnaires were applied, a new evaluation was realized utilizing, in this situation, only the two proposed questions. These questions were:

TABLE 05
NEW QUESTIONS PROPOSED FOR THE EVALUATION

Questions: The professor ……
1) Explains the contents clearly and has a good didactic and pedagogic basis
2) Always answer the students demands, even out of classroom

Although there were small changes in the students, one can consider that the sample space had not changed significantly during this period of time, as well as the professor's behavior. The results were analyzed according to the proposal methodology and the professors were distributed among the clusters.

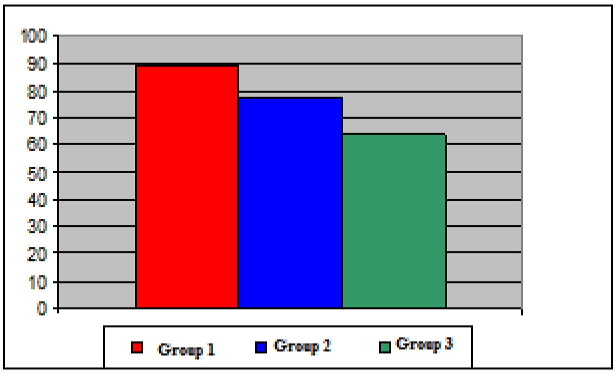

FIGURE 10
FINAL GROUP GRADES – 2 QUESTIONS

The final grades are displayed on Figure 10, where one can see that they had a little variation from the preceding ones, exactly on the same way when a questionnaire with 18 questions was utilized. And more: it is possible to see that, in spite of the small changes that have occurred during this period, the distribution of the professors among the clusters is practically the same that was obtained before, when the first questionnaire was utilized, as shown in Figure 03.

The results obtained gives rise to conclusion that some of the utilized evaluation procedures needs great improvement for facilitating their implementation. It is clear the difficulties associated with the work necessary for answering a questionnaire with 18 questions, each one with 5 options, for 4 or5 disciplines. The same results can be achieved with only two questions facilitating the whole process, which becomes more objective and focused. The results signal also that some educational process evaluation should be considered with more precaution because they can have a lack of appropriate support and technical basis.

REFERENCES

[1] Jonhson, Richard Arnold. "*Applied multivariate statistical analysis*" –5th ed, Prentice-Hall Ed., 2002, NY.

[2] DINIZ, Carlos Alberto R.; LOUZADA NETO, Francisco. "*Data mining*: *uma introdução"*. Associação Brasileira de Estatística, 2000, São Paulo.

[3] Grimm, L.G.; Yarnold, P.R. "*Reading and understanding multivariate statistics*". APA, 2000, Washington, DC.

[4] Kachigan, S.K. *"Multivariate statistical analysis"*. Radius Press, 1991, New York.

[5] Windham, M.P. "Cluster validity c-means for fuzzy clustering algorithm*", IEEE Transactions on Pattern Analyses Machine Intelligence"*, v.4, n.4. 1982, p.354-363.

[6] BUSSAB, Wilton de Oliveira; MIAZAKI, Édina Shizue; ANDRADE, Dalton Francisco, *"Introdução à análise de agrupamentos"*, Associação Brasileira de Estatística, São Paulo ,1990.

[7] *"SPSS 13.0 for Windows User Guide"*, Apache Software Foundation, Chicago, 2004.

[8] M. K. Pakhira, S. Bandyopadhyay, U. Maulik, "Validity index for crisp and fuzzy clusters", *Pattern Recognition*, vol 37, 2004, pp. 487-501.