

ATTRIBUTIVE-SPATIAL TOURIST CLUSTERATION OF REGIONS OF UKRAINE

Oleh VYSOCHAN*

Lviv Polytechnic National University, Department of Accounting and Analysis, Ukraine, e-mail: oleh.s.vysochan@lpnu.ua

Olha VYSOCHAN

Lviv Polytechnic National University, Department of Accounting and Analysis, Ukraine, e-mail: olha.o.vysochan@lpnu.ua

Vasyl HYK

Lviv Polytechnic National University, Department of Accounting and Analysis, Ukraine, e-mail: vasyl.v.hyk@lpnu.ua

Tetiana HRYNIV

Lviv Polytechnic National University, Department of Accounting and Analysis, Ukraine, e-mail: tetiana.t.hryniv@lpnu.ua

Citation: Vysochan, O., Vysochan, O., Hyk, V., & 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>

Abstract: Ukraine is positioned as a country with a strong tourism potential, much of which still remains unrealized. The main task of the study is to segment the regions of Ukraine according to the level of their tourism development. Attribute-cluster analysis using k-means and k-means ++ technologies was used to solve the problem. Territorial distribution showed clear priorities for the development of certain types of tourism in geospatial context: northeast direction – priority development of business tourism; southern direction – priority development of recreational tourism; western direction – priority development of health tourism; center – emphasis on the development of cultural and historical tourism.

Key words: tourist region, cluster analysis, tourist cluster, k-means clustering, k-means++ clustering, spatial clustering

* * * * *

INTRODUCTION

Ukraine has significant natural and recreational (in particular, water, balneological, forest, landscape), as well as historical and cultural (archaeological, architectural, sacred) resources. At the same time, the socio-economic component of Ukraine's tourism potential is at an early stage of development. Transport accessibility, the level of tourist service, infrastructure in the field of hospitality, training of tourism professionals have been criticized by tourists for decades. The financial and economic crisis of 2014-2015, the annexation of the Autonomous Republic of Crimea, hostilities in Donetsk and Luhansk regions, the COVID-19 pandemic led to a decrease in Ukraine's popularity as a tourist destination, reducing the level of tourism potential in domestic and foreign markets. The initial stage on the way to solving systemic problems in the field of tourism in Ukraine is the local positioning and ranking of regions in terms of tourism potential and efficiency of its implementation with reference to the administrative-territorial division of Ukraine into 24 regions, 1 Autonomous Republic of Crimea and 2 cities with special status. Understanding the level of tourism development in the regions of Ukraine is especially important in the context of decentralization reform, which consists in the gradual transfer of powers and finances from public authorities to local governments, and has been taking place in Ukraine since 2014. The regional specificity of the tourist market of Ukraine is due to the difference between natural and recreational and historical and cultural resources, as well as different levels of general economic development of the territories. An effective means of assessing the homogeneity of regions and the proximity of their tourism development is cluster analysis, which is to divide the regions into relatively independent groups – clusters.

Clusters can be considered from two points of view:

1) industrial (according to Michael Porter) approach, in which the cluster is considered as a group of geographically located interconnected companies and organizations in a particular industry. Enright and Roberts (2001) believe that economic reasons for the geographical concentration of certain industries include the availability of unique natural resources, economies of scale, proximity to markets, labor pooling, availability of local suppliers of raw materials and equipment, shared infrastructure, reduced transaction costs and other localized externalities. This approach can be used to form clusters in many sectors of the economy, as demonstrated in (Vysochan et al., 2020);

2) broad approach – a cluster is a set of any interconnected localized objects, united by a number of common features. In this case, the association can take place at the micro, meso and macro levels and does not have to be geospatially organized. Often such groups are informal and are used to empirically study phenomena or processes common to objects as a whole. This approach will be used in our study.

Clustering is a long process, each phase of which requires deep preparation. For this reason, regions wishing to increase their competitiveness should have a strategic plan for managing the clustering process (Yalçınkaya and Güzel, 2019). An element of

* Corresponding author

such a clustering strategy is the analysis, which allows to assess the regions by the level of their tourism development by determining the tourism potential and the level of its use in the administrative-territorial and geospatial sections. The resulting clusters can be the basis for the formation and / or adjustment of tourism development strategies and resorts of the country for the long term, taking into account the characteristics of tourist regions and their priority types of tourism. A separate value of such an analysis is the possibility of using its results to build a regional policy of regulation of the tourist complex to increase the investment attractiveness of a particular area. The main aim of this study is segmentation the regions of Ukraine according to the level of tourism development by some indicators, such as the number of collective accommodation facilities, the number of museum visitors per year, the paid tourist tax and the number of entities engaged in excursion activities. Based on the relevance of the topic and the issues raised in the article, the emphasis is on answering the following practically important questions:

RQ1: Is it possible and, if so, which groups of regions of Ukraine by the level of tourism development can be distinguished?

RQ2: What types of tourism can be recognized as a priority in the geospatial organization of the tourist market of Ukraine, based on the tourism potential of individual regions?

To answer the questions raised, the literature review presents examples of effective research on the topic of spatial clustering of tourist regions in different countries in general and Ukraine in particular. After the literature review, the methodology of primary data processing and their cluster analysis using k-means and k-means++ technologies is considered. The following are the results of clustering the regions of Ukraine by attributes of the number of collective accommodation, the number of museum visitors per year, paid tourist tax and the number of entities engaged in excursion activities using software for statistical data processing R and geospatial visualization QGIS. In the future, the obtained results are compared with the results of research by other scientists on the tourist clustering of regions of Ukraine. At the end, detailed conclusions are presented with the characteristics of regional tourism clusters of Ukraine, as well as the established priorities for the development of certain types of tourism in geospatial context.

LITERATURE REVIEW

The concept of clustering can be successfully applied to tourism, as tourist attractions can be grouped into clusters to optimize the combination of factors for added value (Andria and di Tollo, 2015). In this regard, in recent years in the world scientific community, the topic of research related to tourism clustering has become especially popular. Clustering is carried out in terms of groups of tourists (Xiao-Ting and Bi-Hu, 2012; Munoz et al., 2019), one region or natural object (Sarrion-Gavilan et al., 2015; Rodriguez-Rangel and Sanchez-Rivero, 2020), sets of cities (Yang and Wong, 2013), regions (Morozova et al., 2016; Vieira and Santos, 2017; Litvinova, 2020) or countries (Chavez et al., 2016; Korol, 2017; Sclyarenko and Khanova, 2018; Batista e Silva et al., 2020; Baymenova et al., 2020). This is important because much of the clustering research has been criticized for being applied to the macro-regional environment rather than to the local or micro-environment (Weidenfeld et al., 2011). The generalization of research on attributive-spatial and socio-economic tourism clustering is presented in Table 1.

Table 1. Examples of successful implementation of the procedure of separation of tourist segments by means of cluster analysis in the last decade (Source: author's generalization)

Source	Localization	Basic clustering criteria	Selected clusters
Xiao-Ting and Bi-Hu, 2012	Summer Palace, Beijing, China	Temporal behaviour factors, spatial behaviour factors, activity choice factors, path characteristics factors	Seven clusters of spatial-temporal behaviour patterns
Yang and Wong, 2013	China's cities	Inbound and domestic tourist flows	Beijing-Tianjin cluster, Yangtze River Delta cluster, Fujian coast cluster, Pearl River Delta cluster, Chengdu cluster
Sarrion-Gavilan et al., 2015	Andalusia, Spain	Number of tourist beds available in a destination, Total number of permanent residents	Littoral, Urban interior, Rural interior
Chavez et al., 2016	APEC countries	Total contribution to GDP, tourism contribution to employment, total investment on tourism	4 clusters
Morozova et al., 2016	65 regions of Russia	Performance indicators of the tourism industry by regions	High tourism effectiveness regions, average tourism effectiveness regions, low tourism effectiveness regions,
Korol, 2017	51 world's countries	Tourist arrivals and receipts from inbound tourism	8 clusters
Vieira and Santos, 2017	Portugal	Accommodation capacity, overnight stays	Algarve region, Greater Lisbon, Greater Porto, North Portugal, Centre Portugal
Sclyarenko and Khanova, 2018	28 EU countries	Level of development of the tourism industry	2 clusters with a high level of development and 5 cluster with an average level of development
Munoz et al., 2019	Jotunheimen National Park and Utladalen Protected Landscape, Norway	Set of markers representing different place-based values: biological, clean water, cultural, gathering, hunt/fish, recreation, scenic, social, special place, spiritual, therapeutic, wilderness	13 clusters
Rodriguez-Rangel and Sanchez-Rivero, 2020	Extremadura, Spain	Tourist supply and demand	Badajos, Merida and Caceres (best adjustment between supply and demand), Trujillo-Miajadas-Montanchez and Villuercas-Ibores-Jara Geopark (worst adjustment between supply and demand)
Batista e Silva et al., 2020	EU NUTS3 regions	Nights-spent, tourism intensity, share of foreign tourist, seasonality	Cities, coastal, mountain and nature, rural, urban mix
Baymenova et al., 2020	CIS countries	Incidence rates regarding major classes, levels of healthcare organization, transport infrastructure, tourist services, information and communication technologies	2 large and 4 small clusters considering development of medical tourism

The success of clusters in developed countries has quickly spread to developing countries, attracting the interest of scientists, practitioners and politicians (Ferreira and Esteveao, 2009). With the beginning of the decentralization process in Ukraine, interest in regional clustering in the field of tourism has grown sharply among scientists in the country. The issue of choosing indicators to be used as signs of clustering of tourist regions came to the fore: the volume of tourist flows of inbound and domestic tourism (general or in terms of purpose of stay), as well as indicators of legal entities and individual entrepreneurs – entities tourism (Leontyeva and Vlashchenko, 2012; Yurchenko, 2012; Ocheretin, 2015; Lipjanina, 2016; Nazarova, 2016; Gorina, 2017; Druzhynina, 2017; Neshchadym and Tumchyk, 2017; Pokataieva and Kucherova, 2017; Borblik, 2019; Dutka et al., 2019; Gorina et al., 2020; Zbarsky and Gribova, 2020), index / level of tourist attractiveness (Davydova, 2015; Antonets and Kharchenko, 2016; Bilogur, 2019); inbound and / or outbound tourists by volume of tourist flows (Kondratska, 2019; Kondratska et al., 2019); the level of tourism potential for ecotourism development (Zyma and Holub, 2015). At the same time, the analysis of literature sources showed a significant shortage of completed publications on regional clustering of tourism, taking into account the attributive (in terms of tourism development) and spatial (in terms of regions of Ukraine) components and opportunities to use the results of such analysis to identify priority areas for tourism.

MATERIALS AND METHODS

The sequence of actions in the implementation of attributive-cluster analysis is as follows.

Stage 1. Establish the nomenclature of attributes (descriptive features) by which the clustering, information collection and spatial distribution of clustering objects will be carried out. As the latter, 24 regions representing administrative-territorial units of Ukraine were selected, with the exception of the Autonomous Republic of Crimea, for which data do not seem possible to obtain due to the annexation of the territory. Tourism development of the regions is characterized by a number of indicators, some of which are presented in Table 2. Indicators of the number of collective accommodation facilities, the number of museum visitors per year and the number of entities engaged in excursion activities are obtained from statistical observations conducted by the State Statistics Service of Ukraine (<http://www.ukrstat.gov.ua/>) annually and published in free access. The rate of paid tourist tax is calculated by the Ministry of Economic Development, Trade and Agriculture of Ukraine (<https://www.me.gov.ua/>) according to the State Fiscal Service of Ukraine for a period of six months and a year and are published in free access. Since we do not aim at a clear quantitative measurement of tourism development in each region of Ukraine, but only a qualitative assessment of its level, the list of attributive features is enough to include basic statistics (Table 2). If it is necessary to detail information on individual regions, the list of indicators can be expanded.

Table 2. Indicators for assessing the tourism development of regions (Source: developed by authors)

Indicator	Characteristic	Source of information	Why it is important
Number of collective accommodation facilities, units	Number of hotels and other accommodation in a particular region	Collective accommodation facilities in Ukraine (legal entities, separate divisions of legal entities). Statistical information	The indicator allows you to assess the tourist potential of the region in terms of accommodation for tourists arriving for more than 1 day
Number of museum visitors per year, pers.	Number of visitors (individual and in a group) of research and cultural-educational institutions in order to review the museum collections of a particular region	Institutions of culture, physical culture and sports of Ukraine. Statistical collection	The indicator provides information on the realization of the cultural and historical potential of the region
Paid tourist tax, UAH million	The amount of paid local tax to persons who are temporarily accommodated in the places of residence (overnight stay) of a particular region	Information on financial and economic indicators of tourism and resorts (according to the State Fiscal Service of Ukraine)	The indicator allows to separate the share of tourists from the share of persons who come to the territory of a certain region for non-tourist purposes (eg, on business trips); characterizes the realization of tourist potential
Number of entities engaged in excursion activities, units	Number of excursion bureaus and individuals – entrepreneurs who have the necessary accreditation for the preparation, organization and conduct of excursions in a particular region	Tourist activity in Ukraine. Statistical collection	The indicator provides information on the presence of active players in the market of professional excursion services in the region; characterizes the tourist potential

Stage 2. Gradation of the state of the basic features of clustering.

The grouping of the analyzed set of indicators that characterize the level of tourism development in the regions, on an attribute (descriptive) basis, the question of the number of groups is solved based on gradations of this feature: “low”, “below average”, “above average”, “high”. Given the equality of the intervals, their value depends on the scope of variation of the feature and the number of the population and is determined by formula (see Pedchenko, 2018):

$$int = \frac{x_{max} - x_{min}}{g} \tag{1}$$

where, *int* – the value of the interval;
x_{max} – the maximum value of the feature in the data set;
x_{min} – the minimum value of the attribute in the data set;
g – optimal number of groups.

Stage 3. Establish the optimal number of data clusters.

Modern packages of applied statistical programs allow you to implement several dozen methods for determining the optimal number of clusters. In the work of Charrad et al. (2014) provides a comprehensive description of the following indices

that will be used by us to establish the number of clusters of tourist regions of Ukraine: kl, ch, hartigan, ccc, scott, marriot, trcovw, tracew, friedman, rubin, cindex, db, silhouette, duda, pseudot2, beale, ratkowsky, ball, ptbserial, frey, mcclain, dunn, hubert, sdindex, dindex, sdbw. To solve the question of the optimal number of clusters, you can apply the rule of simple majority or use the tools only as an aid, choosing the number of groups based on the purpose and needs of the analysis.

Stage 4. Direct clustering.

K-means is one of the most commonly used clustering techniques, which is to minimize the root mean square distance between points in a cluster. K-means is implemented using Lloyd's algorithm (Lloyd, 1982). The disadvantages of this technique are to eliminate the technique of k-means++ (Arthur and Vassilvitskii, 2007), which, in contrast to the k-means algorithm implements a specific method of selecting cluster centers. The algorithm is executed in several steps:

Step 1a. The selection of the initial center c_i is sequentially random from the set of data points X .

Step 1b. Select the next center c_i , selecting $c_i = x' \in X$ with probability $\frac{D(x')^2}{\sum_{x \in X} D(x)^2}$.

Step 1c. Repeat step 1b until the total number of centers k is selected.

Step 2. For everyone $i \in \{1, \dots, k\}$, set for the cluster C_i of points in X , which are closer to c_i than to c_j to all $j \neq i$.

Step 3. For everyone $i \in \{1, \dots, k\}$, we install c_i behind the center of mass of all points in C_i : $c_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$.

Step 4. Repeat steps 2 and 3 until C is no longer changed.

Steps 2-4 are similar to the k-means algorithm.

Stage 5. Checking the internal validity of clusters.

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. Internal criteria can be divided into two groups: 1) those that assess the correspondence between the data and the expected structure and 2) those that focus on the stability of the solution (Rendon et al., 2011). The first group includes Connectivity, Silhouette width, Dunn's index, the second – Average proportion of non-overlap (APN), Average distance (AD), Average distance between means (ADM) and Figure of merit (FOM). The use of specialized software facilitates calculations and allows you to present the results of the analysis in a graphical and understandable informative form. Arranging the set of values of objects that characterize the development of tourism, using the methods of k-means and k-means++ can be effectively done in the environment of software products R and QGIS, and visualize the results geospatial using QGIS.

RESULTS

Input data for assessing the tourist development of the regions of Ukraine are obtained from official sources of the State Statistics Service of Ukraine, the Ministry of Economy, Trade and Agriculture of Ukraine and the State Fiscal Service of Ukraine and are reduced to a single measure by normalization (Table 3).

Table 3. Input data for clustering of tourist regions of Ukraine (Source: processed statistical data, 2020)

ID	Region	Number of entities engaged in excursion activities		Number of collective accommodation facilities		Number of museum visitors per year		Paid tourist tax	
		initial	normalized	initial	normalized	initial	normalized	initial	normalized
1	Vinnitsya	2	0.03278689	21	0.046728972	717300	0.159415857	483,8	0.009665421
2	Volyn	7	0.11475410	61	0.233644860	229500	0.035094426	584,6	0.012531777
3	Dnipropetrovsk	0	0	106	0.443925234	726700	0.161811555	2523,5	0.067666479
4	Donetsk	2	0.03278689	88	0.359813084	328300	0.060274741	1224,4	0.030725177
5	Zhytomyr	2	0.03278689	19	0.037383178	368600	0.070545658	544,8	0.011400022
6	Zakarpattia	3	0.04918033	59	0.224299065	694300	0.153554043	3802	0.104021998
7	Zaporizhzhya	7	0.11475410	133	0.570093458	548600	0.116420725	3665,9	0.100151849
8	Ivano-Frankivsk	7	0.11475410	64	0.247663551	470600	0.096541530	3699,4	0.101104457
9	Kyiv and Kyiv	42	0.68852459	225	1	4015500	1	35310,5	1
10	Kirovograd	4	0.06557377	37	0.121495327	259000	0.042612840	259,3	0.003281523
11	Luhansk	2	0.03278689	11	0	91800	0	143,9	0
12	Lviv	61	1	130	0.556074766	1928100	0.468002141	13417,6	0.377451900
13	Mykolayiv	1	0.01639344	103	0.429906542	341100	0.063536968	1832,6	0.048019996
14	Odesa	31	0.50819672	187	0.822429907	576900	0.123633305	11550,7	0.324364596
15	Poltava	0	0	44	0.154205607	605300	0.130871371	1341	0.034040823
16	Rivne	9	0.14754098	18	0.032710280	256100	0.041873742	355,7	0.006022760
17	Sumy	2	0.03278689	21	0.046728972	254900	0.041567908	332,2	0.005354513
18	Ternopil	9	0.14754098	15	0.018691589	479200	0.098733338	402	0.007339350
19	Kharkiv	8	0.13114754	73	0.289719626	830300	0.188215205	2979,6	0.080636172
20	Kherson	9	0.14754098	88	0.359813084	109900	0.004612993	3690,4	0.100848532
21	Khmelnytsky	10	0.16393443	31	0.093457944	523700	0.110074674	731,5	0.016709036
22	Cherkasy	2	0.03278689	55	0.205607477	830700	0.188317150	724,1	0.016498609
23	Chernivtsi	2	0.03278689	13	0.009345794	302100	0.053597370	527,2	0.010899547
24	Chernihiv	1	0.01639344	24	0.060747664	920800	0.211280169	599,4	0.012952631

A cursory review of the indicators of tourism development of the regions of Ukraine makes it possible to draw a conclusion about the significant scatter of their values and the crystallization of areas that can be attributed to the leaders in the field of tourism. For example, the number of hotels and other accommodation facilities, which is a basic indicator for assessing the level of tourist potential of the region, in Kyiv and Kyiv region is 20 times higher than in Luhansk (Figure 1-2).

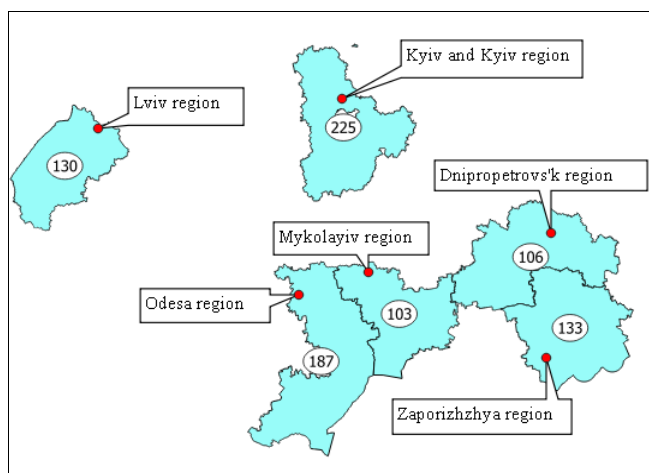


Figure 1. Regions of Ukraine with the largest number of collective accommodation facilities (units) as of 2019 (Source: developed by authors)

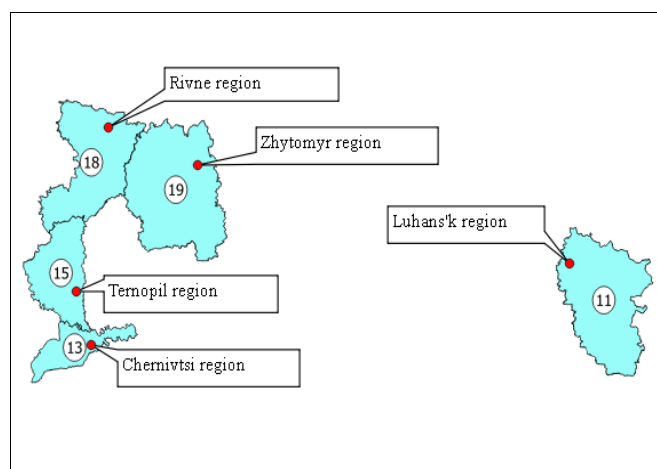


Figure 2. Regions of Ukraine with the lowest number of collective accommodation facilities (units) as of 2019 (Source: developed by authors)

Such an imbalance between the leading regions and regions with medium and low rates of tourism development is especially noticeable on the example of the indicator “Number of entities engaged in excursion activities”.

While in Lviv region this indicator, which characterizes the tourist potential, is at the level of 61 units, in Kyiv and Kyiv region – 42 units, in Odesa region – 31 units, in all other regions it is less than 10 licensed excursion bureaus.

The values of indicators of tourist development of the regions of Ukraine are grouped by distinguishing three groups by descriptive features: “low”, “medium” (combining the attributes “below average” and “above average”), “high”. Since the range of values of the quantitative features of the survey varies widely, unevenly, it is advisable to use unequal intervals of groups, taking into account the following sequence of actions:

Step 1. The flagship region is set for each indicator, the value of this particular indicator of tourism development is significantly higher than the value of a similar indicator of the nearest competitor.

Step 2. Determining the previous number of groups g taking into account the gradations of the attributive feature: “low”, “below average”, “above average”, “high” ($g = 4$).

Step 3. The value of the range of values for each indicator of tourism development is determined by the modified formula (1) minus the value of the indicator of the flagship region (Source: developed by authors):

$$int = \frac{x_{max-1} - x_{min}}{g} \quad (2)$$

where, x_{max-1} – the second largest value of the feature in the data set after the maximum.

Step 4. Clarification and unification of the number of groups and the interval of their data, taking into account the unevenness of the change of features: combining attributive features “below average”, “above average” into one – “average” (Table 4); expanding the range of data on the basis of “low” to 0, on the basis of “high” – to ∞ .

Table 4. Attributive features and data ranges for indicators of tourism development of the regions of Ukraine (Source: author’s own calculations)

Indicator	Range of values		
	Low	Average	High
Number of collective accommodation facilities, units	up to 55 inclusive	56 – 143 inclusive	more than 144
Number of museum visitors per year, million people	up to 0,55 inclusive	0,56 – 1,46 inclusive	more than 1,47
Paid tourist tax, UAH million	up to 3,4 inclusive	3,5 – 10 inclusive	more than 10,1
Number of entities engaged in excursion activities, units	up to 15 inclusive	16 – 45 inclusive	more than 46

Optimal $g = 4$: “low”, “medium”, “high”, “flagship”. The flagship value of the indicator is a benchmark for other regions to improve performance in a particular area of tourism development.

The distribution of regions by the level of achievement of tourism development indicators is presented in Table 5.

The results of in-depth analysis using the software environment R for 26 indices showed: 9 methods offer 2 clusters as the optimal number; 2 methods offer 3 clusters as the optimal number; 10 methods offer 4 clusters as the optimal number; 2 methods offer 6 clusters as the optimal number; 1 method offers 8 clusters as the optimal number; 1 method offers 9 clusters as the optimal number; Method 1 offers 10 clusters as the optimal number. Thus, the closest to describe the real state of affairs with the regional tourism development of Ukraine is the use of two-cluster and four-cluster models (Figure 3-4).

Table 5. Distribution of tourist regions of Ukraine according to the criteria of achieving “low level”, “medium level”, “high level”, “flagship level” in terms of tourism development indicators (Source: developed by authors)

Indicator	Qualitative characteristics of the value of the indicator			Flagship
	Low	Average	High	
Number of collective accommodation facilities, units	Vinnitsya, Zhytomyr, Kirovohrad, Luhansk, Poltava, Rivne, Sumy, Ternopil, Khmelnytsky, Chernivtsi, Chernihiv, Cherkasy	Volyn, Dnipropetrovsk, Donetsk, Zakarpattia, Zaporizhzhya, Ivano-Frankivsk, Lviv, Mykolayiv, Kharkiv, Kherson	Odesa	Kyiv and Kyiv region
Number of museum visitors per year, million people	Volyn, Donetsk, Zhytomyr, Zaporizhzhya, Ivano-Frankivsk, Kirovohrad, Luhansk, Mykolayiv, Rivne, Sumy, Ternopil, Kherson, Khmelnytsky, Chernivtsi	Vinnitsya, Dnipropetrovsk, Zakarpattia, Odesa, Poltava, Kharkiv, Cherkasy, Chernihiv	Lviv	Kyiv and Kyiv region
Paid tourist tax, UAH million	Vinnitsya, Volyn, Dnipropetrovsk, Donetsk, Zhytomyr, Kirovohrad, Luhansk, Mykolayiv, Poltava, Rivne, Sumy, Ternopil, Kharkiv, Khmelnytsky, Cherkasy, Chernivtsi, Chernihiv	Zakarpattia, Zaporizhzhya, Ivano-Frankivsk, Kherson	Lviv and Odesa	Kyiv and Kyiv region
Number of entities engaged in excursion activities, units	Vinnitsya, Volyn, Dnipropetrovsk, Donetsk, Zhytomyr, Zakarpattia, Zaporizhzhya, Ivano-Frankivsk, Kirovohrad, Luhansk, Mykolayiv, Poltava, Rivne, Ternopil, Sumy, Kharkiv, Kherson, Khmelnytsky, Cherkasy, Chernivtsi	Odesa	Kyiv and Kyiv region	Lviv

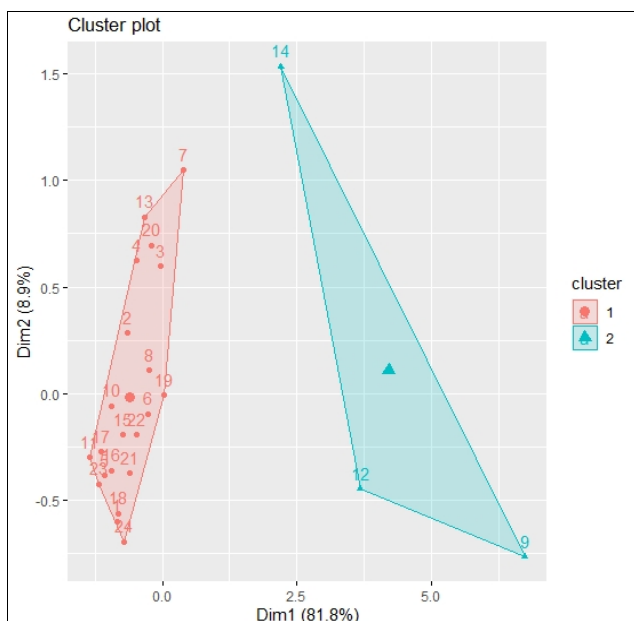


Figure 3. Clustering of regions of Ukraine by level of tourism development using k-means technology (k = 2) (Source: developed by authors)

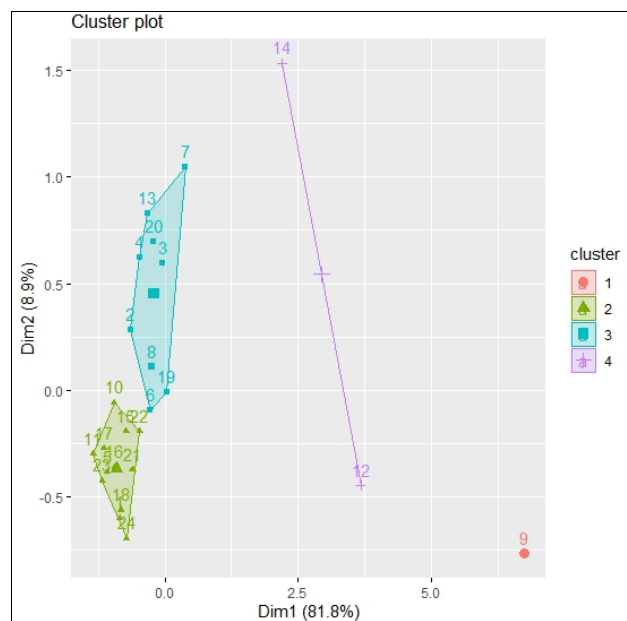


Figure 4. Clustering of regions of Ukraine by level of tourism development using k-means technology (k = 4) (Source: developed by authors)

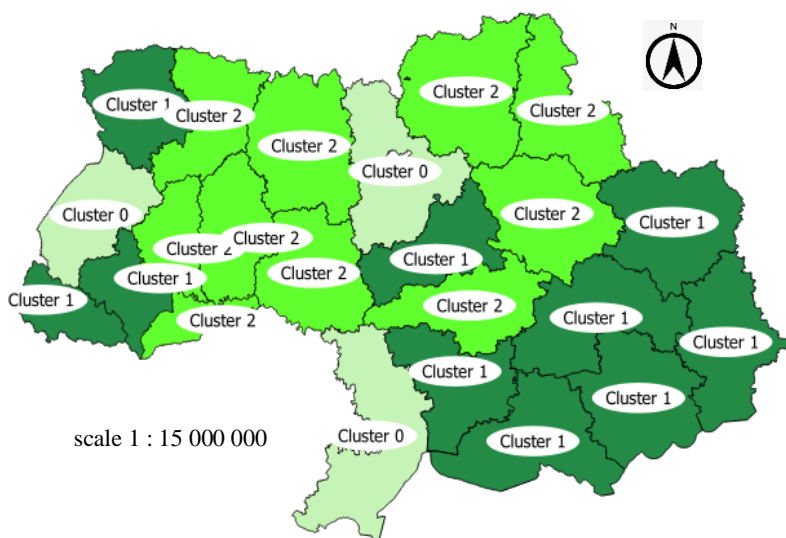


Figure 5. Clustering of regions of Ukraine by level of tourism development using k-mean++ technology (k = 4) (Source: developed by authors)

Note: the map does not show data on the Autonomous Republic of Crimea, reliable information for the calculation of tourism development which could not be obtained and Luhansk region, which forms a separate cluster with the lowest rates of tourism development

Clustering of tourist regions of Ukraine, carried out using the k-means++ algorithm using the Attribute based clustering module of the QGIS system, developed by Kazakov (2016), confirmed the relevance of using a four-cluster model of grouping regions by level of tourism development (Figure 5). Attribute clustering highlights the numerical characteristics of objects (as opposed to the spatial characteristics required for spatial clustering of objects in GIS). The main result of the module is to create a new (or overwrite the existing) attribute of the vector field with an entry in it for each object of the cluster number to which the object was assigned as a result of calculations (Kazakov, 2016).

The imposition of attributive and spatial planes of cluster analysis makes it possible to obtain a visual representation of regional tourism clusters of Ukraine (Figure 6-9).

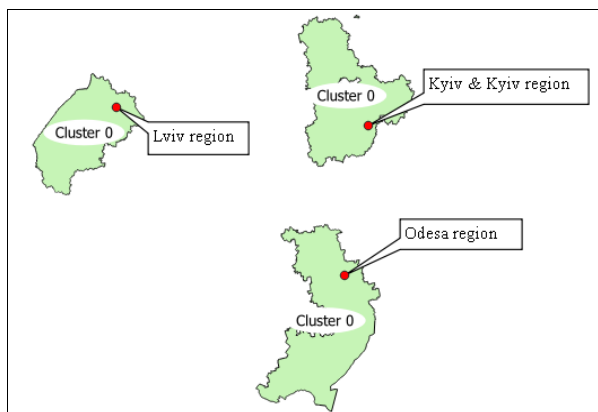


Figure 6. Regions of Ukraine with high tourist potential and high level of its realization (cluster 0) (Source: developed by authors)

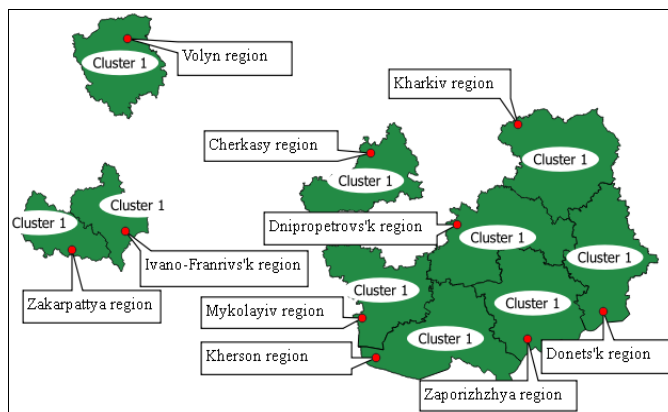


Figure 7. Regions of Ukraine with average tourist potential and high level of its realization (cluster 1) (Source: developed by authors)

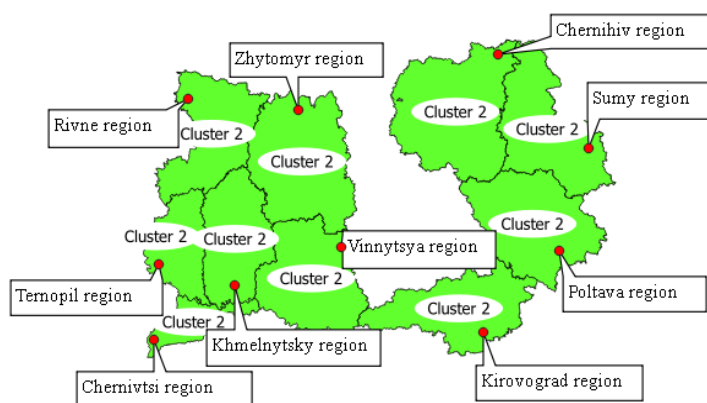


Figure 8. Regions of Ukraine with average tourist potential and low level of its realization (cluster 2) (Source: developed by authors)

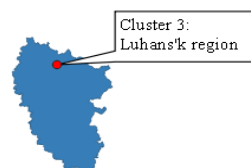


Figure 9. Region of Ukraine with low tourist potential and low level of its realization (cluster 3) (Source: developed by authors)

The results obtained after using both clustering methods are summarized in Table 6. The k-mean++ technology implemented in the QGIS system made it possible to separate the region-outsider of tourism development (Luhans'k region) into a separate cluster, while k-mean (a four-cluster model built in the R environment) provided convincing data on Ukraine's flagship tourism activity – Kyiv and Kyiv region. Comparison of four-cluster models formed using k-means and k-means++ technologies for clusters with average indicators of tourism

development shows comparability of results, except for Cherkasy region, which in the first case is assigned to the lower category, and in the second – to the highest. However, such incomparability is insignificant due to the borderline indicators of tourism development of this region in both clusters and their proximity to the average values by region. Assessment of the validity of the created clusters of tourist regions of Ukraine requires the calculation of a number of indices, summarized in Table 7. We see that the four-cluster model shows better validity, surpassing the two-cluster in terms of stability assessment (AD, ADM and FOM) and slightly inferior to the other two indicators of internal assessment (Connectivity and Silhouette coefficient).

Table 6. Comparison of the results of clustering of tourist regions of Ukraine by k-means k-means++ methods (Source: developed by authors)

k-means		k-means++		Characteristic		
two-cluster model	four-cluster model	four-cluster model				
cluster	ID regions	cluster	ID regions	cluster	ID regions	
1	9, 12, 14	1	9	0	9, 12, 14	Regions of Ukraine with high tourist potential and high level of its realization (flagship region and leading regions)
		2	12, 14			
2	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24	3	2, 3, 4, 6, 7, 8, 13, 19, 20	1	2, 3, 4, 6, 7, 8, 13, 19, 20, 22	Regions of Ukraine with average tourist potential and a high level of its realization
		4	1, 5, 10, 11, 15, 16, 17, 18, 21, 22, 23, 24	2	1, 5, 10, 15, 16, 17, 18, 21, 23, 24	Regions of Ukraine with average tourist potential and low level of its realization
-	-	-	-	3	11	Regions of Ukraine with low tourist potential and low level of its realization (outsider regions)

Table 7. Evaluation of the validity of regional tourist clusters of Ukraine (Source: author's own calculations)

Evaluation indicator	Indicator value		The model for which the indicator is the best	
	for a two-cluster model	for a four-cluster model	a two-cluster model	a four-cluster model
Internal measures				
Connectivity	5.8369	8.7869	+	
Silhouette coefficient	0.7247	0.6054	+	
Dunn's index	0.3988	0.9395		+
Stability measures				
Average proportion of non-overlap (APN)	0.0278	0.1290	+	
Average distance (AD)	1.3256	0.8651		+
Average distance between means (ADM)	0.3075	0.1996		+
Figure of merit (FOM)	0.7480	0.3921		+

DISCUSSION

There are different approaches to the analysis of tourism development in Ukraine using a set of indicators grouped by principle:

1) “top-down” – tourism development is assessed by a system of macro-indicators of the region: GRP per capita, revenues from the payment of tourist tax, price index for transport services, price index for hotel services, the volume of tourism services for the organization domestic tours (Zbarsky and Gribova, 2020); number of administrative-territorial units, natural resource potential, number of monuments of national importance, number of persons who are rehabilitated in sanatoriums, number of persons served in hotels and other places for short-term stay, number of sanatoriums, number of rooms in hotels, restaurant chain (Antonets and Kharchenko, 2016); the level of tourist load on the population (Bilogur, 2019).

2) “bottom-up” – tourism development is assessed by a system of indicators of the functioning of tourism entities – legal entities and individuals: the average number of full-time employees (Gorina, 2017; Gorina et al., 2020; Pokataieva and Kucherova, 2017; Dutka et al., 2019), income from the provision of tourist services (Gorina, 2017; Gorina et al., 2020; Pokataieva and Kucherova, 2017; Borblik, 2019; Dutka et al., 2019), the amount of commissions, agency and other fees (Dutka et al., 2019), the number of sold tourist vouchers (Gorina, 2017; Gorina et al., 2020; Pokataieva and Kucherova, 2017; Borblik, 2019; Dutka et al., 2019), the cost of sold tourist vouchers (Gorina, 2017; Gorina et al., 2020; Pokataieva and Kucherova, 2017; Borblik, 2019; Dutka et al., 2019), number of tourists on realized tourist vouchers (Dutka et al., 2019), number of tourists and excursionists served by tourist enterprises (Gorina, 2017; Gorina et al., 2020; Pokataieva and Kucherova, 2017; Borblik, 2019; Dutka et al., 2019), operating costs (Dutka et al., 2019), costs of tourism entities for the services of third parties (Pokataieva and Kucherova, 2017; Dutka et al., 2019). The number of tourism entities (Gorina, 2017; Gorina et al., 2020; Pokataieva and Kucherova, 2017; Borblik, 2019; Dutka et al., 2019) or the number of tourism entities used as an indicator that ensures the transition to a macro-assessment of tourism development in the region collective accommodation facilities (Borblik, 2019) in terms of regions of Ukraine;

3) a combination of previous approaches – a combination of macro indicators and generalized indicators of tourism enterprises: revenues from tourism services, expenditures of tourism entities and the flow of tourists for the purpose of their trip (Lipjanina, 2016); effective indicators of the region's development in the field of tourism, indicators of demography and labor economy, quantitative and qualitative composition of employees of tourism enterprises, economic indicators of the tourism market (Druzhynina, 2017); performance indicators of tourism enterprises in combination with tourist flows for the purpose of travel, the number of tourism entities, the income of the population of the region (Nazarova, 2016).

At first glance, it is the first approach to assessing the tourism development of the region can be considered basic, but, in our opinion, it contains only indicators that help determine the tourism potential of the region, and not the degree of its implementation. The latter allows us to assess the indicators of financial and economic activity of the tourism market. The second approach, in our opinion, can not be used to assess the tourism development of the region, as many indicators (income from travel services, the amount of commissions, agency and other fees, the number of trips sold, operating costs) involved only indicate about the level of profitability of tourist companies in a particular region, whose activities are aimed not only at the growth of incoming, but also (and sometimes mostly) the growth of outgoing tourist flows.

Table 8. Regions-leaders of tourism of Ukraine according to researches of various scientists (Source: author's generalization)

Authors	Research horizon	Methods of cluster analysis	Sign of clustering	Regions are leaders					
				Dnipropetrovsk	Ivano-Frankivsk	Kyiv and Kyiv region	Lviv	Odesa	Kharkiv
Leontyeva and Vlashchenko, 2012	Not specified	Hierarchical, k-means	Resources, volume of services provided, number of tourism entities	Not installed					
Ocheretin, 2015	2013-2014	K-means	Incoming and outgoing tourist flows, the purpose of the trip	Not installed					
Antonets and Kharchenko, 2016	Not specified	K-means	The level of tourist attraction	-	-	-	+	-	-
Lipjanina, 2016	2011-2014	Self-organizing map	Revenues and expenditures of tourism entities, tourist flows, purpose of the trip	-	-	+	-	-	-
Nazarova, 2016	2014	Hierarchical	Tourist services	-	-	+	-	-	-
Pokataieva and Kucherova, 2017	2014	Hierarchical, k-means	Performance indicators of tourist enterprises	+	-	+	-	-	+
Gorina, 2017	2015	Hierarchical (threshold-05)	Performance indicators of tourist enterprises	-	-	+	-	-	-
Druzhynina, 2017	2015-2016	Hierarchical, k-means	Effective indicators of the region's development, demography and labor economy, quantitative and qualitative composition of employees of tourist enterprises, economic indicators of the tourist market	-	+	+	+	+	-
Neshchadym and Tumchyk, 2017	2016	Self-organizing map	The main indicators of tourism enterprises	Not installed					
Borblik, 2019	2016	Hierarchical, k-means	Indicators of development of markets for tourist services	+	-	-	+	-	+
Dutka et al., 2019	2017	Hierarchical	Performance indicators of tourist enterprises	-	-	-	+	-	-
Gorina et al., 2020	2018	Hierarchical (threshold-05)	Performance indicators of tourist enterprises	-	-	+	-	-	-

The combined approach seems to us to be a priority, as it allows us to combine macro indicators, such as tourist flows and tourist load by region, with indicators calculated by statistics of economic entities – attendance of cultural monuments, congestion of collective accommodation, etc. Statistical processing of the values of indicators of tourism development in the region by means of cluster analysis using hierarchical, k-means and self-organizing maps allowed researchers to identify regions that are leaders in tourism in Ukraine (Table 8).

In general, the results of our study correlate with similar achievements of Ukrainian scientists in the field of tourism, especially in relation to the separation of three regions – Kyiv and Kyiv region, Lviv and Odesa regions as leaders in the tourism industry of Ukraine. At the same time, we managed to obtain empirical evidence of the gradual transformation of the city of Kyiv into the sole flagship of Ukrainian tourism, especially in the field of business tourism. On the one hand, it provides a benchmark for the development of other regions, on the other – creates unnecessary socio-economic risks associated with the disproportion of labor markets, additional pressure on infrastructure, deteriorating environmental situation in the region, which can be viewed through the prism of existing systemic problems in the tourism industry of Ukraine. This was discussed in more detail in Vysochan and Glushko (2011).

CONCLUSION

Attributive-spatial cluster analysis confirmed the existence of four clusters into which the regions of Ukraine are grouped according to the level of their tourism development: cluster 0, which contains the flagship region of tourism in Ukraine in recent decades – Kyiv and two leading regions – Lviv and Odesa; clusters 1 (north and east of Ukraine) and 2 (center of Ukraine and several western region) with an average level of tourism development and cluster 3, which includes the outsider region – Luhansk, which are not least due to hostilities in this direction. We believe that the formation of such clusters, among other things, due to the types of tourism most developed in specific destinations. Despite the fact that the basic type for the economy of Ukraine is cultural and historical tourism, the specialization of the southern regions in recreational tourism, western – in health, and eastern – in business is confirmed by our research and determines the priority of such localization for domestic tourism (Figure 10).

The segmentation of priority types of tourism by geographical destinations has demonstrated the diversity of regional interests:

- North-Eastern direction – priority development of business tourism with centers in Kyiv, Dnipro, Donetsk and Kharkiv (flagship region – Kyiv and Kyiv region);
- Southern direction – priority development of recreational tourism with centers in Odesa, Zaporizhzhya, Mykolayiv and Kherson (leading region – Odesa region);
- Western direction – priority development of health (medical) tourism with centers in Lviv, Uzhgorod, Lutsk and Ivano-Frankivsk (leading region – Lviv region);

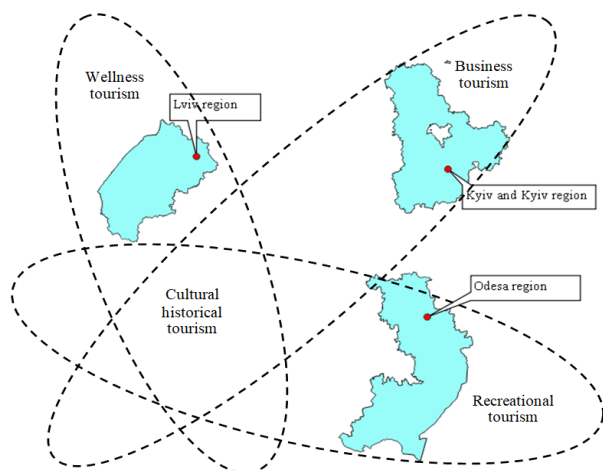


Figure 10. Segmentation of priority types of tourism in the leading tourist regions of Ukraine (Source: developed by authors)

- Center – emphasis on the development of cultural and historical tourism – Cherkasy, Poltava, Chernivtsi and other regions (the leading region – no).

These areas are a priority, but not exclusive, the development of other types of tourism in the regional and national dimensions is to be welcomed. Prospects for further research on the topic presented in the article are an attempt to determine the contribution of each regional tourist cluster in the overall socio-economic development of the country (“bottom-up” assessment), as well as to identify areas for improvement of tourism strategy and resorts in Ukraine. At the same time, it should be borne in mind that the presented study has a number of limitations, which are mainly related to the presence of errors in the obtained input data that arise when compiling tourism entities, according to current legislation in Ukraine, statistical reporting on key indicators own activity, especially Reports on tourist activity and Reports on activity of collective means of accommodation.

REFERENCES

- Andria, J., & di Tollo, G. (2015). Clustering local tourism systems by threshold acceptance. *Computational Management Science*, 12 (4), 559-575. <https://doi.org/10.1007/s10287-015-0238-x>
- Antonets, O., & Kharchenko, A. (2016). Tourist attractiveness estimation of Ukraine’s regions. *Problems and Perspectives of Entrepreneurship Development*, 1, 137-141.
- Arthur, D., & Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. *SODA 2007: Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, New Orleans, Louisiana, USA, January 7-9. <https://doi.org/10.1145/1283383.1283494>
- Batista e Silva, F., Barranco, R., Proietti, P., Pigaiani, C., & Lavallo, C. (2020). A new European regional tourism typology based on hotel location patterns and geographical criteria. *Annals of Tourism Research*, In press. <https://doi.org/10.1016/j.annals.2020.103077>
- Baymenova, A.S., Zhakenova, S.R., & Aydaraliev, A.A. (2020). Cluster analysis of CIS countries regarding development of medical tourism in Kazakhstan. *Social Aspects of Population Health*, 66 (5). <http://vestnik.mednet.ru/content/view/1204/30/lang.ru/> <https://doi.org/10.21045/2071-5021-2020-66-5-11>
- Bilogur, S. (2019). Development of institutional paradigms of cluster concept in tourism. *European Vector of Economic Development*, 2 (27), 7-14. <https://doi.org/10.32342/2074-5362-2019-2-27-1>

- Borblik, K. (2019). Analysis of activities of regional markets of tourist services of Ukraine. *Scientific Notes of Taurida National V. I. Vernadsky University. Series: Economy and Management*, 30 (69), 4, 87-96. <https://doi.org/10.32838/2523-4803/69-4-16>
- Charrad, M., Ghazzali, N., Boiteau, V., & 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>
- Chavez, J.C.L.N., Torres, A.I.Z., & Torres, M.C. (2016). Hierarchical cluster analysis of tourism for Mexico and the Asia-Pacific Economic Cooperation (APEC) countries. *Turismo em Analise*, 27 (2), 235-255. <http://dx.doi.org/10.11606/issn.1984-4867.v27i2p235-255>
- Davydova, O.G. (2015). The evaluation methods of tourist attractiveness in the regions. *Ukrainian Society*, 4 (55), 97-107. <https://doi.org/10.15407/socium2015.04.097>
- Druzhynina, V.V. (2017). Statistical analysis of the activities of regional market of tourism services in Ukraine on the basis of clusterization. *Herald of Khmelnytskyi National University. Economic Sciences*, 5, 292-298.
- Dutka, A., Savitska, O., & Savitska, N. (2019). Cluster analysis of tourism activity subjects: Regional aspects. *Economic Analysis*, 29 (1), 114-125. <https://doi.org/10.35774/econa2019.01.114>
- Enright, M.J., & Roberts, B.H. (2001). Regional clustering in Australia. *Australian Journal of Management*, 26, 65-85. <http://dx.doi.org/10.1177/031289620102601S04>
- Ferreira, J.M., & Estevo, C.M.S. (2009). Regional competitiveness of a tourism cluster: Conceptual model proposal. *Encontros Científicos – Tourist & Management Studies*, 5, 37-51.
- Gorina, G.O. (2017). Cluster analysis of the Ukraine tourist services regional markets. *Bulletin of DonNUET, Series "Economics"*, 1 (66), 70-77.
- Gorina, G.O., Barabanova, V.V., Bohatyryova, G.A., Nikolaichuk, O.A., & Romanukha, A.M. (2020). Clustering of regional tourism service markets according to indicators of the functioning of subjects of tourism activity. *Journal of Geology, Geography and Geoecology*, 29 (4), 684-692. <https://doi.org/10.15421/112061>
- Kazakov, E. (2016). Attribute based clustering plugin for QGIS short manual. http://www.ekazakov.info/projects/abc_tutorial.html
- Kondratska, L. (2019). Approach to the analysis of Ukraine's international tourist flows structure. *Uzhhorod National University Herald. International Economic Relations and World Economy*, 26 (1), 95-99.
- Kondratska, L., Bazhanova, N., & Duma, L. (2019). Econometric analysis of the international destination tourist flow of Ukraine. *Market Infrastructure*, 35, 507-512. <https://doi.org/10.32843/infrastructure35-79>
- Korol, O. (2017). International tourist arrivals and income from foreign tourism: World dynamics and analysis by countries. *Odesa National University Herald. Geography and Geology*, 22 (2), 104-117. [https://doi.org/10.18524/2303-9914.2017.2\(31\).120897](https://doi.org/10.18524/2303-9914.2017.2(31).120897)
- Leontyeva, Y.Yu., & Vlashchenko, N.N. (2012). Applied aspects of the interregional clustering in tourism. *Business Inform*, 9, 177-183.
- Lipjanina, C.V. (2016). Cluster analysis of tourist recreational facilities of the administrative regions of Ukraine. *Scientific Bulletin of UNFU. Economic Series*, 26 (2), 277-282.
- Litvinova, T. (2020). North Caucasus image inside Russia in the context of tourism cluster development. *GeoJournal of Tourism and Geosites*, 28 (1), 275-288. <https://doi.org/10.30892/gtg.28122-469>
- Lloyd, S.P. (1982). Least squares quantization in PCM. *Transactions on Information Theory*, IT-28, 2, 129-137. <https://doi.org/10.1109/TIT.1982.1056489>
- Morozova, L., Morozov, V., Havanova, N., Litvinova, E., & Bokareva, E. (2016). Ensuring the development of tourism in the regions of the Russian Federation, with account of the tourism infrastructure factors. *Indian Journal of Science and Technology*, 9 (5), 1-13. <https://doi.org/10.17485/ijst/2016/v9i5/87599>
- Munoz, L., Hausner, V., Brown, G., Runge, C., & Fauchald, P. (2019). Identifying spatial overlap in the values of local-, domestic- and international tourists to protected areas. *Tourism Management*, 71, 259-271. <https://doi.org/10.1016/j.tourman.2018.07.015>
- Nazarova, O.U. (2016). Statistical study of travel service market. *Young Scientist. Economic Sciences*, 7 (34), 105-108.
- Neshchadym, L.M., & Tumchyk, S.V. (2017). Innovative methods of economic analysis of tourism enterprises in Ukraine. *Business Navigator*, 2 (41), 84-88.
- Ocheretin, D.V. (2015). Cluster analysis of travel flows in Ukraine. *Bulletin of Zaporizhzhia National University. Economic Sciences*, 3 (27), 66-75.
- Pedchenko, H.P. (2018). *Statistics*. Color Print, Melitopol, Ukraine.
- Pokataieva, O., & Kucherova, H. (2017). Innovative approaches to evaluation of tourist potential of Ukraine. *Eastern Europe: Economy, Business and Management*, 1 (06), 154-164.
- Rendon, E., Abundez, I.M., Gutierrez, C., Zagal, S.D., Arizmendi, A., Quiroz, E.M., & Arzate, H.E. (2011). A comparison of internal and external cluster validation indexes. Proceeding of the 5th WSEAS International conference on *Computer Engineering and Applications*, Puerto Morelos, Mexico. 29-31 January.
- Rodriguez-Rangel, M.C., & Sanchez-Rivero, M. (2020). Spatial imbalance between tourist supply and demand: The identification of spatial clusters in Extremadura, Spain. *Sustainability*, 12, 1651, 1-20. <https://doi.org/10.3390/su12041651>
- Sarrion-Gavilan, M.D., Benitez-Marquez, Dolores, M., & Mora-Rangel, E.O. (2015). Spatial distribution of tourism supply in Andalusia. *Tourism Management Perspectives*, 15, 29-45. <http://dx.doi.org/10.1016/j.tmp.2015.03.008>
- Sclyarenko, Ya., & Khanova, E. (2018). Analysis of factors of influence on the development of tourism space of the European Union. *Technology Audit and Production Reserves*, 5 (43), 9-14. <https://doi.org/10.15587/2312-8372.2018.146336>
- Vieira, A.C., & Santos, L.D. (2017). Tourism and regional development: a spatial econometric model for Portugal at municipal level. *FEP Working Papers* 589, Universidade do Porto, Faculdade de Economia do Porto.
- Vysochan, O.S., & Glushko, O.V. (2011). A systematic approach to the formulation and generalization of the problems of the tourism industry in Ukraine. *Bulletin of Donetsk University of Economic and Law*, 1, 165-171.
- Vysochan, O., Hyk, V., & Vysochan, O. (2020). Accounting support of financing IT-clusters activity in Ukraine. *Management Theory and Studies for Rural Business and Infrastructure Development*, 42 (4), 497-503. <https://doi.org/10.15544/mts.2020.51>
- Weidenfeld, A., Butler, R., & Williams, A.W. (2011). The role of clustering, cooperation and complementarities in the visitor attraction sector. *Current Issues in Tourism*, 14, 595-629. <https://doi.org/10.1080/13683500.2010.517312>
- Xiao-Ting, H., & Bi-Hu, W. (2012). Intra-attraction tourist spatial-temporal behaviour patterns. *Tourism Geographies: An International Journal of Tourism Space, Place and Environment*, 14 (4), 625-645. <https://doi.org/10.1080/14616688.2012.647322>
- Yalçinkaya, T., & Güzel, T. (2019). A general overview of tourism clusters. *Journal of Tourism Theory and Research*, 5 (1), 27-39. <https://doi.org/10.24288/jtr.465912>
- Yang, Y., & Wong, K.K.F. (2013). Spatial distribution of tourist flows to China's cities. *Tourism Geographies*, 15 (2), 338-363. <http://dx.doi.org/10.1080/14616688.2012.675511>
- Yurchenko, O. (2012). Ukraine's competitiveness in the world tourism market in the conditions of globalization. *Actual Problems of International Relations*, 111 (1), 307-315.
- Zbarsky, V., & Gribova, D. (2020). Regional features of infrastructure support of tourism development in Ukraine. *Agrosvit*, 8, 23-29. <https://doi.org/10.32702/2306-6792.2020.8.23>
- Zyma, O., & Holub, M. (2015). Ecotourism potential in Ukraine. *Herald of KNUTE*, 5, 53-62.