https://doi.org/10.7251/EMC2601131A
UDK: 330.342:341.217.02(4-672EU)
Datum prijema rada: 4. septembar 2025.
Submission Date: September 4, 2025
Datum prihvatanja rada: 28. februar 2026.
Acceptance Date: February 28, 2026
Časopis za ekonomiju i tržišne komunikacije
Economy and Market Communication Review
Godina/Vol. **XVI** • Br./No. **I**
str./pp. 131-148

ORIGINALNI NAUČNI RAD / ORIGINAL SCIENTIFIC PAPER

CLUSTER ANALYSIS OF THE MEMBER COUNTRIES OF THE EUROPEAN UNION ACCORDING TO ENERGY, ENVIRONMENTAL AND ECONOMIC INDICATORS

Mihovil Anđelinović | Full Professor, Faculty of Economics and Business, University of Zagreb, Zagreb, Croatia, mandelinovic@efzg.hr; ORCID ID 0000-0003-2061-0693
Davor Labaš | Associate Professor, Faculty of Economics and Business, University of Zagreb, Zagreb, Croatia, dlabas2@net.efzg.hr; ORCID ID 0000-0003-2500-4211
Katarina Sičaja | Assistant, Faculty of Economics and Business, University of Zagreb, Zagreb, Croatia, ksicaja@net.efzg.hr; ORCID ID 0009-0009-1631-4534

Abstract: *In this paper, a cluster analysis was conducted to group the European Union member states according to energy, environmental and economic indicators for 2023. The selected variables are GDP per capita, the share of renewable energy sources in total energy consumption and* CO_2 *emissions per capita. The aim of cluster analysis is to group the European Union member states based on economic, energy and environmental indicators in order to identify their mutual similarities and differences. The analysis begins with a hierarchical clustering method using the Ward method with squared Euclidean distance. The dendrogram indicated several potential solutions, and particular attention was paid to solutions with three and five clusters. The observed clusters reflect differences in the level of development and approach to energy policy. Countries with higher levels of economic development are characterized by a lower degree of sustainability due to a lower share of renewable energy sources and high greenhouse gas emissions. Conversely, less developed countries achieve better results according to environmental sustainability indicators, but at the same time have limited economic growth potential. To confirm the stability of the clusters, the K-means method, i.e. a non-hierarchical method, was also applied, which further refined the grouping. The results of the discriminant analysis showed high classification accuracy, confirming the reliability of the model. In conclusion, the cluster analysis provided a deeper insight into energy and environmental patterns within the EU, highlighting the need for harmonization of sustainable policies among members with different levels of economic development. The results obtained can serve as a basis for designing targeted and differentiated policies within the EU that take into account the specificities of each group of countries, with a particularly important balance between economic growth and environmental sustainability.*

Keywords: *European Union, GDP per capita, renewable energy sources,* CO_2 *emissions, cluster analysis*

JEL classification: *C38, Q56, O52*

INTRODUCTION

In the context of current global trends, assessing the sustainability of countries represents a key challenge in achieving sustainable development goals . Economic and environmental indicators are increasingly taking center stage in analyses of sustainable development within the European Union. In the context of global challenges related to climate change and energy transition, differences between member states in terms of economic development, use of renewable energy sources and greenhouse gas emissions are becoming increasingly pronounced. According to an analysis by , 51% of EU regions show increased economic prosperity thanks to the transition to renewable energy sources, with less developed regions being particularly represented in this group. Although some Central and Eastern European countries had a GDP per capita of only about half the EU average in 2004 ; , data from 2023 show visible progress. For example, Bulgaria, which still has the lowest level, reaches 64% of the EU average . According to greenhouse gas emissions in the EU decreased by 19% in the period from 2013 to 2023, but there are large differences between EU member states. While Estonia, Ireland, and Slovenia have achieved reductions in the intensity of CO_2 emissions by more than 40%, other countries have recorded much smaller improvements. This further confirms the unevenness of environmental and economic effects within the EU. The results of the research by indicate diversity among the new EU member states regarding the structure of renewable energy production, which is related to their specific demographic and economic potential. The authors also emphasize the existence of a significant development gap in relation to the old EU members, which is visible in the significant deviations of GDP per capita. Therefore, this further confirms the need to analyze groups of countries with comparable development characteristics. GDP per capita, the share of renewable sources in total energy consumption, and CO_2 emissions per capita are key variables that provide insight into the interrelationships between economic growth, energy policy, and environmental impact. Given the heterogeneity of the European Union member states in the above aspects, cluster analysis is increasingly used as an effective tool for segmenting countries into groups with similar characteristics in order to enable a better understanding of existing sustainability patterns and the design of targeted policies . Such analysis allows the identification of groups of countries that have comparable energy and environmental characteristics and can serve as a basis for the adoption of differentiated and targeted policies at the European Union level. The aim of the research is to identify groups of EU member states that share similar levels of economic development and sustainability in order to determine patterns in energy efficiency and ecological footprint and to enable a better understanding of differences and similarities between EU countries in the context of common sustainable development goals. The specific objectives of the research are as follows:

- To determine whether there are groups of countries that share similar economic and environmental characteristics.
- To analyze which variable contributes most to the formation of clusters.
- To compare the resulting clusters and interpret them in the context of energy policy and sustainability in the EU.

- To provide suggestions for improving policies and research approaches based on the identified groups of countries.

RESEARCH METHODOLOGY

This paper used a quantitative research approach with the aim of grouping the European Union member states according to economic, energy, and environmental indicators. Official statistical data taken from the Eurostat database were used. These are secondary data sources that are reliable and verified. The sample includes all 27 European Union member states, ensuring complete coverage of the population and representativeness of the results. In order to enable comparability between the variables in the analysis, the data were standardized. The methodological framework of the research includes the application of hierarchical and non-hierarchical cluster analysis. Before conducting the cluster analysis, the correlation between the variables was examined to determine whether there was multiple collinearity that could negatively affect the reliability and interpretation of the analysis results. If the VIF value is in the range between 1 and 5, this indicates a moderate correlation between the variables . Since all VIF values in this study are below the threshold value (VIF < 5), it can be concluded that there is no significant multicollinearity between the variables, thus ensuring the reliability of the cluster analysis results. After carrying out the cluster analysis, discrimination analysis was also used to determine the reliability and accuracy of the classification of countries into clusters. Data processing was carried out with the help of *Statistica* statistical software.

HIERARHICAL CLUSTER ANALYSIS

Cluster analysis encompasses a range of multivariate methods aimed at grouping objects into groups based on their shared characteristics . The two basic methods of cluster analysis are hierarchical and non-hierarchical analysis . Hierarchical clustering, also known as hierarchical cluster analysis, is a statistical clustering method that aims to create a hierarchy of clusters . According to , various linkage methods are applied in hierarchical cluster analysis, including the single linkage method, complete linkage method, average linkage method, Ward's method and centroid method. The result of hierarchical cluster analysis is usually presented as a dendrogram . A dendrogram is a graphic representation of the gradual joining of objects into clusters, where the distances between them are read on the horizontal axis, while individual observations are shown on the vertical axis . Furthermore, different methods of hierarchical cluster analysis were performed with the application of different distance measures. The clearest cluster structure is given by the application of Ward's method with the squared Euclidean distance. The dendrogram shown in Figure 1 shows the separation of five clusters, which confirms the choice of that solution as the most suitable for further interpretation.

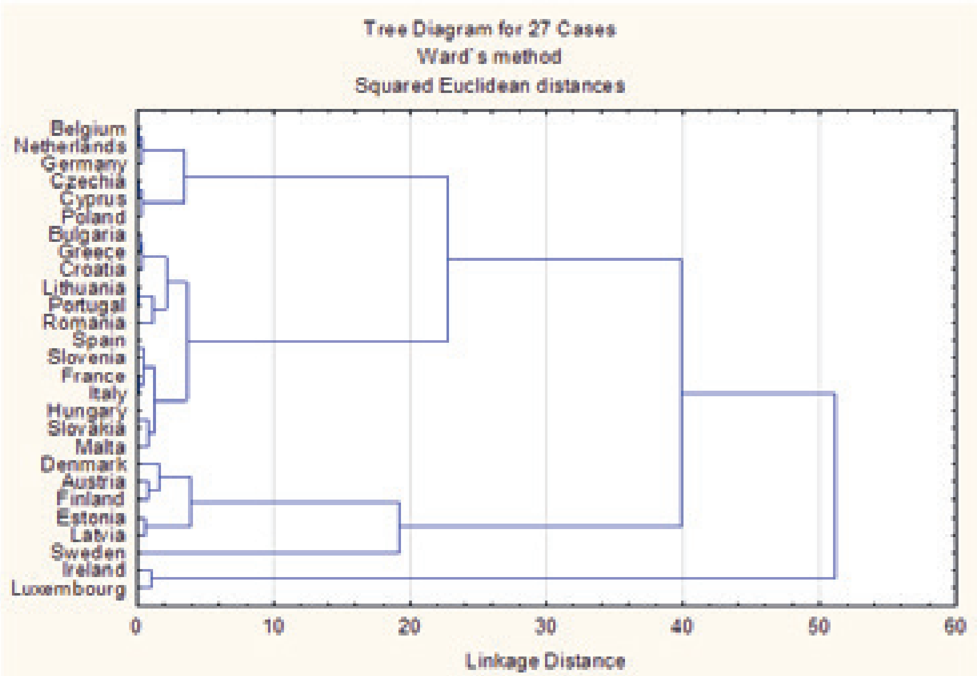


Figure 1. Dendrogram cluster analysis for 27 EU member states

Source: authors' work, result of processing in *Statistica*

At the initial levels of hierarchical clustering, i.e. below level 3, the expected similarities between EU member states are observed, as shown in Figure 2 below. For example, Hungary and Slovakia merge in one of the first iterations (distance 0.0088559), indicating a high degree of similarity in development and environmental factors. This early consolidation is likely a result of their similar economic structure, energy policies and level of adaptation to sustainable standards. Malta joins this cluster later (distance 0.8076450), which may indicate specific common patterns in individual indicators, despite the geographical distance. A similar pattern is visible in Spain and Slovenia, which form an early cluster at a low distance (0.0429791). Later, France and Italy gradually join them at a distance of 0.5145351 and Hungary at a distance of 1.250602, thus forming a group of countries with similar development indicators and energy priorities. Belgium and the Netherlands cluster early (distance 0.1828188), indicating a high degree of concordance in terms of economic development and environmental policies. Later, Germany joins them (distance 0.4013445), thus forming a cluster of highly developed economies with an active approach to reducing CO_2 emissions and increasing the share of renewable energy sources. At intermediate levels of distance (approximately between 3 and 10), larger and more complex groups of countries begin to form, and clustering patterns become increasingly diverse. The Denmark-Austria-Finland cluster is joined by Estonia and Latvia at a distance of 3.937898. Furthermore, Lithuania and Portugal join the Bulgaria–Greece–Croatia cluster at intermediate levels of convergence at a distance of 3.652641, which indicates their high similarity in economic and environmental indicators. The cluster consisting of Belgium-Netherlands-Germany extends to the Czech Republic and Cyprus at a distance of 3.438881. This group of countries shares comparable levels of CO_2 emissions, indicating a similar degree of environmental burden. At the highest levels of distance (above 10), the already formed clusters finally merge into a single set of all countries. At these levels of distance, previously defined clusters such as Belgium-Netherlands-Germany-Czech Republic-Cyprus and Denmark-Austria-Finland-Estonia-Latvia merge.

linkage distance	Amalgamation Schedule Ward`s method Squared Euclidean distances Obj. No. 1	Obj. No. 2	Obj. No. 3	Obj. No. 4	Obj. No. 5
,0088559	Hungary	Slovakia			
,0429791	Spain	Slovenia			
,0781959	Czechia	Cyprus			
,0887239	Lithuania	Portugal			
,1478279	France	Italy			
,1828188	Belgium	Netherlands			
,2243782	Bulgaria	Greece			
,3218806	Bulgaria	Greece	Croatia		
,3454772	Czechia	Cyprus	Poland		
,4013445	Belgium	Netherlands	Germany		
,5145351	Spain	Slovenia	France	Italy	
,5941753	Estonia	Latvia			
,8076450	Hungary	Slovakia	Malta		
,8332978	Austria	Finland			
1,091508	Ireland	Luxembourg			
1,206819	Lithuania	Portugal	Romania		
1,250602	Spain	Slovenia	France	Italy	Hungary
1,665941	Denmark	Austria	Finland		
2,150418	Bulgaria	Greece	Croatia	Lithuania	Portugal
3,438881	Belgium	Netherlands	Germany	Czechia	Cyprus
3,652641	Bulgaria	Greece	Croatia	Lithuania	Portugal
3,937898	Denmark	Austria	Finland	Estonia	Latvia
19,18329	Denmark	Austria	Finland	Estonia	Latvia
22,76650	Belgium	Netherlands	Germany	Czechia	Cyprus
39,92009	Belgium	Netherlands	Germany	Czechia	Cyprus
51,14328	Belgium	Netherlands	Germany	Czechia	Cyprus

Figure 2. Initial cluster merging iterations

Source: authors' work, result of processing in *Statistica*

NON-HIERARCHICAL CLUSTER ANALYSIS

Unlike hierarchical methods, non-hierarchical cluster analysis methods require that the number of clusters be defined in advance . Among the most commonly used non-hierarchical methods, the K-means method stands out, which is often applied after hierarchical analysis to confirm the cluster structure and reduce the impact of possible errors in the initial classification

SOLUTION ANALYSIS WITH THREE CLUSTERS

Using non-hierarchical cluster analysis with three clusters, an analysis of variance was conducted, the results of which are shown in Figure 3. The obtained F-values show to what extent each of the variables contributes to the formation of differences between the formed clusters. The results of the analysis of variance show that the share of renewable energy sources is the variable that contributes the most to the differentiation between clusters, with the highest F-value (F=41.4), which makes this variable a key factor in the clustering process. This is followed by CO_2 emissions with an F-value of 12.2, which also indicates a significant role in the differentiation between clusters, although it is less pronounced than for renewable energy sources. In contrast, GDP per capita shows the lowest F-value (F=8.6), which suggests that economic differences between clusters are not as pronounced as environmental ones. In other words, the share of renewable energy sources and CO_2 emissions played a crucial role in the formation of clusters, while the role of GDP was secondary. The results obtained from the non-hierarchical analysis confirm the patterns observed with the hierarchical method but also more clearly delimit them into three groups. The first group consists of countries with the highest GDP per capita, the lowest shares in renewable energy sources, and the highest level of CO_2 emissions per capita. The second group consists of countries with moderately high GDP per capita, the highest shares in renewable energy sources, and the highest level of CO_2 emissions per capita. The third group consists of countries with the lowest values of GDP per capita and CO_2 emissions per capita and low shares in renewable energy sources.

	Analysis of Variance					
Variable	Between SS	df	Within SS	df	F	signif. p
GDP_per_capita	10,85403	2	15,14597	24	8,59954	0,001527
Renewable_energy_share_pct	20,15891	2	5,84109	24	41,41466	0,000000
CO2_emissions_per_capita	13,09713	2	12,90287	24	12,18067	0,000223

Figure 3. Analysis of variance in non-hierarchical analysis of three selected clusters

Source: authors' work, result of processing in *Statistica*

According to Figure 4, the first group of countries consists of Belgium, the Czech Republic, Germany, Ireland, Cyprus, Luxembourg, and the Netherlands. These countries are characterized by a relatively high level of economic development, sharing medium to high values of GDP per capita and elevated levels of CO_2 emissions per capita. The cluster reflects a group of countries with developed industrial infrastructure and strong economic activity, which is also reflected in a higher emission burden. Although all members of the group are in line with expectations, Luxembourg stands out as the country with the greatest deviation from the center of the cluster, which indicates its specific and extremely high economic indicators that surpass other developed members. The Netherlands, on the other hand, shows the least deviation within the group, making it a typical representative of this cluster.

	Members of Cluster Number 1 and Distances from Respective Cluster Center Cluster contains 7 cases
	Distance
Belgium	0,384117
Czechia	0,795309
Germany	0,371662
Ireland	0,908028
Cyprus	0,736251
Luxembourg	1,424188
Netherlands	0,259156

Figure 4. Members of the first cluster and distance from the center of the cluster for each EU member state

Source: authors' work, result of processing in *Statistica*

It can be seen from Figure 5 that the second group of countries consists of Estonia, Denmark, Latvia, Austria, Finland, and Sweden. These countries are characterized by a strong orientation towards sustainable energy solutions, which is reflected in the highest share of renewable energy sources in total consumption among all clusters. At the same time, they record similar, relatively moderate levels of CO_2 emissions, which indicates successful policies of energy transition and ecological efficiency. These are developed countries with high GDP per capita. Within the cluster, Sweden stands out as the furthest from the center of the group, which results from the highest value of the share of renewable energy sources and the lowest CO_2 emissions in the entire EU. Austria, on the other hand, shows the smallest deviation, which positions it as a balanced representative of the group.

	Members of Cluster Number 2 and Distances from Respective Cluster Center Cluster contains 6 cases
	Distance
Denmark	0,501997
Estonia	0,738487
Latvia	0,601445
Austria	0,451480
Finland	0,584653
Sweden	1,632282

Figure 5. Members of the second cluster and distance from the center of the cluster for each EU member state

Source: authors' work, result of processing in *Statistica*

Figure 6 below shows the third group, which consists of 14 countries, mainly from Central, Eastern, and Southern Europe: Bulgaria, Greece, Spain, France, Croatia, Italy, Lithuania, Hungary, Malta, Poland, Portugal, Romania, Slovenia, and Slovakia. This group shows moderate values of the share of renewable energy sources and lower levels of CO_2 emissions compared to the other clusters. The ecological profile of these countries indicates a balanced, but still transitional, energy structure. This group is characterized by a lower to medium-high GDP per capita, which places them in the less developed group in the context of the European Union. The economic level, although diverse, is on average below the level of the first and second clusters. Within the group, Slovenia is very close to the center of the cluster, which suggests that its profile most accurately reflects the average characteristics of this group. In contrast, Poland shows the largest deviation, which may indicate specific energy or economic characteristics that make it a less typical representative of the group.

	Members of Cluster Number 3 and Distances from Respective Cluster Center Cluster contains 14 cases
	Distance
Bulgaria	0,292725
Greece	0,337694
Spain	0,213459
France	0,369434
Croatia	0,255391
Italy	0,307235
Lithuania	0,425675
Hungary	0,302348
Malta	0,518910
Poland	0,867838
Portugal	0,534344
Romania	0,568229
Slovenia	0,134883
Slovakia	0,290310

Figure 6. Members of the third cluster and distance from the center of the cluster for each EU member state

Source: authors' work, result of processing in *Statistica*

Furthermore, Figure 7 shows the mean values of the three clusters according to three variables: GDP per capita, share of renewable energy sources in total consumption, and CO_2 emissions per capita. Each cluster shows a recognizable profile, which is clearly reflected in the differences shown by the lines in the graph.

- Cluster 1 (blue line)

This cluster is characterized by high GDP per capita, the lowest share of renewable energy sources, and high levels of CO_2 emissions. This cluster suggests a group of economically developed countries with a pronounced industrial infrastructure that still rely on conventional energy sources, resulting in higher greenhouse gas emissions despite high GDP.

- Cluster 2 (red line)

This cluster shows the highest share of renewable energy sources, with GDP per capita and CO_2 emissions around zero (average). This cluster indicates countries that have successfully integrated renewable energy sources into their energy system while maintaining a moderately high level of economic development and a balanced ecological-economic approach.

- Cluster 3 (green line)

Cluster 3 is characterized by low GDP per capita, low share of renewable energy sources, and low CO_2 emissions. This cluster reflects less developed countries in the energy transition phase, where lower economic activity and less developed industry result in lower total CO_2 emissions but also limited capacity to invest in renewable energy sources.

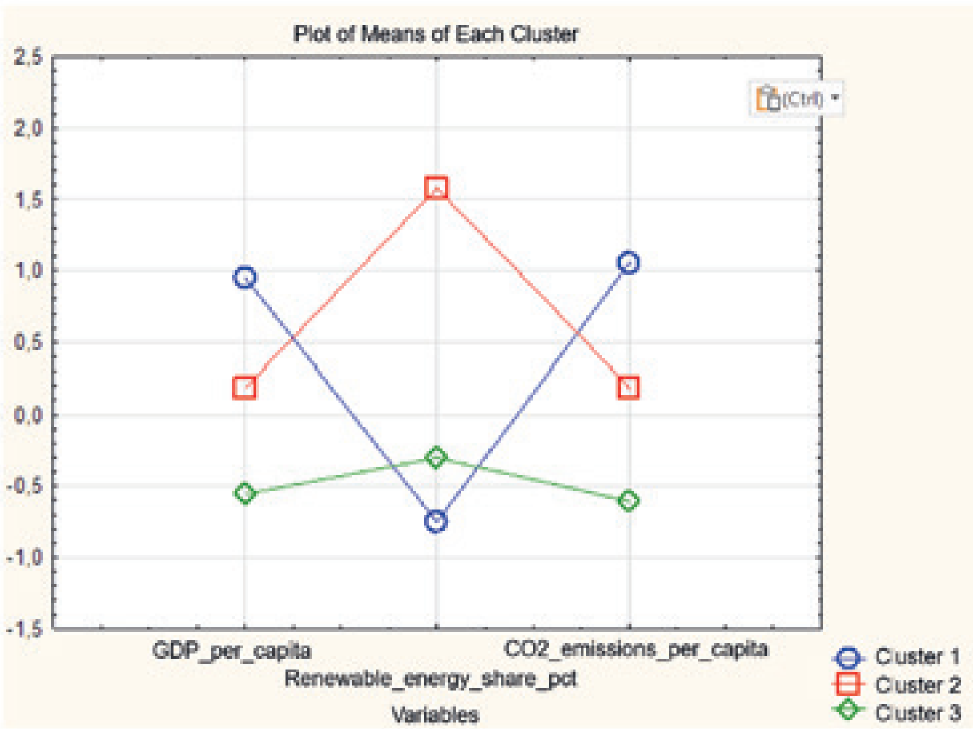


Figure 7. Graphical representation of the means of each cluster for the three selected clusters

Source: author's work, result of processing in *Statistica*

SOLUTION ANALYSIS WITH FIVE CLUSTERS

The results of the analysis of variance obtained by applying non-hierarchical cluster analysis with 5 clusters are shown in Figure 8 below. The obtained F-values show to what extent each of the variables contributes to the formation of differences between the formed clusters. The results of the analysis of variance show that the share of renewable energy sources is the variable that contributes the most to the differentiation between clusters, with the highest F-value (F=30.7), which makes this variable a key factor in the clustering process. It is followed by GDP per capita with an F-value of 19.9, which also indicates a significant role in the differentiation between clusters, although it is less pronounced than for renewable energy sources. In contrast, CO_2 emissions per capita show the lowest F-value (F=9.3). In other words, the share of renewable energy sources and GDP per capita played a crucial role in the formation of clusters, while the role of CO_2 emissions per capita was secondary. The results obtained by the non-hierarchical analysis confirm the patterns observed by the hierarchical method but also more clearly delimit them into three groups. The first group consists of countries with lower GDP per capita and low levels of CO_2 emissions per capita and renewable energy sources. The second group consists of countries with low shares in renewable energy sources and high levels of CO_2 emissions with average GDP. The third group consists of countries with the lowest values of GDP per capita and low levels of CO_2 emissions per capita, with an average share in renewable energy sources. The fourth group consists of countries with the highest values of GDP per capita and CO_2 emissions per capita and the lowest shares in renewable energy sources, while the fifth group consists of countries with extremely high shares in renewable energy sources and average values of GDP per capita and CO_2 emissions per capita.

	Analysis of Variance					
Variable	Between SS	df	Within SS	df	F	signif. p
GDP_per_capita	20,37737	4	5,622631	22	19,93293	0,000000
Renewable_energy_share_pct	22,05468	4	3,945322	22	30,74546	0,000000
CO2_emissions_per_capita	16,36686	4	9,633141	22	9,34459	0,000143

Figure 8. Analysis of variance in non-hierarchical analysis of five selected clusters

Source: authors' work, result of processing in *Statistica*

Compared to the previous solutions with 4 and 5 clusters, France, Italy, Hungary, Malta and Slovakia were separated into a separate cluster, which can be seen in Figure 9.

	Members of Cluster Number 1 and Distances from Respective Cluster Center Cluster contains 5 cases
	Distance
France	0,292614
Italy	0,187833
Hungary	0,264056
Malta	0,285208
Slovakia	0,221793

Figure 9. Members of the first cluster and distance from the center of the cluster for each EU member state

Source: authors' work, result of processing in *Statistica*

Furthermore, Figure 10 shows the same distribution as in the fourth cluster from the previous case.

	Members of Cluster Number 2 and Distances from Respective Cluster Center Cluster contains 6 cases
	Distance
Belgium	0,297046
Czechia	0,286470
Germany	0,272864
Cyprus	0,200173
Netherlands	0,470364
Poland	0,484821

Figure 10. Members of the second cluster and distance from the center of the cluster for each EU member state

Source: authors' work, result of processing in *Statistica*

Since France, Italy, Hungary, Malta and Slovakia separated into a separate cluster, the countries shown in Figure 11 form a new cluster.

	Members of Cluster Number 3 and Distances from Respective Cluster Center Cluster contains 8 cases
	Distance
Bulgaria	0,348233
Greece	0,392904
Spain	0,233883
Croatia	0,112298
Lithuania	0,241795
Portugal	0,352006
Romania	0,472630
Slovenia	0,195854

Figure 11. Members of the third cluster and distance from the center of the cluster for each EU member state

Source: authors' work, result of processing in *Statistica*

Figure 12 shows that the distribution is the same as in the third cluster from the previous case.

	Members of Cluster Number 4 and Distances from Respective Cluster Center Cluster contains 2 cases
	Distance
Ireland	0,301594
Luxembourg	0,301594

Figure 12. Members of the fourth cluster and distance from the center of the cluster for each EU member state

Source: authors' work, result of processing in *Statistica*

Figure 13 also shows the same distribution as in the second cluster from the previous case.

	Members of Cluster Number 5 and Distances from Respective Cluster Center Cluster contains 6 cases
	Distance
Denmark	0,501997
Estonia	0,738487
Latvia	0,601445
Austria	0,451480
Finland	0,584653
Sweden	1,632282

Figure 13. Members of the fifth cluster and distance from the center of the cluster for each EU member state

Source: authors' work, result of processing in *Statistica*

Figure 14 shows the mean values of the three clusters according to three variables: GDP per capita, share of renewable energy sources in total consumption, and CO_2 emissions per capita. Each cluster shows a recognizable profile, which is clearly reflected in the differences shown by the lines in the graph.

- Cluster 1 (blue line)

This cluster is characterized by low GDP per capita, a low share of renewable energy sources, and a low level of CO_2 emissions, i.e. all values are below average. This cluster reflects less developed countries in the energy transition phase with lower total CO_2 emissions due to lower levels of industrialization.

- Cluster 2 (red line)

The cluster is characterized by an average GDP per capita (around zero), a below-average share of renewable energy sources, and a moderately high level of CO_2 emissions (close to 1). Elevated CO_2 emissions indicate energy-inefficient systems and slower introduction of sustainable technologies.

- Cluster 3 (green line)

This cluster is characterized by the lowest GDP values, an average share of renewable energy sources, and low CO_2 emissions. This cluster refers to less economically developed countries, whose industrial capacity is limited, resulting in lower CO_2 emissions. At the same time, the average share of renewable energy sources reflects the initial steps towards a more sustainable energy model.

- Cluster 4 (pink line)

Cluster 4 is characterized by the highest GDP per capita, the lowest share of renewable energy sources, and above-average levels of CO_2 emissions. This cluster indicates highly developed economic systems with intensive industrial production, which results in significant greenhouse gas emissions.

- Cluster 5 (black line)

This cluster shows an above-average share of renewable energy sources, with average GDP per capita and CO_2 emissions per capita hovering around zero. This cluster points to countries that have successfully integrated renewable energy sources while maintaining a balance between economic development and environmental responsibility. These countries exemplify a more sustainable approach, where progress in green energy does not come at the expense of economic stability.

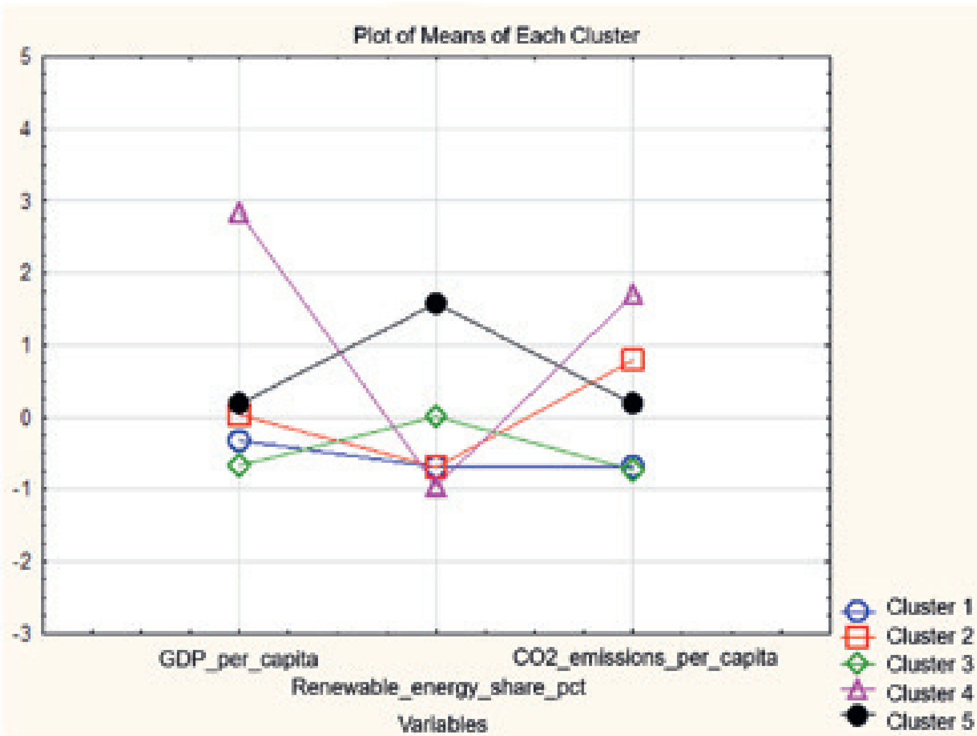


Figure 14. Graphical representation of the means of each cluster for the five selected clusters

Source: authors' work, result of processing in *Statistica*

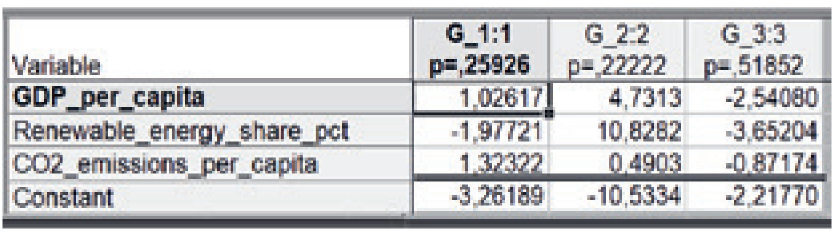
DISCRIMINANT ANALYSIS

Discriminant analysis is a multivariate statistical technique used to determine the relationship between one categorical dependent variable and several independent variables, which are usually continuous . The goal of the analysis is to classify the observations into predefined groups and to identify those variables that contribute the most to the differentiation between these groups. Discriminant analysis is an important statistical tool used to understand differences between groups and to predict, or classify new samples into appropriate groups based on these differences.

SOLUTION ANALYSIS WITH THREE CLUSTERS

The coefficients of the classification function show how much a particular variable contributes to the distinction between clusters. Namely, a higher coefficient means a greater influence of that variable on the separation of groups. From Figure 15, the following can be concluded:

- GDP per capita has the highest positive coefficient in the first cluster (G_1:1 = 1.32322), which indicates that this variable contributes the most to distinguishing this cluster, i.e. that countries in this cluster have extremely high GDP per capita values that clearly distinguish them from other groups.
- The share of renewable energy sources has a very high positive coefficient in the second cluster (G_2:2 = 10.8282), which means that this variable defines this cluster the most. In other words, countries from the second cluster are characterized by a very high share of renewable energy sources in total energy consumption.
- In the third cluster, the share of renewable energy sources also has the largest contribution, but in a negative direction (G_3:3 = -3.65204). This suggests that countries in this cluster have significantly lower values of this variable compared to the other groups, making them easily recognizable through a lower degree of use of renewable energy sources.



Variable	G_1:1 p=,25926	G_2:2 p=,22222	G_3:3 p=,51852
GDP_per_capita	1,02617	4,7313	-2,54080
Renewable_energy_share_pct	-1,97721	10,8282	-3,65204
CO2_emissions_per_capita	1,32322	0,4903	-0,87174
Constant	-3,26189	-10,5334	-2,21770

Figure 15. Classification function coefficients for three clusters

Source: authors' work, result of processing in *Statistica*

Furthermore, the classification matrix serves to assess the success of classifying units into clusters in relation to the previously obtained division from the non-hierarchical analysis. In this case, Figure 16 shows that all observations are located on the main diagonal of the matrix, which means that the model achieved complete classification accuracy. This result indicates a clear demarcation between clusters and confirms the high reliability and consistency of the model.

Group	Classification Matrix (Spreadsheet1) Rows: Observed classifications Columns: Predicted classifications			
	Percent Correct	G_1:1 p=,25926	G_2:2 p=,22222	G_3:3 p=,51852
G_1:1	100,0000	7	0	0
G_2:2	100,0000	0	6	0
G_3:3	100,0000	0	0	14
Total	100,0000	7	6	14

Figure 16. Classification matrix for three clusters
Source: authors' work, result of processing in *Statistica*

SOLUTION ANALYSIS WITH FIVE CLUSTERS

Based on Figure 17, several conclusions can be drawn about the characteristics of individual clusters:

- The first cluster is characterized by a significantly lower share of renewable energy sources compared to the other clusters, given that this variable has the largest contribution in the negative direction (G_1:1 = –9.40766). This indicates that the countries in this cluster are less oriented towards renewable energy sources.
- The second cluster is clearly distinguished by its high GDP per capita, since this variable has the highest positive coefficient (G_2:2 = 0.93805). This suggests that the countries in this cluster are economically more developed compared to the others.
- The third cluster is characterized by lower levels of CO_2 emissions, as this variable has the largest negative contribution (G_3:3 = –3.41577). Countries in this cluster therefore have a lower carbon footprint.
- The fourth cluster also stands out for its high GDP per capita (G_4:4 = 13.4490), which further confirms the economic status of the countries within this group.
- The fifth cluster has an extremely high share of renewable energy sources (G_5:5 = 16.1286), which means that the countries within this group are clearly differentiated from the others by their orientation towards sustainable energy sources.

Variable	Classification Functions; grouping: CLUSTER (Spreadsheet12)				
	G_1:1 p=,18519	G_2:2 p=,22222	G_3:3 p=,29630	G_4:4 p=,07407	G_5:5 p=,22222
GDP_per_capita	-1,55965	0,93805	-3,31426	13,4490	0,2977
Renewable_energy_share_pct	-9,40766	-3,62858	-2,60167	-3,5741	16,1286
CO2_emissions_per_capita	-5,81350	0,47282	-3,41577	5,0601	7,2394
Constant	-7,18408	-2,96654	-3,56171	-27,6710	-14,9312

Figure 17. Classification function coefficients for five clusters
Source: authors' work, result of processing in *Statistica*

According to Figure 18, it can be concluded that in this case, all observations are located on the main diagonal of the matrix, which means that the model achieved complete classification accuracy. The high accuracy of the classification matrix in all three- and five-cluster scenarios examined indicates clearly expressed differences between clusters and consistent patterns in the data. This suggests that the selected set of variables is appropriate for distinguishing EU member states and that the model reliably recognizes mutual similarities and differences between them.

Classification Matrix (Spreadsheet12)
Rows: Observed classifications
Columns: Predicted classifications

Group	Percent Correct	G_1:1 p=,18519	G_2:2 p=,22222	G_3:3 p=,29630	G_4:4 p=,07407	G_5:5 p=,22222
G_1:1	100,0000	5	0	0	0	0
G_2:2	100,0000	0	6	0	0	0
G_3:3	100,0000	0	0	8	0	0
G_4:4	100,0000	0	0	0	2	0
G_5:5	100,0000	0	0	0	0	6
Total	100,0000	5	6	8	2	6

Figure 18. Classification matrix for five clusters
Source: authors' work, result of processing in *Statistica*

CONCLUSION

The cluster analysis was based on a combination of hierarchical and non-hierarchical approaches. First, different methods of hierarchical cluster analysis were conducted with different distance measures. The results obtained by the Ward method with squared Euclidean distances were selected for further interpretation. This method enabled the grouping of European Union member states into distinctive clusters based on three variables: GDP per capita, share of renewable energy sources in total consumption, and CO_2 emissions per capita. Based on the pronounced differences in the coupling distances, models with three and five clusters were identified as potentially stable. Then, a non-hierarchical cluster analysis was conducted to confirm the stability and consistency of the previously identified structures. Although the three-cluster solution offers a simpler structure, it fails to adequately capture the complexity and diversity among the observed countries and did not bring additional interpretative value compared to the five-cluster solution. The five-cluster solution allows for a more precise demarcation between countries and a clearer emphasis on the specificities of countries such as France, Italy, Hungary, Malta, and Slovakia, which would remain unrecognized as a separate group in a smaller number of clusters. Also, each of the five clusters has a clearly recognizable profile, which contributes to the analytical value of the model and allows for more relevant conclusions for the needs of shaping energy and environmental policies. Therefore, the five-cluster solution proves to be more appropriate and informative, since it allows for a deeper and more precise analysis of the differences between the observed European countries. The resulting cluster structure reflects recognizable regional and economic patterns within the European Union, which further confirms the relevance of the variables used and the applied methodology. In conclusion, such a grouping of countries allows for a more precise view of common characteristics and trends within individual groups and can serve as a basis for designing energy and climate strategies that take into account specific patterns of renewable energy consumption and CO_2 emission levels. Based on the identified groups of countries, a differentiated approach to energy and climate policy is recommended, whereby priorities and measures should be adjusted to the level of economic development and the structure of the energy system of each cluster. In order to improve research approaches, it would be useful to conduct longitudinal analyses to monitor the dynamics of the transition and evaluate the effectiveness of existing policies within each cluster. Also, the limitations of this research relate primarily to the choice of variables, the time frame, and the methodological approach. Although key economic, energy, and environmental indicators were analyzed, they do not cover all dimensions of sustainable development, such as social factors. In addition, the research was conducted for one year (2023), which does not allow monitoring changes over time. The results may also be influenced by the choice of clustering method and the number of clusters, while the application of discriminant analysis assumes certain statistical conditions. Finally, secondary data from the Eurostat database were used, which enatils methodological limitations related to the source and structure of the data.

LITERATURE

Babin, B. J., Krey, N., & Cousley, S. (2025.). Overview of Discriminant Analysis, An. *International Encyclopedia of Statistical Science*, pp. 1849.-1853.

Bleischwitz, R., & Bader, N. (2010.). Policies for the transition towards a hydrogen economy: the EU case. *Energy Policy, 38(10)*, pp. 5388-5398.

Eurostat. (2005.). *GDP per capita varied by one to five across the EU25 Member States.*

Eurostat. (2024.). *Greenhouse gas emission accounts.*

Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010.). *Multivariate data analysis (7th ed).* New York: Pearson.

Korotaj, T. (2022.). Cluster analysis of selected european countries in the context of the position of youth in the labour market. *Proceedings of the Faculty of Economics and Business in Zagreb, 20(1)*, pp. 73-90.

Kovačič, A. (2017.). European Union and Sustainable Development Indicators. *Management of Sustainable Development Sibiu, 9(2)*, pp. 9-29.

Krishna, T. V., Babu, Y. A., & Kumar, K. R. (2017.). Determination of Optimal Clusters for a Non-hierarchical Clustering Paradigm K-Means Algorithm. *Proceedings of International Conference on Computational Intelligence and Data Engineering, 9*, pp. 301-316.

Kurnoga Živadinović, N. (2007.). Multivariate classification of Croatian counties. *Proceedings of the Faculty of Economics and Business in Zagreb, 5(1)*, pp. 1-15.

Noble, J. (2024.). What is hierarchical clustering? *IBM.*

Roman, M., Roman, M., & Niedziółka, A. (2020.). Spatial Diversity of Tourism in the Countries of the European Union. *Sustainability, 12(7)*, pp. 2713.

Schwab, T. (2023.). Energising EU Cohesion: The Renewable Energy Transition Powers Up Lagging Regions.

Shetty, P., & Singh, S. (2021.). Hierarchical Clustering: A Survey. *International Journal of Applied Research, 7(4)*.

Shrestha, & Noora. (2020). Detecting Multicollinearity in Regression Analysis. *American Jour-*

nal of Applied Mathematics and Statistics, 8(2), pp. 39-42.

Uchenna, E. O., Olusola, & O, M. (2024.). Overview of Agglomerative Hierarchical Clustering Methods. *British Journal of Computer, Networking and Information Technology, 7(2)*, pp. 14-23.

University of Washington, L. (2024.). Discriminant Analysis.

Wałachowska, A., & Ignasiak-Szulc, A. (2021.). Comparison of Renewable Energy Sources in 'New' EU Member States in the Context of National Energy Transformations. *Energies, 14(23)*, pp. 7963.

Wiper, O., Vuik, S., Cheatley, J., & Cecchini, M. (2022.). *Cluster analysis to assess the transferability of public health interventions.* OECD Health Working Papers No. 133.