Cluster Analysis of Charitable Organizations of Ukraine using K-Means Technology

OLHA VYSOCHAN, OLEH VYSOCHAN, VOLODYMYR CHUBAI, VASYL HYK
Department of Accounting and Analysis
Lviv Polytechnic National University
Bandery St. 12, Lviv
UKRAINE

Abstract: The work is devoted to the issue of segmentation of charitable organizations for structuring the sector of non-profit organizations of Ukraine using cluster analysis tools using software R for automated data processing. The four-cluster and five-cluster models were constructed using the K-means method, the suitability for clustering of which was checked using the Hopkins 'Index (H statistics). The developed four-cluster model demonstrated a significant level of validity in terms of correspondence between data and the stability of their structure. The basic indicators of financial and economic activity of charitable organizations were used as criteria for clustering: the number of staff, charitable assistance received and funds spent on the maintenance of the organization in the reporting period. It was found that the clusters of charitable organizations of Ukraine differ in the scale of activity, the number of funds raised, the number of costs for their own maintenance and the relationship between these indicators. The study demonstrated the existence in Ukraine of the most influential cluster of local charities that address social issues exclusively at the regional level, due to the small financial resources involved to support their activities.

Key-Words: non-governmental organizations, charitable organizations, cluster analysis, classification, segmentation

Received: May 9, 2021. Revised: December 19, 2021. Accepted: January 25, 2022. Published: February 8, 2022.

1 Introduction

The dynamic development of the non-governmental sector of Ukraine testifies to the importance of non-governmental non-profit organizations (NGOs) in the effective functioning of the market economic mechanism. Such organizations are designed to help solve social problems and implement national and local projects of an innovative nature.

Non-governmental organizations should not assume the functions of the state but contribute, within reasonable limits, to the fullest and most effective implementation of the latter. At the same time, like commercial enterprises, they have more opportunities for innovative development in response to the demands of a new type of post-industrial information economy.

The functional and organizational diversity of organizations of the non-profit sector of the economy creates the need for thorough research and financial and economic analysis of their activities, the formation of classification features, their positioning in the network of institutional actors of the state economy. The main indicators that characterize the activities of non-budgetary non-profit organizations and can serve as a basis for the allocation of their individual segments (clusters) are: received charitable assistance; maintenance costs; the number of employees.

The lack of a standardized approach to the species classification of NGOs is often due to the difficulty of implementing scientifically sound means of dividing the set of such organizations into groups according to statistically significant estimates.

2 Literature Review

Clustering is a means of organizing and dividing a set of objects into groups according to a certain feature (or set of features) that can be effectively used to classify and segment charitable organizations.

Generalizations of scientific approaches to the classification of NGOs are presented in table 1.

Table 1. Types of NGOs and the specifics of their activities

Source	Types NGOs	Characteristic
[15]	NGOs with the charitable orientation	They are mainly involved in meeting the food, clothing, shelter and education needs of the poor, as well as in dealing with the effects of natural disasters such as floods and earthquakes
	NGOs with service orientation	They focus on education, health and family planning services
	NGOs with participatory orientation	They use an approach in the form of involving the local population in activities through self-help projects
	NGOs with empowerment	The main goal is to help the poor take control of their own lives by raising their awareness of social, political and economic issues
[11]; [12]	Grassroots support organizations (GSO)	A social development body that provides related support services to local groups of rural or urban households and disadvantaged individuals
	Membership support organizations (MSO)	They also provide services to local groups, however, the MSO represents and is accountable to its core participant, at least in principle
[17]; [10]; [18]; [7]; [22]; [23]; [27]; [21]; [29]	Northern NGOs	Organizations of richer countries, which, calling for organizational and social reforms, work with decision-makers to stimulate such reforms, on a top-down basis
	Southern NGOs	Organizations of poorer countries that use the driving forces of the masses to carry out fundamental social transformations, using a bottom-up approach
[39]; [36]; [28]	Operational:	The primary goal is to develop and implement development projects
	- community-based	They serve the community in a specific geographical area
	- national	Operate in specific developing countries
	- international	Headquartered in developed countries but operating in more than one developing country
	Advocacy	The primary goal is to protect and promote clearly defined cases
[33]	Professional (National and International)	They consist of people with professional experience, skills and special expertise
	Social (National and International)	More related to social and humanitarian activities
[34] for environmental NGOs (ENGOs)	ENGOs as watchdogs	Monitor the implementation of existing environmental regulations, participate in various hearings, conduct legal actions against environmentally harmful projects or campaigns
	ENGOs as value perceivers	Promote environmental values not established by existing legislation, raise public awareness
	ENGOs as field actors and action coordinators	Provide environmental public goods, provide landscape support, coordinate field projects in cooperation with local communities, monitor biodiversity
	ENGOs as knowledge transmitters	Educate stakeholders, cooperate with research organizations, advise on land use issues
	ENGOs as partners in collaborative governance	Influence the implementation of state policy, solve environmental problems and implement projects, organize control over the use of resources and payments for ecosystem services

Source: author's generalization

Thus, the problem raised in the article is the structuring and typification of charitable organizations operating in the non-profit sector of Ukraine (allocation of statistically significant clusters) to identify their most influential segment, based on key performance indicators: charitable assistance received; maintenance costs; the number of employees. The charitable assistance received is the main performance indicator in terms of income, while the number of employees and maintenance costs of the organization characterize the cost part of the activity.

3 Methodology

Cluster analysis is a statistical method of classification. Unlike other statistical classification methods, such as discriminant analysis and automatic interaction detection, it does not make preliminary assumptions about important population differences [30]. This method has been successfully implemented in our other studies [37; 38].

Based on this, the step-by-step implementation of the cluster analysis technique involves the successive implementation of five stages.

Stage 1. Data preparation and standardization.

The implementation of the cluster analysis methodology begins with obtaining data from the available sources and preliminary preparation (generalization, verification, presentation in a tabular editor) for clustering.

Different units of measurement of the characteristics of charitable organizations (natural – for the number of employees; monetary – for the received charitable assistance and expenses for the maintenance of the organization) require unification using the standardization procedure. Standardization is carried out in order to bring the estimated numerical values of the variables that characterize the object of study to a single scale (it is necessary to achieve the same units of measurement or set a dimensionless value for all variables), to correctly interpret the results.

There are a number of approaches to standardizing variables. Social scientists typically assume that a standardized variable is transformed with zero mean value and unit variance, as established by the typical "Z-score" formula [26].

Z-score is a form of standardization used to convert normal variants into a standard form with a score. Based on the input set Y, the standardization formula for Z-estimation is defined as follows:

$$Z(x_{ij}) = \frac{x_{ij}-\bar{x}_j}{\delta_j} \quad (1)$$

where, $\bar{x}_j$– the average value of the j-attribute sample;

δ_j – standard deviation of j-attribute.

The converted variable will have a mean value of 0 and variance 1 [8].

Stage 2. Assessment of suitability for clustering.

A typical distinguishing feature of most clustering algorithms is that they form the clustering structure of the data set X, even if X cannot have essentially any subgroup. In the case where X has a low tendency to divide into subgroups, the results obtained after the application of the clustering algorithm are not real data substructures. The problem of checking whether X has a tendency to cluster (structure clustering) without clear identification is known as a tendency to cluster [20].

One of the possible techniques for assessing the trend towards clustering is the use of a class of remote methods. One of the most powerful distance-based methods is the Hopkins' test [16; 6; 2].

Hopkins' statistic compares the distances between a data set and its nearest neighbors with the distance between a set of pseudo-data that is randomly selected from a complete data set and their nearest neighbors. The technique is effective for small clusters [3].

Hopkins statistic H described by the equation:

$$H = \frac{\sum_{i=1}^{p} w_i}{\sum_{i=1}^{p} u_i + \sum_{i=1}^{p} w_i}) \quad (2)$$

where, p – values that are randomly distributed in the data range of the original set;

w_i – the distance from the original data set to the nearest neighbor of the value sample;

u_i – the distance from the original data set to the nearest neighbor of the artificially generated values.

If the randomly generated values and the sample values from the original data have approximately the same distances to the nearest neighbor, then H will be approximately 0.5. A value of H that is about 1 or 0 indicates high quality data clustering [35].

Step 3. Set the number of data clusters.

There are many indexes for setting the optimal number of clusters for the selected data set. The possibility of implementing most of them is provided in statistical software packages, such as R (Table 2).

Table 2. Overview of the indices implemented in the NbClust package for R

Name of the index in NbClust	Optimal number of clusters
kl	Maximum value of the index
ch	Maximum value of the index
hartigan	Maximum difference between hierarchy levels of the index
ccc	Maximum value of the index
scott	Maximum difference between hierarchy levels of the index
marriot	Maximum value of second differences between levels of the index
trcovw	Maximum difference between hierarchy levels of the index
tracew	Maximum value of second differences between levels
friedman	Maximum difference between hierarchy levels of the index
rubin	Minimum value of second differences between levels
cindex	Minimum value of the index
db	Minimum value of the index

silhouette	Maximum value of the index
duda	Smallest number of clusters such that index > criticalValue
pseudot2	Smallest number of clusters such that index < criticalValue
beale	Number of clusters such that critical value >= alpha
ratkowsky	Maximum value of the index
ball	Maximum difference between hierarchy levels of the index
ptbiserial	Maximum value of the index
frey	Cluster level before index value < 1.00
mcclain	Minimum value of the index
dunn	Maximum value of the index
hubert	Graphical method
sdindex	Minimum value of the index
dindex	Graphical method
sdbw	Minimum value of the index

Source: [13].

Stage 4. Direct clustering

The most widely used optimization criterion for clustering is the clustering error criterion, which for each point calculates its square distance from the corresponding center of the cluster, and then summarizes these distances for all points in the data set. A popular clustering method that minimizes clustering error is the K-mean algorithm [24], which belongs to Partition-based methods [31] and has demonstrated its practical effectiveness in many cases [4]. The K-mean clustering technique is described in [25; 19; 4; 9; 1; 40; 5].

If we have a set of m data points $X = \{x_i | i = 1, \ldots, m\}$, where each of them is an n-dimensional vector, the K-means clustering algorithm allows to divide m data points into k clusters C={c_1,c_2,…,c_k }, in order to minimize the objective function J (V, X) of dissimilarity, which is the intracluster sum of squares. The objective function J based on the Euclidean distance between the vector of the data point xi in the cluster j and the corresponding center of the cluster vj is defined as:

$$J(X,V) = \sum_{j=1}^{k} J_i(x_i, v_j) = \quad \sum_{j=1}^{k} (\sum_{i=1}^{m} u_{ij} . d^2(x_i, v_j)) \tag{3}$$

where, $J_i(x_i, v_j) = \sum_{i=1}^{m} u_{ij} . d^2(x_i, v_j)$ is the target function within the cluster c_i, $u_{ij} = 1$, if $x_i \in c_j$, in another case $u_{ij} = 1$.

$d^2(x_i, v_j)$ is the distance between xi and vj:

$$d^2(x_i, v_j) = \left\| \sum_{k=1}^{n} x_k^i - v_k^j \right\|^2 \tag{4}$$

where, n – the number of measurements of each data point;

x_k^i – the value k-measurement xi;

v_k^j – the value k-measurement vj.

Distributed clusters are defined by $m \times k$ binary membership matrix U, in which the element uij is equal to 1, if the data point xi belongs to the cluster j and 0 – otherwise. Once the centers of the cluster $V = \{v_1, v_2, \ldots, v_k\}$ are fixed, the membership function uij, which minimizes (3), can be obtained as follows:

$$u_{ij} = \begin{cases} 1; if\ d^2(x_i, v_j) \le d^2(x_i, v_{j^*}) j \ne j^*, \forall\ j^* = 1, \ldots, k \\ 0; in\ another\ case \end{cases} \tag{5}$$

Once the membership matrix $U = [u_{ij}]$ is fixed, the optimal center vi, which minimizes (3), is the average value of all vectors of data points in the cluster j. It can be calculated using:

$$v_j = \frac{1}{|c_j|} \sum_{i, x_i \in c_j}^{m} x_i \tag{6}$$

where, $|c_j|$ – cluster size cj, $|c_j| = \sum_{i=1}^{m} u_{ij}$;

Given the initial set of k means or centers of the cluster, $V = \{v_1, v_2, \ldots, v_k\}$, the algorithm is performed by alternating two steps: 1) purpose – assigned to each data point of the cluster with the nearest center; 2) update – the center of the cluster is updated as the average of all data points in this cluster.

Step 5. Validation of clusters.

Despite the same data set, different clustering algorithms can potentially generate very different clusters [41; 32]. Validation makes it possible to answer the question of the acceptability of the configuration of the clusters obtained as a result of the analysis, to solve the tasks.

One approach to validating clusters is to use internal criteria. It makes it possible to evaluate the results of the clustering algorithm using information that includes the vectors of the data sets themselves. The use of specialized software facilitates calculations and allows you to present the results of the analysis in a graphical and understandable informative form. An excellent alternative to many commercial software products in this area is the freely distributable R software environment, which is a dynamically evolving general-purpose statistical platform [14]. In the future, the software environment R will be used by us for statistical processing of data on the activities of charitable organizations in Ukraine.

4 Results and Discussion

The results of normalization of data on the number of employees, charitable assistance received and maintenance costs in terms of individual charitable organizations of Ukraine are presented in table 3.

Table 3. Normalized values of the characteristics of charitable organizations of Ukraine

ID	Name of the charitable organization	Normalized values (dimensionless units):		
		number of employees	charitable assistance received	expenses for the maintenance of the organization
1	CO “Bright kids”	0.0149	0.000860	0.002810
2	CF “Kvitna”	0.0299	0.0109	0.0405
3	CF “Borys Kolesnikov Foundation”	0.0896	0.135	0.151
4	CO “Ukrainian forum of philanthropists”	0.0149	0.00948	0.0666
5	CO “Berezani Community Foundation”	0	0.00245	0.000670
6	CO “CF “Community unity”	0.0149	0.00205	0.00624
7	CO “CF “Svichado”	0.0448	0.0160	0.0253
8	CO “Nechitaylo family foundation”	0.0746	0.0390	0.0594
9	CO ICF “Everyone can”	0.0448	0.00626	0.0745
10	CF “Blagomay”	0.0448	0.0317	0.0346
11	CF “Pediatricians against cancer”	0.0299	0.00396	0.0203
12	A-UCF “Down syndrome”	0.104	0.0184	0.0265
13	WBF “Depol Ukraine”	1	0.0507	0.203
14	A-UCF “Association of Philanthropists of Ukraine”	0.0448	0	0
15	ICF “Caritas Ukraine”	0.955	1	1
16	CO ICF “Ukrainian Charity Exchange”	0.164	0.242	0.640
17	ICF “Life with a surplus”	0.104	0.0328	0.0902
18	ICF “Mission to Ukraine”	0.657	0.0602	0.303
19	CO “CF “Old people”	0.0448	0.00581	0.0289
20	CO “Zahoriy family foundation”	0.104	0.0441	0.109

Source: data of Reports on the use of income (profits) of the non-profit organization, provided on the website of the Ukrainian Forum of Philanthropists (https://rating.ufb.org.ua/) in free access standardized using the normalize function. Abbreviations: CO – charitable organization; CF – charitable foundation; A-UCF – All-Ukrainian Charitable Foundation; ICF – international charitable foundation

To visualize normalized data of large volumes, we use the popular graphical method “Cluster heat map” (Fig. 1).

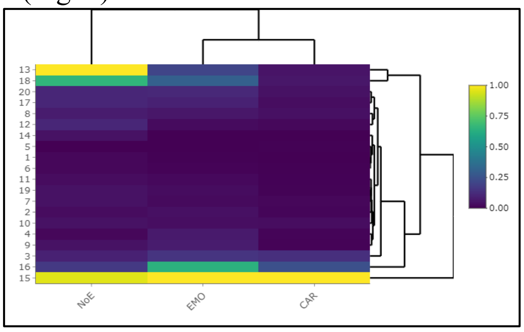

Fig. 1: Thermogram of normalized values of characteristics of activity of charitable organizations of Ukraine

NoE – number of full-time employees, standardized;

EMO – the amount of charitable assistance received, standardized;

CAR – the amount of expenses for the maintenance of the organization (administrative costs), standardized.

Source: data visualization using the heatmaply function of the R software environment

A cursory analysis of heat maps allows us to identify charitable organizations that could potentially serve as basic organizations for data segments: ICF “Caritas Ukraine” (ID 15) for all characteristics; WBF “Depol Ukraine” (ID 13) and

ICF “Mission to Ukraine” (ID 18) by the number of full-time employees and CO ICF “Ukrainian Charity Exchange” (ID 16) by the received charitable assistance. It is clear that ICF “Caritas Ukraine” (ID 15), which has the maximum value for all characteristics will form a separate cluster of charitable organizations in Ukraine.

To assess the suitability for clustering in the software environment R provides a function get_clust_tendency library factoextra. In relation to our task, the indicator H is calculated at the level of 0.8846554 (for 5 points that are randomly distributed in the data range of the original set), which is considered an acceptable value of the quality level of clustering.

Thus, using the rule of simple majority, it was found that the optimal number of clusters for the segmentation of charitable organizations of Ukraine – 4 (Table 4).

Table 4. Positioning methods for establishing the optimal number of clusters

Method of determination	The optimal number of clusters					
	2	3	4	5	6	8
Maximum value of the index	–	ratkowsky	kl, ccc, silhouette, ptbiserial, dunn	ch	–	–
Maximum difference between hierarchy levels of the index	–	scott, trcovw, ball	hartigan	–	–	friedman
Maximum value of second differences between levels of the index	–	marriot, tracew	–	–	–	–
Number of clusters such that critical value >= alpha	beale	–	–	–	–	–
Smallest number of clusters such that index > criticalValue	duda	–	–	–	–	–
Smallest number of clusters such that index < criticalValue	–	–	pseudot2	–	–	–
Minimum value of the index	cindex	–	db, mcclain, sdindex	–	sdbw	–
Minimum value of second differences between levels	–	–	rubin	–	–	–
Cluster level before index value < 1.001	–	–	–	–	–	–
Graphical method	–	–	hubert, dindex	–	–	–
Total	3	6	13	1	1	1

Note: failed to obtain reliable data to establish the optimal number of clusters by the frey method
Source: summarized by the authors

The optimality of the four-cluster division of charitable organizations of Ukraine is confirmed both by methods focused on maximizing the assessment (kl, ccc, silhouette, ptbiserial, dunn) and on its minimization (db, mcclain, sdindex), as well as graphic methods (hubert, dindex).

In the four-cluster model of segmentation of charitable organizations (Fig. 2) identified: cluster 1 (WBF “Depol Ukraine” – ID 13 and ICF “Mission to Ukraine” – ID 18); cluster 2 (ICF “Caritas Ukraine” – ID 15); cluster 3 (CO ICF “Ukrainian Charity Exchange” – ID 16); cluster 4 (other charitable organizations of Ukraine).

In the five-cluster segmentation model (Fig. 3) in a separate group from the cluster “Others” are: CF “Borys Kolesnikov Foundation” (ID 3), ICF “Life with a surplus” (ID 17) and CO “Zahoriy family foundation” (ID 20).

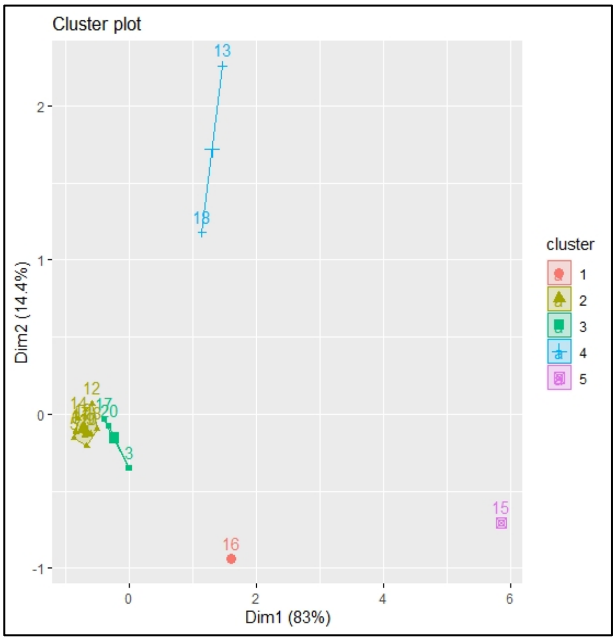

Fig. 2: Visual representation of clusters of charitable organizations of Ukraine (number of clusters – 5)

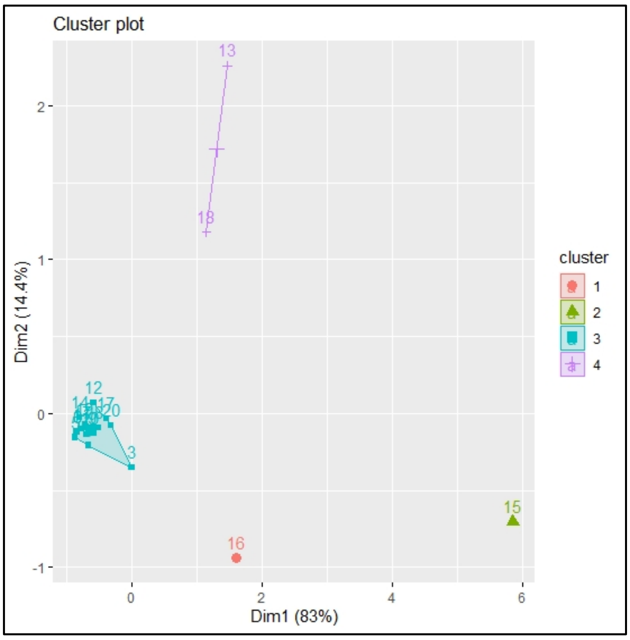

Fig. 3: Visual representation of clusters of charitable organizations of Ukraine (number of clusters – 4)

Source: data visualization using the fviz_cluster function of the R software environment

Assessment of the validity of the created clusters of charitable organizations of Ukraine requires the calculation of a number of indices, summarized in table 5.

Table 5. Assessment of the validity of clusters of charitable organizations of Ukraine

Evaluation indicator	Indicator value		The model for which the indicator is the best	
	for a four-cluster model	for the five-cluster model	four-cluster	five-cluster
Internal measures				
Connectivity	9.7159	11.7159	+	
Silhouette coefficient	0.7598	0.6916	+	
Dunn's index	1.6493	1.3303	+	
Stability measures				
Average proportion of non-overlap (APN)	0.0479	0.0896	+	
Average distance (AD)	0.4025	0.3176		+
Average distance between means (ADM)	0.1713	0.1448		+
Figure of merit (FOM)	0.4286	0.3860		+

Source: generalized by the authors based on calculations using the clValid function of the R software environment

We see that the four-cluster model demonstrates high validity, surpassing the fifth-cluster model in terms of internal assessment and slightly behind the three indicators of stability assessment (AD, ADM and FOM).

5 Conclusions

The cluster analysis showed the existence of 4 fairly clear clusters into which charitable organizations of Ukraine can be united.

The first two clusters (ICF "Caritas Ukraine" and CO ICF "Ukrainian Charity Exchange") are effective charitable organizations that have a good structure and ratio between borrowed and spent funds and, while the first, having significant financial resources and permanently implementing joint projects with various government institutions, have relatively low flexibility in decision-making, others seek to compensate for average funding capacity, speed and consistency in responding to external challenges. The third cluster (WBF "Depol Ukraine" and ICF "Mission to Ukraine") includes organizations that do not fully use the available capacity to attract funding, and also need to improve the existing ratio between borrowed funds and money spent on their own needs. The fourth cluster (CO "Bright kids", CF "Kvitna", CF "Borys Kolesnikov Foundation", CO "Ukrainian forum of philanthropists", CO "Berezani Community Foundation", CO "CF "Community unity", CO "CF "Svichado", CO "Nechitaylo family foundation", CO ICF "Everyone can", CF "Blagomay", CF "Pediatricians against cancer", A-UCF "Down syndrome", A-UCF "Association of Philanthropists of Ukraine", ICF "Life with a surplus", CO "CF "Old people", CO "Zahoriy family foundation") includes a fairly wide range of small charitable organizations, mainly at the regional level, which have limited influence on the formation of state social policy, but are characterized by proximity to the final recipients. This cluster requires further division into smaller segments to establish the effectiveness of their activities. Such organizations are system-creating for the entire non-profit sector in Ukraine, their importance is manifested in the most rapid response to the needs of recipients through the implementation of small charitable projects. Their development in recent years has been caused by a violent volunteer movement related to hostilities in eastern Ukraine. 4th cluster organizations also require additional government support.

Limitations in our study are: the relatively small number of Ukrainian charities that currently provide official reporting in the public domain; frequent cases of improper attitude to the preparation of reports by charitable organizations on the use of income (profits) of a non-profit organization, which are the basic source of information for the implementation of the model presented in the study; The study deals exclusively with economic indicators of charitable organizations, which do not always correlate with the social purpose of their creation.

References:

[1] Abdul Nazeer, K. A. and Sebastian, M. P. (2009). Improving the accuracy and efficiency of the k-means clustering algorithm. WCE 2009: Proceedings of the *World Congress of Engineering*, London, United Kingdom, July 1-3.

[2] Adolfsson A., Ackerman M. and Brownstein N. C. (2017). To cluster, or not to cluster: How to answer the question. TKDD 2017: Proceedings of *Knowledge Discovery from Data*, Halifax, Nova Scotia, Canada, August 13–17.

[3] Adolfsson, A., Ackerman, M. and Brownstein, N. C. (2019). To cluster, or not to cluster: An analysis of clusterability methods. *Pattern Recognition*, 88, 13-26. https://doi.org/10.1016/j.patcog.2018.10.026

[4] Alsabti, K., Ranka, S. and Singh, V. (1997). An efficient k-means clustering algorithm. *Electrical Engineering and Computer Science*, 43, 1-6.

[5] Ansari, Z., Azeem, M. F., Ahmed, W. and Babu, A. V. (2011). Quantitative evaluation of performance and validity indices for clustering the web navigational sessions. *World of Computer Science and Information Technology Journal (WCSIT)*, 1 (5), 217-226.

[6] Banerjee, A. and Loius, S. J. (2007). A recursing clustering methodology using a genetic algorithm. CEC 2007: Conference proceedings of *IEEE Congress on Evolutionary Computation*, Singapore, September 25-28.

[7] Bebbington, A. and Riddell, R. (1995). The direct funding of Southern NGOs by donors: New agendas and old problems. *Journal of International Development*, 7 (6), 879-893. https://doi.org/10.1002/jid.3380070607

[8] Bin Mohamad, I. and Usman, D. (2013). Standardization and its effects on K-means clustering algorithm. *Research Journal of Applied Sciences, Engineering and Technology*, 6 (17), 3299-3303. http://dx.doi.org/10.19026/rjaset.6.3638

[9] Bradley, P. S. and Fayyad, U. M. (1998). Refining initial points for K-means clustering. ICML 1998: Proceedings of the 15th International Conference on *Machine Learning*, Madison, WI, USA, July 24-27.

[10] Brown, D. L. (1993). Social change through collective reflection with Asian nongovernmental development organizations. *Human Relations*, 46 (2), 249-273. https://doi.org/10.1177/001872679304600206

[11] Carroll, T. F. (1992). *Intermediary NGOs: The supporting link in grassroots development*, Kumarian Press, West Hartford, CT, USA.

[12] Carroll, T., Schmidt, M. and Bebbington, T. (1996). Participation through intermediary NGOs. In: *Social Development Papers*, 12, The World Bank, Washington, DC, USA.

[13] Charrad, M., Ghazzali, N., Boiteau, V. and Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistics Software*, 61 (6), 1-36. https://doi.org/10.18637/jss.v061.i06

[14] Cheshkova, A. F., Aleynikov, A. F. and Stepochkin, P. I. (2016). Application of graphical features of the R programming environment for analysis of experimental data on the breeding of triticale. *Computational Technologies*, 21 (1), 104-115.

[15] Cousins, W. (1991). Non-governmental initiatives. In: Report on the Seminar *The Urban Poor and Basic Infrastructure Services in Asia and the Pacific*, Asian Development Bank, Manila, Philippines, 22-28 January, 83-112.

[16] Cross, G. R. and Jain, A. K. (1982). Measurement of clustering tendency. *IFAC Theory and Application of Digital Control*, 15 (1), 315-320. https://doi.org/10.1016/S1474-6670(17)63365-2.

[17] Fowler, A. (1991). Building partnerships between Northern and Southern development NGOs: issues for the 1990s. *Development in Practice*, 1 (1), 5-18. https://doi.org/10.1080/096145249100076011.

[18] Fowler, A. and James, R. (1994). The role of Southern NGOs in development co-operation. *INTRAC Occasional Papers Series*, 2, Oxford, United Kingdom.

[19] Hartigan, J. A. and Wong, M. A. (1979). Algorithm AS 136: A k-means algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28 (1), 100-108.

[20] Kafieh, R. and Mehridehnavi, A. (2013). A comprehensive comparison of different clustering methods for reliability analysis of microarray data. *Journal of Medical Signals and Sensors*, 3 (1), 22-30. http://doi.org/10.4103/2228-7477.114306.

[21] Kajimbwa, M. (2006). NGOs and their role in the global south. *International Journal of Non-for-Profit Law*, 9 (1), 58-64.

[22] Kim, H.-R. (1997). Korean NGOs: Global trend and prospect. *Global Economic Review*, 26 (2), 93-115. http://dx.doi.org/10.1080/12265089708422870.

[23] Lewis, D. (1998). Development NGOs and the challenge of partnership: Changing relations between North and South. *Social Policy & Administration*, 32 (5), 501-512. https://doi.org/10.1111/1467-9515.00111.

[24] Likas, A., Vlassis, N. and Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern Recognition*, 36 (2), 451-461. https://doi.org/10.1016/S0031-3203(02)00060-2.

[25] MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of 5th Berkeley Symposium on *Mathematical Statistics and Probability*, Berkeley, CA, USA, June 21 – July 18, 1965 and December 27, 1965 – January 7, 1966.

[26] Milligan, G. W. and Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification*, 5, 181-204.

[27] Mitlin, D. C. (2003). A study of relations between Northern and Southern NGOs in Kenya. PhD thesis, London School of Economics and Political Science, London, United Kingdom.

[28] Mostashari, A. (2005). *An Introduction to Non-Governmental Organizations (NGO) Management*, Iranian Studies Group, MIT.

[29] Pallas, C. L. and Nguyen, L. (2018). Transnational advocacy without Northern NGO partners: Vietnamese NGOs in the HIV/AIDS sector. *Nonprofit and Voluntary Sector Quarterly*, 47 (4S), 159S-176S. https://doi.org/10.1177/0899764018758462.

[30] Punj, G. and Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for application. *Journal of Marketing Research*, XX, pp. 134-148. https://doi.org/10.1177/002224378302000204.

[31] Qi, J., Yu, Y., Wang, l., Liu, J. and Wang, Y. (2017). An effective and efficient hierarchical K-means clustering algorithm. *International Journal of Distributed Sensor Networks*, 13 (8), 1-17. https://doi.org/10.1177/1550147717728627.

[32] Saha, S. and Bandyopadhyay, S. (2012). Some connectivity based cluster validity indices. *Applied Soft Computing*, 12, 1555-1565. https://doi.org/10.1016/j.asoc.2011.12.013.

[33] Shaw, R. (2003). Role of non-government organizations in earthquake disaster management: An Asian perspective. *Regional Development Dialogue*, 24 (1), 117-129.

[34] Slavikova, L., Syrbe, R.-U., Slavik, J. and Berens, A. (2017). Local environmental NGO roles in biodiversity governance: A Czech-German comparison. *GeoScape*, 11 (1), 1-15. http://doi.org/10.1515/geosc-2017-0001

[35] Tan, P.-N., Steinbach, M. and Kumar, V. (2005). *Introduction to Data Mining*. Boston, MA, United States, Addison-Wesley Longman Publishing Co., Inc.

[36] Teegen, H., Doh, J. P. and Vachani, S. (2004). The importance of nongovernmental organizations (NGOs) in global governance and value creation: an international business research agenda. *Journal of International Business Studies*, 35 (6), 463-483. https://doi.org/10.1057/palgrave.jibs.8400112.

[37] Vysochan, O. and Hyk, V. (2020). Formation of accounting information in strategic cluster management. *Revista Espacios*, 41 (01). http://www.revistaespacios.com/a20v41n01/a20v41n01p25.pdf

[38] Vysochan, O., Vysochan, O., Hyk, V. and Hryniv, T. (2021). Attributive-spatial tourist clusteration of regions of Ukraine. *GeoJournal of Tourism and Geosites*, 35 (2), 480-489. https://doi.org/10.30892/gtg.35228-675.

[39] World Bank (1995). Working with NGO's. *A Practical Guide to Operational Collaboration between the World Bank and Non-governmental Organization* by Carmen Malena, The World Bank, Washington, DC, USA.

[40] Yedla, M., Pathakota, S. R. and Srinivasa, T. M. (2010). Enhancing K-means clustering algorithm with improved initial center. *International Journal of Computer Science and Information Technologies (IJCSIT)*, 2010, 121-125.

[41] Yeung, K. Y., Haynor, D. R. and Ruzzo, W. L. (2001). Validating clustering for gene expression data. *Bioinformatics*, 17 (4), 309-318. https://doi.org/10.1093/bioinformatics/17.4.309