

FACTORS AFFECTING LIFE EXPECTANCY WORLDWIDE**Bazarbayeva L.,***PhD, Assistant-Professor, SDU University, SITAM***Khaval M.,***Bachelor's student in Mathematics***Sarsenbay A.,***Bachelor's student in Mathematics***Yerzhanov M.***Bachelor's student in Mathematics**SDU University**Kaskelen, Kazakhstan*<https://doi.org/10.5281/zenodo.19893252>**Abstract**

The purpose of this study is to identify the factors affecting life expectancy in 207 countries and understanding how these factors effects to human longevity. Using international data from the World Bank, World Population Review, and other open sources, we applied correlation analysis and regression analysis to examine the indicators influencing life expectancy. A two-factor regression model was created by calculating correlation coefficient using Microsoft Excel. The results showed that the influence of various factors on a person's life expectancy varies depending on regional characteristics, the level of development and the social status of countries. Among them, it was found that factors such as gross domestic product, education level and lifestyle play a special role. These indicators were divided into 5 groups for consideration.

Keywords: life expectancy, regression model, corruption, gdp, factor

1. Introduction

Since ancient times, people have striven to live longer. In recent years, increaing life expectancy has become a goal of leading countries, so it is important to accurately assess the socio-economic factors that influence this Indicator. Genetics determine only a quarter of the lifespan. Therefore, people can increase lifespan by taking certain actions. In this regard, life expectancy reflects not only the state of health, but also economic and social well-being, the lifestyle of the population,

the quality of the environment and so on. However, the importance of these factors is very likely to show different results in countries with different levels of development. Globally, a country's gross domestic product, happiness index, and average wage play a significant role in increasing life expectancy. And, as expected, the level of corruption and inflation has a negative impact. This shows that the higher the level of corruption and inflation in a country, the shorter its life expectancy.

1.1 Literature Matrix

Title	Author(s), Year	Methodology	Key Findings	Connection to our work
Factors affecting life expectancy in Kazakhstan	A.Zh. Panzabekova, Ivan E. Digel 2020	The study is based on regression and correlation analysis. Data were selected based on two criteria: availability of long-term statistical data and potential impact on life expectancy. A two-factor power regression model was constructed using Microsoft Excel.	The study used regression models to analyze the impact of socio-economic factors on life expectancy in the regions of Kazakhstan. The results show that economic factors have the most significant influence, while some relationships require further research.	We used the same methods as in the article, but not just for Kazakhstan's regions, but for the entire world. And you can see a trend that within individual countries and within each country, the indicators vary according to certain logics.
Factors Influencing Life Expectancy: Global Evidence from 125 Countries	Gabriel Hoh Teck Ling*, Noradila Rusli, Teh Hock Heng, Tan Kean Poh, Sumaia Faisal, Hu Jianke, Bai Ni, Mohamed S Abdullh Mohamed, 2025	The study uses secondary data from WHO, World Bank, and UN for 125 countries. Life expectancy is modeled as the dependent variable, with socioeconomic, healthcare, lifestyle, and environmental factors as predictors. Multiple linear regression in SPSS analyzes the relationships and significance of these determinants.	The study identified alcohol consumption, happiness, CO2 emissions, and PM2.5 air pollution as significant determinants of life expectancy, emphasizing the need for targeted policies in these areas. While BMI, GDP per capita, corruption perception, and government health expenditure were not significant in the regression model, they may still influence life expectancy in other contexts or analyses.	This work gives us a full understanding of what we're doing and allows us to compare our results with theirs. It's noticeable that the number of countries we work with plays a significant role in correlation indicators.
Modeling of life expectancy in the world	V. A. Shtun, 2021	Econometric analysis using multiple linear regression based on cross-country data (126 countries). Statistical testing of variables and model diagnostics were applied to determine significant factors affecting life expectancy.	The study found that life expectancy is significantly influenced by access to clean drinking water, government health expenditure, vaccination coverage, electricity access, and GDP per capita. Negative effects were found for alcohol consumption and CO2 emissions. Some variables such as unemployment, education level, and smoking were not statistically significant in the global model.	This study supports our diploma research by demonstrating how statistical modeling can be used to analyze determinants of life expectancy across countries. It provides methodological support for our machine learning approach and helps justify the selection of socioeconomic, behavioral, and environmental variables used in our life expectancy prediction model.

2. Methodology

The regression model was used to study the relationship between economic, social, and environmental factors and life expectancy. This approach allows the identification of factors that are most strongly associated with life expectancy and provides a global comparative analysis.

For each independent variable, a simple linear regression model was constructed: $\hat{v}_i = m_1 + m_2 u_i$ (1)

- v_i is the life expectancy for observation i ,

- u_i is an independent variable (factor) that influences life expectancy,

- m_1 is the intercept representing the baseline level when $u_i = 0$,

- m_2 is the slope coefficient indicating the change in life expectancy when u_i increases by one unit.

To examine the relationship between life expectancy and selected factors, the correlation analysis is

applied. In this study, the simple linear correlation coefficient is used to measure the strength and direction of the relationship between variables.

The correlation coefficient is calculated as follows:

$$k_{uv} = \frac{\sum_{i=1}^n u_i v_i - n \cdot \bar{u} \bar{v}}{\sqrt{(\sum_{i=1}^n u_i^2 - n \cdot \bar{u}^2)(\sum_{i=1}^n v_i^2 - n \cdot \bar{v}^2)}} \quad (2)$$

The correlation coefficient ranges from -1 to +1. Values close to +1 indicate a strong positive relationship, values close to -1 indicate a strong negative relationship, and values near 0 indicate no linear relationship.

The coefficient of determination is defined as the proportion of total variation explained by the regression model:

$$K^2 = \frac{\sum_{i=0}^n (\hat{v}_i - \bar{v})^2}{\sum_{i=0}^n (v_i - \bar{v})^2} \quad (3)$$

A higher K^2 indicates a better explanatory power of the independent variable.

3. Regression analysis of Factors Affecting Life Expectancy

A data set was compiled for 207 countries. The following variables were included in the analysis: smoking rate, drug prevalence, obesity, gross domestic product per capita (GDP), happiness index, gender equality index, meat consumption, corruption, education, average salary, inflation, CO2 emissions, and alcohol consumption. Python programming language was used primarily to process the data set. This process included data loading, cleaning missing values, normalizing indicators and combining various sources into a single table containing 207 countries.

Table 1.

Dataset of 207 countries used in the analysis

Country	LE(v)	smo(u ₁)	nar(u ₂)	obe(u ₃)	gdp(u ₄)	happ(u ₅)	geq(u ₆)	mt(u ₇)	corr(u ₈)	ed(u ₉)	aversal(u ₁₀)	infl(u ₁₁)	co2(u ₁₂)	alc(u ₁₃)
Afghanistan	66.29	21.3	1.36	17.7	2202.0	1.72	0.715	9.6	16.0	0.381	0.172	5.7	0.14	0.4
Albania	79.78	20.7	0.7	26.6	23488.0	5.3	0.763	47.9	39.0	0.739	0.511	2.3	1.59	5.9
Algeria	76.48	21.1	2.46	24.2	17553.0	5.36	0.614	22.4	34.0	0.664	0.356	1.5	3.89	0.8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Thailand	76.56	18.3	1.11	14.5	24708.0	5.98	0.728	54.9	33.0	0.426	0.615	-0.88	3.94	5.4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Zambia	66.53	14.1	1.19	11.6	4224.0	3.5	0.707	29.2	37.0	0.313	0.182	7.5	0.46	5.5
Zimbabwe	63.06	10.4	2.31	9.4	3922.0	3.34	0.747	54.9	22.0	0.549	0.169	3.8	0.64	6.4

The correlation coefficients were calculated using formula 2. All calculated correlation coefficients between the studied factors are presented in Table 2.

Table 2.

Correlation coefficients

Indicators	Correlation coefficient (k)
smo	0,167
nar	0,077
obe	0,3
gdp	0,714
happ	0,604
geq	0,357
mt	0,573
corr	-0,561
ed	0,274
aversal	0,63
infl	-0,126
co2	0,314
alc	0,205
pop	-0,017
area	0,02

Since life expectancy can vary depending on continental factors, countries were grouped by continents to analyze which variables have the greatest impact on life expectancy in each country group. Population and Country Area indicators, which showed the lowest correlation, were not included in the analysis. Correlation coefficients were calculated using formula 2. The results for Asian countries are presented in Table 3.

Table 3.

Correlation between life expectancy and factors in Asia

Central Asia	k	k ²	East Asia	k	k ²	West Asia	k	k ²	South-East Asia	k	k ²	South Asia	k	k ²
Smo	0,57	0,33	Smo	-0,55	0,31	Smo	-0,41	0,18	Smo	-0,51	0,26	Smo	0,27	0,08
Nar	0,53	0,28	Nar	-0,19	0,03	Nar	0,07	0,01	Nar	-0,24	0,06	Nar	-	-
Obe	-0,05	0,01	Obe	-0,60	0,37	Obe	0,45	0,20	Obe	0,56	0,31	Obe	-0,21	0,05
Gdp	0,59	0,36	Gdp	0,51	0,26	Gdp	0,79	0,63	Gdp	0,84	0,72	Gdp	0,91	0,84
Happ	0,7	0,5	Happ	0,37	0,01	Happ	0,54	0,30	Happ	0,70	0,50	Happ	0,49	0,24
Geq	-0,15	0,025	Geq	-0,51	0,26	Geq	-0,15	0,02	Geq	-	-	Geq	0,01	0,00
Mt	0,6	0,36	Mt	0,25	0,07	Mt	0,74	0,48	Mt	-0,29	0,08	Mt	0,84	0,71
Corr	-0,91	0,84	Corr	-0,88	0,34	Corr	-0,58	0,34	Corr	-0,53	-0,53	Corr	-0,33	0,11
Ed	0,59	0,36	Ed	-0,66	0,44	Ed	0,26	0,07	Ed	-0,49	0,24	Ed	0,36	0,13
Aversal	0,16	0,026	Aversal	0,97	0,94	Aversal	0,59	0,35	Aversal	0,83	0,70	Aversal	-	-
Infl	0,69	0,5	Infl	-0,80	0,65	Infl	0,03	0,01	Infl	-0,67	0,46	Infl	-0,56	0,31
Co2	0,3	0,093	Co2	0,35	0,12	Co2	0,68	0,46	Co2	0,59	0,34	Co2	0,70	0,49
Alc	0,11	0,02	Alc	0,16	0,03	Alc	-	-	Alc	-0,14	0,02	Alc	0,43	0,18
Central Asia: Kazakhstan - 74,53 Kyrgyzstan - 71,82 Tajikistan - 71,93 Turkmenistan - 70,2 Uzbekistan - 72,53			East Asia: China - 78,02 Hong Kong - 85,63 Japan - 84,85 Macau - 83,23 Mongolia - 71,99 North Korea - 73,74 South Korea - 84,43			West Asia: Bahrain - 81,42 Turkey - 77,42 Kuwait - 80,6 Oman - 80,25 Saudi Arabia - 78,98 Qatar - 82,55 United Arab Emirates - 83,07 Yemen - 69,44 Armenia - 75,85 Azerbaijan - 74,58 Georgia - 74,66 Iran - 77,85 Iraq - 72,42 Israel - 82,73 Jordan - 77,98 Lebanon - 77,95 Palestine - 69,21 Syria - 72,56 Cyprus - 81,82 Egypt - 71,81			South-East Asia: Brunei - 75,5 Timor-Leste - 67,9 Vietnam - 74,74 Indonesia - 71,29 Cambodia - 70,82 Laos - 69,22 Malaysia - 76,82 Myanmar - 67,1 Singapore - 83,86 Thailand - 76,56 Philippines - 69,95			South Asia: India - 72,24 Pakistan - 67,8 Bangladesh - 74,93 Afghanistan - 66,29 Bhutan - 73,26 Maldives - 81,29 Nepal - 70,64 Sri Lanka - 77,67		

Table 4.

Correlation between life expectancy and factors in Europe

West Europe	k	k ²	East Europe	k	k ²	North Europe	k	k ²	South Europe	k	k ²
Smo	0,14	0,02	Smo	0,06	0,04	Smo	-0,88	0,77	Smo	0,07	0,01
Nar	-0,27	0,07	Nar	-0,52	0,27	Nar	0,35	0,12	Nar	0,29	0,08
Obe	-0,83	0,68	Obe	0,47	0,22	Obe	-0,85	0,72	Obe	-0,42	0,18
Gdp	0,65	0,43	Gdp	0,76	0,57	Gdp	0,81	0,66	Gdp	0,87	0,76
Happ	0,13	0,02	Happ	0,31	0,09	Happ	0,81	0,66	Happ	-	-
Geq	-0,58	0,34	Geq	-0,45	0,20	Geq	0,67	0,45	Geq	-0,08	0,01
Mt	-0,52	0,27	Mt	0,20	0,04	Mt	0,24	0,06	Mt	0,32	0,10
Corr	-0,73	0,53	Corr	0,54	0,29	Corr	-0,83	0,69	Corr	-0,53	0,28
Ed	0,50	0,25	Ed	0,49	0,24	Ed	0,27	0,07	Ed	0,03	0,01
Aversal	-0,16	0,03	Aversal	0,46	0,21	Aversal	0,88	0,77	Aversal	0,76	0,58
Infl	0,45	0,20	Infl	-0,31	0,09	Infl	-0,19	0,04	Infl	0,25	0,06
Co2	-0,48	0,23	Co2	-	-	Co2	0,41	0,17	Co2	-0,04	0,01
Alc	-0,17	0,03	Alc	0,05	0,01	Alc	-0,68	0,46	Alc	-0,46	0,21
West Europe: Belgium - 82,27 United Kingdom - 81,45 Germany - 81,54 Ireland - 82,58 Luxembourg - 82,36 Monaco - 86,5 Netherlands - 82,3 France - 83,46 Switzerland - 84,09 Austria - 82,12			East Europe: Belarus - 74,62 Bulgaria - 75,8 Hungary - 77,18 Moldova - 71,33 Poland - 78,81 Russia - 73,34 Romania - 76,09 Slovakia - 78,49 Ukraine - 74,69 Czechia - 79,97			North Europe: Denmark - 82,09 Iceland - 83,01 Latvia - 76,33 Lithuania - 76,17 Norway - 83,46 Finland - 82,08 Sweden - 83,42 Estonia - 79,31			South Europe: Albania - 79,78 Andorra - 84,19 Bosnia and Herzegovina - 78,04 Greece - 82,03 Italy - 83,87 Spain - 83,8 Malta - 83,47 Portugal - 82,55 San Marino - 85,82 Serbia - 76,94 Montenegro - 77,25 Croatia - 78,75		

Table 5.

Correlation between life expectancy and factors in America and Oceania

Central America	k	k ²	North America	k	k ²	South America	k	k ²	Oceania	k	k ²
Smo	-0,43	0,18	Smo	0,23	0,05	Smo	-0,24	0,06	Smo	-0,64	0,41
Nar	-0,38	0,14	Nar	0,48	0,23	Nar	-0,56	0,31	Nar	0,56	0,32
Obe	0,3	0,19	Obe	0,34	0,12	Obe	0,48	0,23	Obe	-0,29	0,08
Gdp	0,86	0,73	Gdp	0,49	0,24	Gdp	0,60	0,37	Gdp	0,83	0,70
Happ	0,58	0,34	Happ	-	-	Happ	0,60	0,36	Happ	-	-
Geq	0,7	0,49	Geq	-	-	Geq	-	-	Geq	-	-
Mt	0,81	0,65	Mt	0,44	0,20	Mt	0,27	0,07	Mt	0,21	0,04
Corr	-0,63	0,40	Corr	-0,43	0,18	Corr	-0,55	0,31	Corr	-	-
Ed	-0,14	0,02	Ed	0,37	0,14	Ed	-0,13	0,02	Ed	0,52	0,27
Aversal	0,73	0,54	Aversal	0,12	0,01	Aversal	0,52	0,27	Aversal	0,87	0,76
Infl	-0,62	0,38	Infl	-0,52	0,27	Infl	-0,27	0,07	Infl	-	-
Co2	0,74	0,55	Co2	0,41	0,17	Co2	-	-	Co2	0,07	0,01
Alc	0,39	0,15	Alc	0,35	0,12	Alc	-0,09	0,01	Alc	-0,86	0,74
Central America: Belize - 73,74 Costa Rica - 81 El Salvador - 72,3 Guatemala - 72,75 Honduras - 73,04 Nicaragua - 75,1 Panama - 79,77			North America: Greenland - 70,19 Canada - 82,72 United States - 79,46 Mexico - 75,26 Antigua and Barbuda - 77,77 Bahamas - 74,71 Barbados - 76,33 Cuba - 78,26 Dominica - 71,29 Dominican Republic - 73,87 Grenada - 75,37 Haiti - 65,12 Honduras - 73,04 Jamaica - 71,61 Saint Kitts and Nevis - 72,28 Saint Lucia - 72,85 Saint Vincent and the Grenadines - 71,38 Trinidad and Tobago - 73,62			South America: Argentina - 77,54 Bolivia - 68,74 Brazil - 76,02 Chile - 81,36 Colombia - 77,91 Ecuador - 77,58 Guyana - 70,32 Paraguay - 73,98 Peru - 77,94 Suriname - 73,76 Uruguay - 78,29 Venezuela - 72,67			Oceania: Australia - 84,07 Fiji - 67,46 Kiribati - 66,6 Marshall Islands - 67,08 Micronesia - 67,36 Nauru - 62,28 New Zealand - 82,24 Palau - 69,38 Papua New Guinea - 66,26 Samoa - 71,83 Solomon Islands - 70,69 Tonga - 73,07 Tuvalu - 67,26 Vanuatu - 71,65		

Table 5.

Correlation between life expectancy and factors in Africa

North Africa	k	k ²	West Africa	k	k ²	Central Africa	k	k ²	East Africa	k	k ²	South Africa	k	k ²
Smo	-0,05	0,01	Smo	0,15	0,02	Smo	0,51	0,26	Smo	-	-	Smo	-0,46	0,21
Nar	-0,37	0,14	Nar	-0,43	0,19	Nar	0,41	0,17	Nar	0,17	0,03	Nar	-0,19	0,04
Obe	0,10	0,01	Obe	0,44	0,20	Obe	0,97	0,94	Obe	0,43	0,19	Obe	-0,08	0,01
Gdp	0,62	0,39	Gdp	-	-	Gdp	0,84	0,71	Gdp	0,71	0,49	Gdp	0,92	0,85
Happ	-0,35	0,13	Happ	0,04	0,01	Happ	0,30	0,10	Happ	0,05	0,01	Happ	0,42	0,18
Geq	0,22	0,05	Geq	0,32	0,10	Geq	0,60	0,36	Geq	-0,11	0,01	Geq	0,66	0,43
Mt	0,19	0,04	Mt	0,50	0,26	Mt	0,56	0,31	Mt	0,42	0,18	Mt	0,57	0,33
Corr	-0,87	0,76	Corr	-0,33	0,11	Corr	-0,51	0,27	Corr	-0,65	0,43	Corr	-0,56	0,32
Ed	0,23	0,06	Ed	0,72	0,53	Ed	0,93	0,87	Ed	-0,15	0,02	Ed	-0,16	0,03
Aversal	0,42	0,18	Aversal	-0,22	0,05	Aversal	0,11	0,01	Aversal	0,28	0,08	Aversal	0,34	0,12
Infl	-0,39	0,16	Infl	-0,22	0,05	Infl	0,79	0,63	Infl	-0,50	0,25	Infl	-0,28	0,08
Co2	0,11	0,01	Co2	0,02	0,01	Co2	0,73	0,54	Co2	0,29	0,09	Co2	0,52	0,28
Alc	-0,51	0,27	Alc	-0,08	0,01	Alc	0,91	0,83	Alc	0,26	0,07	Alc	0,80	0,64
North Africa: Algeria - 76,48 Egypt - 71,81 Libya - 71,12 Morocco - 75,49 Sudan - 66,52 Tunisia - 76,71			West Africa: Benin - 60,96 Burkina Faso - 61,29 Gambia - 66,06 Ghana - 65,7 Guinea - 60,9 Guinea-Bissau - 64,25 Ivory Coast - 62,11 Liberia - 62,32 Mali - 60,67 Mauritania - 68,71 Niger - 61,43 Nigeria - 54,63 Senegal - 68,92 Sierra Leone - 61,96 Togo - 62,93 Burkina Faso - 61,29 Gambia - 66,06 Ghana - 65,7 Guinea - 60,9 Guinea-Bissau - 64,25 Ivory Coast - 62,11			Central Africa: Cameroon - 63,97 Central African Republic - 57,67 Chad - 55,24 Equatorial Guinea - 63,91 Gabon - 68,51			East Africa: Burundi - 63,82 Comoros - 67,02 Djibouti - 66,2 Eritrea - 68,89 Ethiopia - 67,6 Kenya - 63,83 Madagascar - 63,83 Malawi - 67,56 Mauritius - 75,1 Mozambique - 63,8 Rwanda - 68,02 Seychelles - 72,99 Somalia - 58,97 South Sudan - 57,74 Tanzania - 67,21 Uganda - 68,49 Zambia - 66,53 Zimbabwe - 63,06			South Africa: Botswana - 69,29 Eswatini - 64,26 Lesotho - 57,8 Namibia - 67,52 South Africa - 66,31		

4. Results and Discussion

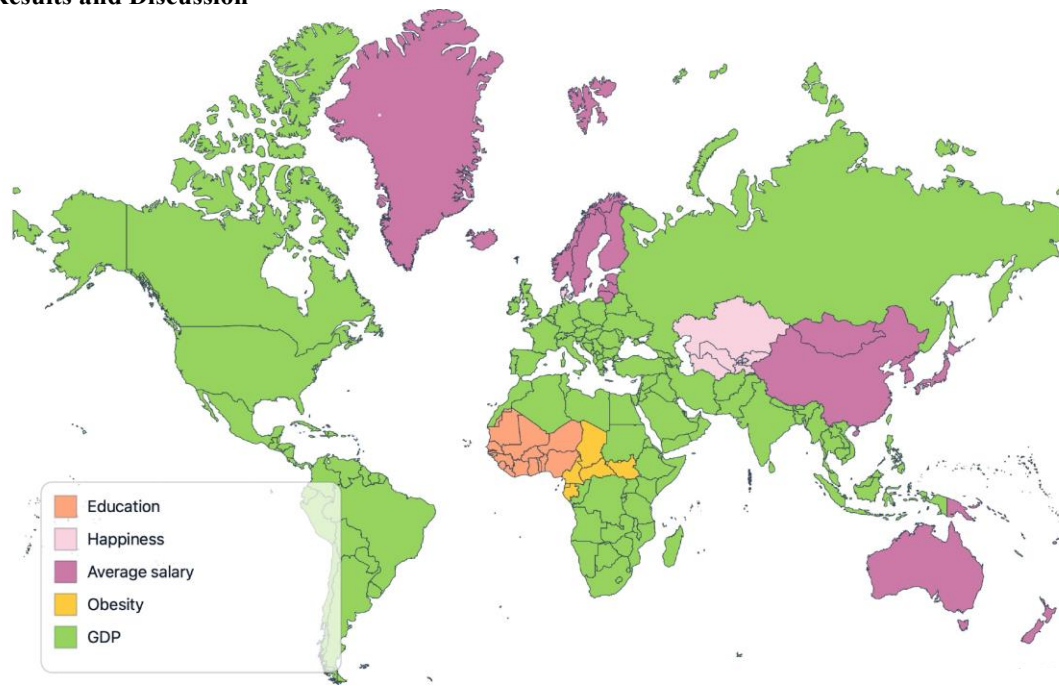


Figure 1. The factors that have the greatest positive impact on life expectancy Worldwide

As figure 1 illustrates, GDP has the most greatest positive impact worldwide. Here in most Asian and North American countries, life expectancy depends on it. The next factors that also have a great impact is the average salary, which can be seen in North Europe and some East Asian countries, and the education index, which can be seen in West Africa.

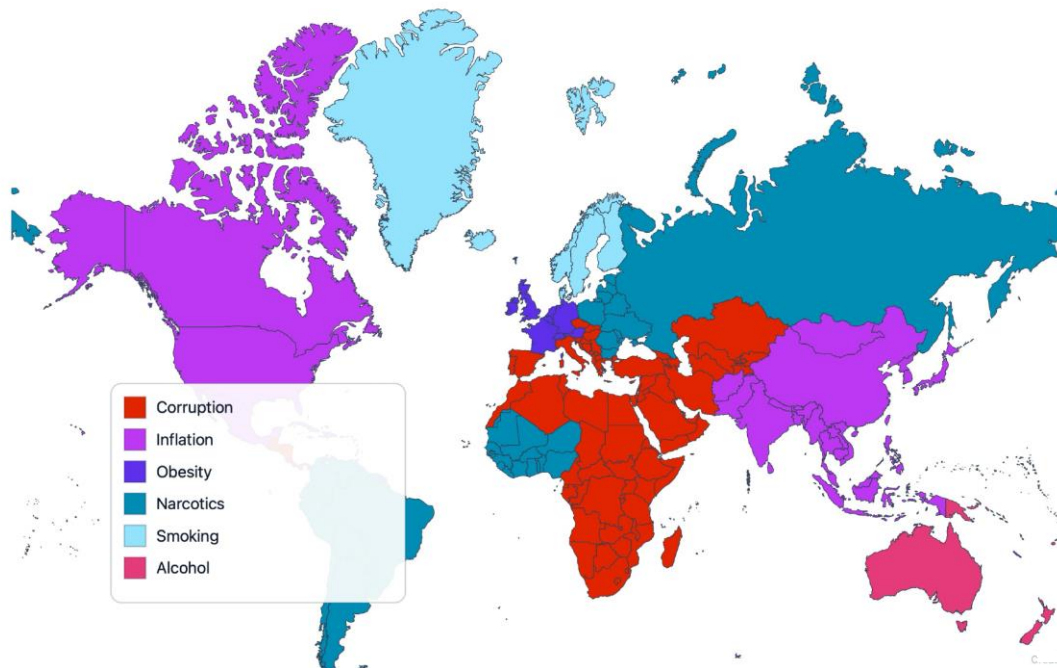


Figure 2. The factors that have the most negative impact on life expectancy Worldwide

In figure 2 it shows that factor which has the most negative impact on life expectancy is inflation. In all countries of North America, South and East Asia, this factor is negative which affects the reduction of their life expectancy. The second most negative factor that affects life expectancy is corruption and the rarest one is obesity. Obesity only influences West Europe as an adverse cause.

Conclusion

A multiple regression model was used to evaluate the factors affecting life expectancy in World countries. To find out the relationship between life expectancy and socio-economic factors, correlation coefficients for different countries and continents were calculated by selecting several indicators. Factors with the lowest coefficients such as population size and area of countries were excluded. Overall, the results of this study con-

firm that life expectancy depends primarily on the overall state of a country's healthcare system, particularly on its level of economic development, social security, and the quality of governance. In Asia and North America, a high GDP has a more significant positive impact.

The negative impact of factors such as inflation and corruption is manifested in different ways in different regions of the world. For example, in North America, South Asia, and Eastern Europe, high inflation reduces economic stability and limits opportunities for health and social programs. Additional factors such as population size, employment rate, access to educational institutions and market infrastructure have a moderate positive impact in all regions, confirming the complex nature of the life expectancy formation process

Using a life expectancy model, it is possible to predict the implications of various changes in life expectancy for Kazakhstan and the world as a whole.

References:

1. Humbat Nabi oglu Aliyev. STATISTICS SOLVED PROBLEMS AND EXERCISES. PART 2. Recommended for publication by the Academic Council of Suleyman Demirel University. "KazGASA Publishing House 'Construction and Architecture'", 2006. isbn:9965960577.

2. CO2 Emissions by Country. url: <https://www.worldometers.info/co2-emissions/>

3. CORRUPTION PERCEPTIONS INDEX 2025. url: <https://www.transparency.org/en/cpi/2025>

4. Happiest Countries in the World 2025. url: <https://worldpopulationreview.com/country-rankings/happiest-countries-in-the-world>

5. I.E.Digel and A.Zh.Panzabekova. "Factors affecting life expectancy in Kazakhstan". In: R-Economy 6.4 (2020), pp. 261–270. url: <https://r-economy.com/article/view/4899>

6. Life Expectancy has increased across the World. url: <https://ourworldindata.org/life-expectancy>

7. List of countries by life expectancy. url: https://en.wikipedia.org/wiki/List_of_countries_by_life_expectancy

8. Ranked: Countries with The Highest and Lowest Average Salaries, 2025. url: <https://ceoworld.biz/2025/08/08/ranked-countries-with-the-highest-and-lowest-average-salaries-2025/>

9. Weng Shin Teh, Gabriel Ling Hoh Teck Ling, and Noradila Rusli. "Factors Influencing Life Expectancy: Global Evidence from 125 Countries". In: Penerbit UTM Press 12.2 (2025), pp. 17–28. doi: <https://doi.org/10.11113/ijbes.v12.n2.1294>.