



**Studia i Materiały. Miscellanea Oeconomicae**  
Rok 18, Nr 4/2014  
Wydział Zarządzania i Administracji  
Uniwersytetu Jana Kochanowskiego w Kielcach

**Gospodarka a rozwój regionalny**

**Robert Kowal, Yaroslav I. Yeleyko, Vasylyna I. Kharkhalis<sup>1</sup>**

## **APPLICATION OF CLUSTER ANALYSIS TO ASSESS THE REGIONAL DEVELOPMENT OF FOREIGN BANKING IN UKRAINE**

### **1. Introduction**

The development of the banking system and the economy of Ukraine is in close interdependence of globalization processes. Under the influence of global financial system and the movement of capitals between countries in the domestic banking sector there is the consolidation of bank capitals, mergers and acquisitions of banks with foreign capital, and the concentration of banking institutions in the regions where there is a possibility of obtaining the greatest profitability from their activities. At the same time the regional distribution of banking institutions, the distribution of their profits, and attachments in regional economic development have a significant impact on the prospects of the economy development as a whole [1].

A detailed analysis of the role of banks with foreign capital at the regional level stays away. Thus, the purpose of this work is the development of the classification of Ukrainian regions from the positions of operations realization and the analysis of the total foreign banking institutions that operate in this country, for the efficient analysis of incomes and expenses using the methods of multivariate statistical analysis including the methods of cluster analysis. This will help to develop the ways of domestic banking business development in the regional aspect on the basis of analysis of banks with foreign capital.

Cluster analysis is a set of techniques which allows to classify multivariate observations, each of which is described by a set of output variables. Unlike the combinational groups, cluster analysis leads to splitting into groups regarding all

---

<sup>1</sup> Dr Robert Kowal, wykładowca, Politechnika Świętokrzyska w Kielcach; prof. zw. dr hab. Yaroslav I. Yeleyko, profesor, Uniwersytet Jana Kochanowskiego w Kielcach, Lwowski Uniwersytet Narodowy im. Iwana Franki, mgr Vasylyna I. Kharkhalis, asystent, Lwowski Uniwersytet Narodowy im. Iwana Franki.

group-formed signs simultaneously. It should be noted that different clustering methods generate different clustering solutions for the same data. Therefore, appropriate methods should be pre-defined depending on the input data and the desired results.

## 2. The main material

For further analysis one should develop clustering of regions of Ukraine with the most active lending activities of foreign banks and which are the most profitable for banks with foreign capital by the average weighted calculated values using the method of tree (hierarchical) clustering (the essence of this algorithm is to combine objects in large enough clusters, using some measure of similarity or distance between objects. Distance in this case can be Euclidean distance:

$d_{ij}(x, y) = \sqrt{\sum_{k=1}^n (x_{ik} - y_{jk})^2}$ , where  $d_{ij}(x, y)$  is a distance between objects

$i$  and  $j$ ,  $x_{ik}$  is value of  $k^{\text{th}}$  variable for the  $i^{\text{th}}$  object,  $y_{jk}$  is value of  $k^{\text{th}}$  variable for the  $j^{\text{th}}$  object) and the method of k-means (a researcher specifies in advance the number of  $k$  clusters and the clustering method allows to find these clusters so that they are maximally different from each other).

For the analysis 27 regions of Ukraine were selected by sample of clustering, as a clustering criterion we selected number of bank divisions (of the banking system, banks with foreign capital, banks with 100% foreign capital), population (the number of population, population density), the average parameter of foreign banks activity in the region (average volume of assets, average volume of liabilities, the average income (loss)).

All the calculations are carried out according to the data of National Bank of Ukraine and the State Committee of Statistics of Ukraine as of 01.01.2013 [3; 4].

The analysis of the activity of banks with foreign capital by regions of Ukraine was conducted using the software package for statistical analysis STATISTICA 8.0.

Thus, for conducting clustering we have a eight-measured space of outgoing signs for 27 regions of Ukraine (Table 1).

### 1) The method of tree (hierarchical) clustering.

The purpose of this algorithm is to combine objects in large enough clusters, using some measure of similarity or distance between objects. A typical result of this clustering is a hierarchical tree. The results of all hierarchical procedures are usually processed as a so-called dendrogram.

This method is the most illustrative and understandable, and is used when forming of clusters.

The result of cluster analysis is a visual representation of the structure of clusters of regions that form the classifications created by combining them into larger clusters using a certain measure of similarity. Thus, using this method we received the hierarchical structure of clusters of regions shown in dendrogram in Figure 1.

Table 1. Baseline data for Cluster Analysis.

	1 The whole amount of banks	2 Banks with foreign capital	3 Banks with 100% foreign capital	4 Population	5 The population density	6 The average volume of assets	7 The average volume of co mmittments	8 Averaged amount of income (loss)
Autonomy Republic of Crimea	1037	530	37	1965,2	75,3	33287,43	29189,55	231,97
the city of Sevastopol	213	101	16	383,4	426	8518,22	7397,49	64,84
Wynnytsia	564	277	32	1627	61,4	17095,3	14858,54	81,96
Volyn	357	188	23	1040	51,74	12195,97	10580,82	37,66
Dnipropetrovs'k	1507	810	152	3307,8	103,69	56483,43	49225,77	295,27
Donets'k	1679	762	139	4375,4	165,11	61155,34	52969,88	288,17
Zhytomyr	358	210	43	1268,9	42,58	14550,48	12557,71	50,19
Transcarpatian	464	198	27	1254,4	98	12871,09	11120,06	61,06
Zaporizhia	729	377	54	1785,2	65,63	29282,36	25303,52	169,59
Ivano-Frankivsk	455	223	13	1381,8	99,41	13883,71	12027,6	64,23
Kiev	764	208	39	1722,1	61,28	15244,22	13193,36	56,67
the city of Kiev	2073	900	254	2845	3556,25	79766,86	68125,17	188,52
Kirovograd	363	175	15	995,2	40,46	11548,09	10065,8	55,44
Luhans'k	768	512	42	2256,5	84,51	32215,2	28124,15	157,57
Lviv	1143	490	59	2540,7	116,55	31657,64	27212,9	127,66
Mykolayiv	582	234	34	1173,5	47,7	16171,1	14027,97	73,22
Odesa	1256	480	69	2395,2	71,93	32775,66	28259,43	146,22
Poltava	739	294	37	1467,8	50,97	19574,2	16963,02	83,4
Rivne	378	167	13	1156,9	57,56	10785,8	9395,21	32,08
Sumy	467	214	22	1143,2	48,03	13397,25	11644,22	42,55
Ternopil	288	172	13	1077,3	78,07	12597,87	10898,53	34,03
Kharkiv	1209	640	100	2744,4	87,4	40644,9	34823,58	188,99
Kherson	438	245	24	1078,2	37,83	15277,97	13243,7	54,64
Khmel'nyts'k	426	195	21	1314	63,79	13989,43	12145,73	43,16
Cherkasy	692	208	18	1268,9	60,71	13798,08	12049,27	67,55
Chernivtsi	385	159	12	907,2	112	9407,3	8167,97	36,55
Chernihiv	526	161	17	1077,8	33,79	10888,27	9476,46	21,81

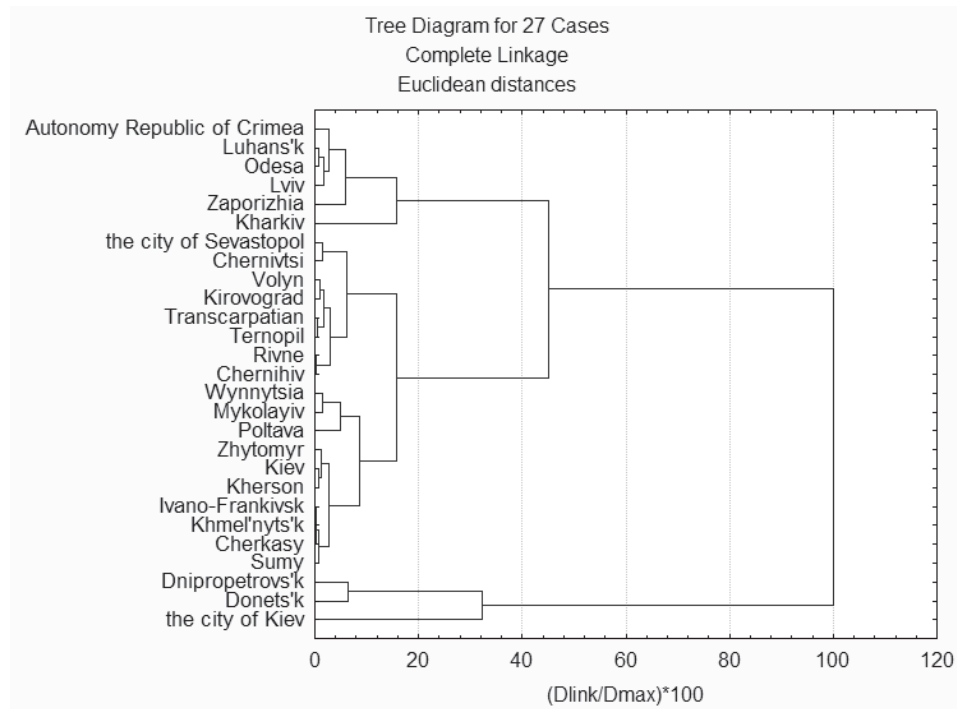


Figure1. Hierarchical cluster tree.

Along the vertical axis there are the areas that are subject to classification, and along the horizontal axis there is a postponed distance between combined objects.

According to the results of clustering using tree clustering method the following 3 groups should be allocated (Table 2).

Table 2. Results of cluster analysis.

Number of cluster	Name of region
1 cluster	Dnipropetrovs'k, Donetsk, the city of Kiev
2 cluster	Autonomy Republic of Crimea, Zaporizhia, Luhans'k, Lviv, Odesa, Kharkiv
3 cluster	the city of Sevastopol, Wynytsia, Volyn, Zhytomyr, Transcarpatian, Ivano-Frankivsk, Kiev, Kirovograd, Mykolayiv, Poltava, Rivne, Sumy, Ternopil, Kherson, Khmel'nyts'k, Cherkasy, Chernivtsi, Chernihiv

Regions with the highest credit activity of the banks with foreign capital belong to cluster 1. This is explained, in our opinion, by the greatest number of people in these regions, the highest population density and the concentration of large banks with foreign capital in these regions as the possibility of consumption of banking services. Accordingly, cluster 2 includes regions with lower volumes of assets, liabilities and earnings (losses) of banking institutions. The 3<sup>rd</sup> cluster corresponds to the regions of Ukraine which are less profitable for banks with foreign capital.

2) K-means method is used only when a researcher has already priori hypotheses regarding the number of clusters.

A researcher specifies in advance the number of k clusters and the clustering method allows to find these clusters so that they are maximally differed from each other. The problem is solved by breaking down n objects into k ( $k < n$ ) of homogeneous, in a certain sense, clusters. At the initial stage of its implementation the original objects are arranged and the first k objects are subsequently treated as separate clusters, which provide individual weighting coefficients. Then a researcher takes the  $(k + 1)^{th}$  object and finds out which of the existing cluster is the closest. This new cluster is replaced, located in the center weight of the baseline cluster and  $(k + 1)^{th}$  object. If the  $(k + 1)^{th}$  object point is equidistant from several clusters, it is placed to the cluster with the lowest number. Then the objects that are in turn are added to the existing cluster [3].

The advantage of this method is the possibility to test the statistical significance of the differences between the selected clusters.

Having the analysis of variance table (Table 3) according to this method we received value intergroup (Between SS) and intra group (Within SS) dispersion characteristics. The smaller is the value of intra group variance and more important intergroup variance, the better is the feature characterizing regions belonging to the cluster and the better is the clustering. The value of F-test (the higher value means the better) and the significance level (p-value) indicate the quality of clustering.

Table 3. Analysis of variance.

Variable	Analysis of Variance					
	Between SS	df	Within SS	df	F	signif. p
The whole amount of banks	4,873514E+06	2	810783	24	72,1305	0,000000
Banks with foreign capital	1,210422E+06	2	80475	24	180,4913	0,000000
Banks with 100% foreign capital	6,545874E+04	2	12173	24	64,5279	0,000000
Population	1,661031E+07	2	3325429	24	59,9393	0,000000
The population density	3,794222E+06	2	7943363	24	5,7319	0,009228
The average volume of assets	7,769248E+09	2	504827800	24	184,6788	0,000000
The average volume of commitments	5,770334E+09	2	347208100	24	199,4309	0,000000
Averaged amount of income (loss)	1,429204E+05	2	19063	24	89,9690	0,000000

Let us repeat the clustering statistics from Table 1 by the method of k-means using Euclidean distance. This method can be called "analysis of variance on the contrary" in the sense that the criterion of significance in the analysis of variance compares intergroup variability of intra group when checking the hypothesis that the average group differ from each other. Objects are moved from one cluster to another in order to obtain the most meaningful results at carrying out analysis of variance in this clustering method.

Let us use the results of the clustering tree. We set the number of clusters in advance, that we will classify the regions into three groups: high, medium and low levels of profitability of banks with foreign capital.

Table 4. Results of cluster analysis.

Number of cluster	The number of regions in the cluster	The area included in the cluster	The value of the cluster
1	3	Dnipropetrovs'k, Donets'k, the city of Kiev	Regions with high levels of dispersion (average values of variance 4287.1773)
2	6	Autonomy Republic of Crimea, Zaporizhia, Luhans'k, Lviv, Odesa, Kharkiv	Regions with an average level of dispersion (average values of variance 1168.3373)
3	18	the city of Sevastopol, Wynnytsia, Volyn, Zhytomyr, Transcarpatian, Ivano-Frankivsk, Kiev, Kirovograd, Mykolayiv, Poltava, Rivne, Sumy, Ternopil, Kherson, Khmel'nyts'k, Cherkasy, Chernivtsi, Chernihiv	Regions with low dispersion (average values of variance 977.97644)

As we can see from Table 4, the results of the cluster analysis by two methods are the same, that is, all areas that are in clusters 1, 2 and 3 using the method of k-means are in the first, second and third hierarchical cluster method, as well.

Graphical interpretation of the obtained results criteria that characterize the regions for each cluster is depicted on a linear graph (Figure 2):

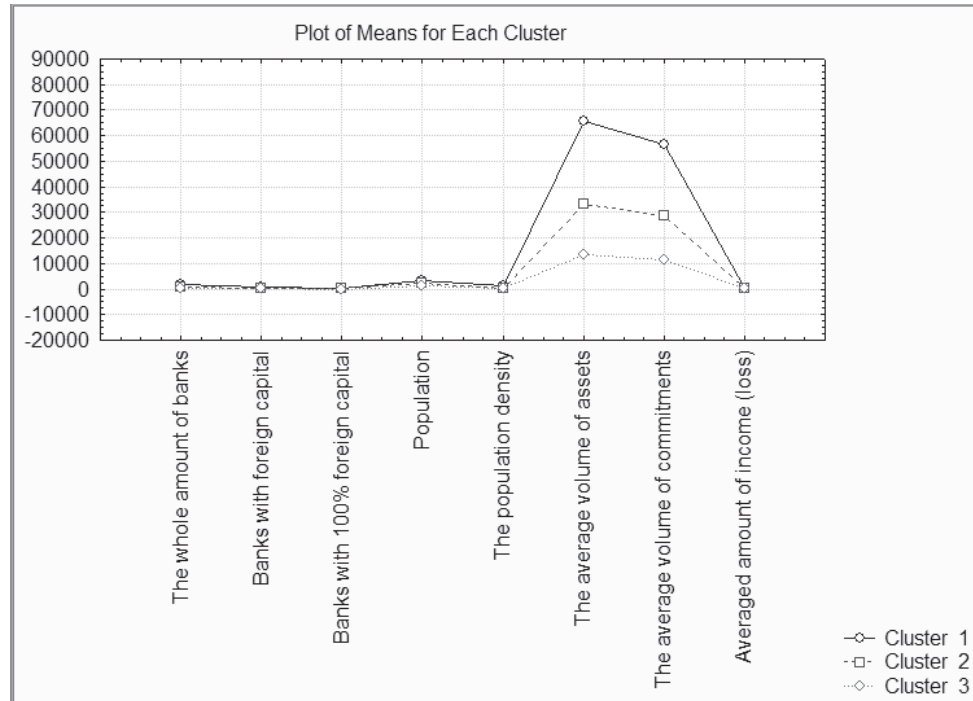


Figure 2. Cluster analysis using k-means.

After analyzing the results of the clustering of regions where there are banks with foreign capital by various methods, we can conclude that although the grouping by different methods gives similar results, the method of k-means is the most accurate as the values of between group and intra group variance are better by this method.

### 3. Conclusions

Application of hierarchical classification method allows getting dendrograms cluster areas. It enables to establish the structure of the objects inside the cluster and interrelation cluster among themselves that allows taking management decisions for homogeneous groups of regions.

Application of the theory of cluster analysis provides an opportunity to choose the extent of determining the distance between objects and rules of cluster association among themselves.

Therefore the resulting distribution of regions by such methods provides an opportunity to work out ways of increasing the efficiency of the domestic banking

business. For example, concerning 3 cluster expansion of regional foreign banking will allow promoting quality of banking services in the remotest corners of the country and ensuring the expansion of banking services.

It is possible to tell from the results that the leaders among the regions of Ukraine by the number, profitability and other characteristics of banks with foreign capital are the city of Kyiv, Donetsk and Dnipropetrovs'k regions, which predominate over other regions in terms of population in the region.

This means that the modern methods of cluster analysis can be used not only for comparing statistical data, but also for detecting problem situations that is why cluster analysis as an effective way of classification of objects by their attributes performed well in all spheres of social activity, especially in the economics .

### **Literature**

1. Vladychn U., *Regional peculiarities of foreign banking in Ukraine* / Scientific journal "The regional economy" - 2013. - № 2. - P. 158-167.
2. Stavisky G. *Distribution of banks in Ukraine by means of methods of cluster analysis.* – 2013 – p. 1-3.
3. *Basic indicators of banks in Ukraine as of 01.01.2013* / The National Bank of Ukraine. - Mode of access: <http://www.bank.gov.ua>.
4. *State Statistics Committee of Ukraine* — Mode of access: <http://www.ukrstat.gov.ua/>

### **Abstract**

Clustering of Ukrainian regions with the most active credit activities of foreign banks that are the most favorable for the development of banks with foreign capital by the average weighted calculated values was developed and illustrated for 27 regions of Ukraine, using two methods of cluster analysis, namely: a method of tree (hierarchical) clustering and a k-means method.

### **Zastosowanie analizy klastrow do regionalnej oceny stopnia rozwoju bankowości zagranicznej na Ukrainie**

Na Ukrainie w stosunkowo dużym tempie rozwija się działalność banków z kapitałem zagranicznym. Wywierają one coraz większy wpływ na rozwój gospodarczy poszczególnych regionów tego kraju. W pracy dokonano klasteryzacji 27 Regionów Ukrainy ze względu na największą aktywność kredytową banków zagranicznych oraz stopnia ich rozwoju. Wykorzystano do tego hierarchiczną metodę drzewa klastrow oraz metodę k-średnich.

PhD Robert Kowal, lecturer, Kielce University of Technology.

Professor Yaroslav I. Yeleyko, professor, University of Jan Kochanowski Kielce, Ivan Franko National University of Lviv,

MBA Vasylyna I. Kharkhalis, junior lecturer, Ivan Franko National University of Lviv.