A STUDY OF THE STRUCTURE OF FINANCING AND THE CONDITION OF THE HEALTH CARE SYSTEM OF THE EUROPEAN UNION IN THE CONTEXT OF THE 2020 PANDEMIC

Gybkina N., Sidorov M., Storozhenko O.

The problem of classification of the member states of the European Union in the XXI century on some indicators of financing and a state of their systems of public health services, and also on some indicators of demographic structure of the population is considered. The principal components method and the cluster analysis method were used for the solution. This allowed to compress the space of the studied features without losing informativeness and to reflect on the plane of the division of the studied countries into clusters in terms of highlighting the general trends in the state of their medical field. To study the nature of the impact of the studied indicators on the course of the COVID-19 pandemic, the method of regression analysis was used. The ratio of deaths from COVID-19 to those cured of the disease was used as an indicator of the development of the pandemic. Among the many factors using stepwise regression those that have the greatest impact on this indicator were selected.

The state of the health care system as an indicator of quality of life

The level of development of the health care system of each population is one of the characteristics of quality of life. The United Nations and the World Health Organization recommend that the level of funding for this system be closely monitored with a view to reaching as many sections of the population as possible. [25, 27].

The state of the medical sector is characterized by a large number of socio-economic indicators that describe the amount and structure of funding, the peculiarities of national legislation and the organization of the health care system in each country. In the developed countries, the stable state of the health care system allows maintaining the volume and quality of health care at an acceptable level. But in the early 2020, health systems around the world faced an unprecedented challenge - the COVID-19 coronavirus pandemic, which can be seen as testing their

viability in critical conditions. According to the events of the first half of 2020, in a few days the spread of the infection in some countries has reached incredible proportions. Such a sharp increase in infections has jeopardized the state-of-the-art health care systems, including the most developed countries in the world, including the United States, France, Britain, Italy, Spain, and so on. The additional burden on the medical system is that COVID-19 often leads to complications and fatalities in the elderly and people with chronic diseases [27].

For this reason, research has recently become relevant in which various methods, in particular mathematical, are used to analyze the dynamics of the spread of COVID-19. Among the works on this topic, a large group consists of articles that use mathematical models based on differential and difference equations of epidemic distribution, time series analysis, autoregression [3, 4, 10, 17, 18, 21 - 23, 26]. In addition, some authors consider the problem of a pandemic in terms of its impact on various spheres of life, especially the economy [5, 15].

Therefore, it is of interest to conduct a statistical analysis of indicators that characterize the level of funding of the medical sector and the age structure of the population, in relation to the statistics of coronavirus incidence. Note that the absolute number of infections or deaths from COVID-19 is not an adequate indicator of the complexity of the situation in the country, as it does not take into account the total population. Therefore, relative indicators will be more significant. One of these is the ratio of deaths caused by COVID-19 to the number of patients treated for the disease.

Since the factors that characterize the state of funding, the level of medical care and quality of life are usually very many, it is advisable to use methods to reduce the dimension (principal components method), classification methods (cluster analysis) and regression analysis methods (stepwise regression).

Here is a brief description of each of these methods.

Description of some methods of multidimensional statistical analysis

When working with large data sets and when analyzing systems that are

characterized by a large number of parameters, there is a need to compress the data without losing their informativeness about the description of the system. This problem is solved by methods of reducing the dimension, in particular, the method of principal components. The essence of the principal components method is to construct on the basis of the whole set of initial data some subset of generalized indicators, and the number of generalized indicators used in the future is usually chosen much less than the dimension of the space of initial factors. The method analyzes the structure of the covariance (correlation) matrix of indicators and on the basis of this analysis identifies linear combinations of initial features with the largest contribution to the total variance. The algorithm of the principal components method is given, for example, in [1, 2, 7, 8, 19, 20].

Cluster analysis solves the problem of classifying the analyzed objects and allows you to divide the set of these objects into groups (clusters). Cluster analysis algorithms are based on processing a matrix of distances between objects calculated from a selected metric. One of the most frequently used methods of cluster analysis are probabilistic methods, in particular, the k-means method, the k-median method, and so on. Algorithms of these methods of cluster analysis are considered, for example, in [9, 12 - 14].

Another important task of statistical analysis is to study the presence and type of relationship between the initial factors and the response given by the system in the experiment. The fact of the relationship of variables can be established using correlation indicators, and the construction of a specific type of relationship between these variables is the task of regression analysis. The most commonly used regression analysis procedure is based on the application of the least squares method and allows to construct the dependence (in particular, linear) of the response on all factors that were considered in the experiment. The quality of this dependence can be described by several indicators, among which the multiple coefficient of determination is preferred. In the presence of a large number of factors, some of them may be dependent or related to each other, so in this case it is advisable to use a step-by-step regression algorithm [6, 16], which allows to choose only the most significant for the studied response.

Application of the principal components method and clustering to the analysis of the financing structure of health care systems in the countries of the European Union

The European Union is made up of a large number of countries whose political, economic and social status varies greatly from country to country. Therefore, we will further focus on the analysis of the problem described in the introduction for EU member states.

First, let's study the specifics of financial investment in the EU health care system through the consistent application of component and cluster analysis procedures. As initial data for this analysis we will consider some economic indicators that characterize the distribution of funds, and demographic indicators, in particular:

X_1 – Current health expenditure (% of GDP);

X_2 – Current health expenditure per capita (at current prices);

X_3 – Domestic government health expenditure (% of current health expenditure);

X_4 – Domestic public health expenditure (% of GDP);

X_5 – Domestic public expenditure on health (% of public expenditure);

X_6 – Domestic public health expenditure per capita (current prices);

X_7 – Domestic private health expenditure (% of current health expenditure);

X_8 – Domestic private health expenditure per capita (current prices);

X_9 – External health expenditure (% of current health expenditure);

X_{10} – External health expenditure per capita (current prices);

X_{11} – Actual expenses (% of current health care expenses);

X_{12} – Actual expenses per capita (at current prices);

X_{13} – Number of doctors (per 1000 people);

X_{14} – Number of nurses and midwives (per 1,000 people);

X_{15} – GDP per capita (at current prices);

X_{16} – GDP at current prices);

X_{17} – Population aged 0-14 years (% of the total population);

X_{18} – Population aged 15-64 years (% of the total population);

X_{19} – Population aged 65 and over (% of the total population);

X_{20} – Population, total.

Here $n = 27$ is the number of EU member states in 2017; $p = 20$ - the number of indicators that were selected for the study. The values of the selected indicators for each EU country in 2017 are taken from [24].

Cluster analysis aims to identify similarities in the financing systems of the medical sector, and the previous application of the principal components method will make it more visible. The calculation formulas of the principal components method are given in [7, 19]. The initial data are pre-normalized in order to bring them to one scale, because the initial information is presented in different units (US dollars, interest, quantity, etc.), and on the basis of normalized data is built a correlation matrix (table 1). Since the correlation matrix is symmetric, table 1 shows only the diagonal values and the values above the main diagonal.

As in previous works [7, 19, 20], we will display the results of the analysis in two-dimensional space, so for further clustering we will use two main components of the principal components method, which describe 59.77% of the contribution to the total variance. The first two maximum eigenvalues λ of the correlation matrix and their corresponding eigenvectors **L** have the form:

$$\lambda_1 = 7.64004 ;$$

$$\mathbf{L}^{(1)} = (-0.265927; -0.338643; -0.226118; -0.319254; -0.292187;$$
$$-0.340892; 0.225778; -0.258257; 0.0102735; -0.0389634; 0.262345;$$
$$-0.219748; -0.0649413; -0.285099; -0.271841; -0.145929;$$
$$-0.120987; 0.11647; -0.0126471; -0.120437)^T ;$$

$$\lambda_2 = 4.31387 ;$$

$$\mathbf{L}^{(2)} = (-0.35066; 0.109332; 0.0335511; -0.178047; -0.123892; 0.117308;$$
$$-0.0441984; 0.0426999; 0.407052; 0.398625; -0.0606547; 0.0266198;$$
$$-0.226555; 0.0899394; 0.289816; -0.192328; 0.191783; 0.335022;$$
$$-0.411905; -0.197509)^T.$$

Then the values of the first two main components (which are integral indicators based on the original set of indicators) take the form:

$$y^{(1)} = -0.265927x_1 - 0.338643x_2 - 0.226118x_3 - 0.319254x_4 - 0.292187x_5 -$$
$$-0.340892x_6 + 0.225778x_7 - 0.258257x_8 + 0.010274x_9 - 0.038963x_{10} +$$
$$+0.262345x_{11} - 0.219748x_{12} - 0.064941x_{13} - 0.285099x_{14} - 0.271841x_{15} -$$
$$-0.145929x_{16} - 0.120987x_{17} + 0.116470x_{18} - 0.012647x_{19} - 0.120437x_{20};$$

$$y^{(2)} = -0.235066x_1 + 0.109332x_2 + 0.033551x_3 - 0.178047x_4 - 0.123892x_5 +$$
$$+0.117308x_6 - 0.044198x_7 + 0.042700x_8 + 0.407052x_9 + 0.398625x_{10} -$$
$$-0.060655x_{11} + 0.026620x_{12} - 0.226555x_{13} + 0.089939x_{14} + 0.289816x_{15} -$$
$$-0.192328x_{16} + 0.191783x_{17} + 0.335022x_{18} - 0.411905x_{19} - 0.197509x_{20}.$$

According to the obtained ratios, the position of the EU member states in the coordinate system is calculated $(y^{(1)}, y^{(2)})$:

Austria: (-2,327; -0,318);

Belgium: (-3,309; 0.264);

Bulgaria: (3,934; -1,460);

Croatia: (1,479; -0,457);

Czech Republic: (0.568; -0.499);

Denmark: (-3,873; -0,071);

Estonia: (1,739; -0,233);

Finland: (-2,713; -1,180);

France: (-3,188; -1,208);

Greece: (2,525; -1,381);

Hungary: (2,795; -0,171);

Italy: (–0.613; –2.297);

Ireland: (-3,186; 2,580);

Latvia: (3,979; 0,079);

Lithuania: (2,524; -0,396);

Luxembourg: (-2,120; 8,510);

Malta: (0.256; –0.580);

The Netherlands: (-2,151; 0.147);

Poland: (2,787; 0.372);

Portugal: (0.979; –1.871);

Romania: (2,583; 0.225);

Slovakia: (1,835; 1,037);

Slovenia: (0.658; –0.255);

Spain: (–0.093; –1.099);

Sweden: (-4,693; -0,507).

Table 1

Correlation matrix of indicators for the main components method

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}
X_1	1	0,659	0,171	0,893	0,667	0,620	–0,161	0,679	–0,390	–0,319	–0,311	0,621	0,355	0,433	0,272	0,375	–0,017	–0,505	0,431	0,449
X_2		1	0,438	0,739	0,653	0,989	–0,443	0,823	0,206	0,298	–0,577	0,732	0,128	0,717	0,885	0,266	0,380	–0,143	–0,146	0,196
X_3			1	0,588	0,502	0,537	–0,999	–0,029	0,0005	0,190	–0,855	–0,037	0,064	0,488	0,421	0,221	0,150	–0,160	0,029	0,174
X_4				1	0,765	0,755	–0,580	0,531	–0,327	–0,194	–0,627	0,499	0,303	0,574	0,406	0,408	0,086	–0,500	0,356	0,447
X_5					1	0,650	–0,492	0,538	–0,381	–0,241	–0,584	0,389	0,239	0,624	0,413	0,333	0,312	–0,386	0,104	0,338
X_6						1	–0,543	0,728	0,231	0,338	–0,619	0,680	0,127	0,719	0,891	0,265	0,377	–0,151	–0,137	0,190
X_7							1	0,028	–0,027	–0,215	0,854	0,034	–0,059	–0,489	–0,436	–0,216	–0,151	0,145	–0,015	–0,168
X_8								1	0,036	0,050	–0,294	0,771	0,112	0,556	0,655	0,217	0,312	–0,103	–0,131	0,181
X_9									1	0,967	0,014	0,111	–0,191	0,034	0,578	–0,180	0,029	0,599	–0,519	–0,219
X_{10}										1	–0,160	0,097	–0,170	0,140	0,662	–0,130	0,021	0,554	–0,475	–0,165
X_{11}											1	0,018	0,031	–0,605	–0,525	–0,184	–0,274	0,104	0,104	–0,269
X_{12}												1	0,196	0,412	0,585	0,284	0,124	–0,095	–0,007	0,017
X_{13}													1	–0,143	–0,017	0,236	–0,383	–0,323	0,535	0,076
X_{14}														1	0,648	0,269	0,535	–0,259	–0,156	0,055
X_{15}															1	0,125	0,414	0,154	–0,416	0,008
X_{16}																1	–0,286	–0,243	0,401	0,614
X_{17}																	1	–0,151	–0,570	–0,206
X_{18}																		1	–0,726	–0,188
X_{19}																			1	0,300
X_{20}																				1

The application of cluster analysis (k-median method with 5 clusters based on Euclidean metrics) to the results of the principal components method allowed to visually display the location of countries and identify separate groups of countries that can be considered somewhat similar in the studied indicators. A graphical illustration of the location of countries in the two-dimensional plane and their division into clusters is shown in fig. 1.
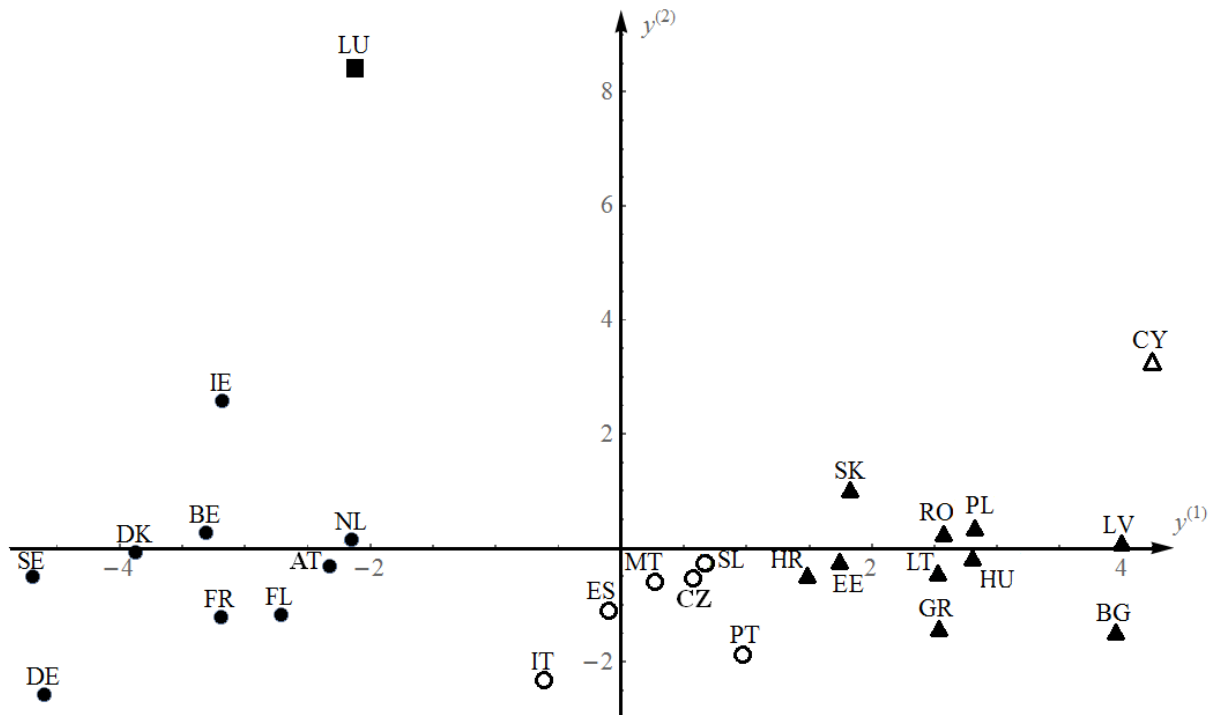

Fig. 1. Clustering by the method of k-medians (5 clusters) of EU countries according to some indicators that characterize the financing and condition of their health care systems

To indicate the countries in fig. 1 used two-letter abbreviations of their names in accordance with the system ISO 3166-1 alpha-2, defined in ISO 3166-1, which is part of the international standard ISO 3166: AT - Austria, BE - Belgium, BG - Bulgaria, CY - Cyprus, CZ - Czech Republic, DE - Germany, DK - Denmark, EE - Estonia, ES - Spain, FL - Finland, FR - France, GR - Greece, HR - Croatia, HU - Hungary, IE - Ireland, IT - Italy, LV - Latvia, LT - Lithuania, LU - Luxembourg, MT - Malta, NL - Netherlands, PL - Poland, PT - Portugal, RO - Romania, SE – Sweden, SL – Slovenia, SK - Slovakia.

As you can see, the results of component and cluster analysis take into account the characteristics of individual EU countries. In particular, Luxembourg stands out in a separate cluster due to the peculiarities of its political and economic situation, which affects all spheres of life in the country, including the medical sphere. Countries that have been in a difficult economic and political situation in recent years (Italy, Portugal, Spain) are also in a separate group, so it can be considered that this affects their health care systems. Consistent application of component analysis and cluster analysis allowed to establish the similarity of the state of health care systems and other EU countries according to the analyzed indicators.

Next, we will examine the effectiveness of the medical industry in the context of the coronavirus pandemic COVID-19 for EU member states. In response, we will consider the ratio of the total number of deaths from coronavirus in each country to the number of people who were cured of the disease (hereinafter "lethality/cure"). Both of these indicators are taken on the selected date (17.07.2020) for the entire period of observation since the beginning of the pandemic, assuming that in the future the mortality and morbidity rates will change slightly over time, i.e. the system has reached a saturated state.

We will use the factors $X_1, X_2,..., X_{20}$ given above as input data. Since their total number is large, we will use step-by-step linear regression to build a regression model. Statistics for selected factors and response are taken from [11, 24]. Still, $p = 20$, $n = 27$.

We will perform calculations according to the algorithm given, for example, in [6, 16]. Before applying the method, all initial data were normalized in order to bring them to one measurement scale. The correlation matrix of factors $X_1, X_2,..., X_{20}$ used to construct stepwise regression is given above in table 1. The correlation coefficients between the response and the factors $X_1, X_2,..., X_{20}$ have the form:

$(0.223; 0.239; -0.128; 0.091; 0.115; 0.155; 0.129; 0.502; -0.053; -0.034;$
$-0.216; 0.036; -0.010; 0.135; 0.149; -0.126; 0.049; -0.053; 0.010; 0.014).$

The step-by-step regression procedure at the significance level $\alpha = 0.15$ revealed that the greatest influence on the "lethality/cure" response was due to factors:

X_8 – Domestic private health expenses per capita (at current prices);

X_{12} – Actual expenses per capita (at current prices);

X_{17} – Population aged 0-14 (% of total population);

X_{20} – Population, total.

The model of linear multiple regression corresponding to this dependence has the form:

$$\hat{y} = -2.22045 \cdot 10^{-16} + 1.47999 x_8 - 1.05681 x_{12} - 0.34544 x_{17} - 0.30775 x_{20}.$$

The quality of the constructed regression model can be assessed using the table of regression analysis (table 2).

Table 2

Table of variance analysis

Source of variance	Sum of squares	Degrees of freedom	Mean square	F - ratio
Regression	18.48458	4	4.62115	11.939
Remainder	8.51542	22	0.38706	
Full	27.00000	26		

The multiple coefficient of determination is $R^2 = 0.684614$, therefore, in the constructed model 68.46% of the variation of the response is explained by the variables X_8, X_{12}, X_{17}, X_{20} selected by the stepwise regression procedure, which indicates the acceptable quality of the model.

Analyzing the obtained results, we can see that the coefficients for the variables X_{12}, X_{17}, X_{20} are included in the constructed regression model with a

73

negative sign. This means that an increase in each of these variables with a constant value of all other variables leads to a decrease in response, i.e. in terms of "lethality/curability" either the proportion of patients who die will decrease or the proportion of those treated will increase. Thus, it follows from the model that the variables X_{12}, X_{17}, X_{20} positively affect the dynamics of the process. The variable X_8 is included in the regression model with a positive sign, i.e. its increase will negatively affect the response. This can be explained by a complex and multilevel relationship between the variables analyzed.

It should be noted that the selected variables are linearly related to other variables, in particular, $X_2 = X_6 + X_8 + X_{10}$ where $X_8 = X_2 - X_6 - X_{10}$; and $X_{17} + X_{18} + X_{19} = 100\%$ where from $X_{17} = 100\% - X_{18} - X_{19}$. Therefore, the constructed regression can also be written through the variables X_2, X_6, X_{10}, X_{18}, X_{19}. In this case, the sign of the coefficient before all new variables, except X_2, will change to the opposite, which will correspond to the change of the direction of influence on the response from positive to negative (or vice versa, respectively).

Note that the multiple linear regression, built on all factors $X_1, X_2, ..., X_{20}$, will look like:

$$\hat{y}(\vec{x}) = 4.01236 \cdot 10^{-7} + 2.30722 x_1 - 18756.72041 x_2 - 1538.94463 x_3 -$$
$$-4.00017 x_4 - 0.21200 x_5 + 15451.06976 x_6 - 1540.15232 x_7 + 4183.44722 x_8 -$$
$$-37.19776 x_9 + 147.58759 x_{10} - 1.00525 x_{11} - 0.31234 x_{12} - 0.57789 x_{13} +$$
$$+0.03213 x_{14} - 3.50467 x_{15} - 0.14067 x_{16} - 2.31391 \cdot 10^7 x_{17} - 2.76483 \cdot 10^7 x_{18} -$$
$$-3.32613 \cdot 10^7 x_{19} - 0.07258 x_{20}.$$

This model corresponds to the value of the multiple coefficient of determination $R^2 = 0.928916$, therefore, selected by stepwise regression 4 variables X_8, X_{12}, X_{17}, X_{20} are almost 74% of the coefficient of determination of the general regression model, i.e. the other 16 variables in the general model account for only 26% of the response.

Analysis of the obtained results

The statistical analysis procedures used in the work today are quite promising methods of processing and analysis of complex, multidimensional, multidimensional and multidimensional data. The authors attempted to apply these methods to identify patterns in the data describing the state and financing of the medical sector, in particular, in the crisis, which was the pandemic of the coronavirus COVID-19 for all countries.

The integrated application of component analysis and cluster analysis has made it possible to group EU countries by the level of funding for the health care system, some indicators of its state and demographic indicators, as well as to establish similarities in these indicators for different countries. The obtained results are in good agreement with the general conclusions made by the authors in previous works [7, 20] and are in line with generally accepted trends. Preliminary application of component analysis additionally allows to significantly reduce the number of analyzed variables without losing informativeness, and its results can be used in other statistical procedures.

Another way to reduce the set of features is not to combine them into integral indicators, but to choose the most significant of the whole set of initial features under consideration. The use of stepwise regression allowed many variables describing the financing and state of the health care system, as well as the age structure of the population, to select those that have the most significant impact on the medical response to COVID-19 and to analyze their significance. Note that the choice as a response to an indicator different from that considered in this paper, will lead to the selection of another group of the most significant initial features. In the case of a large set of initial features, stepwise regression allows to reduce the time for further work with the constructed model by significantly reducing the variables used in it.

Summarizing the results of the study, it should be noted that the analysis is not exhaustive. Rather, it can be seen as an initial step in a comprehensive (involving experts not only in applied mathematics but also in medicine, economics, sociology,

etc.) analysis of the causes and consequences of the COVID-19 pandemic in 2020, in particular in terms of responsiveness health systems of individual countries to critical situations caused by mass morbidity and complications of diseases with a high rate of spread.

REFERENCES

1. Aivazyan, S. A., Bezhaeva, Z. I., Staroverov, O. V. (1974), *Classification of multivariate observations* [*Klassifikatsiya mnogomernykh nablyudeniy*], Moscow, Statistics, 240 p.

2. Anderson, T. (1963), *Introduction to multivariate statistical analysis* [*Vvedeniye v mnogomernyy statisticheskiy analiz*], Moscow, GIFML, 500 p.

3. Palmin, R. S., Palmin, S. A., Pertsev, S. F., Boyarintsev, V. V. (2020), "Nique for predicting parameters of the epidemic process caused by COVID-19", *Kremlin Medicine Journal*, No. 2, P. 14–21. DOI: 10.26269/z81m-6646

4. Brovchenko, I. (2020), "Development of a mathematical model of the spread of the COVID-19 epidemic in Ukraine" ["Rozrobka matematychnoyi modeli poshyrennya epidemiyi COVID-19 v Ukrayini"], *Worldview*,No. 2 (82), P. 2–14.

5. Vasiev, M., Bi, K., Denisov, A., Bocharnikov, V. (2020), How COVID-19 Pandemics Influences Chinese Economic Sustainability, *Foresight and STI Governance*, Vol. 14, No. 2, P. 7–22. DOI: 10.17323/2500-2597.2020.2.7.22

6. Gibkina, N. V., Sidorov, M. V., Storozhenko, O. V. (2016), "Econometric modeling of inflation in Ukraine in 2000 – 2015" ["Ekonometrychne modelyuvannya rivnya inflyatsiyi v Ukrayini u 2000 – 2015 rokakh"], *Information technologies and innovations in economics, project and program management*: monograph for the general ed. by V. Timofeeva, I. Chumachenko, Kharkiv, P. 104–120.

7. Gibkina, N. V., Sidorov, M. V., Storozhenko, O. V. (2017), "Classification of the European Union by the main socio-economic indicators by the method of main components" ["Klasyfikatsiya krayin Yevropeysʹkoho Soyuzu za osnovnymy sotsialʹno-ekonomichnymy pokaznykamy metodom holovnykh component"], *Mathematical models and new technologies for managing economic and technical systems*: monograph for the general ed. by V. Timofeeva, I. Chumachenko, Kharkiv, P. 116–133.

8. Dubrov, A. M., Mkhitaryan, V. S., Troshin, L. I. (2003), *Multivariate statistical methods* [*Mnogomernyye statisticheskiye metody*], Moscow, Finance and statistics, 352 p.

9. Kim, J. O., Mueller, C. W., Klekka, W. R. (1989), *Factorial, discriminant and cluster analysis*, Moscow, Finance and statistics, 215 p.

10. Koltsova, E. M., Kurkina, E. S., Vasetsky, A. M. (2020), "Mathematical modeling of the spread of the COVID-19 coronavirus epidemic in a number of European, Asian countries, Israel and Russia" ["Matematicheskoye modelirovaniye rasprostraneniya epidemii koronavirusa COVID-19 v ryade yevropeyskikh, aziatskikh stran, Izraile i Rossii"], *Problems of Economics and Legal Practice*, Vol. XVI, No. 2, P. 154–165.

11. Coronavirus - Statistics by country (17.07.2020) - Map of infections, graphs, available at : https://index.minfin.com.ua/reference/ coronavirus/geography/

12. Lytvyn, V. V., Pasichnyk, V. V., Nikolsky, Yu. V. (2015), *Data and knowledge analysis* [*Analiz danykh ta znan'*], Lviv, "Magnolia 2006", 276 p.

13. Mandel, I. D. (1988), *Cluster Analysis* [*Klasternyy analiz*], Moscow, Finance and Statistics, 176 p.

14. Aivazyan, S. A., Buchstaber, V. M., Enyukov, I. S., Meshalkin, L. D. (1989), *Applied Statistics: Classification and Dimension Reduction* [Prikladnaya statistika: Klassifikatsiya i snizheniye razmernosti], Moscow, Finance and Statistics, 607 p.

15. Stoyanova, R. (2020), "Analysis of unemployment before the pandemic of KOVID 19 in Bulgaria", *Management and education*, Vol. 16 (1), P. 20–24.

16. Afifi, A. A., Azen, S. P. (1979), *Statistical Analysis: A Computer Oriented Approach*, 2nd edition, New York, Academic Press, 450 p.

17. Anastassopoulou, C., Russo, L., Tsakris, A., Siettos, C. (2020), "Data-based analysis, modelling and forecasting of the COVID-19 outbreak", *PloS ONE*, Vol. 15, No. 3, P. 1–21. DOI: 10.1371/journal.pone.0230405

18. Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., Colaneri, M. (2020), "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy", *Nature Medicine*., Vol. 26, P. 855–860. DOI: 10.1038/s41591-020-0883-7

19. Gybkina, N., Sidorov, M., Storozhenko, O. (2019), "Research of the economic development structure of the regions of Ukraine with the methods of multi-dimensional statistical analysis", *Information systems and innovative technologies in project and program management* : Collective monograph edited by I. Linde, I. Chumachenko, V. Timofeyev, Riga, ISMA, P. 280–293.

20. Gybkina, N., Sidorov, M., Storozhenko, O. (2019), "To classification of the countries of the European Union for some social indicators", *Big Data processing: methods, models and information technologies* : monograph, edited by Oleg I. Pursky, Steyr, Austria, Shioda GmbH, P. 49–58.

21. Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R. M. (2020), "Early dynamics of transmission and control of COVID-19: a mathematical modelling study", *The Lancet Infectious Diseases*, Vol. 20, P. 553–558. DOI: 10.1016/S1473-3099(20)30144-4

22. Peng, L., Yang, W., Zhang, D., Zhuge, C., Hong, L. (2020), *Epidemic analysis of COVID-19 in China by dynamical modeling*, arXiv preprint arXiv:2002.06563, 11 p.

23. Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N. (2020), "The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study", *The Lancet Public Health*, Vol. 5, P. e261–e270. DOI: 10.1016/S2468-2667(20)30073-6

24. The World Bank, available at : http://www.worldbank.org

25. United Nations "Peace, dignity and equality on a healthy planet", available at : https://www.un.org/en/.

26. Unlu, E., Leger, H., Motornyi, O., Rukubayihunga, A., Ishacian, T., Chouiten, M. (2020), *Epidemic analysis of COVID-19 Outbreak and Counter-Measures in France*, medRxiv 2020.04.27.20079962, 10 p. DOI: 10.1101/2020.04.27.20079962

27. WHO – World Health Organization, available at : https://www.who.int/en/