JURNAL STATISTIKA DAN APLIKASINYA

e-ISSN 2620 – 8369
VOLUME 9 (ISSUE 2) DECEMBER 2025: 93-103
https://doi.org/10.21009/JSA.09209

CLUSTERING OF COUNTRIES BASED ON WORLD HAPPINESS INDICATORS USING K-MEANS

Fahira Puti Adylla[1], Dian Lestari[2*], Yekti Widyaningsih[3]

[1,2,3] Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia
Jl. Margonda Raya, Pondok Cina, Depok, 16424, Indonesia

Corresponding author's e-mail: ***dian.lestari@sci.ui.ac.id***

ABSTRACT

Article History:
Received: 13, November 2025
Revised: 15, December 2025
Accepted: 29, December 2025
Published: 31, December 2025
Available online.

Keywords:

Biplot; Multivariate Analysis; Silhouette Index; Subjective Well-being

Happiness is a multidimensional concept encompassing emotional well-being, life satisfaction, and perceived quality of life. The increasing use of happiness indicators as complementary measures of development beyond economic growth has attracted growing attention in statistical and applied research. This study aims to classify countries based on a comprehensive set of world happiness indicators using the K-Means clustering method. The indicators include the Happiness Index (subjective), gross domestic product (GDP) per capita, social support, healthy life expectancy, freedom to make life choices, generosity, negative perceptions of corruption, crime index, and cost of living. The optimal number of clusters is determined using the Silhouette Index, while Biplot analysis is employed to visualize cluster characteristics and relationships among indicators. The results identify three distinct clusters. Cluster 1 is dominated by countries with low happiness levels, Cluster 2 represents countries with moderate happiness profiles, and Cluster 3 consists of countries with high happiness levels. The findings demonstrate the effectiveness of multivariate clustering techniques in revealing structural patterns in happiness data and provide empirical evidence that may support comparative statistical analysis and policy-oriented applications.

How to cite this article:
F. P. Adylla, D. Lestari, Y. Widyaningsih, "CLUSTERING OF COUNTRIES BASED ON WORLD HAPPINESS INDICATORS USING K-MEANS", Jurnal Statistika dan Aplikasinya, vol. 9, iss. 2, pp. 93 – 103, December 2025

Journal e-mail: *jsa@unj.ac.id*
Research Article · Open Access

1. INTRODUCTION

The measurement and analysis of happiness have become increasingly important topics in applied statistics and social science research. Research by Rozario in 2018 shows that happiness and economic growth are positively related and more significant in developed countries than developing countries [1]. As conventional economic indicators such as gross domestic product (GDP) provide only a partial description of societal progress, statisticians and applied researchers have sought alternative indicators that better capture human well-being. Happiness, measured through subjective evaluations and objective socio-economic indicators, offers a rich multidimensional framework that can be explored using multivariate statistical methods. Rapid advances in economic development, technology, and globalization have profoundly transformed societies worldwide. Although these advances have improved material living standards, they have also generated new social challenges, including inequality, environmental stress, declining social cohesion, and increasing mental health concerns. As a result, there is growing consensus that traditional economic indicators, such as gross domestic product (GDP), are insufficient to fully capture societal progress and human well-being. This limitation has motivated the increasing use of happiness and subjective well-being as complementary measures of development.

Happiness is commonly defined as a state of positive emotional experience combined with a cognitive evaluation of life satisfaction and meaning. Lyubomirsky emphasizes that happiness encompasses joy, affection, contentment, and the belief that one's life is valuable and worthwhile [2]. In empirical research, happiness is typically measured through both subjective and objective indicators. Subjective indicators capture individuals' self-assessments of their lives, while objective indicators describe material, social, and institutional conditions that shape well-being, including income, health, governance quality, and social relationships [3].

One of the most influential global efforts to measure happiness is the World Happiness Report (WHR), published annually by the United Nations. The WHR combines survey-based happiness scores with key socio-economic indicators such as GDP per capita, social support, healthy life expectancy, freedom to make life choices, generosity, and perceptions of corruption. These indicators reflect the multidimensional nature of happiness and provide a comprehensive framework for cross-country comparisons. However, much of the existing literature relies on regression-based or ranking-oriented analyses, which may obscure underlying multivariate structures and latent similarities among countries.

Grouping countries based on happiness indicators is important because it enables the identification of clusters of nations that share similar well-being profiles. Such clustering moves beyond simple rankings and allows policymakers to benchmark countries against structurally comparable peers rather than global averages. Moreover, clustering can reveal complex interactions among happiness determinants that are not captured by univariate or bivariate analyses. Therefore, this study applies K-Means clustering to classify countries based on a broad set of happiness-related indicators, complemented by Biplot analysis to enhance interpretability and visualization.

The theoretical foundation of happiness research is inherently interdisciplinary, drawing from economics, psychology, sociology, and public policy. In economics, traditional welfare theory equates well-being with utility derived from income and consumption. However, empirical findings such as the Easterlin Paradox demonstrate that long-term economic growth does not necessarily translate into sustained increases in happiness, prompting a broader conceptualization of welfare [4]. Sen's capability approach further argues that well-being should be evaluated based on individuals' capabilities and freedoms rather than solely on material resources [5].

From a psychological perspective, happiness is closely related to the concept of subjective well-being, which consists of cognitive components (life satisfaction) and affective components (positive and negative emotions) [6]. Social and cultural contexts play a critical role in shaping happiness, as social support, trust, and freedom influence individuals' perceptions of life quality. These perspectives collectively justify the use of multiple indicators to capture happiness at the national level [7].

Methodologically, clustering analysis originates from multivariate statistics and unsupervised learning. K-Means clustering aims to partition observations into homogeneous groups by minimizing within-cluster variance and maximizing between-cluster differences. When applied to happiness data, clustering facilitates the identification of latent structures and country groupings that reflect similar socio-economic and institutional conditions. Visualization tools such as Biplot analysis further support interpretation by jointly representing variables and observations in a reduced-dimensional space.

A substantial body of literature has examined the determinants of happiness across countries. Early contributions by Easterlin in 1974 [4] questioned the direct relationship between income growth and happiness, while subsequent studies highlighted the importance of social capital, trust, and institutional quality [8][9]. Health outcomes and life expectancy have also been shown to significantly influence happiness levels [10]. Healthy life expectancy at birth in this study is based on data taken from the World Health Organization (WHO). According to WHO , there has been a significant increase in life expectancy worldwide, with global healthy life expectancy at birth increasing by 6 years since 1990. However, there are large differences in life expectancy between countries and regions [11].

The World Happiness Report has become a primary data source for empirical happiness research. Many studies employ regression-based approaches to assess the contribution of individual indicators to happiness scores. For example, Helliwell et al. (2019) demonstrate that social support and freedom of choice are consistently strong predictors of life satisfaction across countries [12]. Other studies emphasize the negative effects of corruption and crime on well-being [13][14].

Despite these advances, several gaps remain in the literature. First, most studies focus on marginal effects of individual indicators rather than the multivariate structure of happiness. Second, country rankings dominate the discourse, offering limited insight into latent groupings of nations with similar happiness profiles. Third, relatively few studies integrate additional contextual indicators, such as crime and cost of living, into global happiness clustering frameworks. Addressing these gaps requires the application of unsupervised multivariate techniques that can uncover hidden patterns in happiness data.

This study contributes to the statistical literature and its applications in three main ways. First, it extends conventional analyses of world happiness by incorporating additional contextual indicators—namely crime index and cost of living—into a unified multivariate framework. Second, the study applies K-Means clustering combined with Silhouette-based validation and Biplot visualization, demonstrating a practical application of unsupervised statistical learning methods to real-world socio-economic data. Third, by focusing on latent group structures rather than rankings, the study provides an alternative statistical perspective for comparative analysis of countries. This study offers several novel contributions to the literature on global happiness. First, it integrates both conventional World Happiness Report indicators and additional contextual variables—namely crime index and cost of living—providing a more comprehensive representation of national well-being. Second, the study applies a combined K-Means clustering and Biplot visualization framework, allowing for both rigorous classification and intuitive interpretation of happiness patterns. Third, by focusing on latent groupings rather than rankings, the study advances comparative happiness analysis and offers policy-relevant insights for countries facing similar well-being challenges. These contributions extend existing happiness research by emphasizing multidimensional structure, methodological integration, and practical interpretability.

2. METHODS

Material and Data

The analysis employs cross-sectional country-level data derived primarily from the World Happiness Report (www.worldhappiness.report), supplemented with crime index and cost-of-living data from international databases (www.numbeo.com). The subjective indicator is the Happiness Index based on the Cantril Ladder [15]. Objective indicators include GDP per capita, social support, healthy life expectancy, freedom to make life choices, generosity, negative perceptions of corruption, crime index, and cost of living. All variables are standardized prior to analysis. All calculations in this study were carried out with the help of RStudio software [16].

Research Method

K-Means Clustering and Cluster Validation

K-Means clustering is used to partition countries into K clusters based on Euclidean distance. The optimal number of clusters is determined using the Silhouette Index, which evaluates cluster cohesion and separation [17][18][19][20][21].

K-Means is a hard clustering method in which each observation is assigned to exactly one cluster. The K-Means algorithm groups observations by assigning each observation to the cluster with the nearest centroid (mean). The clustering procedure using the K-Means method is described as follows.

1. Determine the number of clusters, c, and initialize the cluster centroids randomly from the observed data. In this study, the optimal number of clusters is determined using the Silhouette index.
2. Assign each observation to the nearest cluster centroid based on a distance measure. In this study, the Euclidean distance is used, defined as follows:

$$d_{ik} = \sqrt{\sum_{j=1}^{p} \left(x_{ij} - v_{kj}\right)^2} \tag{1}$$

where.
d_{ik} : Euclidean distance between the *i*-th observation and the *k*-th cluster centroid
x_{ij} : the value of the *i*-th observation in the k-th cluster for the *j*-th variable
v_{kj} : centroid of the *k*-th cluster for the *j*-th variable
i : observation index, $i = 1, 2, \dots, n$
j : variable index where $j = 1, 2, \dots, p$
k : cluster index where $k = 1, 2, \dots, c$

3. Update the cluster centroids using the following equation:

$$v_{kj} = \frac{\sum_{i=1}^{n_k} x_{ij}}{n_k} \tag{2}$$

where:
v_{kj} : centroid of the *k*-th cluster for the *j*-th variable
x_{ij} : the value of the *i*-th observation in the *k*-th cluster for the *j*-th variable
n_k : the number of observation of the *k*-th cluster

4. Recalculate the distance between each observation and the update centroids, then reassign each observation to the nearest centroid.
5. Repeat Steps 3 and 4 until convergence is achieved, that is, until there are no changes in cluster membership.

The determination of the optimal number of clusters in Step 1 is based on the Silhouette value. The Silhouette value measures how well an observation fits within its assigned cluster compared to other clusters and is defined as follows:

$$s_i = \frac{[b_i - a_i]}{\max\{b_i, a_i\}}, \qquad i = 1, 2, \dots, n \tag{3}$$

where:
a_i : average distance between the *i*-th observation and all other observations within the same cluster
b_i : minimum average distance between the *i*-th observation and observations in other clusters

Finally, the overall Silhouette index is computed as:

$$S = \frac{1}{n}\sum_{i=1}^{n} s_i \tag{4}$$

where:
n = total number of observations
s_i = silhouette for observation i

Biplot Analysis

Biplot analysis based on Singular Value Decomposition is employed to visualize the relationships among variables and country clusters in a two-dimensional space [21][22].

A biplot is an exploratory data analysis technique used to enhance clustering analysis by simultaneously
displaying observations and variables in a two-dimensional graphical representation [22]. The procedure for conducting a biplot analysis is described as follows [21]:

1. Construct a data matrix $\mathbf{X}$ consisting of n observations and p variables.

2. Decompose the matrix $\mathbf{X}$ using Singular Value Decomposition (SVD):

$$\mathbf{X} = \mathbf{UDV'} \tag{5}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices. $\mathbf{D}$ is a diagonal matrix whose diagonal elements are the square roots of the eigenvalues of $\mathbf{X'X}$, and $\mathbf{V}$ is a matrix containing the eigenvectors corresponding to the eigenvalues of $\mathbf{X'X}$. The matrix $\mathbf{U}$ is obtained using $\mathbf{U} = \mathbf{XVD^{-1}}$.

3. Form matrices $\mathbf{G}$ and $\mathbf{H}$ such that:

$\mathbf{G} = \mathbf{UD}^{\alpha}$ and $\mathbf{H}' = \mathbf{D}^{1-\alpha}\mathbf{V}'$ where $0 \leq \alpha \leq 1$.

Thus, Equation (5) can be rewritten as:

$$\mathbf{X} = \mathbf{GH'} \tag{6}$$

4. The *(i,j)*-th element of the matrix $\mathbf{X}$ can be expressed as:

$$x_{ij} = \mathbf{g}_i'\mathbf{h}_j\ ;\ i = 1,2,\dots,n; j = 1,2,\dots,p \tag{7}$$

where $\mathbf{g}_i'$ is the *i*-th row vector of matrix $\mathbf{G}$ and $\mathbf{h}_j$ is the *j*-th row vector of matrix $\mathbf{H}$, both having dimension r.

5. Select the first two columns of matrices $\mathbf{G}$ and $\mathbf{H}$, denoted as $\mathbf{G_2}$ and $\mathbf{H_2}$, respectively.

6. Construct a two-dimensional coordinate plot using matrices $\mathbf{G_2}$ and $\mathbf{H_2}$. Each row of matrix $\mathbf{G_2}$ represents the coordinates of an observation, while each row of matrix $\mathbf{H_2}$ represents the coordinates of a variable.

7. The goodness of fit of the biplot representation of matrix $\mathbf{X}$ is measured using the following equation:

$$\rho^2 = \frac{\lambda_1 + \lambda_2}{\sum_{k=1}^{r} \lambda_k} \tag{8}$$

λ_1 : the largest eigenvalue
λ_2: the second largest eigenvalue
λ_k : the *k*-th eigenvalue (k = 1, 2, ..., r)

If the value of ρ^2 is close to one, the biplot representation provides a better approximation of the original data structure.

3. RESULTS

The clustering analysis using the K-Means algorithm indicates that the optimal number of clusters is three, as determined by the Silhouette Index with a value of 0.441. This value suggests a moderately strong clustering structure, indicating that countries within the same cluster share meaningful similarities in their happiness profiles, while remaining sufficiently distinct from countries in other clusters.

Cluster Characteristics

The characteristics of each cluster and its member countries can be seen in Table 1. Cluster 1 comprises primarily countries from Sub-Saharan Africa and South Asia. This cluster is characterized by the lowest average Happiness Index among all groups. Countries in this cluster exhibit low GDP per capita, weak social support systems, low healthy life expectancy, and limited freedom to make life choices. In addition, this cluster records the highest crime index and the most negative public perceptions of corruption. Interestingly, generosity levels in this cluster are relatively high, and the cost of living is comparatively low, suggesting that strong interpersonal solidarity may partially offset structural disadvantages.

Cluster 2 consists mainly of countries from Central and Eastern Europe, the Commonwealth of Independent States, and Latin America and the Caribbean. This group represents a moderate happiness profile. Countries in Cluster 2 display higher GDP per capita and healthier life expectancy than those in Cluster 1, accompanied by stronger social support networks. However, high crime rates and negative perceptions of corruption persist, and generosity levels are relatively low. These characteristics indicate that economic and health improvements alone may be insufficient to substantially enhance happiness without parallel improvements in governance and public safety.

Cluster 3 is dominated by countries from Western Europe, North America, and Australia–New Zealand. This cluster exhibits the highest average Happiness Index. Countries in this group benefit from high GDP per capita, strong social support, long healthy life expectancy, and high levels of freedom to make life choices. Moreover, crime rates and perceived corruption are the lowest among all clusters. The primary trade-off observed in this cluster is a relatively high cost of living, which does not appear to diminish overall happiness levels due to compensating for institutional and social advantages.

Table 1: Regions (Countries) and Characteristics of each cluster

Number Cluster	Region (Countries)	Characteristics
1	- Central and Eastern Europe (Kosovo) - Commonwealth of Independent States (Kyrgyzstan) - Latin America and Caribbean (Honduras) - Middle East and North Africa (Iran, Iraq) - South Asia (Afghanistan, Bangladesh, India, Nepal, Pakistan) - Southeast Asia (Cambodia, Indonesia, Laos, Myanmar) - Sub-Saharan Africa (Burkina Faso, Cameroon, Congo (Brazzaville), Ethiopia, Gambia, Ghana, Ivory Coast, Kenya, Liberia, Madagascar,Malawi, Mozambique, Nigeria, Rwanda, Senegal, South Africa, Tanzania, Uganda, Zimbabwe)	- Low GDP per capita (USD 8313/month) - Weak social support systems (71.3%) - Low healthy life expectancy (58.977 years old) - Limited freedom to make life choices (76.6%) - Highest crime index (54.842) - Moderate public perceptions of corruption (77.5%) - High generosity levels (0.082) - Low cost of living (Rp7.974.000/month)

Number Cluster	Region (Countries)	Characteristics
2	- Central and Eastern Europe (Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Hungary, Latvia, Lithuania, Montenegro, North Macedonia, Poland, Romania, Serbia Slovakia) - Commonwealth of Independent States (Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Moldova, Russia, Tajikistan, Ukraine) - East Asia (China, Mongolia, South Korea) - Latin America and Caribbean (Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, Venezuela) - Middle East and North Africa (Algeria, Egypt, Jordan, Kuwait, Lebanon, Libya, Palestinian Territories, Tunisia, Turkey) - South Asia (Maldives, Sri Lanka) - Southeast Asia (Malaysia, Philippines, Thailand, Vietnam) - Sub-Saharan Africa (Botswana, Gabon, Mauritius) - Western Europe (Belgium, Cyprus, Greece, Italy, Portugal, Spain)	- Moderate GDP per capita (USD 9770/month) - Moderate social support systems (85.3%) - Moderate healthy life expectancy (67.451 years old) - Moderate freedom to make life choices (78.8%) - Moderate crime index (45.870%) - Highest negative public perceptions of corruption (80.1%) - Moderate generosity levels (-0.099) - Moderate cost of living (Rp8.060.000/month)
3	- Central and Eastern Europe (Estonia, Slovenia) - Commonwealth of Independent States (Turkmenistan, Uzbekistan) - East Asia (Hong Kong S.A.R. of China, Japan, Taiwan Province of China) - Middle East and North Africa (Bahrain, Israel, Saudi Arabia, United Arab Emirates) - North America and ANZ (Australia, Canada, New Zealand, United States) - Western Europe (Austria, Denmark, Finland, France, Germany, Iceland, Ireland, Luxembourg, Malta, Netherlands, Norway, Sweden, Switzerland, United Kingdom)	- High GDP per capita (USD 10.772/ month) - Strong social support systems (92.5%) - High healthy life expectancy (71.752 years old) - High freedom to make life choices (89.4%) - Lowest crime index (32.28%) - Lowest negative public perceptions of corruption (48.9%) - Moderate generosity levels (0.047) - High cost of living (Rp13.963.000/month)

Biplot Interpretation

The Biplot analysis explains 64.2% of the total variance in the data, indicating a satisfactory two-dimensional representation of the multivariate structure as can be seen in Figure 1.

GDP per capita has the longest vector, signifying the greatest variability and strongest contribution to differentiating countries. Happiness Index, social support, healthy life expectancy, and freedom to make life choices form acute angles with GDP per capita, suggesting strong positive associations.

In contrast, crime index and negative perceptions of corruption form obtuse angles with the Happiness Index, indicating strong negative relationships. Generosity appears approximately orthogonal to the Happiness Index, implying a weaker direct association at the cross-country level. The spatial distribution of clusters in the Biplot confirms clear separation: Cluster 1 aligns closely with crime and corruption indicators, Cluster 2 occupies an intermediate position, and Cluster 3 aligns strongly with economic prosperity, institutional quality, and subjective well-being.

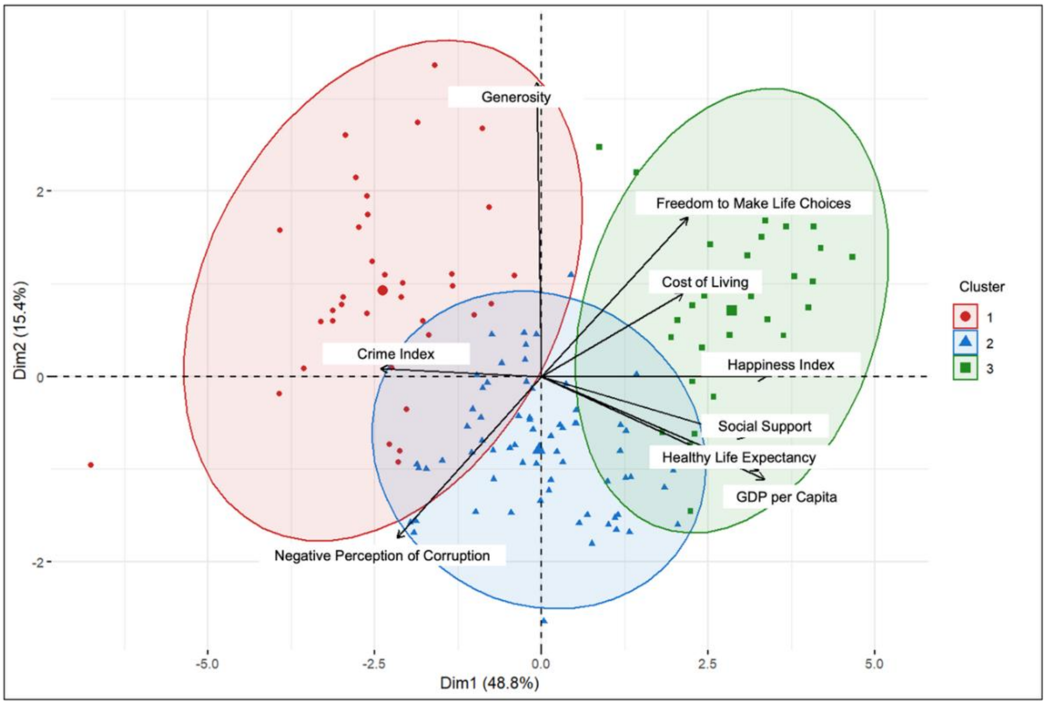

Figure 1. The map of the clustering results of the objective indicators and world happiness index

4. DISCUSSION

The clustering results reveal pronounced global disparities in the determinants of happiness and underscore the usefulness of K-Means clustering as an applied statistical tool for cross-country analysis. By grouping countries based on similarities across multiple happiness-related indicators, the analysis moves beyond simple rankings and highlights latent structures that characterize national well-being profiles. The three clusters identified in this study reflect distinct combinations of economic conditions, social support, institutional quality, and public safety.

Countries in the lowest happiness cluster face persistent structural constraints related to income, health outcomes, governance quality, and public security. The strong association between high crime levels, negative perceptions of corruption, and low happiness scores in this group reinforces existing evidence that institutional weakness and insecurity substantially undermine well-being. In contrast, countries in the highest happiness cluster benefit from a combination of economic prosperity, effective institutions, strong social support networks, and low levels of crime and corruption. These results emphasize that high income alone is insufficient; rather, happiness emerges from the interaction of economic, social, and institutional factors.

From a methodological perspective, the use of K-Means clustering offers several advantages for the present study. The method produces clearly defined and mutually exclusive clusters, which facilitates interpretation and comparison at the country level. This property is particularly valuable in applied statistical and policy-oriented contexts, where clear group membership supports benchmarking and comparative evaluation. The Silhouette Index further confirms that the selected clustering structure provides a reasonable balance between within-cluster homogeneity and between-cluster separation.

Nevertheless, the results should be interpreted with caution. K-Means clustering assumes spherical cluster shapes and equal importance of variables after standardization, which may oversimplify complex relationships in happiness data. Despite these limitations, the clustering outcomes provide meaningful insights into global happiness patterns and offer a solid foundation for subsequent methodological extensions and policy-relevant analysis.

Statistical and Policy Implications

From a statistical perspective, the results demonstrate that K-Means clustering, when combined with appropriate validation and visualization techniques, is effective for exploring complex, multidimensional social indicators. The use of the Silhouette Index provides an objective criterion for selecting the number of clusters, while Biplot analysis enhances interpretability by revealing relationships among variables and clusters.

From a policy perspective, the clustering results indicate that economic indicators alone are insufficient to explain national happiness levels. Countries with similar income levels may fall into different clusters due to differences in social support, governance quality, and public safety. Policymakers may therefore benefit from benchmarking their countries against statistically similar peers rather than relying solely on global rankings. For countries in lower-happiness clusters, improvements in institutional quality and public safety may yield substantial gains in well-being, while higher-happiness countries illustrate the role of strong social and institutional frameworks in sustaining well-being despite higher living costs. The findings of this study have several important policy implications. First, the clustering results demonstrate that economic growth alone is insufficient to guarantee high levels of happiness. Countries in Cluster 2 illustrate that relatively high income and health outcomes can coexist with moderate happiness when crime and corruption remain prevalent. Policymakers should therefore prioritize institutional quality, public safety, and trust-building measures alongside economic development strategies.

Second, the experience of countries in Cluster 3 highlights the importance of social support systems, freedom of choice, and effective governance. Investments in healthcare, education, social protection, and transparent institutions appear to yield substantial returns in terms of national happiness. Although the cost of living is high in these countries, strong institutional frameworks mitigate its negative effects on well-being.

Third, for countries in Cluster 1, policy interventions should focus on improving basic living conditions, public security, and governance while preserving existing strengths in social cohesion and generosity. Targeted international cooperation and development assistance may play a crucial role in addressing structural constraints faced by these countries.

Overall, grouping countries based on happiness indicators allows policymakers to design context-specific strategies and learn from peer countries facing similar challenges, rather than relying solely on global rankings.

5. CONCLUSION

This study applies K-Means clustering and Biplot analysis to classify countries based on global happiness indicators, resulting in three distinct clusters representing low, moderate, and high happiness levels. Countries in the high-happiness cluster are associated with strong social support, higher income, better health outcomes, and lower levels of crime and perceived corruption, while countries in the low-happiness cluster face structural challenges across these dimensions. These results highlight the multidimensional nature of happiness and demonstrate the usefulness of multivariate statistical methods in exploring complex social phenomena.

However, the study is limited by its reliance on cross-sectional and self-reported data, potential measurement inconsistencies across countries, and methodological constraints inherent to K-Means clustering. Future research may benefit from longitudinal data, alternative clustering techniques, and the inclusion of additional indicators to further enhance the robustness and interpretability of global happiness classifications.

6. REFERENCES

[1] M. M. Rozario*, Happiness and Economic Growth: A cross sectional analysis*, 2018.

[2] S. Lyubomirsky, *The how of happiness: A scientific approach to getting the life you want*. Penguin Press, 2008.

[3] E. Diener, *The science of well-being: The collected works of Ed Diener*. Springer., 2009. https://doi.org/10.1007/978-90-481-2350-6

[4] R. A. Easterlin, "Does economic growth improve the human lot? Some empirical evidence", In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramovitz* (pp. 89–125). Academic Press. 1974.

[5] A. Sen, *Development as freedom*, Oxford University Press, 1999.

[6] P. A. Linley, J. Maltby, A. M. Wood, G. Osborne, R. Hurling, "Measuring happiness: The higher order factor structure of subjective and psychological well-being measures", *Personality and Individual Differences*, Volume 47, Issue 8, December 2009, Pages 878-884.

[7] E. Diener, "Subjective well-being. The science of happiness and a proposal for a national index", *Am Psychol*, 2000 Jan; 55(1):34-43. PMID: 11392863.

[8] J. F. Helliwell & R. D. Putnam, "The social context of well-being", *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1449), 1435–1446., 2004. https://doi.org/10.1098/rstb.2004.1522

[9] J. C. Ott, "Government and happiness in 130 nations: Good governance fosters higher level and more equality of happiness", *Social Indicators Research*, 102(1), 3–22, 2011. https://doi.org/10.1007/s11205-010-9719-z

[10] A. Deaton, "Income, health, and well-being around the world: Evidence from the Gallup World Poll", *Journal of Economic Perspectives*, 22(2), 53–72, 2008. https://doi.org/10.1257/jep.22.2.53

[11] R. Moreno-Serra and P. C. Smith, "Does progress towards universal health coverage improve population health?," *Lancet*, vol. 380, no. 9845, pp. 917–923, 2012.

[12] J. F. Helliwell, R. Layard and J. D. Sachs (Eds.), "World happiness report 2019", *Sustainable Development Solutions Network*. https://worldhappiness.report. 2019.

[13] Z. Li and L. An, "Corruption, crime, and subjective well-being: Evidence from international panel data", *Social Indicators Research*, 149(3), 1137–1161, 2020. https://doi.org/10.1007/s11205-019-02227-1

[14] S. Staubli, M. Killias and B. S. Frey, "Happiness and victimization: An empirical study for Switzerland", *European Journal of Criminology*, 11(1), 57–72, 2014. https://doi.org/10.1177/1477370813490758

[15] H. Cantril, *Pattern of human concerns*, 1965.

[16] H. Wickham and G. Grolemund, *R for data science: Import, tidy, transform, visualize, and model data*, O'Reilly Media, 2017.

[17] B. S. Everitt, S. Landau, M. Leese and D. Stahl, *Cluster analysis* (5th ed.). Wiley, 2011. https://doi.org/10.1002/9780470977811

[18] J. F. Hair, W. C. Black, B. J. Babin and R. E. Anderson, *Multivariate Data Analysis* (8th ed.), Cengage Learning, 2019.

[19] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: An introduction to cluster analysis*. Wiley, 2009.

[20] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis", *Journal of Computational and Applied Mathematics*, 20, 53–65, 1987. https://doi.org/10.1016/0377-0427(87)90125-7

[21] A. C. Rencher , W. F. Christensen, *Methods of multivariate analysis* (3rd ed.), Wiley, 2012.

[22] I. T. Jolliffe and J. Cadima, "Principal component analysis: A review and recent developments", *Philosophical Transactions of the Royal Society A*, 374(2065), 20150202, 2016. https://doi.org/10.1098/rsta.2015.0202