МАТЕМАТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ ЭКОНОМИКИ

УДК: 91:338(575.1)

ТАВАКАЛОВ ЖОНИБЕК ТАВАКАЛОВИЧ
младший научный сотрудник, Научно-исследовательский центр «Научные основы и проблемы развития экономики Узбекистана» при ТГЭУ,
e-mail: tsue_economics@mail.ru

МНОГОМЕРНЫЕ КЛАССИФИКАЦИИ УРОВНЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО РАЗВИТИЯ РЕГИОНОВ

Аннотация. Цель работы. *Определение уровня промышленного производства в регионах республики.* ***Метод или методология проведения работы.*** *В исследовании при классификации развития регионов использован метод многофакторного статистического анализа (кластерный анализ).* ***Результаты.*** *Полученные результаты показали, что, согласно расчетам, полученным с помощью многофакторного статистического анализа, г. Ташкент в период исследования стоит на самом высоком кластерном этапе, Ташкентская и Навоийская области — на этапе среднего развития, а остальные регионы находятся на этапе слабого развития.* ***Область применения результатов.*** *Полученные результаты исследования могут быть использованы при разработке стратегии социально-экономического развития регионов Республики Узбекистан, а также региональных и отраслевых программ.* ***Выводы.*** *В большинстве регионов республики целесообразно осуществить структурные изменения в промышленном производстве и повысить долю перерабатывающей сферы в общей структуре промышленности.*
Ключевые слова: *промышленный продукт, производство, многомерная классификация, кластерный анализ.*

TAVAKALOV JONIBEK TAVAKALOVICH
Junior Research Associate, Scientific and Research Center "Scientific Foundations and Problems of Development of the Economics of Uzbekistan" associated with Tashkent State Economic University,
e-mail: tsue_economics@mail.ru

MULTI-FACTORIAL CLASSIFICATIONS OF THE LEVEL OF SOCIO-ECONOMIC DEVELOPMENT OF REGIONS

Abstract. The goal of the study. *Determining the level of manufacturing production in the regions of the republic.* ***The method or methodology of performing the study.*** *In this study, when classifying the development of the regions, we have used the method of a multi-factorial statistical analysis (a cluster analysis).* ***The results.*** *The results obtained showed that, according to calculations obtained using a multi-factorial statistical analysis, the city of Tashkent for the study duration is at the highest cluster level, while Tashkent and Navoisk regions are at the stage of medium development, and the rest of the regions are at the stage of weak development.* ***The area of application of the results.*** *The results obtained may be used when developing a strategy of the socio-economic development of the regions of the Republic of Uzbekistan, as well as regional and industry-specific programs.* ***The conclusions.*** *In the majority of regions of the Republic it is advisable to perform structural changes in the industrial production and to increase the percentage share of the processing industry in the overall structure of manufacturing.*
Keywords: *an industrial product, production, a multi-factorial classification, a cluster analysis.*

1. Introduction

Classification of multidimensional observations is a complex process which requires a lot of efforts and complex calculations. In turn, a classification is carried out by the use of cluster analysis methods. The main aim of clustering is to form similar groups between the objects and analyze interrelations between them.

The relevance of clustering is specified with the nature of practical problems which it can solve, and with its economically actuality. In our research we elucidate the issues of classification of multidimensional observations, according to which Uzbekistan regions are divided into lower, medium and higher classes by their "development level of manufacturing industries". Classifying these regions by certain development level and identifying the regularity is an important issue where every regions is characterized with a set of specific indicators. Therefore, it is impossible to classify them by the same level. This, in turn, puts forward the new scientific problems to the issue of classification of objects, and a special attention is paid to the solution of this issue in the research.

A need for scientific research to the proposed methodology is that there studied not only the issue of classification of objects but also the issue of determining a disposition of objects to development in classified groups

For classification of objects, a parallel clustering method is used in the research. The idea of parallel cluster operations is that in each step of algorithm all the indicators are compared and calculated simultaneously. The main idea to develop different algorithms of parallel classification is to specify the method which provides a rapid solution of the objectives being set and reduces the options of selection.

2. Literature review

Although the first publications on cluster analysis started to emerge in late of 1930s of the last century, however, this method actively developed at the end of 1960s and beginning of 1970s (Soshnikova, Tamashevich, Uebe and Shefer, 1999). Later on, as a result of development of the information technologies there appeared new methods, modifications and algorithms of clustering, and emerged opportunities to massively process the labor-intensive and large-scale data (Everitt, Landau, Leese, et al., 2011).

A specific aspect of cluster analysis is that it is a common research method for almost all sciences. At present, this method is being widely used in medical science to classify the organisms and determine the hereditary diseases, in chemistry to classify the similar properties of a substance, and in environmental science to classify the changes and processes of the nature (Greenacre and Primicerio, 2013).

As theoretical analyses show that in many scientific sources the hierarchical, parallel and sequence clustering operations have been widely investigated (Rencher, 2002; Manly, 2005).

In scientific literature related to theoretical aspects of our research there have been investigated the issues such as clustering methods and their specific features, problems of interpretation and clustering of the results of cluster analysis (Yegorova and Khachatrian, et al., 2001; Milligan and Cooper, 1987), clustering methods and algorithms, particularly issues of clustering in conditions of uncertainty (Jain, Murty and Flynn, 1996), selection of variables and measuring the distance between objects and the issues of evaluation of a quality of clustering (Qin, 1999), problems emerging with optimal grouping of objects and alternative algorithms serving for overcoming these problems (Vakharia and Mahajan, 2000), hierarchical and nonhierarchical methods of cluster analysis and optimization of a number of clusters (Cornish, 2007), as well as contemporary and classical approaches to cluster analysis, specifically problem statement and algorithms of cluster analysis in conditions of uncertainty (Řezanková, 2014).

In scientific literature (Šipilova, 2015) directly related to practical aspects of our research there have been investigated the issue of classification of European countries by structural changes in industrial production.

Compared to researches mentioned above, our research has two specific aspects. Firstly, in our research we consider not only the issue of classification of objects but the issue of determining their disposition to development in classified groups. Secondly, a classification algorithm enabling "transferring the objects from class to class" is used in our classification issue.

3. Methodology

Aforementioned differing development level of manufacturing industrial products by sectors in the regions of the republic requires relatively specific approaches and this is done by multidimensional classification i. e. by cluster analysis method.

Cluster analysis —— multidimensional statistical method which allows classifying the objects with multiple signs of character. Also, cluster analysis, based on identifying the level similarities in appearances and signs related to each specific cluster units [2]. From here we derive a concept called "distance" and it is measured by the differences between signs and values of studied object. Certain distance limit can be measured differently. Usually, "Euclidean space" is used and it may come in two different ways.

– "simple" Euclidean space [3]:

$$\rho_E\left(X_i, X_j\right) = \sqrt{\left(x_i^{(1)} - x_j^{(1)}\right)^2 + \left(x_i^{(2)} - x_j^{(2)}\right)^2 + \ldots + \left(x_i^{(p)} - x_j^{(p)}\right)^2}$$

– "pulled" Euclidean space

$$\rho_{BE}\left(X_i, X_j\right) = \sqrt{\omega_1\left(x_i^{(1)} - x_j^{(1)}\right)^2 + \omega_2\left(x_i^{(2)} - x_j^{(2)}\right)^2 + \ldots + \omega_p\left(x_i^{(p)} - x_j^{(p)}\right)^2}$$

Also, in the cluster analysis method, concept such as the distance between the groups of the objects and the measurement of the gap of objects' two group is used. Namely, the measurement and the distance of mostly used relationship between objects' classes consists the following:

– "close neighbor" principle changing the distance [4]:

$$\rho_{\min}\left(S_l, S_m\right) = \min_{X_i \in S_l, X_j \in S_m} \rho\left(X_i, X_j\right)$$

– "far neighbor" principle changing the distance:

$$\rho_{\max}\left(S_l, S_m\right) = \max_{X_i \in S_l, X_j \in S_m} \rho\left(X_i, X_j\right)$$

It must be noted though that due to impractical possibility of adding the squared values between two points denoted with different quality signs, and in order to separate differences between each signs of the sample size and denoting it with relative measurement indicator, "normalized difference" formula is applied [5]:

$$z_{ij} = \frac{x_{ij} - \overline{x}_j}{\sigma_j}$$

Hence, with the help of [5] normalized difference matrix is formed and on the basis of average normalized Euclidean space objects will be attached to clusters, first, we take closest distances then objects moving away from each other will be taken. In the first step the object with the closest Euclidean space will be joined with the first object in the first cluster. Afterwards, we calculate the Euclidian space distances from clusters to other objects and the average normalized differences for this cluster on every unit. In this order, once the first average Euclidian space cluster reaches the boundary limit the second cluster will be created. Thus, step by step cluster attachment and a process of new cluster creation will be continued until all the predetermined critical Euclidian spaces limit $P_{critical}$ is reached and the attachment process is ended.

4. Empirical results

Currently, modernization of industrial sectors, fostering their technological updates calls for special attention in order to increasing the volumes of manufacturing in a country scale. Hence it is considered to be highly topical subject of the national economy to develop sectors of the economy and diversifying the range of products produced in this sector.

The changes during the recent years in the form quality and quantity of produced good in the regions of the country is the result of thoroughly planned economic reforms that were seen as the

development of industrial sectors of regions. Namely, in the electric power industry Tashkent region (795 bln.soum), Syrdarya region (684 bln.soum), Tashkent city (544,58 bln.soum), Navoi regions (399,2 bln.soum) performed with high figures in the country scale whereas chemical and oil chemistry industry Tashkent city (714,77 bln.soum), Tashkent region (646 bln.soum), Navoi region (523,2 bln.soum), Fergana region (396,54 bln.soum) and Andijan regions (229,8 bln.soum), in the light industry Fergana region (900,88 bln.soum), Tashkent city (726,11 bln.soum), Andijan region (664 bln.soum) and Tashkent regions (603 bln.soum) also, in the food industry Tashkent city (2257,8 bln.soum), Samarkand region (1094bln.soum) and Tashkent regions (1036 bln.soum) registered high results. (Table 1.) [1]

Here: $X^{(1)}$ — electric energy, $X^{(2)}$ — fuel, $X^{(3)}$ — black nonferrous metals, $X^{(4)}$ — chemistry and oil chemistry, $X^{(5)}$ — mechanical engineering and metals recycling, $X^{(6)}$ — construction materials industry, $X^{(7)}$ — light industry, $X^{(8)}$ — food, $X^{(9)}$ — flour and miller products.

It is however, development of manufacturing industrial products by sectors in the regions of the republic is not exactly the same. Because there are many factors that can affect the development of certain industries of the region. These are advantageous geographical placement of the region, level of natural resources abundance, number of labors and its quality, climate, nature, placement of the population and others.

Table 1

Sectorial structure of industrial product manufacturing (bln. soum)[1]

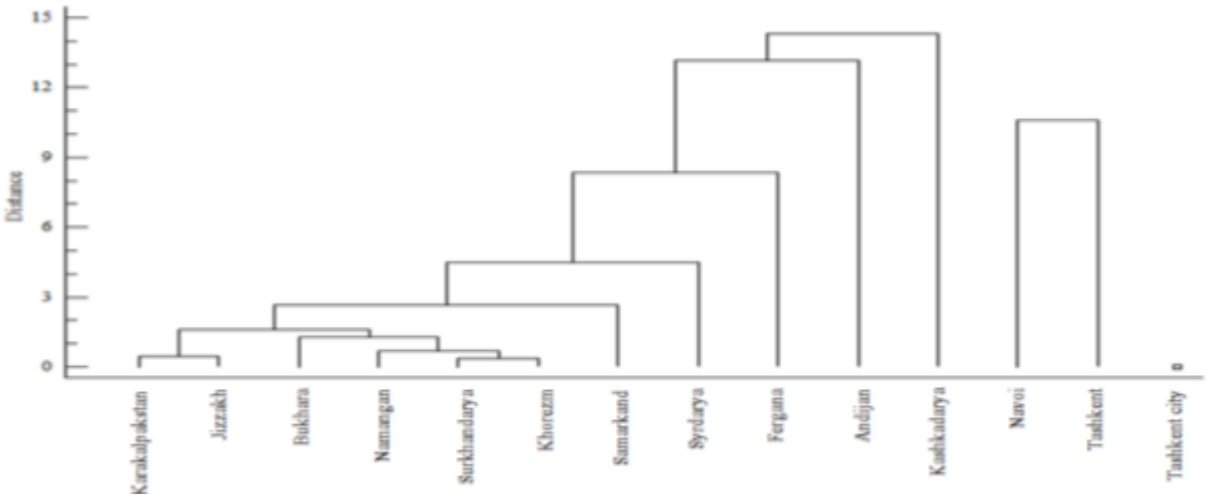
№	Regions	Volume production of economic sectors by regions								
		$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$	$X^{(5)}$	$X^{(6)}$	$X^{(7)}$	$X^{(8)}$	$X^{(9)}$
1	Karakalpakstan	197,3	75,7	9,2	66,6	6,7	54,1	203,9	126,5	59,9
2	Andijan	121,3	51,07	19,15	229,8	4852,0	70,23	664,0	210,7	127,7
3	Bukhara	47,9	878,0	8,32	10,4	62,4	135,0	531,0	277,0	91,6
4	Jizzakh	42,36	7,169	5,865	9,124	75,6	46,92	248,3	142,1	59,96
5	Kashkadarya	335,39	4108,6	11,18	128,57	39,13	83,85	419,24	240,37	201,24
6	Navoi	399,2	0,539	3291,0	523,2	102,5	690,5	188,8	107,9	64,73
7	Namangan	78,44	4,129	8,257	19,27	141,7	49,54	483,0	406,0	125,2
8	Samarkand	93,0	1,368	5,47	68,38	500,5	150,4	547,0	1094,0	172,3
9	Surkhandarya	53,34	148,59	2,85	0,95	34,29	76,2	355,28	140,02	114,3
10	Syrdarya	684,0	2,7	1,35	1,35	8,09	12,1	375,0	166,0	71,5
11	Tashkent	795,0	156,0	2583,0	646,0	419,0	525,0	603,0	1036,0	99,4
12	Fergana	142,45	1228,1	0,385	396,54	84,69	400,39	900,88	338,79	146,3
13	Khorezm	70,04	2,501	0,334	7,5	35,02	44,19	326,8	190,9	140,9
14	Tashkent city	544,58	2314,5	453,82	714,77	2552,7	487,86	726,11	2257,8	192,87

Now, we will be separating the industrial manufacturing in the regions by clusters via “simple” Euclidian space that is changing through “close neighbor” principle (Picture 1.).

This graph is called a dendrogram where manufacturing industrial products by sectors in the regions of the republic is divided into 3 clusters. In the following table below full details of each cluster and related information is given (Table 2.).

Thus, according to results after analyzing the resulting clusters via “simple” Euclidian space that is changing through “close neighbor” principle on manufacturing industrial products by sectors in the regions of the republic, in the first cluster we have 11 regions i.e.: Republic of Karakalpakstan, Andijan, Bukhara, Jizzakh, Kashkadarya, Namangan, Samarkand, Surkhandarya, Sirdarya, Fergana and Khorezm were included which makes up 78,57 % of territories from the total related clustered regions.

[1] Devised by author in 2012.



Picture 1. *Clustering the manufacturing industrial products by sectors in the regions of the republic*[2]

In the second cluster two regions of the republic i.e.: Tashkent and Navoi have been included, their share represent 14,29 % from total cluster related regions.

Table 2

Results of clustering the manufacturing industrial products by sectors in the regions[3]

Cluster serial number	Regions belonging to a calaster	Percentage of regions belonging to the cluster	Cluster structure
1	11	78,57	Republic of Karakalpakstan, Andijan, Bukhara, Jizzakh, Kashkadarya, Namangan, Samarkand, Surkhandarya, Sirdarya, Fergana and Khorezm
2	2	14,29	Tashkent and Navoi
3	1	7,14	Tashkent city

In the third cluster we have only one region of the republic which is Tashkent city. Its share represents 7,14 % of the total cluster related regions.

Now, we construct the following table after calculating average values for all the clusters that are being studied (Table 3.) [6].

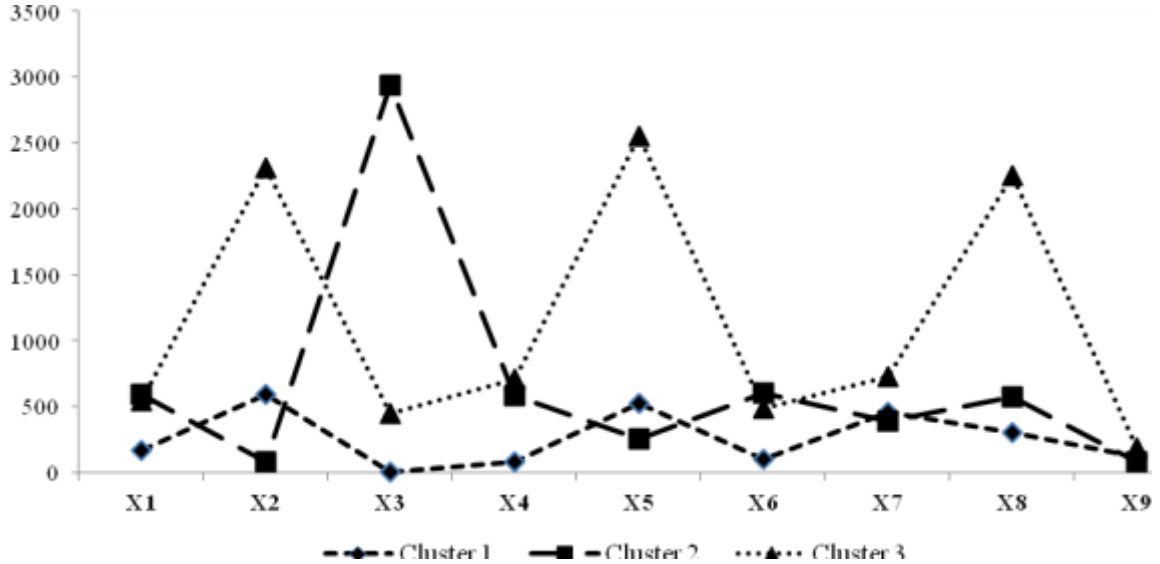
Table 3

Average value of results in clusters (bln.soum)[4]

Index	Cluster		
	$S^{(1)}$	$S^{(2)}$	$S^{(3)}$
$X^{(1)}$	169,59	597,1	544,58
$X^{(2)}$	591,66	78,27	2314,5
$X^{(3)}$	6,57	2937,0	453,82
$X^{(4)}$	85,32	584,6	714,77
$X^{(5)}$	530,93	260,75	2552,7
$X^{(6)}$	102,08	607,75	487,86
$X^{(7)}$	459,49	395,9	726,11
$X^{(8)}$	302,94	571,95	2257,8
$X^{(9)}$	119,17	82,06	192,87

[2] Conducted by author with Stat Graphics Centurion 16.1 software program.
[3] Devised by author.
[4] Conducted by author with Stat Graphics Centurion 16.1 software program.

Hence, the analysis of average results by clusters shows that $X^{(1)}$, $X^{(3)}$ and $X^{(6)}$ is in the $S^{(2)}$ cluster, and $X^{(2)}$, $X^{(4)}$, $X^{(5)}$, $X^{(7)}$, $X^{(8)}$ and $X^{(9)}$ are in the $S^{(3)}$ cluster representing higher values whereas $X^{(1)}$, $X^{(3)}$, $X^{(4)}$, $X^{(6)}$, $X^{(8)}$is in the $S^{(1)}$ cluster and $X^{(2)}$, $X^{(5)}$, $X^{(7)}$, $X^{(9)}$ results are represented their minimum values in $S^{(2)}$ cluster. This situation can be clearly seen in the picture below (Picture 2).



Picture 2. *Graphical representation of the average values of results in the clusters*[5]

Analysis of average values of parameters considered in each cluster shows that, $S^{(2)}$ — cluster electric energy, black nonferrous metals and construction materials industries are characterized by higher values, $S^{(3)}$ — cluster fuels, chemistry and oil chemistry, mechanical engineering and metals recycling, light industry, food, also flour and miller products industries are characterized by their higher values here.

Also, taking into account $S^{(1)}$ — cluster electric energy, black nonferrous metals and construction materials industries and food sectors' lower average values than $S^{(2)}$ and $S^{(3)}$ clusters, $S^{(2)}$ cluster with fuels, mechanical engineering and metals recycling, light industry, flour and miller products industries characterized by having lower average values than $S^{(3)}$ cluster.

In concluding remarks it is important to note that multidimensional classification of manufacturing industrial products by sectors in the regions allows making key scientific decisions with regards to development of country's industrial potential and conducting this process with the help of cluster analysis will serve the purpose well.

5. Consulations

In addition, multidimensional classification of manufacturing industrial products by sectors in the regions by cluster analysis serves the purpose of identifying least developed regions as well as the assessment of the sectors of industries that can help to make objective decisions by government. Apart from this, on the basis of multidimensional classification it would be possible to assess the specialization of regions by industrial sectors, future perspectives will be set and through disclosure of certain qualities of classified cluster, there will be separate solutions made by the government on each clusters.

Литература

1. Айвазян С. А. Прикладная статистика. Основы эконометрики: Учебник для вузов: В 2 т. 2-е изд., испр. — Т. 2: Айвазян С. А. Основы эконометрики. — М.: ЮНИТИ-ДАНА, 2001. — 432 с.
2. Айвазян С. А., Бежаева З. И., Староверов О. В. Классификация многомерных наблюдений. М.: "Статистика" 1974. 78 с.
3. Айвазян С. А., Мхитарян В.С. Прикладная статистика в задачах и упражнениях: Учебник для вузов. — М.: ЮНИТИ-ДАНА, 2001. — 197 с.
4. Б. Кобб 1989. Индекс устойчивого экономического благосостояния. (Х. Дэли & Джей Кобб, ЭЦП. Для

[5] Conducted by author with Stat Graphics Centurion 16.1 software program.

общего блага. Бостон: Beacon Press с. 401–455.)
5. Бурхардт Т., Ле Гранд Дж. и Piachaud Д. (2002) Степени исключения: разработка динамической многомерной меры, в Дж. Хиллс, Ле Гранд Дж. Piachaud (ЭЦП), Понимание социальной изоляции, Оксфорд: Оксфордский университет Пресс. С. 30–43
6. Данфорд М., Смит А. Догоняет или отстает? Экономические показатели и региональные траектории в "Новой Европе" Экон. Геогр. 2000; 76(2): 169–195.
7. Дубров А. М., Мхитарян В. С., Трошин Л. И. Многомерные статистические методы: Учебник. М.: Финансы и статистика, 2003. — 248 с.
8. Дюран Б., Оделл П. Кластерный анализ. Пер. с англ. М.: "Статистика", 1977. — 18 с.
9. Жамбю М. Иерархический кластер — анализ и соответствия. Пер. с фр. — М.: Финансы и статистика, 1988. — 82 с.
10. Зарова Е. В. Сбалансированная система показателей развития региона: статистическое обоснование и эконометрическое моделирование // Вопросы статистики. 2008. № 8. С. 59–66.
11. Кругман П. Возрастающая доходы и Экономическая география // Журнал Политической Экономии. — 1991. — №3. — С. 483–499.
12. Мандель И.Д. Кластерный анализ. — М.: Финансы и статистика, 1988. — 176 с.
13. Программы Stat graphics Centurion 16.1.
14. Рябцев В. М. Многомерный статистический анализ экономического развития регионов Российской Федерации / В.М. Рябцев, Е.И. Тихомирова, С. И. Чаплыгин. — Самара: Изд-во Самарск. гос. экон. акад., 2002. — 356 с.
15. Смирнов В. В. Повышение эффективности социально-экономического развития регионов: методология, оценка, эффективность / Аудит и финансовый анализ. — 2007. — № 6.
16. Соатов Н. М. Статистика. Т.: ТДИУ. 2003. С. 217–220.
17. Сошникова Л. А., Тамашевич В. Н., Уебе Г., Шефер М. Многомерный статистический анализ в экономике: Учеб. пособие для вузов. М.: ЮНИТИ-ДАНА, 1999. — 472 с.
18. Статистический ежегодник регионов Узбекистана. Ташкент. — 2013. С. 71–73.
19. Хабек П. Отчет об устойчивом развитии. Раскрытие влияния бизнеса на общество и окружающую среду. Sci. J. Maritime Univ. Szczecin. 2012; 31(103): 83–88.
20. Хохлова О. А. Методология статистического исследования экономики региона. — Иркутск: изд-во БГУЭП, 2006. — 276 с.
21. Эльдяева Н. А. Многомерный анализ влияния миграции на социально-экономическое развитие регионов // Экономика, статистика и информатика. Вестник УМО. — 2013. — № 6. — С. 121–126.

References:

1. Statistical yearbook of the regions of Uzbekistan. Tashkent. 2013. p. 71–73/
2. Soatov N. M. Statistics. (Tashkent, 2003). TSUE.
3. Ayvazyan S. A., Bejaeva Z. I., Staroverov O. V. Classification of multidimensional observations. (Moscow, 1974). Statistics.
4. Dubrov A. M., Mxitaryan V. S., Troshin L. I. Multidimensional Statistical Methods. Textbook. (Moscow, 2003). Finance and statistics.
5. Soshnikova L. A., Tamashevich V. N., Uebe G., Shefer M. Multidimensional statistical analysis in economics. Textbook for high schools. (Moscow, 1999). UNITY-DANA.
6. Krugman P. Increasing Returns and Economic Geography // The Journal of Political Economy. 1991. No. 3. P. 483–499.
7. Ayvazyan S. A., Mxitaryan V. S. Applied statistics in tasks and exercises. Textbook for high schools. (Moscow, 2001).
8. Duran B., Odell P. Cluster analysis. Translation from English. (Moscow, 1977). Statistics.
9. Mandel I.D. Cluster analysis. (Moscow, 1988). Finance and statistics.
10. Zhambyu M. Hierarchical cluster — analysis and conformity. Translation from French. (Moscow,1988). Finance and Statistics.
11. Stat Graphics Centurion 16.1 software program.
12. Dunford M, Smith A. Catching up or falling behind? Economic performance and regional trajectories in the "New Europe" Econ Geogr. 2000; 76(2): 169–195.
13. Habek P. Sustainability report. Disclosure the impact of business on society and the environment. Sci. J. Maritime Univ. Szczecin. 2012; 31(103): 83–88.
14. Burchardt T., Le Grand, J. and Piachaud D. (2002) Degrees of exclusion: developing a dynamic multidimensional measure, in J. Hills, J. Le Grand and D. Piachaud (eds) Understanding social exclusion, Oxford: Oxford University Press. P. 30–43.
15. Cobb J. B. 1989. The Index of Sustainable Economic Welfare. (In H.E. Daly & J.B. Cobb, eds. For the common good. Boston: Beacon Press. P. 401–455.
16. Ryabtsev V. M. Multivariate statistical analysis of economic development of the regions of the Russian Federation // V. M. Ryabtsev, H. R. Tikhomirov, S. I. Chaplygin. Samara: Izd-vo Samarsk. gos. ekon. akad., 2002. 356 s.
17. Aivazyan S. A. Applied Statistics. Basics Econometrics: Textbook for Universities: In 2 volumes, 2nd ed, M.: UNITY-DANA 2001. 432.
18. Smirnov V. V. Increase of efficiency of socio-economic development of regions: methodology, evaluation, efficiency / Audit and financial analysis, 2007. No. 6.
19. Khokhlova O. A. Methodology of statistical research of the economy of the region — Irkutsk: publishing house bguep, 2006. 276 p.

20. Zarova E. V. Balanced system of indicators of regional development: statistical reasoning and econometric modeling // Problems of statistics. 2008. No. 8. P. 59–66.
21. Eldaeva N. Multivariate analysis of the impact of migration on the socio-economic development of regions // Economy, statistics and Informatics. Bulletin UMO. 2013. No. 6. P. 121–126.