

Exploring Happiness Across Europe: A Cluster Analysis of EU Countries through Well-Being Indicators

Adina-Iuliana JIGANI

Bucharest University of Economic Studies, Bucharest, Romania
adina.jigani@csie.ase.ro

Alexandra-Nicoleta CIUCU (DURNOI)

Bucharest University of Economic Studies, Bucharest, Romania
alexandra.ciucu@csie.ase.ro

Anca-Elena VLAD*

Bucharest University of Economic Studies, Bucharest, Romania
**Corresponding author, vladanca18@stud.ase.ro*

Cătălin ȚIBULCĂ

Bucharest University of Economic Studies, Bucharest, Romania
catalin.tibulca@csie.ase.ro

Abstract. *Happiness is a concept that everyone wants to achieve. Although it is an emotion, various variables have been determined to describe it. In this sense, the most famous document that studies this notion is the World Happiness Report, created in partnership with the UN Sustainable Development Solutions Network, a sign that the well-being of citizens is very important in the process of achieving sustainability. Given the notoriety of this Report, the current research aims to take into analysis the indicators used in this paper to describe happiness, with the aim of grouping the 27 EU member states according to the degree of similarity identified between them. From a methodological point of view, the current study initially involves the application of factor analysis to determine a matrix of comprehensive factor scores, which is used as the starting point of the cluster analysis. Comparing the clustering of countries with the associated Life ladder values, it is found that there is a link between the results obtained and people's opinion about their happiness. In this sense, cluster 1 seems to contain countries with high Life ladder values, while the second cluster encapsulates countries with lower values of this indicator.*

Keywords: well-being, cluster analysis, panel data, SDGs.

Introduction

The concept of happiness has acquired many meanings over the decades and has been defined in different ways, based on psychological concepts, interdisciplinary approaches, and differentiated according to geographical location. Today, in a broader sense, happiness is viewed as a concept that cannot be measured, but rather describes a way of life rather than a state (Rozario, 2018). According to a 2016 study, out of 785 people interviewed, most associated the word "happiness" with health and interpersonal relationships (Bojanowska, 2016). This concept is considered to have a multidimensional character, as it encompasses aspects related to emotional well-being, life satisfaction, and the way in which the quality of life is perceived (Adylla, 2025).

It is important to mention that the concept of happiness is increasingly used in recent studies as an indicator for measuring the degree of development of a country, as well as its economic status. The results of Rozario's (2018) research showed that there is a positive correlation between happiness and economic growth. The author studied 15 developed and 15 developing countries, using cross-sectional regression, and found that this correlation is significant in developed countries, compared to developing countries. According to another study, happiness was chosen by researchers as an alternative indicator in their studies in order to better capture the progress of society and the well-being of citizens. They consider that conventional indicators are limiting and cannot provide a clear picture for comparisons between countries (Adylla, 2025).

A first paper to mention is the World Happiness Report (Helliwell et al., 2025), which has been contributing significantly, since 2012, to the measurement and comparison of the level of happiness of populations in different countries, using global statistical data. The purpose of this annual report is to highlight that happiness represents a significant indicator, complementary to the Gross Domestic Product (GDP), regarding the economic growth of a country, analyzed from the perspective of the impact it has on its citizens. Also, this indicator can represent a solid basis for shaping public policies in any state. The indicators used in the World Happiness Report indirectly capture the well-being of individuals at work, since factors such as social support, freedom to make choices, and positive states lead to both personal and professional well-being and a reduction in the risk of burnout. Another significant study, conducted on the same set of indicators, used k-means clustering as the analysis method. After going through the specific steps of the analysis, it was observed that there was a high similarity between countries belonging to the same cluster. It was also concluded that high incomes in certain countries are not sufficient to increase the happiness fence, but rather, it occurs when there is a combination of factors such as economic, social, and institutional (Adylla, 2025).

However, based on the studies identified so far, there are significantly fewer studies that use clustering than those that use regression (Adylla, 2025); Helliwell et al., 2025; Ulkhaq & Adyatama, 2021; (Pelka, 2019); Banerjee et al., 2019). Regression is seen as a standard method that helps researchers determine whether there is a link between happiness and other economic indicators, while clustering is seen as a modern technique that helps identify patterns in the dataset used in the World Happiness Report (Helliwell et al., 2025). Therefore, this paper brings an element of novelty through the analysis used, since numerous studies based on happiness indicators have used regression as the analysis method.

This study aims to classify European countries based on a set of happiness indicators, which are found in the World Happiness Report, using the cluster method on panel data. The indicators used include life ladder, log of GDP per capita, social support, healthy life expectancy at birth, freedom to make life choices, generosity, perceptions of corruption, positive affect, and negative affect. Life ladder is one of the main indicators used in the World Happiness Report (Helliwell et al., 2025) and represents a subjective scale for the quality of life, log GDP per capita represents the Gross Domestic Product per capita, logarithmized in order to reduce variation and normalize the distribution. Social support is the indicator that measures the degree to which each citizen believes that they can be helped by friends in case of need, healthy life expectancy at birth illustrates the healthy life expectancy at birth, while freedom to make life choices records binary responses for the freedom of each citizen in a country to make choices regarding their own life. Generosity shows the proportion of people who have recently made a charitable gesture, perceptions of corruption represents the general perception that citizens have regarding the

degree of corruption within the Government and public institutions, positive affect quantifies the average of positive emotions from the day before the questionnaire, while the last indicator in the data set used in the analysis, negative affect, quantifies the average of negative emotions experienced on the day previous (Helliwell et al., 2025).

This paper is structured into four sections. The first one, entitled "Literature review", is intended to present similar articles and highlight the main results obtained, the second section presents the cluster analysis applied to the set of indicators also used in the World Happiness Report, because this method helps to position countries according to certain patterns. The third section is dedicated to the interpretation of the results obtained following the application of the cluster analysis. The article ends by highlighting the relevant conclusions.

Literature review

Happiness is multidimensional and has been defined in the literature in various ways over time, based on aspects related to individuals' subjective perceptions of life, emotional state, as well as economic, social, and institutional factors (Helliwell et al., 2012). In another study, happiness is described as a combination of several factors, including life satisfaction and the balance between positive and negative affects (Diener, 2003). This idea is also supported by one of the most important researchers in psychology, who states that happiness is a way of life obtained at the intersection of positive experiences and other aspects of a psychological nature, such as social relationships (Seligman, 2011). When it comes to comparative analyses between countries, happiness is measured through a series of indicators, as can be seen in the World Happiness Report (Helliwell et al., 2025), the most important being life satisfaction and freedom to make life choices, as they reflect each individual's perception of the quality of life and the freedom to make their own choices. At the same time, generosity and social support are indicators that reflect the social and relational dimensions of happiness. Both indicators are able to describe the degree of solidarity within a country and define subjective well-being and life satisfaction (Helliwell et al., 2019). They focus on each person's behavior in relation to others, thus going beyond the individual sphere, contributing to shaping a higher level of life satisfaction, as the feeling of belonging to a group is strengthened. Thus, the concept of happiness is shaped from two perspectives, one subjective, emotional, and the other objective, determined based on indicators such as income or health.

The World Happiness Report (WHR), published annually since 2012, is the most important source of data for analyzing happiness at a global level (Helliwell et al., 2025). It ranks countries according to their level of happiness. The report contains a set of subjective and objective indicators that capture economic, social, and institutional dimensions of well-being, along with assessments of life satisfaction and emotional experiences. These indicators include log of GDP per capita, social support, healthy life expectancy at birth, freedom to make life choices, generosity, perceptions of corruption, positive affect, and negative affect. The study conducted within the report is based on a regression model, which explains how general well-being is influenced by the variables taken into account. Gross Domestic Product (GDP), the indicator used in numerous studies in the specialized literature to measure the level of happiness of individuals in a country (Easterlin, 1974; Helliwell et al., 2025), is considered a determinant, because income significantly influences access to quality education and health services. Generosity and social support are also indicators that, although they represent an affective component, vary depending on the economic context of a country and describe the state of average income.

In the specialized literature, there are numerous studies based on the indicators used in the World Happiness Report to measure the degree of happiness at the level of countries of the world. A significant part of them focused on regression-based analysis models (Boarini et al., 2012; Deaton, 2008), while another part used the novelty of clustering-based analysis to group countries according to common characteristics that define the concept of well-being. It is worth mentioning the study of (Easterlin, 1974), because he demonstrated, based on an analysis, that although richer countries tend to be happier than poorer ones, long-term economic growth within a country does not necessarily lead to an increase in the average level of happiness. This discovery is today called Easterlin's paradox. Another relevant study in relation to the studied topic is the article written by Behera (2024), which highlights, based on a regression analysis, that the indicators per capita income, social support, and freedom to make choices in life have a positive impact on happiness, while gender inequality does not have a significant impact on happiness. This article confirms the paradox discovered by (Easterlin, 1974), highlighting the fact that income can mediate happiness by promoting emotional well-being, gender equality, and a clean environment.

A significant part of the articles found in the specialized literature is based on analyses of happiness indicators using econometric approaches and regression methods; this fact can also be observed in the World Happiness Report (Helliwell et al., 2023). Although these approaches offer valuable insights into the causal relationships between socio-economic factors and the level of well-being of individuals, they offer a low degree of identification of structural similarities between countries. Thus, cluster analysis proposes a complementary approach, offering the possibility of grouping countries according to the degree of well-being (Chakraborty & Tsokos, 2021). However, current studies in the specialized literature use clustering methods in the analysis of happiness on cross-sectional data. These techniques are characterized by their limited ability to capture temporal dynamics and long-term developments (Bartolini et al., 2016). To our knowledge, the use of cluster analysis on panel data in studies describing well-being at the country level remains underexplored in the literature. Therefore, the present study contributes to the development of the literature by using a clustering approach on panel data for European Union countries, based on the indicators also used in the World Happiness Report. It offers a dynamic and comparative perspective on the evolution of well-being.

It is important to mention that in the specialized literature, there are results that indicate that certain indicators may be much more relevant in one region than in another region. These differences indicate that complex analyses are needed to identify typologies of countries, since happiness at a global level cannot be measured through a single classic model.

In order to perform the cluster analysis, the R environment was used, as it is capable of analyzing a significant volume of statistical data, is prepared to offer advanced packages for data visualization (Wickham, 2016), and, most importantly, offers support in establishing the level of relevance by offering quantifiable indicators, such as silhouette (Kaufman & Rousseeuw, 1990) or elbow method. In addition, the results obtained from clustering can be analyzed through the prism of dendograms, which provide a clear picture of the groups of countries and the distances between them. Cluster analysis was chosen as the research method because, unlike regression, which is able to highlight the individual impact of each variable on happiness, it can identify complex patterns and combinations of factors that are found in certain region (Everitt et al., 2011). This approach is in line with the specialized literature, which affirms the complexity of the concepts of well-being and happiness and emphasizes the need for multiple methods for an adequate understanding.

Methodology

The steps taken for the cluster analysis of happiness data for the period 2015-2022 have been presented in this section. Thus, starting from the determination of the database, it should be noted that the two indicators analyzed in the current research were taken from the World Happiness Report (Helliwell et al., 2023) for the year 2023 because this was the last year in which the authors published the data.

Given that one of the partners helping to publish the World Happiness Report is the UN Sustainable Development Solutions Network, Table 1 presents the indicators investigated, as well as the link between them and the Sustainable Development Goals (SDGs).

Table 1. Indicators

Label	Name	SDG
LL	Life ladder	SDG 3
IGDP	Log of GDP per capita	SDG 8
SS	Social support	SDG 3, SDG 10
HLEB	Healthy life expectancy at birth	SDG 3
FMLC	Freedom to make life choices	SDG 16
G	Generosity	SDG 10, SDG 17
PC	Perceptions of corruption	SDG 16
PA	Positive affect	SDG 3
NEA	Negative affect	SDG 3

Source: Authors' own research.

As can be seen, SDG 3 "Good health and well-being" and SDG 16 "Peace, justice and strong institutions" were most often associated with indicators intended to explain the variation in happiness. To these two is added SDG 8 "Decent work and economic growth", which is closely related to IGDP. Also, SDG 10 "Reduced inequalities" can be considered, which can be associated with SS and G; the last indicator is also part of SDG 17 "Partnerships for the goals" because it talks about helping less developed countries, as citizens help less advantaged social categories through donations.

The research follows the steps implemented in the work published by Wang and Lu (2021), which provides the code in the R programming language. Moreover, the authors explain, in detail, each step that must be taken to perform the cluster analysis, starting from the construction of a three-dimensional matrix specific to panel data. Therefore, the first methodological step is to determine the factorability of the data, using the Kaiser–Meyer–Olkin (KMO) test, for which the Sampling Adequacy Measure (MSA) is calculated at the individual level, but also overall. The closer this total value is to one, the higher the factorability (Kaiser & Rice, 1974).

Once the factorability of the indicators is established, Principal Component Analysis (PCA) is applied in order to determine the number of common factors. Through PCA, information redundancy is reduced, each component taking over the maximum variance of the original variables, provided that the components are independent, i.e. the correlation between them is non-existent (Jolliffe, 2002). It should be noted that a step preceding the analysis is the standardization of the data to reduce the possible differences that may arise due to differences in scale. To determine the number of principal components taken into account in the analysis, two criteria were taken into account: Kaiser's criterion (Kaiser, 1960) (according to which those components whose eigenvalues are greater than one are taken into account in the analysis) and the covered variance criterion (Wang & Lu, 2021) (which indicates that the first n components

that cover approximately 80% of the variance of the original variables should be taken into account). Next, to determine the common factors, an orthogonal rotation is performed, performed using the Varimax method, in order to maximize the variance of the factors. The new eigenvalues obtained, together with the scores, are taken into account to form the comprehensive factor score matrix, with the help of which the cluster analysis is implemented. The comprehensive factor score is calculated as the ratio of the sum of each common factor score multiplied by its eigenvalue to the sum of the eigenvalues (Wang & Lu, 2021).

Cluster analysis begins by calculating the Euclidean distances between objects, using hierarchical algorithms, which are one of the most popular methods for unsupervised clustering of shapes (Arutyunova & Röglin, 2025). Since the number of clusters must be specified by the researcher, several methods were used, including Within Sum of Square (Zhao et al., 2009), Average Silhouette Width (Rousseeuw, 1987), and Gap Statistics (Tibshirani et al., 2001). To ensure greater accuracy in determining the number of clusters, the NbClust package was used, which contains a number of 30 indicators designed to recommend the optimal number of clusters, including the three methods listed above (Charrad et al., 2014). The function called in RStudio needs certain parameters, in this case the Euclidean distance is used, and the complete linking and Ward methods are used (Amhimmid et al., 2025); the number of clusters that can be recommended is in the range of two to five. After determining the number of clusters, the Average silhouette width was calculated, in order to ensure the robustness of the results, a value of this indicator greater than 0.5 indicating the achievement of an efficient clustering method.

Subsequently, the dendrogram of the best clustering method determined was presented, creating a map to help visualize the clusters. Subsequently, the results of the cluster analysis were compared with those of LL. The last step of the analysis consisted of presenting the dynamics for each indicator, from the analyzed period for the determined clusters.

Results and discussions

This section begins with determining the factorability of the indicators, subsequently applying Principal Component Analysis and Factorial Analysis in order to determine the comprehensive factor score. Next, various methods for identifying the number of clusters are presented, implementing the hierarchical algorithms method. For the clusters obtained, the results are compared with the Life Ladder, and the dynamics of the clusters for each indicator are also displayed.

As explained in the methodology, the pre-analysis stage involves data standardization. A first step in implementing factor analysis is to determine the factorability of the data. In this regard, Table 2 presents the MSA at the individual and overall levels. The value of 0.81 indicates very good factorability, which is why this analysis can be applied to these variables.

Table 2. KMO - Measure of Sampling Adequacy

Indicator	IGDP	SS	HLEB	FMLC	G	PC	PA	NEA	Total
MSA	0.89	0.88	0.65	0.91	0.88	0.81	0.91	0.48	0.81

Source: Authors' own research.

To determine the number of common factors, PCA is applied, the results of which are illustrated in Table 3. As only two of the components had eigenvalues greater than one, and the percentage taken up by them is quite close to 80%, they were taken into account in the analysis.

Table 3. Principal Component Analysis

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
SS	4.21	1.52	0.69	0.51	0.41	0.30	0.22	0.14
Proportion	0.53	0.19	0.09	0.06	0.05	0.04	0.03	0.02
Cumulative	0.53	0.72	0.80	0.87	0.92	0.95	0.98	1.00

Source: Authors' own research.

In the case of common factors, an orthogonal rotation is performed using the Varimax method, in order to maximize the variance. In Table 4, the new results related to FA are highlighted, the two eigenvalues being taken into account in the construction of the comprehensive factor score.

Table 4. Factor analysis

	RC1	RC2
SS	3.72	2.00
Proportion	0.47	0.25
Cumulative	0.47	0.72

Source: Authors' own research.

The loadings are presented in Figure 1. In the case of the first principal component, the strongest links are identified with the variables IGDP, FMLC, HLEB, G and PA in the direct sense and PC in the reverse sense. As for the second component, it is very clear that the links with the variables are weaker compared to those highlighted for the first component, this is due to the fact that the first component takes the maximum of the variance of the original variables. Thus, a direct link with NEA and an inverse one with SS is found.

As for the relationship with the variables viewed from the perspective of factor analysis (Figure 1 b) the value associated with the first factor and the HLEB variable is much higher compared to that obtained by the first component and the same variable. However, the meaning of the relationships between the first factor and the original variables is the same as in the case of the first component; with the mention that the relationship between the first factor and NEA is insignificant. In the case of the second factor the relationship with SS is positive, and that with NEA is negative, exactly the opposite compared to the second main component.

	PC1	PC2		RC1	RC2
IGDP	0.84	0.22	IGDP	0.85	0.16
SS	0.49	-0.54	SS	0.22	0.70
HLEB	0.59	0.69	HLEB	0.83	-0.38
FMLC	0.79	-0.14	FMLC	0.66	0.46
G	0.75	0.23	G	0.78	0.11
PC	-0.91	0.06	PC	-0.80	-0.44
PA	0.86	-0.03	PA	0.77	0.39
NEA	-0.39	0.79	NEA		-0.88

(a)

(b)

Figure 1. Loadings: PCA (a) and FA (b)

The comprehensive factor score is calculated using common factor scores and eigenvalues for common factors, forming a matrix with which to perform cluster analysis. Because the number of clusters is not known and must be defined by the researcher, several methods were chosen to help identify the number of classes.

Figure 2 encapsulates three methods, the first being the Within Sum of Square (a), it being desirable that its value be as high as possible, the choice of the optimal number of classes being made by the Elbow method, thus, two classes being taken into account. In Figure 2 (b) the Average Silhouette Width is calculated for a maximum number of 10 clusters, it is clearly observed that the highest value was attributed to the formation of two clusters. The Gap Statistic method (Figure 2 c), also indicates the grouping of countries into two classes.

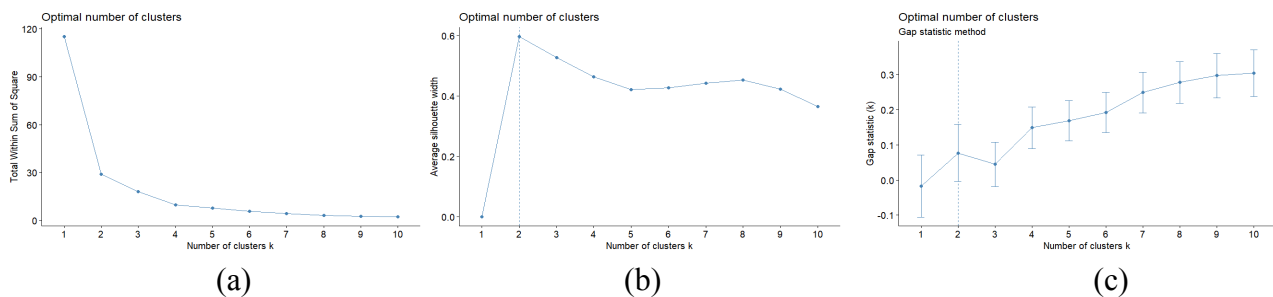


Figure 2. Methods for determining the number of clusters: Within Sum of Square (a), Average Silhouette Width (b), Gap Statistic (c)

Source: Authors' own research.

Although all these methods would have been sufficient to conclude that the optimal number is two clusters, the NbClust package was used, which includes 30 indices, including the three presented previously, in order to once again verify the optimal number of clusters. In the R function, the type of distance, the linking method, and the range in which to find the number of clusters were specified. Thus, the Euclidean distance was chosen, based on linking methods, the complete distance, and Ward's distance, with the number of clusters being a minimum of two and a maximum of five. In this sense, by implementing the complete method, most indicators recommend two clusters, while Ward's method recommends four classes (Table 5).

Table 5. Identifying the number of clusters using the NbClust package

Linkage method	Number of indicators	Number of recommended clusters
complete	10	2
	2	3
	7	4
	4	5
Ward	8	2
	2	3
	10	4
	3	5

Source: Authors' own research.

Because different numbers of classes were recommended, Average silhouette widths were calculated for the two methods of grouping countries into classes, finding that for the complete

method the value was 0.6, a value of 0.5 indicating a good way of grouping. Thus, the countries were grouped into two classes using the complete method.



Figure 3. Average silhouette width: complete (a) and Ward's method (b)

Source: Authors' own research.

Figure 4 illustrates the dendrogram according to which Denmark, Sweden, Ireland, Luxembourg, Finland, the Netherlands, Estonia, Austria, Germany, Belgium, France and Malta were placed in one class, while the other EU member states were included in the second cluster.

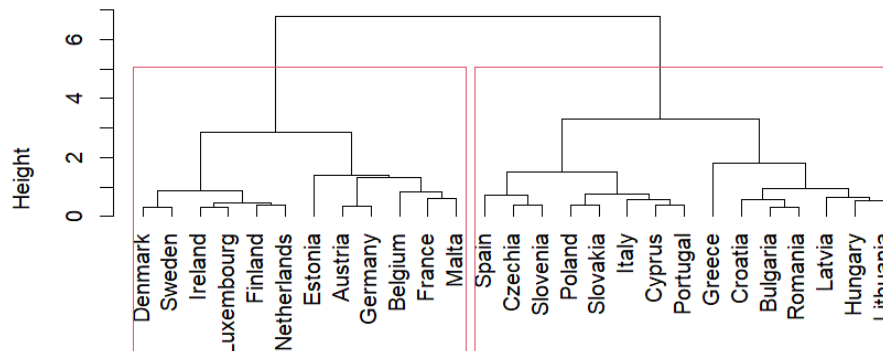


Figure 4. Dendrogram

Source: Authors' own research.

As can be seen from Figure 5, there is some influence of geographical location in the formation of clusters, with countries in the north and west being assigned to cluster 1, while countries in the south and east were placed in cluster 2.

The eight indicators taken in the cluster analysis are used to explain the happiness of citizens, this being expressed by the variable LL. As can be seen by comparing Figure 5 and Figure 6, countries with high values assigned to happiness were grouped in cluster 1.

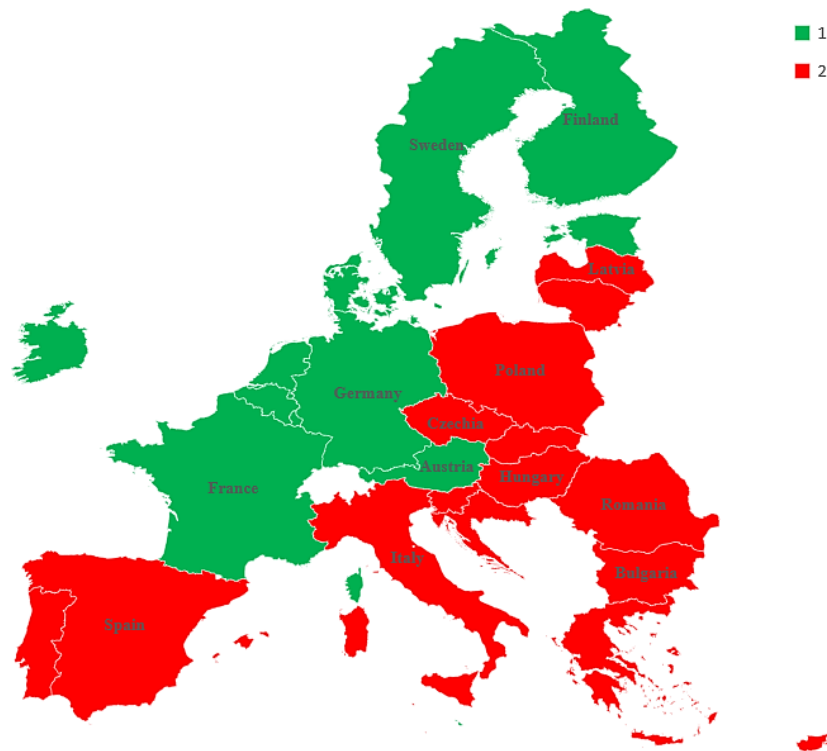


Figure 5. Cluster map

Source: Authors' own research.

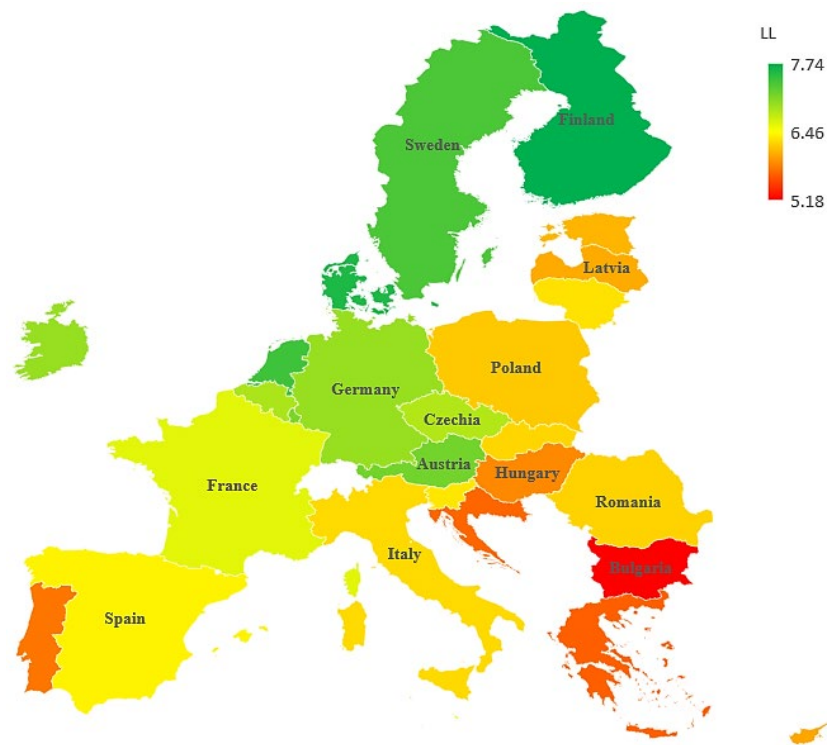


Figure 6. Life ladder

Source: Authors' own research.

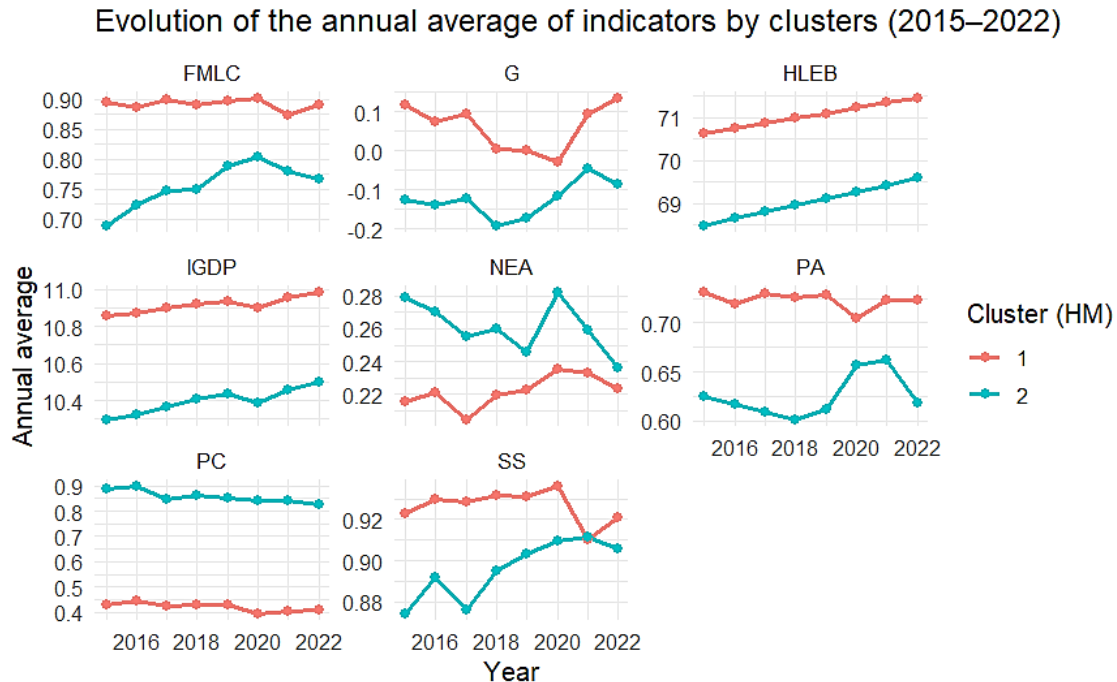
Table 6 presents the averages obtained by the two clusters for the eight indicators analyzed. In this regard, it can be noted that cluster 1 presents better values for all indicators. The biggest difference highlighted is that regarding the PC, a sign that citizens in the second cluster believe that corruption is very high in the countries they come from. Cluster 2 is characterized by sadder people (higher NEA) and less generous (negative G) although there are no significant differences between the perceived social support (SS) and the average wealth level (IGDP) of the clusters.

Table 6. Cluster averages

Cluster	IGDP	SS	HLEB	FMLC	G	PC	PA	NEA
1	10.91	0.93	71.05	0.89	0.06	0.42	0.72	0.22
2	10.40	0.90	69.04	0.76	-0.13	0.86	0.63	0.26

Source: Authors' own research.

Figure 7 illustrates the dynamics of each indicator from the perspective of the two clusters in the period 2015-2022. Thus, for the FMLC indicator, throughout the analyzed period, a much better opinion was observed among those in cluster 1, however, a decrease was noted in 2021, most likely as a result of the pandemic. In the case of cluster 2, the trend was upward until 2020, since then continuing to decrease until the end of the analyzed period. In case G, until 2020, the trend was downward in cluster 1, with a spectacular increase in 2021, the value continuing to increase in 2022. For the second cluster, a decrease is noted in 2018 and another in 2022, otherwise, the trend is upward. For HLEB, a constant growth trend is highlighted from one year to the next, although the value assigned to cluster 1 was always higher compared to that in cluster 2. For IGPD, the trends for both clusters are similar, being characterized by an upward trend with a slight decrease in 2020. Again, cluster 2 fell below cluster 1, although, until 2019, the growth trend was higher in the case of cluster 2. PC had a slight decreasing trend in the case of the second cluster; values remained relatively constant during the period 2015-2019 for cluster 1, followed by a decrease in 2020, with values increasing slightly in the last two years investigated. Although SS was much higher in cluster 1 during the period 2015-2020, 2021 was characterized by a much lower value; even after the increase in the value in 2022, the level of the variable did not return to that of 2020. 2021 was the year for which the average in cluster 1 was lower than that of cluster 2, which was characterized by an upward trend of this indicator during the period 2017-2021, followed by a slight decrease in SS in 2022. Finally, the affects were left, in the case of the positive ones (PA), noting that in 2020 a decrease was noted in the case of the first cluster, otherwise, the values being quite similar for this class. In the case of cluster 2, the trend was one of decrease in the period 2015-2028, followed by an increase until 2021, and a significant decrease in 2022. In the case of NEA, despite the decrease in 2017 and 2022, the trend was one of increase for this indicator, a sign that people in cluster 1 are experiencing more negative emotions. Except for the slight increase in 2018, until 2019, the trend associated with cluster 2 was a downward one, followed by a sudden increase in 2020 and an equally sudden decrease in the following period, managing to reach a minimum value in 2022. Although in cluster 2 the trend was one of decrease, it still did not reach the level of cluster 1, although in 2022 the values were quite close.



The results present the existence of differences between the countries of Northern and Western Europe and those of Central and Eastern Europe for subjective well-being. The cluster analysis highlighted two homogeneous groups for the two geographic regions. These discrepancies strengthen the hypothesis that the indicators used – logarithm of GDP per capita, social support, healthy life expectancy, freedom of choice, generosity, perception of corruption and affective dimensions – reflect the level of perceived happiness in Europe. Similar results were obtained in the study conducted by Ciucu-Durnoi et al. (2025) which highlights the importance of GDP per capita, social support and life expectancy on happiness in the European Union. The study shows that Western and Nordic countries have higher levels of subjective well-being compared to Eastern European countries, thus reinforcing the idea that the existence of the two clusters identified in the present research is confirmed in Europe. Bartolini and Sarracino (2014) demonstrate that social relationships can explain long-term variations in well-being better than traditional economic indicators that often consider only economic implications, not perceived emotions. This study confirms the results obtained for Cluster 1, where high levels of social support are associated with higher levels of well-being. In contrast, the study conducted by Akgun et. al (2022) revealed that there is a positive relationship between traditional economic indicators such as inflation rate or current taxes and happiness index ratings. Conversely, the importance of the role of individual freedom in determining well-being is also mentioned by Rahman and Veenhoven (2018), who show that freedom of choice is a significant determinant of well-being.

In addition to the theoretical and empirical contribution, the results obtained can contribute to the development of European public policies to reduce the differences observed between the two clusters. The results of this research show discrepancies in economic and social development between the states of Western Europe and those of Central and Eastern Europe.

These discrepancies are attempted to be reduced by the cohesion policies of the European Union which aim to reduce the differences between regions and promote economic and social convergence. The reports of the European Commission claim that the European Structural and Investment Funds have contributed and will contribute to improving the standard of living in the states of Central and Eastern Europe through investments in infrastructure, education and the labor market, supporting the process of economic convergence until 2050 (European Commission, 2022). Another initiative that aims to strengthen social inclusion and improve access to education and employment is the European Pillar of Social Rights, which has as its main objectives the reduction of at least 15 million people at risk of poverty or social exclusion, the annual participation in vocational training activities of at least 60% of adults and ensuring a job for 78% of the population between 20 and 64 years old (European Union, 2017). For the period 2021 - 2027, the European Union proposes that through the European Social Fund Plus (ESF) it will support young people to obtain a qualification, find a job and improve their level of education and skills; retraining and upgrading people for a digital and green economy by increasing the quality of education and vocational training systems. Supporting social innovation, allocating resources to support children in difficulty and combating poverty (Delhey et al., 2018; European Union, 2021).

Conclusion

The current study uses panel data analysis and cluster analysis through NbClust package in R to examine key indicators of happiness drawn from the World Happiness Report for European countries between 2015 and 2022. These indicators: log of GDP per capita, social support, healthy life expectancy at birth, freedom to make life choices, generosity, perceptions of corruption, positive affect, and negative affect offers a profound understanding of perceived happiness across the European continent. Following the analysis, two clusters were identified that divided the countries into the following clusters. Cluster 1: Denmark, Sweden, Ireland, Luxembourg, Finland, Netherlands, Estonia, Austria, Germany, Belgium, France and Malta and cluster 2: Spain, Czech Republic, Slovakia, Romania, Poland, Italy, Slovenia, Cyprus, Portugal, Greece, Croatia, Lithuania, Latvia and Hungary. Indicators considered in the cluster analysis had values close to the results from The World Happiness Report for the Life Ladder classification. It was observed that for the two clusters in the analyzed period the Healthy life expectancy at birth indicator registered a constant increase. It is worth mentioning that Perceptions of corruption increased slightly in 2016 for both clusters, and starting with 2017 it registered a slight decrease for both clusters. Major changes were recorded in the case of Negative affect, which fluctuated considerably for cluster 2 in 2020 with the pandemic, then began to decrease. Regarding cluster 1, an increase was recorded from 2018 to 2021, and it will only decrease in 2022. Moreover, the findings of the current study highlight a discrepancy between the northern and western European countries (Cluster 1) and eastern European nations (Cluster 2). This enforces the importance of the role of social support, freedom and health in promoting resilience as well as the significance of each country's economic status.

While the current research contributes to the existing literature and offers valuable insights regarding happiness across Europe by providing insights for policymakers aiming to improve well-being through targeted improvements in health, governance, or social policies, the authors considered some limitations. One significant constraint is the reduced number of analyzed indicators. Although the indicators were collected from World Happiness Report, there

are factors that were not covered by the current analysis such as environmental quality, education access, or cultural variables that could improve the results of the study. Additionally, another limitation may be considered the novelty of the data presented. For this research the used data was collected from The World Happiness Report published between 2015 and 2022. Future researchers could extend this work by considering most recent that that will be available in March 2026 for 2025 year on the report webpage. Moreover, the use of panel data and clustering methods assumes certain statistical properties that may not fully account for unobserved heterogeneity or endogeneity, suggesting opportunities for more advanced econometric techniques in future studies.

Acknowledgment

This paper was co-financed by The Bucharest University of Economic Studies during the PhD program.

References

- Abdur Rahman, A., & Veenhoven, R. (2018). Freedom and Happiness in Nations: A Research Synthesis. *Applied Research in Quality of Life*, 13(2), 435-456. <https://doi.org/10.1007/s11482-017-9543-6>
- Adylla, F. P. (2025). *View of CLUSTERING OF COUNTRIES BASED ON WORLD HAPPINESS INDICATORS USING K-MEANS*. <https://journal.unj.ac.id/unj/index.php/statistika/article/view/62022/22353>
- Akgun, A. İ., Türkoğlu, S. P., & Erikli, S. (2022). Investigating the determinants of happiness index in EU-27 countries: A quantile regression approach. *International Journal of Sociology and Social Policy*, 43(1-2), 156-177. <https://doi.org/10.1108/IJSSP-01-2022-0005>
- Amhimmid, Fatma Alzahra A. Aljehany, Mohamed A. Mohamed, & Kasem A. Farag. (2025). Comparative Study of Four Methods in Hierarchical Cluster Analysis. *Libyan Journal of Medical and Applied Sciences*, 11-16. <https://doi.org/10.64943/ljmas.v3i4.183>
- Arutyunova, A., & Röglin, H. (2025). The Price of Hierarchical Clustering. *Algorithmica*, 87(10), 1420-1452. <https://doi.org/10.1007/s00453-025-01327-7>
- Banerjee, A., Chakraborty, S., & Biswas, A. (2019). *Selection of link function in binary regression: A case-study with world happiness report on immigration* (arXiv:1910.07748). arXiv. <https://doi.org/10.48550/arXiv.1910.07748>
- Bartolini, S., Bilancini, E., & Sarracino, F. (2016). Social Capital Predicts Happiness Over Time. In S. Bartolini, E. Bilancini, L. Bruni, & P. L. Porta (Ed.), *Policies for Happiness* (pp. 175-198). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198758730.003.0009>
- Bartolini, S., & Sarracino, F. (2014). *It's not the economy, stupid! How social capital and GDP relate to happiness over time* (arXiv:1411.2138). arXiv. <https://doi.org/10.48550/arXiv.1411.2138>
- Behera, D. K., Rahut, D. B., Padmaja, M., & Dash, A. K. (2024). Socioeconomic determinants of happiness: Empirical evidence from developed and developing countries. *Journal of Behavioral and Experimental Economics*, 109, 102187. <https://doi.org/10.1016/j.socec.2024.102187>
- Boarini, R., Comola, M., Smith, C., Manchin, R., & De Keulenaer, F. (2012). *What Makes for a Better Life?: The Determinants of Subjective Well-Being in OECD Countries – Evidence*

- from the Gallup World Poll (OECD Statistics Working Papers Nr. 2012/03; OECD Statistics Working Papers, Vol. 2012/03). <https://doi.org/10.1787/5k9b9ltjm937-en>
- Bojanowska, A. (2016). *Lay Understanding of Happiness and the Experience of Well-Being: Are Some Conceptions of Happiness More Beneficial than Others?* | *Journal of Happiness Studies* | Springer Nature Link. <https://link.springer.com/article/10.1007/s10902-015-9620-1>
- Chakraborty, A., & Tsokos, C. P. (2021). A Real Data-Driven Clustering Approach for Countries Based on Happiness Score. *Amfiteatru Economic*, 23(S15), 1031. <https://doi.org/10.24818/EA/2021/S15/1031>
- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). **NbClust**: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, 61(6). <https://doi.org/10.18637/jss.v061.i06>
- Ciucu-Durnoi, A.-N., Jigani, A.-I., Chiriță, N., & Delcea, C. (2025). Happiness in the European Union: A Panel Regression Model and TOPSIS Analysis. *Proceedings of the International Conference on Business Excellence*, 19(1), 2650-2667. <https://doi.org/10.2478/picbe-2025-0204>
- Deaton, A. (2008). Income, Health, and Well-Being around the World: Evidence from the Gallup World Poll. *Journal of Economic Perspectives*, 22(2), 53-72. <https://doi.org/10.1257/jep.22.2.53>
- Delhey, J., Dragolov, G., & Boehnke, K. (2018). *Social cohesion and well-being in Europe*. Eurofound. <https://doi.org/10.2806/261816>
- Diener, E. (2003). The evolving concept of subjective well-being: The multifaceted nature of happiness. In *Advances in Cell Aging and Gerontology* (Vol. 15, pp. 187-219). Elsevier. [https://doi.org/10.1016/S1566-3124\(03\)15007-9](https://doi.org/10.1016/S1566-3124(03)15007-9)
- Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In *Nations and Households in Economic Growth* (pp. 89-125). Elsevier. <https://doi.org/10.1016/B978-0-12-205050-3.50008-7>
- European Commission. (2022, februarie 9). *8th Cohesion Report: Cohesion in Europe towards 2050* [Text]. European Commission - European Commission. https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_763
- European Union. (2017). *Pilonul european al drepturilor sociale—EUR-Lex* [Text]. European Union. <https://eur-lex.europa.eu/RO/legal-content/glossary/european-pillar-of-social-rights.html>
- European Union. (2021, ianuarie 1). *Fondul social european Plus – FSE+ (2021-2027)* | EUR-Lex. European Union. <https://eur-lex.europa.eu/RO/legal-content/summary/european-social-fund-plus-esf-2021-2027.html>
- Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis*.
- Helliwell, J. F., Layard, R., & Sachs, J. D. (2012, aprilie 4). *World Happiness Report 2012*. <https://worldhappiness.report/ed/2012/>
- Helliwell, J. F., Layard, R., & Sachs, J. D. (2019, martie 20). *World Happiness Report 2019*. <https://worldhappiness.report/ed/2019/>
- Helliwell, J. F., Layard, R., Sachs, J. D., Aknin, L. B., De Neve, J.-E., & Wang, S. (2023). *World Happiness Report 2023 (11th ed.)*. Sustainable Development Solutions Network. 978-1-7348080-5-6
- Helliwell, J. F., Layard, R., Sachs, J. D., Neve, J.-E. D., Aknin, L. B., & Wang, S. (2025, martie 20). *World Happiness Report 2025*. <https://worldhappiness.report/ed/2025/>

- Jolliffe, I. T. (2002). *Principal Component Analysis*. Springer-Verlag. <https://doi.org/10.1007/b98835>
- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141-151. <https://doi.org/10.1177/001316446002000116>
- Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, 34(1), 111-117. <https://doi.org/10.1177/001316447403400115>
- Kaufman, L., & Rousseeuw, P. (1990). Frontmatter. In *Finding Groups in Data* (pp. i-xiv). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470316801.fmatter>
- Pelka, M. (2019). *Analysis of happiness in EU countries using the multi-model classification based on models of symbolic data—Lower Silesian Digital Library*. <https://www.dbc.wroc.pl/dlibra/publication/140897/edition/72797>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Rozario, M. M. (2018). *Happiness and Economic Growth: A cross sectional analysis*. <https://dspace.uiu.ac.bd/handle/52243/301>
- Seligman. (2011). *The PERMA Model: A Scientific Theory of Happiness*.
- Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the Number of Clusters in a Data Set Via the Gap Statistic. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 63(2), 411-423. <https://doi.org/10.1111/1467-9868.00293>
- Ulkhag, M. M., & Adyatama, A. (2021). Clustering countries according to the world happiness report 2019. *Engineering and Applied Science Research*, 48(2), 137-150.
- Wang, W., & Lu, Y. (2021). Application of Clustering Analysis of Panel Data in Economic and Social Research Based on R Software. *Academic Journal of Business & Management*, 3(10). <https://doi.org/10.25236/AJBM.2021.031018>
- Wickham, H. (2016). *Preface to the third edition – ggplot2: Elegant Graphics for Data Analysis (3e)*. <https://ggplot2-book.org/preface-3e.html>
- Zhao, Q., Xu, M., & Fränti, P. (2009). P. Sum-of-Squares Based Cluster Validity Index and Significance Analysis. In *Adaptive and Natural Computing Algorithms* (pp. 313-322). Springer Berlin Heidelberg.