Analysis of life expectancy in Russian regions

Olga Kapitanova[1,*], *Iurii* Kuznetsov[1], *Alexey* Semenov[1], *Veronika* Kisova[2], and *Irina* Koneva[2]

[1]National Research Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russia
[2]Minin Nizhny Novgorod State Pedagogical University (Minin University), Nizhny Novgorod, Russia

Abstract. This study deals with the analysis of the factors that determine life expectancy in Russia at the macro level. Among the available statistical indicators, those were selected that can be considered as a macroeconomic alternative to the basic microeconomic determinants affecting the life expectancy of an individual. As a result of the regression analysis, significant indicators were identified, which included the marriage rate, the unemployment rate, the number of students, the number of hospital beds, the gross regional product, budget expenditures on health care, and the region's subtropical climate. A cluster analysis of the regions was also carried out for these indicators, with the exception of the last one, and it was shown that the regions are distributed into 5 clusters. The cluster, which includes Moscow and St. Petersburg, where there is a high life expectancy due to the metropolitan area of these regions and good medicine, is singled out separately. The regions of the North Caucasus are also characterized by high life expectancy, but here it is due to completely different reasons: religious and climatic features. The Tyumen, Magadan, Sakhalin regions and the Chukotka Autonomous Okrug are united into a separate cluster. These regions are characterized by the highest GRP, but low life expectancy and the lowest number of students.

1 Introduction

The Government of the Russian Federation and regional authorities should be guided by those strategic goals that are formulated in Decree No. 474 of the President of the Russian Federation on the approval of national development goals for the period up to 2030, which was issued in July 2020 [1]. This document contains the main indicators of the quantitative assessment of the degree of goal implementation. The goals and indicators are deciphered in a unified plan to achieve the national development goals of the Russian Federation for the period up to 2024 and for the planned period up to 2030 (hereinafter referred to as the Plan), which was prepared by the Government of the Russian Federation [2].

Life expectancy of the population, which, according to the document, should be increased to 78 years by 2030, is considered to be one of the most important indicators in the plan. One should note here that this indicator has grown in Russia by 8 years over the past 20 years and amounted to 73.3 years in 2019. Naturally, the pandemic led to a number of negative consequences, so that life expectancy was reduced to 71.5 years. The

[*] Corresponding author: kapitanova@iee.unn.ru

importance of life expectancy in describing the economic development of a country is undeniable. For example, studies [3, 4] suggest that it is a significant explanatory factor in describing the relationship between human capital and economic growth.

On the scale of the Russian economy, not only the average life expectancy on a national scale is of great importance, but also similar indicators for the regions of Russia., This is particularly due to the fact that the territory of the country has a large extent and is located in different climatic zones and natural conditions, with different economic, cultural and social prerequisites, so that territorial differentiation is of great importance. Therefore, any measures that are developed to improve the situation should take into account the difference in the development of regions. This can be confirmed, for example, in [5].

It is often difficult to determine which factor affects a particular process. Most often, a number of factors influence, different combinations of which give a different effect on the result. Therefore, micro level factors (per person) are determined. Determination of such factors at the macro level – on a national scale, using the available statistical indicators is a very difficult research task.

There is a fairly large number of researches devoted to the study of factors affecting life expectancy in different countries: most often these are OECD countries (for example, [6-8], or USA (e.g. [9]). Health care spending, GDP per capita, labor productivity, population, infant mortality rates, cancer deaths, suicide rates, physician density, hospital bed density, social spending, the prevalence of chronic respiratory diseases and others are distinguished among the factors.

In recent years one can mention the studies of [10-13] among Russian scientists. They consider various aspects that affect the life expectancy of Russia and the differentiation of regions in terms of this indicator, but this paper offers a slightly different point of view on the problem.

The objectives of the study are to analyze the macroeconomic indicators for 2019 that could affect the life expectancy of the population in 2020, and highlight the most significant of them. Based on the identified factors, this study performs clustering of regions and studying the results. The goals set by the authors seem to be very relevant, which is confirmed by the analysis of sources. It shows that similar studies have not been previously conducted.

Section 2 of this paper describes the data on the basis of which the study was carried out and the methods that were used in its implementation. Section 3 is devoted to describing the results of regression and cluster analysis. Section 4 contains a discussion allowing to analyze the results and draw conclusions. At the end of the study a conclusion and a list of references are presented.

2 Materials and methods

This paper is devoted to the study of a number of indicators, which, to a certain extent, are macroeconomic analogues of the factors that determine the life expectancy of a person. The objectives of the study are to identify the determinants that are associated with life expectancy in the macroeconomic aspect, and on their basis to analyze clusters consisting of regions in which similar development trends are observed.

To achieve these goals, the following research objectives were formulated:

1) identify statistical indicators that can be used as macroeconomic analogues of factors affecting human life expectancy;

2) identify among them those that correlate with life expectancy in the regions of the Russian Federation;

3) build a regression model of life expectancy in the regions of the Russian Federation;

4) cluster the regions according to the selected indicators;

5) analyze the results.

To solve the tasks set, such methods of scientific research as analysis, synthesis, comparison and generalization are used, as well as economic and statistical methods that allow processing statistical data on various indicators, in particular, correlation, regression and cluster analyses are applied. Correlation and regression types of analysis are aimed at determining and describing the closeness of the relationship between statistical values. Cluster analysis is the division of a set of objects into groups, so that the objects have similar characteristics within the group, and between groups they differ greatly.

Statistical data for 2019 was obtained from the website of the Federal State Statistics Service [14]. The consideration included those macroeconomic indicators of the activities of the regions, which, to a certain extent, can be an analogue for the factors that determine the life expectancy of a person: health, education, marital status, income, eating habits, bad habits, climate, degree of environmental pollution, etc.:

population, at the end of the year, in thousand people, X_1;

total marriage rates, pcs. per 1000 people population, X_2;

general divorce rates, pcs. per 1000 people population, X_3;

unemployment rate, %, X_4;

median per capita money income of the population, rub. per month, X_5;

the number of people with incomes below the subsistence level, %, X_6;

consumption of meat and meat products per capita, kg per year, X_7;

consumption of milk and dairy products per capita, kg per year, X_8;

sugar consumption per capita, kg per year, X_9;

number of students enrolled in bachelor's, specialist's, and master's programs, pers. per 10,000 people population, X_{10};

number of graduate students, persons, X_{11};

number of hospital beds, pcs. per 10,000 people population, X_{12};

the number of theater spectators, pers. per 1000 people population, X_{13};

gyms, pcs., X_{14};

number of library users, thousand people, X_{15};

emissions of pollutants into the atmospheric air from stationary sources, thousand tons, X_{16};

fresh water use, million cubic meters, X_{17};

gross regional product, million rubles, X_{18};

population using the Internet every day or almost every day, %, X_{19};

the number of researchers with academic degrees, persons, X_{20};

the level of innovative activity of organizations, %, X_{21};

expenditures of the consolidated budgets on healthcare, billion rubles, X_{22}.

Two dummy variables were also formed to describe the climatic zone of the regions. Variable X_{23} is responsible for belonging to the arctic and subarctic climate, X_{24} – to the subtropical climate zone.

3 Results

Table 1 shows the results of the regression analysis. First, model 1 was created, the choice of factors for which is determined [15]. However, the quality of the constructed model is not too high, although the variables under consideration are significant. Next, a correlation analysis was performed and the most correlated factors with Y were selected (X_1, X_2, X_3, X_4, X_9, X_{10}, X_{11}, X_{12}, X_{18}, X_{20}, X_{22}, X_{23}, X_{24}), which were included in the regression. Some of the coefficients in the new model turned out to be statistically insignificant, so the corresponding factors were successively excluded, resulting in the specification Model 2.

Table 1. Results of regression analysis.

	Model 1	Model 2
Constant	71,516 (0,368)	76,489 (1,412)
X_2		-0,537 (0,173)
X_4		0,125 (0,050)
X_{10}		0,006 (0,002)
X_{12}		-0,063 (0,012)
X_{18}	-0,184 (0,060)	0,087 (0,042)
X_{22}	0,033 (0,008)	0,014 (0,006)
X_{24}		2,301 (0,468)
R^2	0,159	0,724
F- criterion	8,681	31,294
DW	0,954	1,688
Standard errors are given in parentheses; all parameters are significant at the 5% level. n=82		

Source: compiled by the authors.

To identify trends in the development of the regions of the Russian Federation according to the parameters of model 2 (except X_{24}), clustering was carried out. All regions were divided into 5 clusters using the k-means method in the Statistica package. All data were standardized to avoid the negative impact of variable scales. Figure 1 shows the dynamics of average values by clusters.

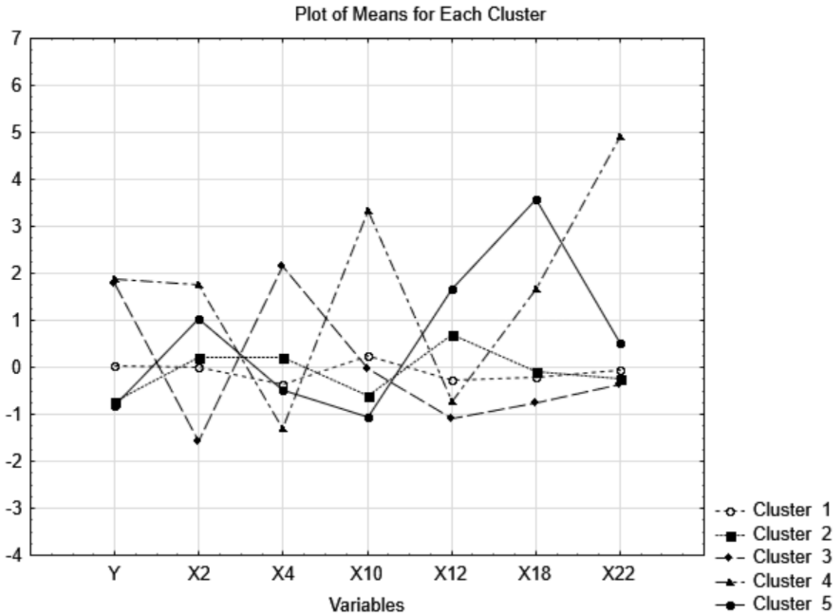

Fig. 1. Dynamics of average values of indicators by clusters. *Source*: compiled by the authors.

The first cluster (46 regions) consists of "middle peasants" in almost all indicators. The second (22 regions) and fifth (4 regions) clusters are characterized by low life expectancy. The fifth cluster includes the large northern regions and is characterized by the lowest number of students, the highest number of hospital beds and the highest GRP per capita. The remaining two clusters are characterized by relatively high life expectancy. The third cluster included 8 regions of the North Caucasus, the fourth – Moscow and St. Petersburg. Corresponding differences for all indicators are evident from Figure 1. It should be noted that climate variables were not included in the cluster analysis, but their indirect influence is obvious. The distribution of regions by clusters is shown in Figure 2.

Fig. 2. Distribution of regions by clusters. *Source*: compiled by the authors.

4 Discussion

The regression model shows that life expectancy is inversely related to the number of marriages and the number of hospital beds. While the first fact is easily explained taking into account the current trends in marriage, the second one most likely shows that in those regions where life expectancy is high (in the Caucasus) there are fewer hospitals. The positive impact of the subtropical climate, GRP and health care costs is obvious, and the increase in life expectancy with rising unemployment is due to the fact that the North Caucasus has the highest values of the indicator in Russia. The positive impact of the number of students is consistent with the microeconomic thesis that better educated people live longer. Thus, this regression model includes both theoretically obvious factors and features associated with the development of individual regions in Russia.

Cluster analysis makes it possible to identify regions where life expectancy inspires some concern. These are not only the northern regions with a harsh climate for living, but also other territories where low life expectancy is a consequence of both the unfavorable economic situation and the insufficient development of medical infrastructure.

5 Conclusion

On the basis of the study conducted, a regression model of the influence of a number of factors on life expectancy in the regions of Russia was created. A cluster analysis of the regions was performed according to the above indicators and a high degree of regional differentiation was revealed.

It should be noted that achievable goal of increasing life expectancy raises certain concerns as most of the regions are at the average level and below the average. A favorable situation is observed only in the Caucasus, where this is due to climatic conditions and a small number of mining enterprises, as well as in Moscow and St. Petersburg, where the country's best medical institutions are concentrated. Therefore, support measures and regional programs should primarily affect the regions from the second and fifth clusters.

Comparison of the results with other years, as well as the additional inclusion of such factors as alcohol consumption or the number of prisoners per capita, can be considered as a topic for further research.

Acknowledgments

The work of Iu. Kuznetsov is partially supported by the agreement No SSZ-1771 dated 22.04.2021 for the implementation of R&D on the topic: "Creation of high-tech sugar production on the basis of JSC "Sergach Sugar Factory", within the framework of the Agreement on the provision of subsidies from the federal budget for the development of cooperation between the Russian educational organization of higher education and the organization of the real sector of the economy in order to implement a comprehensive project to create high-tech production No. 075-11-2021-038 dated 06/24/2021 (IGC 000000S407521QLA0002).

References

1. Decree of the President of the Russian Federation No. 474 dated 21.07.2020 "On the National Development Goals of the Russian Federation for the period up to 2030". Accessed on: December, 14, 2022. [Online]. Available: http://www.kremlin.ru/acts/bank/45726

2. Unified plan for achieving the national development goals of the Russian Federation for the period up to 2024 and for the planning period up to 2030. Accessed on: December, 14, 2022. [Online]. Available: https://www.economy.gov.ru/material/dokumenty/edinyy_plan_po_dostizheniyu_nacionalnyh_celey_razvitiya_rossiyskoy_federacii_na_period_do_2024_goda_i_na_planovyy_period_do_2030_goda.html
3. A. Castelló-Climent, R. Doménech, Econ. J. **118(528)**, 653–677 (2008)
4. A.A. Erakhtina, Probl.f Econ. Trans. **63(1-3)**, 20-33 (2022)
5. O.V. Kapitanova, I.A. Kuznetsov, A.V. Semenov, V.V. Kisova, I.A. Koneva, Dela Press Conf. Ser.: Econ., Bus. Manag. **001**, 028 (2022)
6. J.W. Shaw, W.C. Horrace, R.J. Vogel, Southern Econ. J. **71(4)**, 768-783 (2005)
7. F. Morina, A. Komoni, D. Kilaj, D. Selmonaj, S. Grima, Int. J. Sust. Devel. Plan. **17(5)**, 1389-1401 (2022)
8. P. Roffia, A. Bucciol, S. Hashlamoun, Int. J. Health Econ. Manag. (2022). Accessed on: December, 14, 2022. [Online]. Available: http://doi.org/10.1007/s10754-022-09338-5
9. X. Qiu, S. Wu, Appl. Geography, **147**, 102786 (2022)
10. A. Shchur, S. Timonin, Demog. Rew. **7(5)**, 63-83 (2021)
11. L.A. Rodionova, E.D. Kopnova, Stat. Iss. **27(1)**, 106-120 (2020)
12. A. Pyankova, T.A. Fattahov, Prev. Med. **23(2)**, 89 (2020)
13. T.V. Kossova, Stat. Iss. **27(5)**, 76-86 (2020)
14. Federal State Statistics Service. Accessed on: December 02, 2022. [Online]. Available: https://rosstat.gov.ru/
15. V.K. Faltsman, Stud. Rus. Econ. Devel. **2(32)**, 194-198 (2021)