CLUSTER ANALYSIS IN THE STUDY OF MIGRATION

OLENA OVCHINNIKOVA[1]

Klastrová analýza v štúdiu migrácie

Abstract: *The article covers the application of cluster analysis methods in the study of external population movements in the Khmelnytskyi region, which allowed to reveal the features of migration processes in the region's districts and group them by similar characteristics. Cluster analysis is a method of multivariate statistical analysis, which includes data collection containing information about the sample objects, and organizes them into relatively homogeneous and similar groups. Migration process is a kind of demographic and demo-economic "investment" in recipient regions and, conversely, it causes ageing and deterioration of labour resource capacity in donor regions. Thus, information about the migratory clusters should help regional managers carry out immigration policy. Therefore, the application of cluster analysis in the study on management of migration is a relevant investigation today.*

Keywords: *migration, external displacement, cluster analysis, territory, region, Euclidean distance*

JEL Classification: C 50, C 82, J 61, R 10, F 22

1 Introduction

The analysis of migration processes and theoretical aspects of migration management problems in Ukraine has shown that it is necessary to solve quite a large number of social and economic issues and identify priority areas in management of "critical" regions of the country in terms of negative migration trends for a purposeful effective migration policy [1]. Under modern conditions, the problems of determining the real scale

[1] Olena Ovchinnikova, PhD, Ass. prof., 29016, Ukraine, Khmelnitsky, Str. Instytutska 11 Khmelnytsky National University, Department of ASME, e-mail: mouseinlove@mail.ru

of migration, establishing full information support in the study of this phenomenon, the study of formation process of volumes and directions of migration movements, optimisation of migration flows, improvement of social security measures for workers, forecasting future scale migratory movements, etc. become particularly acute.

Indigenous population attribute is the ability to move in space, resulting in individual states, and their subdivisions are capable to exchange population among themselves through migration processes. Migration (from the Latin "Migratio" – movement, resettlement) – is the process of movement of individuals in space associated with the relocation forever or for a very long time [2]. Researchers have always been interested in the study of migration – from demographers to economists, sociologists and political scientists. Managing the flow of migrants can solve many issues.

2 Materials and Methods

Formation of migration flows occurs under specific laws. The toughest ones are:

- Priority of economic factors in the formation of voluntary migration flows – bulk movement in space caused by the desire of individuals to improve their economic status;
- Differentiation of the level in the migration mobility depending on the quality characteristics of individuals, while an increased degree of participation in the migration movement is inherent to young, well-educated individuals who are not married, are in good health, have certain personality traits (initiative, courage);
- Operation of two oppositely directed flows – rather continued existence of migration flow as a result of the return of migrants who failed to gain a foothold in a new place;
- Inertia of migration flows – movement of population in the prescribed direction continues for some time after the termination of factors that led to formation of the flow;
- Decisive influence on the formation of volume displacement migration opportunities in the region, release and migration capacity of the region, which are indicators of the population – the more populous the region is through which the migration flow passes, the more popular the migration flow is;
- Dependence of migration on a distance – the closer the regions are, the

more intense migratory ties between them, other things are equal;
- The relationship of migration and mental characteristics of the population – the greater the degree of affinity of the inhabitants mentality in two regions are (the similarity of languages, customs, religion, prevailing values, and reproductive plants), the more intensive migration contacts are, and the presence of long and intensive exchange of population contributes to the convergence of social features.

Migration processes influence the natural reproduction by altering the age structure of the population and region output, transforming reproductive behaviour contingents of the population, changes in the conditions of life of migrants, which are reflected in the state of their health. Migration affects the functioning and development of the labour market; migration processes contribute to the development of individuals with increasing adaptability to different conditions of life, and willing to share skills and experience between residents of different regions, spread of useful knowledge. Migration processes lead to a redistribution of producers and consumers of goods and paid services between regions and their subdivisions, changing the quality characteristics of the labour force in different regions due to uneven participation in migration processes from different socio-demographic groups [3].

Migration processes often lead to filling jobs that are not in demand among the indigenous population – where individual professionals migrate abroad and work there for ordinary work (especially women – housewives, maids, caring for the elderly, etc.). The set of socio-economic conditions, the specificity of geographical location, non-organizational, legal preparedness for migration led to the emergence of the phenomenon of human trafficking and contributed to the fact that Ukraine is a donor country and country of transit [4].

The migration process, therefore, is a kind of demographic and demo-economic "investment" in recipient regions and, conversely, causes aging and deterioration of labour resource capacity in donor regions. It is especially the factor of urbanisation that is of a great economic importance, and it is an essential source of quantitative and qualitative growth. Thus, the migration process is always a form of expression for geographical and socio-economic content.

It is obvious by the fact that to secure a balanced policy in the field of migration management, there has to be developed a science-based distribution of regions and districts for the typical group of narrowly directed government policy and its specific characteristics and distinctive features.

When analyzing and forecasting social and economic phenomena, we often face the multidimensionality of their description. Therefore, methods of multidimensional analysis are the most effective tools for quantitative research on the socio-economic processes described by a large number of characteristics. These methods also include cluster analysis.

Cluster analysis appeared relatively recently – in 1939. An especially rapid development of cluster analysis took place in the sixties of the previous century. The prerequisites for this were the emergence of high-speed computers and recognition of classification by fundamental research.

Cluster analysis is a method of multivariate statistical analysis, which includes data collection, containing information about the sample objects, and organizes them into relatively homogeneous and similar groups [5]. Thus, the essence of cluster analysis is to implement object classification study using multiple computing procedures. As a result, "clusters" or groups of very similar objects are formed. Unlike other methods, this type of analysis makes it possible to classify objects not on one basis, but on several bases simultaneously. To do this, enter the appropriate parameters characterizing a degree of similiarity, by classification parameters.

The goal of cluster analysis is to find existing structures, resulting in the formation of groups of similar objects – clusters. However, its performance is also in bringing the structure to the objects. This means that clustering methods are necessary to identify patterns in the data, which is not easy to find by visual inspection or by experts.

Clustering can be done in two main ways (methods), including hierarchical and iterative procedures.

Hierarchical procedures are consistent actions to form clusters of different rank subordinated to each other by clearly established hierarchy. Most procedures are carried out by hierarchical agglomerative (unifying) action [6].

Iterative procedure is the formation of the original single-level data (same rank) rather than clusters hierarchically subordinate to each other.

One of the most common ways of iterative procedures serves k-means method (developed in 1967 by J. McQueen). Its application requires the following steps:

- Dividing the output data of the studied population into a given number of clusters;
- Calculation of multidimensional medium (centre of gravity) of selected clusters;
- Calculating Euclidean distances of each unit of totality to certain centres of cluster gravity and developing the distance matrix, based on the metric distances. Different metric distances are used, eg Euclidean distance (simple and considered), Manhattan, by Chebyshev, Minkowski, Mahalonobisa, etc.;
- Identifying new centres of gravity and new clusters.
- The most famous and widely used methods of clusters formation are also:
- A single method of communication;
- A complete method of communication;
- The method of Ward.

A single method of communication (a close neighbour method) involves joining the cluster units together if it is close (on a par similarity) to at least one object of this cluster. A complete method of communication (distant neighbour) requires a level of similarity facility (not less than limit level), which is expected to include in a cluster with any other cluster. According to the method of Ward, joining objects to clusters is carried out in case of the minimum rate of intra-sum of squared deviations. Due to this method, there are formed clusters of about the same size, which have the shape of hypersphere [7].

3 Clustering region for migration features

For a more detailed assessment of the structure and dynamics of the number of migrants in the Khmelnitskyi region, we will conduct its analysis in the context of individual cities and regions. The information indicates a trend that is characterized by a predominance of the number of arrivals to the region over the number of outgoers. This trend is typical for most cities and districts. Moreover, every year the gap between these categories of workers is significantly reduced.

Let us analyze, for example, the dynamics of the number of people who dropped out of the cities and districts of the Khmelnitsky region during the period 2011 – 2013 (Table 1)

Table 1

Assessment of the dynamics of migrants who left the Khmelnytskyi region *

Cities and villages of the region	The number of departures over the years, people			Deviation					
				absolute, people			relative, %		
	2011	2012	2013	2012 / 2011	2013 / 2012	2013 / 2011	2012 / 2011	2013 / 2012	2013 / 2011
Khmelnytskyi region (total)	26357	25120	23287	-1237	-1833	-3070	-4,69	-7,30	-11,65
including the city: Khmelnitskyi	4360	4043	3766	-317	-277	-594	-7,27	-6,85	-13,62
Kamianets-Podilskyi	3072	2787	2463	-285	-324	-609	-9,28	-11,63	-19,82
Netishyn	844	772	689	-72	-83	-155	-8,53	-10,75	-18,36
Slavuta	810	782	639	-28	-143	-171	-3,46	-18,29	-21,11
Starokostiantyniv	989	879	658	-110	-221	-331	-11,12	-25,14	-33,47
Shepetivka	862	775	760	-87	-15	-102	-10,09	-1,94	-11,83
districts: Bilogirskyi	515	441	446	-74	5	-69	-14,37	1,13	-13,40
Vinkovetskyi	454	443	396	-11	-47	-58	-2,42	-10,61	-12,78
Volochysskyi	1038	904	909	-134	5	-129	-12,91	0,55	-12,43
Hopodotskyi	899	924	812	25	-112	-87	2,78	-12,12	-9,68
Derazhnianskyi	768	699	647	-69	-52	-121	-8,98	-7,44	-15,76
Dunaevetskyi	1335	1353	1190	18	-163	-145	1,35	-12,05	-10,86
Iziaslavskyi	1055	930	961	-125	31	-94	-11,85	3,33	-8,91
Kamianets-Podilskyi	1259	1236	1201	-23	-35	-58	-1,83	-2,83	-4,61
Krasylivskyi	1088	942	992	-146	50	-96	-13,42	5,31	-8,82
Letychivskyi	535	535	486	0	-49	-49	0	-9,16	-9,16
Novoushytskyi	558	630	621	72	-9	63	12,90	-1,43	11,29
Polonskyi	715	795	791	80	-4	76	11,19	-0,50	10,63
Slavutskyi	748	723	678	-25	-45	-70	-3,34	-6,22	-9,36
Starokonstantinivskyi	769	748	644	-21	-104	-125	-2,73	-13,90	-16,25
Staposynyavskyi	415	439	416	24	-23	1	5,78	-5,24	0,24
Teofipolskyi	452	534	450	82	-84	-2	18,14	-15,73	-0,44
Khmelnytskyi	844	770	795	-74	25	-49	-8,77	3,25	-5,81
Chemepovetskyy	838	946	817	108	-129	-21	12,89	-13,64	-2,51
Shepetivskyi	476	490	495	14	5	19	2,94	1,02	3,99
Yapmolynetskyi	659	600	565	-59	-35	-94	-8,95	-5,83	-14,26

Source: based on to the Central Statistical Office in Khmelnytskyi region [8].

Any act of migration is connected with complex motives. They depend on subjective and objective reasons: on the one hand, migration taking place by personal desire – every citizen has the right and opportunity to choose for a residence and work that meet their diverse needs of financial, professional qualification and spiritual nature.

It should be emphasized that migration is reduced significantly in most cities and districts. The highest growth of migration reduction was from cities of Starokostiantyniv, Slavuta and Kamenets-Podilskyi as well as Vinkovetskyi, Derazhnianskyi, and Starokonstantynivskyi districts; but there are districts where the rate increased over the years. Thus, such districts include Shepetivskyi and Staposynyavskyi. That is, there is uneven distribution of districts migratory signs confirming the possibility of clustering.

It should be noted that the data that will be used in subsequent calculations are formed so that in the distribution areas of the group there was taken into account the general trend of population movement (arrivals and departures) for several years, and this factor appeared to be a significant addition to the total clustering task.

To eliminate the heterogeneity measurement of output data, all of them mentioned previously are normalized, i.e. expressed in terms of the ratio of these values to some value that reflects certain properties of the indicator. Rationing of input data for cluster analysis is limited to calculating the so-called standardized payment (1):

$$z_{ij} = \frac{x_{ij} - \overline{x_j}}{\sigma_j} \qquad (1)$$

where
x_{ij} – value of this observation;
xj – average value;
σj – standard deviation.
Z – contribution shows how many standard deviations separate this observation from the mean (Table 2).

Table 2

An array of standardized baseline data migration performance in the Khmelnitskyi region

Districts Regions (Objects of clustering)	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
Bilogirskyi	-0,7175	-0,9319	-0,9995	-1,2203	-1,0584	-1,0979
Vinkovetskyi	-0,7750	-1,1540	-0,9147	-1,2125	-1,2990	-1,3015
Volochisskiyi	0,6898	0,9720	0,4918	0,5842	0,4798	0,7876
Horodotskyi	0,0625	0,4660	0,3757	0,6622	0,3220	0,3926
Derazhnianskyi	0,2092	-0,0109	-0,2774	-0,2148	-0,2972	-0,2794
Dunaevetskyi	0,9546	2,0532	1,2924	2,3343	1,5565	1,9320
Iziaslavskyi	0,3819	1,0339	0,4290	0,6856	0,4009	0,9994
Kamiaanets-Podilskyi	2,4396	1,7765	2,2846	1,8782	1,8799	1,9768
Krasylivskiy	0,6841	1,1540	0,5954	0,7324	0,8190	1,1256
Letychivskyi	-0,7031	-0,8591	-0,6950	-0,8540	-0,8651	-0,9350
Novoushitskyi	-0,7203	-0,7754	-0,6228	-0,4837	-0,4589	-0,3853
Polonskyi	-0,2973	-0,2039	-0,2334	0,1594	0,0972	0,3071
Slavutskyi	-0,2311	-0,0837	-0,0984	-0,1212	-0,3327	-0,1531
Starokonstantynivskyi	-0,0930	-0,0073	-0,2837	-0,0238	-0,6600	-0,2916
Starosyniavskyi	-1,0110	-1,2960	-0,8739	-1,2281	-1,1412	-1,2201
Teofipolskyi	-1,6441	-1,1613	-1,5678	-0,8579	0,6533	-1,0816
Khmelnytskyi	1,9503	0,2657	2,0617	0,0620	1,9430	0,3233
Chemerovetskyi	0,1977	0,2439	0,1496	0,7479	-0,0921	0,4129
Shepetivskyi	-1,0081	-1,0739	-0,8268	-1,0294	-1,1451	-0,8984
Yarmolynetskyi	-0,3692	-0,4077	-0,2868	-0,6006	-0,8020	-0,6133

Source: own processing.

Table 2 – Z1 – Z6 – is the symbol of each object with a set of standardized data.

To determine the proximity of elements (points in multidimensional space) Euclidean distance is used, according to which the distance between the objects is the sum of the squares of the differences between the values of similar indicators combined for a single pair of objects (2):

$$d_{ij} = \sqrt{\sum_{k=1}^{m} (x_{ik} - x_{jk})^2} \tag{2}$$

where d_{ij} – Euclidean distance between *i-th* and *j-th* objects;
x_{ik} – value of *k-th* indicator for the *i-th* object;
x_{jk} – value of *k-th* indicator for the *j-th* object.

During the implementation process of payments we accept symbol of each district with the letter K index, which means its serial number and is subsequently adopted as the cluster (it includes all of the objects in one area). Here is Table 3, in which the initial calculations of distances matrix task clustering are performed.

A couple of the closest clusters is determined by the minimum value of the coefficient of Euclidean distance, which determines the closest objects, and after this operation combining them in a single cluster is performed. (In this case, it is the intersection of clusters K2 and K15 with the value of 0.0546, and these two clusters are subjects to association.) Next, using the method of "nearest neighbor" fusion of these clusters is performed and cluster with the highest serial number is removed from the matrix (i_e, K15).

Table 3

Initial matrix of distances between clusters

Кластер	$K_1=\{n_1\}$	$K_2=\{n_2\}$	$K_3=\{n_3\}$	$K_4=\{n_4\}$	$K_5=\{n_5\}$	$K_6=\{n_6\}$	$K_7=\{n_7\}$	$K_8=\{n_8\}$	$K_9=\{n_9\}$	$K_{10}=\{n_{10}\}$	$K_{11}=\{n_{11}\}$	$K_{12}=\{n_{12}\}$	$K_{13}=\{n_{13}\}$	$K_{14}=\{n_{14}\}$	$K_{15}=\{n_{15}\}$	$K_{16}=\{n_{16}\}$	$K_{17}=\{n_{17}\}$	$K_{18}=\{n_{18}\}$	$K_{19}=\{n_{19}\}$	$K_{20}=\{n_{20}\}$
K_1	0	0,0796	8,5035	6,0623	2,2445	22,8062	8,6372	27,8880	10,5704	0,1482	0,7882	3,2533	2,1976	1,9990	0,1282	2,1479	15,2975	5,3156	0,1091	0,7943
K_2	0,0796	0	9,7004	7,0015	2,8628	24,6690	9,8588	29,7812	11,9479	0,2958	1,1541	4,0072	2,7756	2,5099	0,0546	2,5837	16,5386	6,1363	0,1441	1,1055
K_3	8,5035	9,7004	0	0,4251	2,0846	3,7057	0,0820	5,9858	0,1476	6,7731	4,8407	1,7204	2,1769	2,5033	9,9224	9,9220	3,5907	0,6918	8,4471	4,3205
K_4	6,0623	7,0015	0,4251	0	1,1397	5,4225	0,4012	8,7138	0,8487	4,4794	2,8392	0,6299	0,9766	1,2930	7,0666	6,9655	4,7197	0,1490	5,8219	2,6293
K_5	2,2445	2,8628	2,0846	1,1397	0	11,0520	2,2768	14,4717	3,2306	1,4436	0,8387	0,4676	0,1286	0,1298	3,0602	4,1920	7,0187	0,8478	2,3398	0,5037
K_6	22,8062	24,6690	3,7057	5,4225	11,0520	0	3,5178	1,7903	2,5639	19,7138	15,9231	9,2447	10,9263	11,6222	24,8385	22,6770	6,3392	6,3470	22,3728	15,4785
K_7	8,6372	9,8588	0,0820	0,4012	2,2768	3,5178	0	6,3969	0,1632	6,8694	4,8091	1,6403	2,2102	2,5564	10,0162	9,8436	4,4697	0,6634	8,4424	4,4284
K_8	27,8880	29,7812	5,9858	8,7138	14,4717	1,7903	6,3969	0	4,7428	24,5895	20,7888	13,3364	14,8507	15,7029	30,1596	29,2463	4,3040	9,7724	27,7819	19,6591
K_9	10,5704	11,9479	0,1476	0,8487	3,2306	2,5639	0,1632	4,7428	0	8,6204	6,2867	2,5065	3,2705	3,7464	12,1114	11,4444	3,4494	1,3009	10,4539	5,8775
K_{10}	0,1482	0,2958	6,7731	4,4794	1,4436	19,7138	6,8694	24,5895	8,6204	0	0,3085	2,1514	1,3058	1,2062	0,3076	2,0329	13,1065	3,8610	0,1335	0,3267
K_{11}	0,7882	1,1541	4,8407	2,8392	0,8387	15,9231	4,8091	20,7888	6,2867	0,3085	0	0,9297	0,5969	0,6797	1,0676	1,8787	10,9958	2,3834	0,6228	0,2774
K_{12}	3,2533	4,0072	1,7204	0,6299	0,4676	9,2447	1,6403	13,3364	2,5065	2,1514	0,9297	0	0,2562	0,5241	3,9517	3,8919	6,9783	0,4928	3,0120	1,1415
K_{13}	2,1976	2,7756	2,1769	0,9766	0,1286	10,9263	2,2102	14,8507	3,2705	1,3058	0,5969	0,2562	0	0,0975	2,8483	3,8469	7,4930	0,7432	2,0775	0,4107
K_{14}	1,9990	2,5099	2,5033	1,2930	0,1298	11,6222	2,5564	15,7029	3,7464	1,2062	0,6797	0,5241	0,0975	0	2,6980	4,2158	8,4557	0,8749	1,9425	0,3466
K_{15}	0,1282	0,0546	9,9224	7,0666	3,0602	24,8385	10,0162	30,1596	12,1114	0,3076	1,0676	3,9517	2,8483	2,6980	0	2,1385	16,6925	6,2761	0,0973	1,2113
K_{16}	2,1479	2,5837	9,9220	6,9655	4,1920	22,6770	9,8436	29,2463	11,4444	2,0329	1,8787	3,8919	3,8469	4,2158	2,1385	0	16,3065	6,8422	2,1293	3,1188
K_{17}	15,2975	16,5386	3,5907	4,7197	7,0187	6,3392	4,4697	4,3040	3,4494	13,1065	10,9958	6,9783	7,4930	8,4557	16,6925	16,3065	0	5,6742	15,5554	10,1004
K_{18}	5,3156	6,1363	0,6918	0,1490	0,8478	6,3470	0,6634	9,7724	1,3009	3,8610	2,3834	0,4928	0,7432	0,8749	6,2761	6,8422	5,6742	0	5,0657	2,1561
K_{19}	0,1091	0,1441	8,4471	5,8219	2,3398	22,3728	8,4424	27,7819	10,4539	0,1335	0,6228	3,0120	2,0775	1,9425	0,0973	2,1293	15,5554	5,0657	0	0,7632
K_{20}	0,7943	1,1055	4,3205	2,6293	0,5037	15,4785	4,4284	19,6591	5,8775	0,3267	0,2774	1,1415	0,4107	0,3466	1,2113	3,1188	10,1004	2,1561	0,7632	0

Source: own processing.

The next step is performed repeatedly. Payments end when the matrix consists of two clusters as hierarchical clustering method requires combining of all incoming cluster in one. The final form of the matrix of distances is shown in Table 4.

Table 4

The ultimate matrix of distances between clusters

CLUSTER	$K_1=\{n_{1;2;15;10;19;11;20;16;3;7;9;4;18;5;13;14;12}\}$	$K_6=\{n_{6;8;17}\}$
K_1	0	30,1596
K_6	30,1596	0

Source: own processing.

The last step of clustering distribution of districts identified the level of migration activity in two groups, but we need to determine the optimal number of clusters, which boil down all the smallest ratios of distances between clusters of each step of the algorithm (Table 5).

Table 5

Minimum distance ratios between clusters of districts

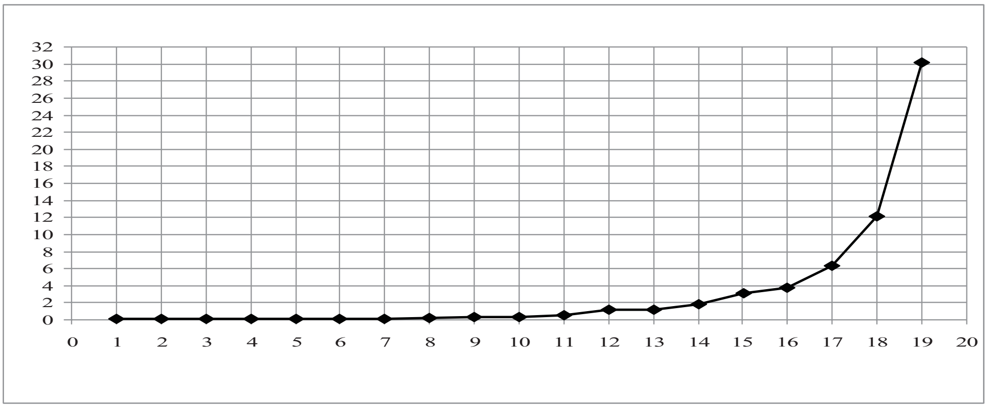
Step algorithm	Minimum ratio
1	0,0546
2	0,0796
3	0,0820
4	0,0975
5	0,1298
6	0,1335
7	0,1490
8	0,1632
9	0,2774
10	0,3076
11	0,5241
12	1,1397
13	1,2113
14	1,7903
15	3,1188
16	3,7464
17	6,3392
18	12,1114
19	30,1596

Source: own processing.

For visual perception of trends in change of minimum distances, we display the data in the graph (Figure 1).

Figure 1

The trend of changing minimum distance ratio depending on the steps of the algorithm



Source: own processing.

Since the optimal number of clusters is considered to be the amount that is determined as the difference between the number of observations and the total number of steps of the algorithm, after which the distance increases abruptly association – it is advisable to take the optimal distribution into four clusters (after step 16 there is a sharp increase, and the next steps show a clear and significant increase) [9].

One of the major factors that influence the reproduction of the population is environmental problems. Along with this indicator, it is closely linked urbanization. In recent years, the acute problem of urbanization has risen while maintaining favourable natural living conditions. The dependence of the physical condition of the person as well as the way of its activity on the characteristics of the natural environment is very high. Recent changes in natural conditions are related to the territorial organization of production and urbanization. In these regions, the level of pollution of air, surface water and ground exceeds the capacity of purification. This leads to the degradation of the environment that adversely affect health. Adverse environmental conditions cause about 20% of direct diseases.

4 Results and Discussion

Thus, the result of clustering districts of the Khmelnytskyi region is typing them into four groups (Table 6).

Table 6

Clustering districts of the Khmelnytskyi region

No	Clusters	Districts of the Khmelnitsky region
1	$K_1=\{n_{1;2;15;10;19;11;20;16}\}$	Bilogirskyi Vinkovetskyi Letychivskyi Novoushitskiy Starosynyavskyi Teofipolskyi Shepetivskyi Yapmolynetskyi
2	$K_3=\{n_{3;7;9;4;18;5;13;14;12}\}$	Volochisskyi Horodotskyi Derazhnianskyi Iziaslavskyi Krasylivskyi Polonskyi Slavutskyi Starokonstantynivskyi Chemerovetskyi
3	$K_6=\{n_{6;8}\}$	Dunaevetskyi Kamianets-Podilskyi
4	$K_{17}=\{n_{17}\}$	Khmelnytskyi

Source: own processing.

Thus, the groups, which were divided into districts by main indicators of migration, reflect a kind of vital activity of the region. The area where the regional center is situated is generally separated and cannot be equalized with other areas in terms of migration, because it is not typical for the entire region. Table 6 clearly shows the uneven distribution areas by groups, which means a number of features and factors that reflect the state of social and political groups of the region. Accordingly, we can state that regional authorities of the Khmelnytskyi region should explore and analyze four different types of migration and active compounds and form appropriate policies narrowly directed management of migration flows to achieve uniformity and positive results of the future status of population movement in the districts.

The feasibility of the use of cluster analysis methods in migration studies is dictated primarily by the fact that they help build scientific and reasonable classification, identify internal connections between population units that are observed. In addition, cluster analysis methods can be used for compression of information, which is an important factor in continuous and increasing complexity of flow statistics.

Population migration typically occurs under some constraints, which can deeply affect the structure of society and some other related aspects. The calculated results indicate that the distribution of the relative migration strength is governed by a shifted power-law relationship, and that the distribution of migration distances is dominated by a truncated power-law ship [10].

The leadership of the Ukrainian state has to solve a question of developing a new migration policy that would be aimed at reducing immigration flows of population and protection of workers outside the country.

References

[1] GUMENIUK, Yu. P.: The theoretical aspects of the study of the influence of international labor migration on the development of the economic system / Yu. P. Gumeniuk, G. P. Gogol // Proceedings of the National University "Lviv Polytechnic". *Management and Entrepreneurship in Ukraine: stages of formation and problems of development*. – № 22 (657), p. 484 – 497, 2009. ISBN 0542-33-10-39.

[2] PRYBYTKOVA, I. M.: *Fundamentals of Demography*: posib. dlya studentiv humanitarnykh i suspil'nykh fakul'tetiv vyshchykh navchal'nykh zakladiv / I. M. Prybytkova. – K. : ArtEk, 1995. – 256 p.

[3] Complex demographic forecast on Ukraine period to 2050 p. (kolektiv avtoriv) / for Ed. Corresponding Member. NASU, Doctor of Economics, Professor. E. M. Libanova. - K .: Ukrainian Centre of Social Reforms, 2006. 138 p.

[4] PUNDA, A. V.: The main causes and conditions of trafficking in Ukraine. Available at: http://www.dy.nayka.com.ua/?op=1&z=668 (access date February 12, 2016)

[5] Cluster analysis //Internet resource. Available at: http://www.statsoft.ru/home/textbook/modules/stcluan.html

[6] GRIGORUK, P. M.: Ekonometrichne modelyuvannya. Abstract lektsiy / P.M. Grigoruk. Hmelnitsky: KNU, 2005. 126 p.

[7] CHUBUKOVA, I. A.: Lecture: Methods of cluster analysis. Interactive methods. Online university of information technology // Internet resource. Available at: http://www.uran.donetsk.ua/~masters/2008/fvti/koshelyeva/library/index9.htm

[8] Official site of the Khmelnytsky Department of Statistics // Internet resource. Available at: http://www.km.ukrstat.gov.ua/ukr/index.htm

[9] OVCHINNIKOVA, O. R.: *Stochastic and fuzzy modeling migration*. Khmelnytsky, KhNU, 2014, 252 p. ISSN 978-966-330-215-7

[10] Statistical characteristics of dynamics for population migration driven by the economic interests. In: *Physica A: Statistical Mechanics and its Applications*, Volume 451, 1 June 2016, 123–134pp. / Jie Huo, Xu-Ming Wang, Ning Zhao, Rui Hao.