Professor Georgeta VINTILĂ, PhD
E-mail: vintilageorgeta@yahoo.fr
Ştefan Cristian GHERGHINA, PhD Candidate
E-mail: stefan.gherghina@fin.ase.ro
Bucharest University of Economic Studies

PATTERN RECOGNITION TECHNIQUES TO CLASSIFY THE EUROPEAN EMERGING MARKETS COMPANIES FROM THE VALUATION PERSPECTIVE

Abstract. *In this research our aim is to classify a sample of companies which belong to five European emerging countries, respectively Hungary, Poland, Russia, Slovakia, and Ukraine, from the valuation perspective, by using pattern recognition techniques. The classification of the selected companies was realized according to ten indicators: the debt to equity ratio, the debt to total assets ratio, the financial debt to equity ratio, earnings per share, price/earnings ratio, return on sales, current ratio, quick ratio, return on assets, and return on equity. Thus, by employing Ward's method as agglomerative hierarchical clustering there resulted three classes of companies. Subsequently, we identified the discriminant functions based on which we could accomplish predictions regarding the companies' membership to the three previously established classes. The usefulness of both techniques in financial field is remarkable in order to set out the membership of certain objects to several classes, thus being taken the best decisions.*

Keywords: *cluster analysis, discriminant analysis, Ward's method, Fisher's linear discriminant, emerging markets, firm value.*

JEL Classification: C38, G32

1 Introduction

Many times, during the companies' valuation approach there come out the need to fit, to differentiate, to join, or to classify several companies in certain groups or classes whose delimitation should be very clear and very reasonably (Sumathi & Sivanandam, 2006). However, clustering and discrimination are the activities of organizing or joining the companies in certain groups, categories, or classes, depending on the rate of similarity or the rate of disparity between the companies (Hand et al., 2001). The classes are distinct informational entities which comprise all the companies similarly valued or the companies which are barely different as regards their value, but which are significantly different from the companies' value belonging to other groups. Thus, all the classification techniques, respectively clustering and discrimination, are also known as the pattern recognition theory.

The rest of this paper is organized as follows. In Section 2 we review the prior studies which used pattern recognition techniques in financial field, while Section 3 describes the data and research methodology, as well as the fundamentals of cluster analysis and discriminant analysis. The results of the empirical research are presented in Section 4, while Section 5 concludes the paper.

2 Literature Review

Da Costa Jr et al. (2005) employed cluster analysis in order to group the companies from North and South America according to the risk-return criterion. Sori & Jalil (2009) used discriminant analysis to develop a failure prediction model for Singaporean companies. Likewise, by using discriminant analysis, Oz et al. (2011) predicted the stock returns for 30 companies listed on the Istanbul Stock Exchange (ISE). Rashid & Abbas (2011) created a discriminant model with the aim of classification the companies from Pakistan according to their bankruptcy risk. As well, Armeanu et al. (2012) used cluster analysis and discriminant analysis in order to classify a sample of companies listed on the Bucharest Stock Exchange depending on their bankruptcy risk.

3 Data and Research Methodology

The aim of this empirical research consist in the determination of certain classes or groups of companies which belong to European emerging countries according to their value. However, we will employ pattern recognition techniques, respectively cluster analysis as unsupervized classification technique and discriminant analysis as supervized classification technique (Sharma, 1996). The distinction between both classification techniques is represented by the fact that within cluster analysis the classification of the companies is gradually achieved without knowing aprioristic the number of classes, while within the dicriminant analysis the number of classes is aprioristic known.

The initial sample comprised 310 companies from five European emerging countries as follows: 11 companies from Hungary, 125 companies from Poland, 89 companies from Russia, 5 companies from Slovakia, and 80 companies from Ukraine.

However, by considering the existence of several extreme values we decided to remove 62 obsevations from the initial sample (two companies from Hungary, 11 companies from Poland, 20 companies from Russia, 4 companies from Slovakia, and 25 companies from Ukraine), thus resulting a final sample consisting of 248 companies. The classification of the selected companies will be accomplished according to ten indicators. The indicators employed and their computation method will be described below. Also, the values of the indicators correspond to 2009. Besides, before to begin the empirical analysis, the data were standardized having the mean equal to one, respectively the variance equal to zero. The data were provided by ISI Emerging Markets. Also, we will use the software instrument SAS 9.2 in order to apply the pattern recognition techniques, according to Khattree & Naik (2000), Der & Everitt (2001), Delwiche & Slaughter (2008), and SAS INSTITUTE INC (2008).

I_1: DE = the debt to equity ratio, calculated by dividing total liabilities by stockholders' equity. Also known as global financial autonomy, this indicator assesses the size of external funds compared with the funds from shareholders. As the value of debt to equity ratio is higher, the business depends more on its creditors, respectively the related risk is higher, because all the liabilities from the balance sheet gives rights to third parties on the company. A higher debt to equity ratio involves a higher risk for creditors. The latter will take account for the current banking rules and for the regulations specific to this field of activity. Usually, a satisfactory value for the most activities is lower than 0.5. However, a lower debt to equity ratio shows the ability of the company to raise the size of its credits, under the reserve of existence of suitable cash flows which could allow the company to bear the future debt service;

I_2: DTA = the debt to total assets ratio. Also known as the general indebtedness ratio, this indicator reflects the means in which the company's assets are financed by debt. In ordinary activity conditions, the indebtedness level should be near 50 percent. A limit under 30 percent signifies a reserve in contracting credits and loans, while a value of the debt to total assets ratio over 80 percent supports evidence for a credit dependency, respectively an alarming scenario;

I_3: LEV = the financial debt to equity ratio, representing financial leverage ratio, through which is reflected the financial managers' ability to collect outside resources in order to stimulate the equity' efficiency;

I_4: EPS = earnings per share or the internal return of a certain share in terms of the income which is generated by that share in a financial year, is computed by dividing net income by total number of capital stock shares. This ratio allows the investors to compare the results recorded by the company in order to decide if the owned capital stock shares will be kept, cleared, or raised;

I_5: PER = price/earnings ratio, computed by dividing the company's current share price by its per-share earnings, is showing the market return of a certain share in terms of the amount which the investors are willing to pay per dollar of company's earnings. Also, this ratio shows the period required to a shareholder in order to recover the invested capital. Thus, price/earnings ratio reflects a proxy of investors' reliance in the company, in the sense that a higher value of this ratio indicates a higher level of investors' expectations towards the evolution of the company's earnings. The values of this indicator should be interpreted in the context of the companies from the same field of activity. Thereby, as PER is lower, the share is considered more interesting, thus the investors being advised to buy and hold that share. However, a lower value of PER could be associated to the companies characterized through risky businesses, while the companies characterized through better development perspectives could record a higher level of PER;

I_6: ROS = return on sales is the ratio of net income before interest and tax divided by net sales, usually presented in percent. On the one hand, ROS highlights the part of each dollar of sales that the company is able to turn into income. On the other hand, ROS shows the contribution of company's income in order to strengthen the self-financing ability of the company;

I_7: CR = current ratio is the ratio of current assets recorded in the balance sheet of a particular company for a given period of time to its current liabilities (short-term liabilities). This indicator reflects the possibility of current patrimonial elements to transform into liquidities in a short time in order to pay the current liabilities. CR is considered satisfactory for values between 1.2 and 1.9;

I_8: QR = quick ratio, also known the Acid-test ratio is calculated as the difference between current assets and inventory divided by current liabilities. This indicator reflects the possibility of current assets represented by accounts receivable and short-term investments to cover the current liabilities. QR is considered satisfactory for values between 0.65 and 1;

I_9: ROA = return on assets is computed by dividing the company's income after interest and tax by its total assets. This indicator shows the efficiency recorded in company's assets utilization;

I_{10}: ROE = return on equity is equal to net income divided by shareholders' equity. The contribution of shareholders in order to finance the company is measured through shareholders' equity, thus return on equity reflecting the efficiency of the company at generating profits from every unit of shareholders' equity. This indicator is specific to shareholders which assess based on the value of ROE if their investment is justified. Thus, the investors could decide to persist in supporting the company development through new paid in capital or they could waive for a limited period at due dividends. Nevertheless, the investors could not carry on to support the company development.

3.1 Cluster Analysis Description

We could divide the clustering methods in two categories according to their nature, operating manner, and the type of solution which it provide. Thus we distinguish hierarchical clustering methods and iterative clustering methods (Witten & Frank, 2005). Likewise, the hierarchical clustering algorithms could be divided in two categories, respectively agglomerative clustering algorithms and divisive clustering algorithms (Han & Kamber, 2006; Ruxanda, 2001). Within the current research we will employ Ward's method, as agglomerative hierarchical clustering. Thus, by the instrumentality of Ward's method, in each classification stage there are merged those two clusters for which the sum of squared deviations corresponding to the resulted cluster after merging is the smallest, compared with other pairs of clusters. However, the Ward's method evaluates the distance between two clusters as sum of the squared deviations for the cluster configuration resulted from merging the two clusters for which the distance is valued.

The sum of the squared deviations is defined as follows, where y_{ij} represent the jth object from the ith cluster, and n_i is the number of objects which belong to the ith cluster:

$$SSE = \sum_{i=1}^{K} \sum_{i=1}^{n_i} (y_{ij} - \overline{y_i})^2 \qquad (1)$$

In fact, in order to construct the hierarchical cluster tree there will be covered the following stages (Hastie et al., 2009):

Baseline, we will consider a number of clusters equal to the number of companies, respectively 248. Thus, there result the fact that each cluster consists of a single object:

$$\omega_1^{(1)} = \{C_1\}, \omega_2^{(1)} = \{C\}, \ldots, \omega_{247}^{(1)} = \{C_{247}\}, \omega_{248}^{(1)} = \{C_{248}\} \qquad (2)$$

Further, during several stages the initial clusters are gradually aggregated in order to obtain certain classes which are increasingly complex. Thus, in each stage which will be marked with t, there will be aggregated only two clusters, respectively those clusters for which the aggregation distance is minimal compared with the distances between any two clusters existing at that stage. However, the aggregation distance could be defined as below:

$$d_{aggregation}^{(t)} = min\left\{d_{\omega_i^{(t)},\omega_j^{(t)}}\right\} \quad (3)$$

Therefore, in the last stage of the aggregation process, all the objects are included in a single cluster:

$$\omega^{(247)} = \left\{C_1, C_2, \dots, C_{247}, C_{248}\right\} \quad (4)$$

3.2 Discriminant Analysis Description

The purpose of the discriminant analysis is to separate the prediction classes inside of Ω. Subsequently, there is required to establish the membership of new objects out of Ω to the K classes. However, there may be necessary to realize predictions regarding the membership of the new considered objects (Ruxanda, 2001). Thus, for each company, knowing the vector x which contains the values of the selected indicators, the aim is to determine the membership of every company to one out of the K possible classes from Ω.

Therefore, in order to achieve this purpose there will be formed the classification criteria according to which we will make predictions regarding the membership of new companies. However, the membership is initially unknown. The classification criteria are also known as classifiers. We mention the fact that the number of discriminant functions is determined by the number of descriptor variables and by the number of existent classes for the researched sample.

Likewise, usually, the discriminant functions are linear functions as follows:

$$d_i = \beta_1^{(i)} * x_1 + \beta_2^{(i)} * x_2 + \cdots + \beta_9^{(i)} * x_9 + \beta_{10}^{(i)} * x_{10} \quad (5)$$

Besides, the Fisher's linear discriminant functions which will be employed in this research are linear functions having the following form:

$$D(x) = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \dots + \beta_9 * x_9 + \beta_{10} * x_{10} \quad (6)$$

4 Empirical Research Results

4.1 The Results of the Cluster Analysis

Table 1 shows the cluster history for the last twenty generations out of a history of 248. The first column exhibits the number of clusters (NCL), while the second column displays the clusters joined. However, the observations are identified either through their identifier or through the number corresponding to the cluster. The column entitled 'FREQ' exhibits the number of observations in the new established cluster, while the column entitled 'SPRSQ' shows the value of Semipartial R-Squared

(*SPRSQ* = B_{KL}/T, in order to join the clusters K and L). Besides, the column entitled 'RSQ' displays the value of R-Squared (*RSQ* = *1* - P_G/T), respectively the proportion of the variance recorded by each cluster. Thus, if our selected sample of companies is classified in two clusters, the proportion of the variance incorporated by the clusters should be 16.5 percent. 'ERSQ' represents the Approximate Expected R-Squared, this expectation being approximated under the null hypothesis that the data have a uniform distribution instead of forming distinct clusters. 'CCC' describes the Cubic Clustering Criterion. 'PSF' represents the Pseudo-F Statistic. 'PST2' describes the Pseudo-T^2 Statistic. The criteria mentioned above are very useful in order to estimate the number of clusters.

Table 1 Cluster History

NCL	Clusters Joined		FREQ	SPRSQ	RSQ	ERSQ	CCC	PSF	PST2
20	**CL66**	**CL39**	5	0.0097	.766	.705	9.64	39.3	5.4
19	**CL26**	**CL34**	38	0.0111	.755	.698	8.78	39.2	12.1
18	**CL49**	**CL40**	14	0.0114	.743	.690	7.97	39.2	13.8
17	**CL52**	**CL25**	9	0.0128	.731	.683	7.03	39.2	4.8
16	**CL29**	**CL27**	32	0.0129	.718	.675	6.21	39.3	11.8
15	**CL60**	**NFI Octava**	6	0.0152	.702	.666	5.16	39.3	20.5
14	**CL18**	**CL110**	16	0.0153	.687	.656	4.25	39.5	9.8
13	**CL19**	**CL43**	70	0.0166	.671	.646	3.31	39.9	20.1
12	**CL23**	**CL36**	17	0.0212	.649	.635	1.75	39.7	14.5
11	**CL45**	**CL14**	25	0.0222	.627	.622	0.56	39.9	11.8
10	**CL22**	**CL16**	78	0.0280	.599	.609	-1.1	39.5	28.7
9	**CL21**	**CL12**	21	0.0287	.571	.593	-2.5	39.7	9.3
8	**CL24**	**CL15**	34	0.0288	.542	.575	-3.6	40.5	21.5
7	**CL13**	**CL30**	76	0.0290	.513	.554	-4.5	42.3	25.3
6	**CL17**	**CL10**	87	0.0523	.461	.529	-7.2	41.3	33.2
5	**CL8**	**CL20**	39	0.0602	.400	.498	-9.1	40.6	25.7
4	**CL11**	**CL9**	46	0.0623	.338	.456	-9.7	41.5	17.9
3	**CL7**	**CL6**	163	0.0640	.274	.393	-8.6	46.2	34.4
2	**CL5**	**CL3**	202	0.1088	.165	.264	-6.9	48.7	42.6
1	**CL2**	**CL4**	248	0.1653	.000	.000	0.00	.	48.7

Source: Author's computations using SAS 9.2

The criteria for the number of clusters are showed in Figure 1, respectively CCC, PSF, and PST2. Their computation method will be described below.

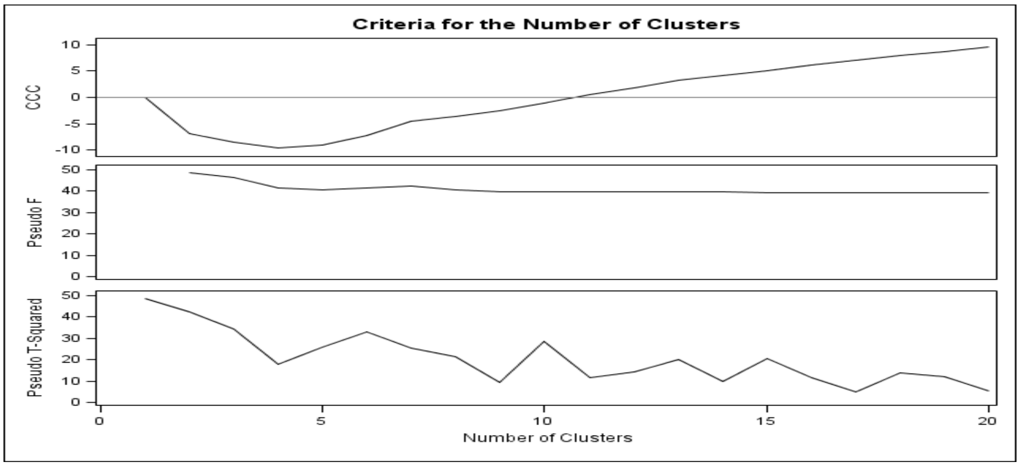

Figure 1 Criteria for the Number of Clusters
Source: Output SAS 9.2

The Cubic Clustering Criterion (CCC) was developed by SAS company (Sarle, 1983), while the performance of this criterion in order to estimate the number of clusters was examined by Milligan & Cooper (1985) and Cooper & Milligan (1988). This criterion represents a comparative metric of the deviations between the value of R^2 if the data were obtained from a uniform distribution and the actual value of R^2 for every number of clustering options. Thus, CCC is computed as follows, where $E(R^2)$ is the expected value of R^2, n is the number of observations, and p is an estimation of the dimensionality of the between-cluster variation:

$$CCC = ln\left[\frac{1-E(R^2)}{1-R^2}\right] * \frac{\sqrt{(np/2)}}{(0.001+E(R^2))^{1/2}} \quad (7)$$

The Pseudo-F statistic was suggested by Caliński & Harabasz (1974), being computed as follows, where G is the number of clusters, T is the total sum of squares, and P_G is the within-group sum of squares:

$$Pseudo\text{-}F = \frac{(T-P_G)/(G-1)}{P_G/(n-G)},\ T= \sum_{i=1}^{n}\|x_i - \bar{x}\|^2 \quad (8)$$

The Pseudo-T^2 statistic is computed as below, N being the number of observations from the clusters K and L, while C_M is the cluster M:

$$Pseudo\text{-}T^2 = \frac{B_{KL}}{(W_K+W_L)/(N_K+N_L-2)} \quad (9)$$

$$B_{KL} = W_M - W_K - W_L \text{ , if } C_M = C_K \cup C_L \tag{10}$$

$$W_K = \sum_{i \in C_K} \|x_i - \bar{x}_K\|^2, W_L = \sum_{i \in C_L} \|x_i - \bar{x}_L\|^2 \tag{11}$$

Besides, the Pseudo-T^2 statistic is associated with the statistic $J_e(2)/J_e(1)$, proposed by Duda & Hart (1973):

$$\frac{J_e(2)}{J_e(1)} = \frac{W_K + W_L}{W_M} = \frac{1}{1 + \frac{t^2}{N_K + N_L - 2}} \tag{12}$$

Therefore, we will use CCC criterion, Pseudo-F statistic, and Pseudo-T^2 statistic in order to establish the number of clusters. Thus, the peaks of the CCC plot, with values greater than two or three shows good levels of clustering. The peaks of the CCC plot with values between zero and two indicate possible levels of clustering, while negative values shows the possibility of outliers' existence. However, according to the CCC criterion we could set out 13 clusters, but we are considering the fact that this number of clusters is too high. Likewise, another method to ascertain the number of clusters is represented by the Pseudo-F statistic. Thus, the values relatively large of Pseudo-F statistic signifies a good number of clusters. Within current research, the Pseudo-F statistic implicates two clusters (Pseudo-F = 48.7) or three clusters (Pseudo-F = 46.2). The number of clusters according to Pseudo-T^2 statistic will be given by the significant changes of its value. Thus, we could distinguish possible levels of clustering at two clusters, three clusters, four clusters, five clusters, nine clusters, 15 clusters, 17 clusters, or 20 clusters. Taking into consideration the Pseudo-F statistic and Pseudo-T^2 statistic we will retain three clusters.

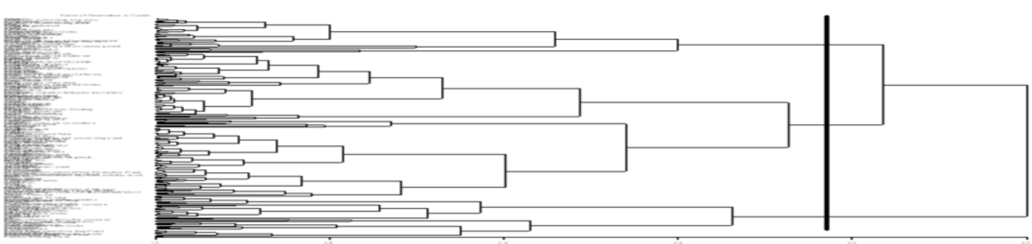

Figure 2 Tree Diagram of Clusters versus R-Square Values

Source: Output SAS 9.2

The hierarchical clustering could be represented through a bidimensional diagram entitled dendogram which illustrates the merger realized at each successive stage of the analysis. Thus, the dendogram showed in Figure 2 could support us in order to decide on the optimum number of classes which must be retained within the analysis through the occurrence of certain gaps (Spircu, 2005). A such gap is evidenced in Figure 2 through the black vertical line which intersect the dendogram in three points. Thereby, this fact implies the existence of three clusters: medium valued companies, lower valued companies, and highly valued companies.

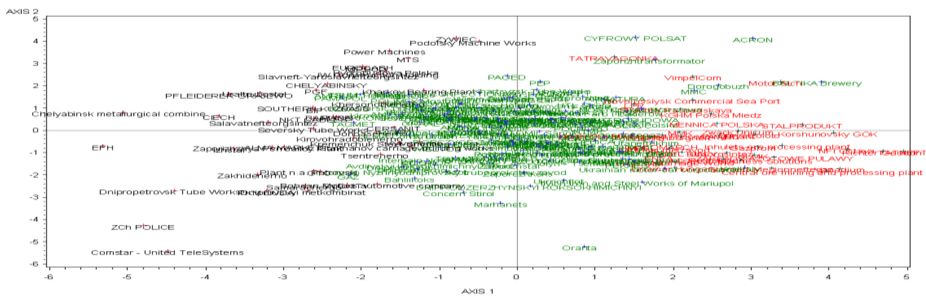

Figure 3 Graphical representation of the companies in the cluster plot
Source: Output SAS 9.2

Figure 3 and Figure 4 show the classification of the selected companies in the three classes according to their value. The companies from the first cluster (medium valued companies) are marked with blue colour, the companies from the second cluster (lower valued companies) are marked with red colour, and the companies from the third cluster (highly valued companies) are marked with green color.

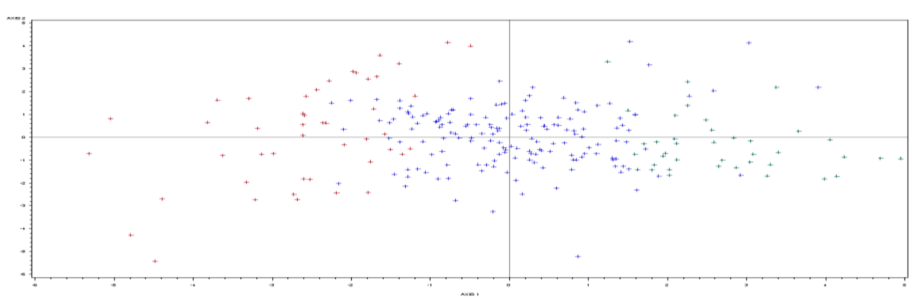

Figure 4 Scatter plot of clusters
Source: Output SAS 9.2

4.2 The Results of the Discriminant Analysis

Table 2 displays information regarding the frequency of each cluster out of the three clusters, their weight and their proportion, thus resulting 163 companies assigned in the first cluster (medium valued companies), 46 companies assigned in the second cluster (lower valued companies), and 39 companies assigned in the third cluster (highly valued companies).

Table 2 Class level information

CLUSTER	Variable Name	Frequency	Weight	Proportion
1	_1	163	163.0000	0.657258
2	_2	46	46.0000	0.185484
3	_3	39	39.0000	0.157258

Source: Author's computations using SAS 9.2

Table 3 shows the canonical correlations. However, the canonical correlation represents a metric of the association' degree between the score of discrimination and the membership at certain clusters.

Table 3 Canonical Correlation Analysis

Canonical Corr.	Adj. Canonical Corr.	Approx. Std. Err.	Squared Canonical Corr.	Eigenvalues of CanRsq/(1-CanRsq)			
				Eigenvalue	Diff.	Prop.	Cum.
0.8577	0.8513	0.0168	0.7357	2.7838	1.8737	0.7536	0.7536
0.6902	0.6802	0.0333	0.4764	0.9101		0.2464	1.0000
Test of H0: The canonical correlations in the current row and all that follow are zero			**Likelihood Ratio**	**Approx. F Value**	**Num D F**	**Den DF**	**Pr > F**
			0.1383	39.85	20	472	<.0001
			0.5235	23.97	9	237	<.0001

Source: Author's computations using SAS 9.2

The first canonical correlation (0.8577) is the highest possible multiple correlation which could be obtained by using a linear combination of quantitative variables. Likewise, Table 3 displays a test of probability ratio, being established the null hypothesis according to which the last q correlation ratios are zero. Besides, the squared canonical correlation between Can1 and the class variable (0.7357) is higher than the squared canonical correlation corresponding to Can2 (0.4764). The adjusted canonical correlations are less influenced by errors than raw canonical correlations. Approximate Standard Error represent approximations of standard errors for the canonical correlations. Squared Canonical Correlation are the square of the canonical correlation, and it could be interpreted similarly to the coefficient of determination denoted R^2 out of the OLS regression, respectively the proportion of the variance in the canonical variate of one set of variables explained by the canonical variate of the other set of variables.

The eigenvalues could be computed by using the squared canonical correlations. Thus, the highest eigenvalue is equal to the highest squared canonical corelation divided by 1- the highest squared canonical correlation, as following:

$$0.7357/(1-0.7357^2) = 1.6037$$

The column 'Difference' from Table 3 represents the difference between the given eigenvalue and the following highest eigenvalue:

$$2.7838 - 0.9101 = 1.8737$$

Likewise, the column 'Proportion' is the proportion out of the sum of eigenvalues, represented by an eigenvalue, while the column 'Cumulative' represents the cumulative sum of proportions.

According to the data from Table 4, we will define the following linear discriminant functions:

Table 4 Linear Discriminant Function for Cluster

	1	2	3
Constant	-17.23801	-22.41222	-20.22793
DE	-15.83324	-4.45666	-13.72033
DTA	93.74706	72.62429	75.30576
LEV	0.87795	0.60461	1.14422
EPS	-7.72364	-5.86363	26.51465
PER	2.21890	-0.72643	1.69524
ROS	-8.05217	-13.98001	-5.24113
CR	1.95915	1.31820	3.43058
QR	2.21210	2.10563	3.32770
ROA	65.05895	89.07578	103.67873
ROE	-21.75527	-42.50495	-38.49709

Source: Author's computations using SAS 9.2

$\hat{f}_1$ = -17.23801 - 15.83324*DE + 93.74706*DTA + 0.87795*LEV - 7.72364*EPS + 2.21890*PER -8.05217*ROS + 1.95915*CR + 2.21210*QR + 65.05895*ROA - 21.75527*ROE

$$\widehat{f_2} = -22.41222 - 4.45666*DE + 72.62429*DTA + 0.60461*LEV - 5.86363*EPS - 0.72643*PER - 13.98001*ROS + 1.31820*CR + 2.10563*QR + 89.07578*ROA - 42.50495*ROE$$

$$\widehat{f_3} = -20.22793 - 13.72033*DE + 75.30576*DTA + 1.14422*LEV - 26.51465*EPS - 1.69524*PER - 5.24113*ROS + 3.43058*CR + 3.32770*QR + 103.67873*ROA - 38.49709*ROE$$

From Table 5 we could notice that for the first company out of the sample, the probability of membership in the first cluster (0.9733) is higher than the probability of membership in the second cluster (0.0261) or than the probability of membership in the third cluster (0.0005). Thus, the first selected company belongs to the first cluster. Also, for the third company out of the sample, the probability of membership in the second cluster (0.9816) is higher than the probability of membership in the first cluster (0.0184) or than the probability of membership in the third cluster (0.0000). Further, the third selected company belongs to the second cluster.

Table 5 Posterior probability of membership in cluster for the first ten companies

Obs	From CLUSTER	Classified into CLUSTER	1	2	3
1	**1**	**1**	0.9733	0.0261	0.0005
2	**1**	**1**	0.9670	0.0326	0.0005
3	**2**	**2**	0.0184	0.9816	0.0000
4	**2**	**2**	0.0369	0.9631	0.0000
5	**1**	**1**	0.9853	0.0001	0.0146
6	**1**	**1**	0.9886	0.0001	0.0113
7	**1**	**1**	0.9956	0.0013	0.0031
8	**1**	**1**	0.9963	0.0011	0.0025
9	**1**	**1**	0.9877	0.0116	0.0007
10	**1**	**1**	0.9935	0.0056	0.0008

Source: Author's computations using SAS 9.2

Table 6 highlights a table of classification. However, the values from the diagonal of the classification' table reflects the right classification of the companies in clusters, based on the scores related to the discriminant dimension.

Table 6 Number of observations and percent classified into cluster

From CLUSTER	1	2	3	Total
1	148 90.80	5 3.07	10 6.13	163 100.00
2	2 4.35	44 95.65	0 .00	46 100.00
3	4 10.26	0 0.00	35 89.74	39 100.00
Total	154 62.10	49 19.76	45 18.15	248 100.00
Priors	0.33333	0.33333	0.33333	
Error Count Estimates for CLUSTER				
	1	**2**	**3**	**Total**
Rate	0.0920	0.0435	0.1026	0.0794
Priors	0.3333	0.3333	0.3333	

Source: Author's computations using SAS 9.2

Therefore, 148 companies which belong to the first cluster were allocated rightly, while five companies out of the first cluster were allocated wrong in the second cluster, and ten companies out of the first cluster were allocated wrong in the third cluster. Likewise, 44 companies which belong to the second cluster were assigned correctly, while two companies out of the second cluster were assigned wrong in the first cluster. As regards the third cluster, 35 companies were distributed rightly, while four companies out of the third cluster were distributed wrong in the first cluster. Besides, the total probability to allocate wrong a company in a certain cluster is 7.94 percent. The probability to distribute wrong a company which belongs to the first cluster in another cluster is 9.20 percent. The probability to allocate wrong a company out of the second cluster in the first cluster is 4.35 percent, while the probability to assign wrong a company which belongs to the third cluster in another cluster is 10.26 percent.

Further, taking into consideration the classifier represented by the identified linear discriminant functions, our aim is to predict the membership to the three established clusters of two new companies, Umpo and Gostomel Glass Plant.

The values of the indicators corresponding to both companies are showed in Table 7.

Table 7 New companies in order to predict their membership in cluster

New Selected Companies → Indicator ↓	Umpo	Gostomel Glass Plant
DE	3.9448	6.2447
DTA	0.7978	0.8587
LEV	4.9448	0.138
EPS	0.0001	0.0191
PER	0.3577	0.0034
ROS	0.005	0.07
CR	2.4402	2.1827
QR	1.5959	1.2713
ROA	0.0031	0.0149
ROE	0.0154	0.1078

Source: ISI Emerging Markets

For the company Umpo we distinguish the fact that it belongs to the second class of companies (lower valued companies), considering the following criterion: $max(\hat{f}_1^{Umpo}, \hat{f}_2^{Umpo}, \hat{f}_3^{Umpo})$

$\hat{f}_1$ = -17.23801 - 15.83324*3.9448 + 93.74706*0.7978 + 0.87795*4.9448 - 7.72364*0.0001 + 2.21890*0.3577 - 8.05217*0.005 + 1.95915*2.4402 + 2.21210*1.5959 + 65.05895*0.0031 - 21.75527*0.0154 = 8.366043599

$\hat{f}_2$ = -22.41222 - 4.45666*3.9448 + 72.62429*0.7978 + 0.60461*4.9448 - 5.86363*0.0001 - 0.72643*0.3577 -13.98001*0.005 + 1.31820*2.4402 + 2.10563*1.5959 + 89.07578*0.0031 - 42.50495*0.0154 = 26.80475654

$\hat{f}_3$ = -20.22793 -13.72033*3.9448 + 75.30576*0.7978 + 1.14422*4.9448 - 26.51465*0.0001 - 1.69524*0.3577 - 5.24113*0.005 +3.43058*2.4402 + 3.32770*1.5959 + 103.67873*0.0031 - 38.49709*0.0154 = 5.378346386

Likewise, for the company Gostomel Glass Plant, $\hat{f}_1$ = -30.48107579, $\hat{f}_2$ = 13.40944148, $\hat{f}_3$ = -31.825743, thus resulting the fact that this firm belongs to the second cluster of companies (lower valued companies).

5 Concluding Remarks

By selecting a sample of companies which belong to five European emerging countries and by employing cluster analysis we have realized a classification of the selected companies in three classes according to their value, respectively medium valued companies, lower valued companies, and highly valued companies. In addition, by applying Ward's method as agglomerative hierarchical clustering, there resulted the fact that most of the companies were classified in the medium valued companies class. Subsequently, by employing discriminant analysis there were identified the discriminant functions. Therefore, by the instrumentality of the identified discriminant functions we could accomplish predictions regarding the membership of new European emerging markets' companies to the three previously established classes pursuant to their value.

REFERENCES

[1] **Armeanu, Ș. D., Vintilă, G., Moscalu, M., Filipescu, M. O., Lazăr, P. (2012),** ***Using Quantitative Data Analysis Techniques for Bankruptcy Risk Estimation for Corporations***; *Theoretical and Applied Economics*, 19(1), 86-102;
[2] **Caliński, T., Harabasz, J. (1974),** ***A Dendrite Method for Cluster Analysis***; *Communications in Statistics*, 3(1), 1-27;
[3] **Cooper, M. C., Milligan, G. W. (1988),** ***The Effect of Error on Determining the Number of Clusters***; published in *Data, Expert Knowledge and Decisions: An Interdisciplinary Approach with Emphasis on Marketing Applications*; Springer, 319-328;
[4] **Da Costa Jr, N., Cunha, J., Da Silva, S., (2005),** ***Stock Selection Based on Cluster Analysis***; *Economics Bulletin*, 13(1), 1-9;
[5] **Delwiche, L. D., Slaughter, S. J. (2008),** ***The Little SAS Book: A Primer***, 4th Edition; Cary, NC: SAS Institute Inc.;
[6] **Der, G., Everitt, B. S. (2001),** ***A Handbook of Statistical Analyses Using SAS***, 2nd Edition; Boca Raton, FL: Chapman & Hall/CRC;
[7] **Duda, R. O., Hart, P. E. (1973),** ***Pattern Classification and Scene Analysis***; New York: John Wiley & Sons;
[8] **Han, J., Kamber, M. (2006),** ***Data Mining: Concepts and Techniques***, 2nd Edition; San Francisco: Morgan Kaufmann Publishers;
[9] **Hand, D. J., Mannila, H., Smyth, P. (2001),** ***Principles of Data Mining***; Cambridge, MA and London, England: The MIT Press: A Bradford Book;

[10] **Hastie, T., Tibshirani, R., Friedman, J. (2009),** ***The Elements of Statistical Learning. Data Mining, Inference and Prediction***, 2nd Edition; Springer;
[11] **Khattree, D. R., Naik, D. N. (2000),** ***Multivariate Data Reduction and Discrimination with SAS Software***; NC: SAS Institute Inc.;
[12] **Milligan, G. W., Cooper, M. C. (1985),** ***An Examination of Procedures for Determining the Number of Clusters in a Data Set***; *Psychometrika*, 50(2), 159-179;
[13] **Oz, B., Ayricay, Y., Kalkan, G. (2011),** ***Predicting Stock Returns with Financial Ratios: A Discriminant Analysis Application on the ISE 30 Index Stocks***; *Anadolu University Journal of Social Sciences*, 11(3), 51-64;
[14] **Rashid, A., Abbas, Q. (2011),** ***Predicting Bankruptcy in Pakistan***; *Theoretical and Applied Economics*, 18(9), 103-128;
[15] **Ruxanda, G. (2001),** ***Analiza datelor***; Bucharest: ASE Publishing House;
[16] **Sarle, W. S. (1983),** ***Cubic Clustering Criterion***; SAS Technical Report A-108, Cary, NC: SAS Institute Inc.;
[17] **SAS INSTITUTE INC (2008),** ***SAS/STAT 9.2 User's Guide***; Cary, NC: SAS Institute Inc.;
[18] **Sharma, S. (1996),** ***Applied Multivariate Techniques***; John Wiley & Sons, Inc.;
[19] **Sori, Z. M., Jalil, H. A. (2009),** ***Financial Ratios, Discriminant Analysis and the Prediction of Corporate Distress***; *Journal of Money, Investment and Banking*, 11, 1-15;
[20] **Spircu, L. (2005),** ***Analiza Datelor: Aplicații Economice***; Bucharest: ASE Publishing House;
[21] **Sumathi, S., Sivanandam, S. N. (2006),** ***Introduction to Data Mining and its Applications***; Springer;
[22] **Witten, I. H., Frank, E. (2005),** ***Data Mining. Practical Machine Learning Tools and Techniques***, 2nd Edition; San Francisco: Morgan Kaufmann Publishers.