Ivana Unukić
Josip Juraj Strossmayer University of Osijek
Faculty of Economics and Business in Osijek
31000 Osijek, Croatia
iunukic@efos.hr

Nataša Nater Drvenkar
Josip Juraj Strossmayer University of Osijek
Faculty of Economics and Business in Osijek
31000 Osijek, Croatia
natasa.drvenkar@efos.hr

JEL: E00, R11
Preliminary communication
https://doi.org/10.51680/ev.38.2.12

Received: June 16, 2025
Revision received: October 7, 2025
Accepted for publishing: October 11, 2025

CLUSTERING REGIONAL COMPETITIVENESS IN CENTRAL AND EASTERN EUROPE: INSIGHTS FROM THE K-MEANS METHOD

Abstract:

Purpose: This paper investigates regional competitiveness of NUTS 2 regions in eleven post-transition EU Member States in Central and Eastern Europe (CEE) from 2011 to 2021. It applies Martin's (2004) "Regional Competitiveness Hat" model to identify whether distinct regional profiles, knowledge hubs, production locations, and regions with growing yields, can be empirically validated using clustering techniques.

Methodology: The study utilises the k-means clustering method to classify 61 NUTS 2 regions based on three key indicators: GDP per capita, population density, and gross domestic expenditure on R&D. Data were standardised and tested for outliers using Mahalanobis distance. ANOVA and post-hoc Games-Howell tests were conducted to verify statistical significance and interpret the stability and movement of regions between clusters over time.

Results: The analysis produced three statistically robust and theoretically consistent clusters: knowledge hubs (e.g., Zagreb and Bucharest), production locations (regions with low GDP per capita and population density), and regions with growing yields (moderate GDP per capita and lower density). The results affirm the utility of Martin's model in the CEE context and reveal stability among clusters, with notable mobility only between production locations and growing yield regions.

Conclusion: This study confirms the applicability of Martin's framework to post-transition CEE regions and offers a dynamic, data-driven classification tool for regional development policy. It highlights GDP per capita, population density, and R&D investment as critical competitiveness indicators. The findings support targeted EU policy-making and suggest future inclusion of digitalisation and sustainability metrics.

Keywords: Regional competitiveness, CEE countries, k-means clustering, GDP per capita, NUTS 2 regions

1. Introduction

Although the European Union (EU) has one of the highest standards of living in the world, there are significant economic, social and territorial disparities between its Member States and, in particular, between its different regions. These disparities can hinder the EU's global competitiveness and pose challenges for balanced development, as pointed out by Melecký and Staníčková (2014). In this context, understanding the factors behind regional competitiveness is vital for designing policies that promote cohesion and growth. This paper focuses on NUTS 2 regions in Central and Eastern European (CEE) countries, nations that have undergone significant transformation since joining the EU. Using Martin's (2004) regional competitiveness hat model as a framework, this paper examines whether these regions can be grouped into three categories: knowledge hubs, production locations and regions with increasing yields. By applying k-means clustering, this paper aims to detect meaningful patterns and provide insights that could inform both academic understanding and practical policymaking.

The authors of this paper examine regional competitiveness clusters of the NUTS 2 regions of the eleven post-transition EU Member States, commonly referred to as CEE countries (Bulgaria, the Czech Republic, Estonia, Hungary, Lithuania, Latvia, Poland, Romania, Slovenia, Slovakia and Croatia), for the period from 2011 to 2021. Based on the literature review and before conducting the empirical part of the paper, the authors put forward a research hypothesis:

There is no difference in the number of clusters resulting from k-means clustering of the regions with respect to the Regional Competitiveness Hat model.

Accordingly, this study aims to answer the following research question: To what extent can Martin's (2004) Regional Competitiveness Hat model be empirically validated within the post-transition context of Central and Eastern Europe (CEE)? The central hypothesis is that NUTS 2 regions across CEE countries can be classified into three internally consistent clusters corresponding to the model's theoretical dimensions: knowledge hubs, production-oriented regions, and areas characterised by increasing returns.

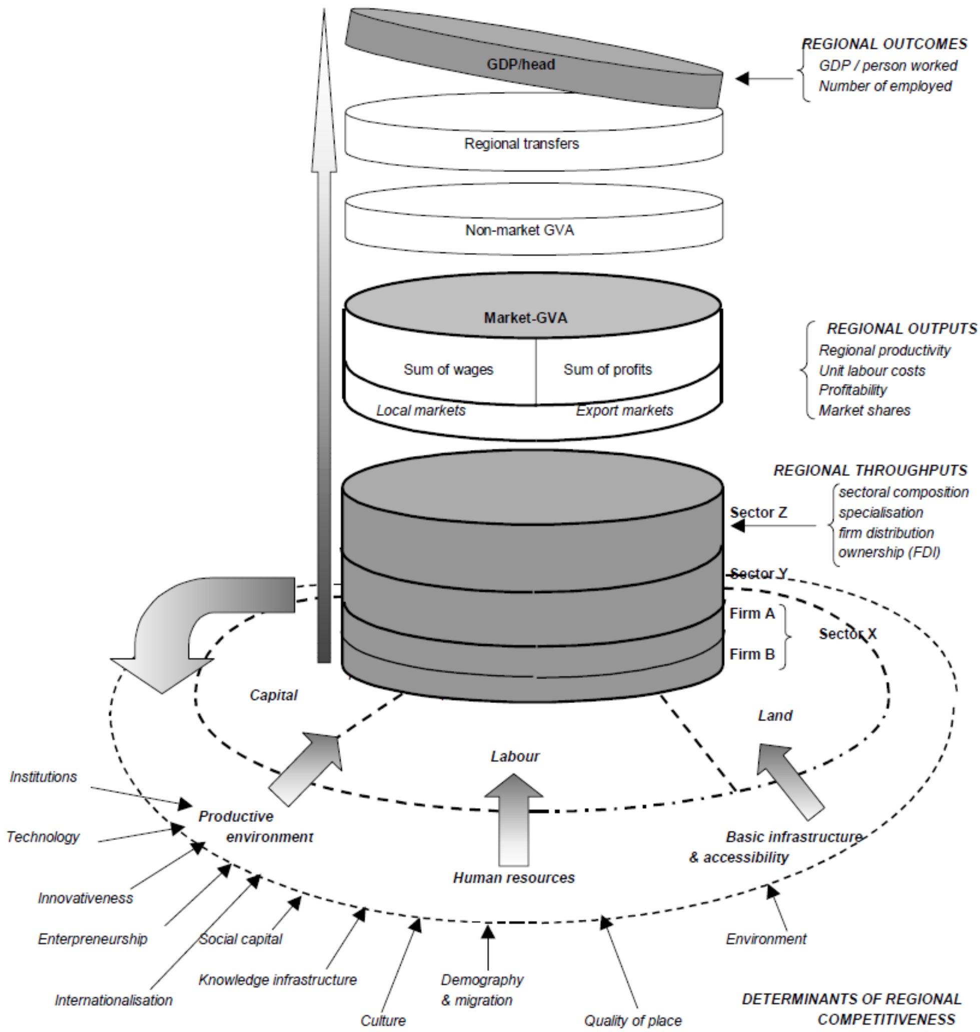
The article is structured as follows. Section 1 introduces the study's background and objectives. Section 2 reviews the evolution of the regional competitiveness concept and surveys empirical applications of clustering approaches. Section 3 describes the methodological framework and presents the empirical results. Section 4 discusses the main challenges and perspectives arising from the analysis, while Section 5 concludes with a synthesis of findings, policy implications, and suggestions for further research.

2. Literature review

The competitiveness criteria serve as a basis for measuring regional competitiveness. The "national diamond" model, the "double diamond" model, the "nine factors" model, the "regional competitiveness hat" model, the "pyramid of regional competitiveness" model and the "tree of regional competitiveness" model are some of the traditional competitiveness models that can serve as a methodological basis for determining competitiveness factors (Snieška and Bruneckienė, 2009). Each model distinguishes between different methods of selecting and combining the elements of competitiveness in a general system. Over the years, policy makers and academics have paid great attention to competitiveness, especially its economic aspect. Focusing on increasingly broad strata based on companies, industries or entire countries, regional competitiveness hat has been studied from both micro and macro perspectives (Shivindina, 2020).

Martin's (2004) Regional Competitiveness Hat model identifies the key determinants of regional competitiveness, placing infrastructure within its second analytical layer (Martin, 2004; Rahmat and Sen, 2021; Ferrarini et al., 2024). As a conceptual framework, it offers a systematic approach to identifying and linking factors that enhance regional performance and has informed subsequent analytical methodologies at the European scale. For example, Ferrarini et al. (2024) use the TOPSIS multi-criteria method to assess competitiveness among EU regions, demonstrating how composite indicator techniques can effectively complement cluster-based models in revealing structural heterogeneity. Regional outcomes, regional outputs, regional throughputs and factors influencing regional competitiveness are represented by different levels of the hat. To account for the presence of external opportunities and threats, the determinants are presented on the way to the crown of the hat and then they return to the base. In many rings surrounding the production cylinder, the fundamental determinants of regional competitiveness are located at the bottom of the hat. The first ring contains the components of production (labour, capital and land). Since labour and land are less mobile, the regional powers have more influence on them. The most important determinants of the regional investment climate, including infrastructure and accessibility, human resources and the productive environment, are found in the second ring. Institutions, internationalisation, technology, demographics, location quality and environment are examples of secondary factors. Figure 1 shows the regional competitiveness hat.

Figure 1 Regional competitiveness hat



Source: Martin, 2004

These factors are all linked to regional competitiveness. The regional competitiveness hat includes a variety of flexible activities, such as the influence that multinational companies have on the availability, price and standard of determinants. The regional competitiveness hat also draws attention to how companies have become more competitive in industries that are more specialised (economic structure), gain market share, are more productive and profitable, consolidate areas with high per capita production, and promote an improved quality of life. At this level, the economy is able to deal with the opportunities and challenges of the global market. Ručinska and Ručinsky (2007) use the concept of regional competitiveness hat for their study. Considering all hierarchical and classifying divisions/types of regions, the authors choose the one that classifies them on the basis of regional competitiveness. According to Martin (2004), regions can be Regions as places of export specialisation, Regions as sources of increasing returns, and Regions as knowledge hubs. It can be said that the factors of regional competitiveness are different for each of these types of regions, but they overlap.

The three classifications of regions created by Martin (2004) are described below:

- **Regions as places of export specialisation (production locations)**

According to Martin (2004), regions compete to attract economic activity by exploiting their comparative advantages resulting from locational factors such as the availability of resources, labour and access to markets. Low- to middle-income regions specialise in industries that rely on cheap inputs and thus avoid the disadvantages of urbanisation. They attract vertical foreign investment and focus on cheaper production, often as production hubs. Factors such as labour costs, population density and access to transport hubs also influence their status as manufacturing regions.

- **Regions as a source of increasing yields (with increasing yields)**

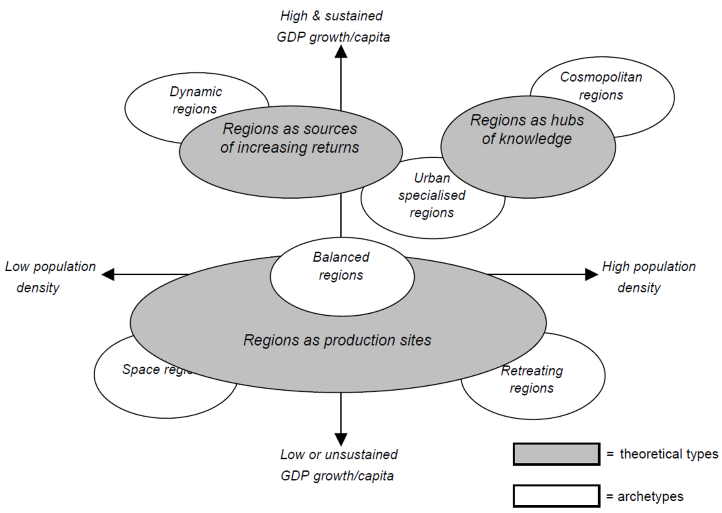
In recent years, the concept of increasing returns has regained importance in the economy. The growth of a region is linked to the demand for its exports, with the expansion of production promoting technological change and increasing labour productivity. Martin (2004) claims that dynamic regions with an average population density and a specific economic structure become sources of increasing returns. These regions, such as Zuid-Oost-Brabant, Oost-Vlaanderen, Rhône-Alpes and Toulouse, benefit from agglomeration effects, specialised industries, labour and technological spillover effects that enable them to achieve high and sustainable incomes and long-term competitiveness.

- **Regions as hubs of knowledge**

Martin (2004) points out that Marshall emphasised the importance of three key economies of localisation: specialised labour, supporting companies, and division of labour between companies. These interactions create a "local industrial climate" that fosters innovation and entrepreneurship. Regions that are knowledge hubs often rely on local innovation networks that include businesses, universities and other institutions, as well as the economic benefits of urbanisation. Porter's concept of geographical clusters also emphasises the importance of regions for competitiveness (Porter, 2012). Examples of innovative regions include Baden-Württemberg and Emilia-Romagna, which are characterised by high research intensity, high wages and a strong presence of universities.

In addition to the regions mentioned above, Martin (2004) also lists the following regions: space regions, balanced regions, dynamic regions, retreating regions, regions of urban specialisation and cosmopolitan regions, which lie at the interface between the three main regions mentioned above and define the three main types. The graphical representation of Martin's division into regions can be found in Figure 2 below. It is determined by two axes on which the population density and the (growth) GDP per capita lie.

Figure 2 Typology of regions



Source: Martin, 2004

From the above it can be concluded, at least in part, that the only thing that is certain for regions as hubs of knowledge is a high and growing GDP per capita and a high population density. Regions as production centres have a low GDP per capita, they tend to be on the low population density side, but not necessarily. Regions with growing yields are assigned moderately high GDP growth per capita and moderate to low population density. This confirms how difficult it is to categorise regions into a specific (theoretical) group.

The research conducted by Martin (2004) was based on the NUTS 2 regions of the then EU Member States, but also of the EU candidate countries (which include the countries that are the subject of this research, i.e., the CEE countries that are EU Member States), but the author did not have the data for the candidate countries (except for some at NUTS 0 and 1 levels). Martin (2004) draws on seven case studies, each looking at a NUTS 2 region from Belgium, Spain, France, Italy, Finland, the UK and Hungary. The only region he examined that belongs to the regions that are the subject of this paper is the Hungarian region of Nyugat (the then HU03), which belonged to the group of regions defined as production locations.

A slightly different model of the regional competitiveness hat was used by Cambridge Econometrics ECORYS-NEI and provides a methodological framework for analysing and assessing regional competitiveness (Nowak, 2011). The process involves several important steps. First, regional performance is assessed using economic indicators such as GDP per capita and gross value added to gain a basic understanding of regional success. Second, the analysis shifts to evaluating how efficiently regions use their resources, which involves looking at indicators like regional value added, profitability, market share, and unit labour costs, all of which offer insight into how productive and competitive a region really is. Third, the structure of the regional economy is explored by examining how sectors are distributed, the level of specialisation, how firms are spread out, and what types of ownership are present. The aforementioned helps reveal the underlying economic makeup of each region. Finally, the study identifies a region's potential for future competitiveness. This includes foundational elements such as infrastructure, the quality and availability of human capital, and the broader production environment. A wide range of influencing factors are considered, among them labour costs, profitability, sectoral structure, institutional support, access to technology, innovation capacity, entrepreneurial activity, international engagement, social capital, knowledge exchange, cultural dynamics, demographic trends, migration, locational advantages, and environmental quality. Taken together, this broad and multidimensional approach offers a more complete picture of what drives regional success, integrating economic, social, and institutional factors (Nowak, 2011).

Clustering methods have previously been applied to CEE countries, and the findings of this study align closely with those of Psycharis et al. (2020), who also used a cluster-based approach to examine regional competitiveness across Europe. By classifying the regions in their study according to economic and social indicators, emerged clusters reveal differences in competitiveness, particularly between the Central and Eastern European (CEE) regions and their Western counterparts. Both studies emphasise that the most important factors for regional differentiation are GDP per capita and investments related to innovation. Conceptually, the clusters in this study (knowledge hubs, production locations and regions with increasing yields) correspond to the clusters identified by Psycharis et al. (2020), which include innovation-driven regions, transition economies and structurally disadvantaged areas. The parallels make it clear how useful clustering is as an analytical tool for analysing regional economic systems.

Although regional competitiveness has been extensively examined, recent studies highlight the need for more dynamic and data-driven approaches to capture the evolving nature of disparities within post-transition economies. Conceptual frameworks such as Martin's (2004) "Regional Competitiveness Hat" were primarily developed for Western European contexts, leaving both conceptual and empirical uncertainties regarding their relevance for Central and Eastern Europe (CEE). This research gap has been reiterated in several recent studies (Grassia et al., 2024). Grassia et al. (2024) also conducted a structure-based topic analysis of contemporary literature on regional competitiveness, revealing a growing emphasis on incorporating institutional and sustainability dimensions into competitiveness models. Their findings reinforce the need for continuous methodological refinement in evaluating regional performance, particularly in post-transition settings. Chrobocińska (2023), for example, used RCI-type indicators and clustering analyses for CEE regions, identifying persistent heterogeneity between post-transition and Western European economies. Similarly, Kouskoura et al. (2024) highlighted the increasing significance of sustainability and digitalisation metrics in competitiveness assessments. Moreover, the European Commission's Regional Competitiveness Index 2.0 (Dijkstra et al., 2023) integrates updated methodological elements that better capture innovation dynamics and institutional quality. Collectively, these contributions underscore the need to empirically revisit classical frameworks such as Martin's within the specific socioeconomic conditions of CEE countries, a task that constitutes the principal contribution of this paper.

However, this study focuses on post-transition CEE areas and provides a more thorough examination of their particular development problems and trajectories, while Psycharis et al. (2020) present a national analysis. This study highlights the institutional and spatial elements that influence regional competitiveness and provides a solid theoretical framework for interpreting the cluster results by incorporating Martin's (2004) regional competitiveness hat model. Examining the stability and mobility of regions over time is one of the key differentiators of this study. In contrast to Psycharis et al. (2020), who mainly focus on static cluster characteristics, this paper presents a transition analysis that observes the movement of regions between clusters over the observation period. The policy-relevant component that this dynamic perspective entails enables targeted recommendations for transitional zones. In addition to validating the cluster technique used in this study, a comparison with Psycharis et al. (2020) shows its particular contribution to understanding the changing regional competitive landscape of CEE regions within the wider European framework.

3. K-means cluster analysis

In this paper, a cluster analysis using the k-means method was used to group regions from the sample into clusters of these regions, which can be seen later in the paper. The research sample includes 11 post-transition EU countries (Bulgaria, Croatia, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia, as mentioned above, CEE countries) and the corresponding 61 NUTS 2 regions (according to the NUTS 2021 terminology). The data were collected and processed for the period from 2011 to 2021 for all regions mentioned.

Methodology

One of the common techniques for statistical data analysis that is used in many areas is clustering. This is the process of grouping similar objects into different groups, i.e., dividing a data set into subsets, with the data in each subset separated by a certain distance measure (Everitt et al., 2011; Madhulatha, 2012; Wieland, 2019; Eva et al., 2022; Gamidullaeva et al., 2022; Giełczewski et al., 2022). In clustering, the algorithms can be hierarchical or partitioned. Most studies use one of two popular heuristic methods, such as the k-means algorithm and the k-medoids algorithm method.

Clustering is a descriptive technique, and the clustering solution is not unique and depends largely on the researcher's choice. Clustering always results in groups, even if there is no group structure (Landau and Everitt, 2003; Madhulatha, 2012). Cluster analysis aims to discover groups of observations from originally unclassified data (Landau and Everitt, 2003). The k-means method assigns each point to the cluster whose centre, also called the centroid, is closest to the cluster (Landau and Everitt, 2003). The centre is the average of all points in the cluster, i.e., its coordinates are the arithmetic mean for each dimension separately for all points in the cluster (Madhulatha, 2012). As Devčić et al. (2012) stated, "the main advantage of cluster analysis lies in the objective data reduction based on the reduction of information from the entire population and the reduction of population characteristics to those of representative groups with minimal loss of information".

The preliminary analyses include the deflation of monetary variables using the consumer price index (CPI) to present the values in constant prices, the standardisation of the data by calculating z-values and the outlier analysis using the Mahalanobis distance. Deflation is performed on variables related to GDP, labour productivity and foreign investment, while standardisation allows the comparison of data expressed in different units of measurement. Outliers were identified using the Mahalanobis test and mainly relate to the capitals of CEE countries.

Results

Clustering with the k-means method set to 10 iterations has shown that by the tenth iteration there are no more differences in the cluster centres, which means that the data have reached stability and three clusters have been formed using the above method (Table 1). The number of clusters was determined according to Martin's (2004) theoretical framework, which conceptually distinguishes three regional types: knowledge hubs, production locations, and regions with growing yields. Although statistical validation methods such as the elbow or silhouette approach were considered, they were not applied, as the analysis was intended to be primarily exploratory and theory-driven, rather than to identify a purely data-driven optimal solution. Given the limited number of standardised variables and the relatively small NUTS 2 sample, such indices would likely yield unstable or misleading results. Therefore, the selection of three clusters was based on theoretical coherence and the demonstrated stability of the iterative process, rather than on additional statistical criteria. All other k-means clustering attempts did not provide good results in less than 15 to 20 iterations, showing that 3 clusters are the most stable, as clustering ended in the tenth iteration.

During the clustering process, the k-means algorithm iteratively adjusts the cluster centres to minimise the differences within each cluster. By the tenth iteration, the algorithm reached a point where no significant changes in cluster centres were observed, indicating that the grouping had stabilised. This stability means that the division into three clusters accurately reflects the inherent structure of the data.

Table 1 History of cluster creation iterations

History of iterations[a]			
Iterations	**Change in cluster centres**		
	1	**2**	**3**
1	1.044	2.123	2.134
2	.424	.041	.022
3	.458	.149	.257
4	.151	.224	.551
5	.147	.072	.145
6	.285	.008	.036
7	.437	.000	.046
8	.126	.007	.027
9	.000	.005	.011
10	.000	.000	.000

a. Convergence achieved by little or no change in cluster centres. The largest absolute coordinate change for each centre is 0.000. The current iteration is 10. The minimum distance between the initial centres is 3.277.

Source: Unukić, 2024

When determining the number of clusters using the k-means method for the entire period from 2011 to 2021, three variables were used, of which the results of the ANOVA test show that all variables have a statistically significant influence on the classification of the regions into the identified clusters (Table 2). The results of the ANOVA test show that all three variables have the same statistically significant influence, as shown in the table below. This shows that there are significant differences in the mean values of these variables between the different clusters. Population density (F = 5038.279), followed by GDP per capita (F = 440.425) and gross domestic expenditure on research and experimental development (F = 424.267) provide the greatest separation between the clusters. ANOVA tests showed statistically significant differences in the mean values of these variables between the clusters, underlining the appropriateness of three different cluster groups. The decision to group the units into three clusters is based on a combination of statistical rigour, theoretical basis and the observed stability of the clustering process.

Table 2 Results of the ANOVA test of clusters for the observed period

<table>
<tr><th colspan="7">ANOVA</th></tr>
<tr><th rowspan="2"></th><th colspan="2">Cluster</th><th colspan="2">Error</th><th rowspan="2">F</th><th rowspan="2">Sig.</th></tr>
<tr><th>Mean Square</th><th>df</th><th>Mean Square</th><th>df</th></tr>
<tr><td>Population density (gusnas)</td><td>59.166</td><td>2</td><td>.012</td><td>624</td><td>5038.279</td><td>.000</td></tr>
<tr><td>Gross domestic expenditure on research and experimental development (dBI1)</td><td>77.390</td><td>2</td><td>.182</td><td>624</td><td>424.267</td><td>.000</td></tr>
<tr><td>GDP per capita (dBDPpc)</td><td>77.914</td><td>2</td><td>.177</td><td>624</td><td>440.425</td><td>.000</td></tr>
<tr><td colspan="7">* F-tests should only be used for descriptive purposes, as the clusters were selected to maximise the differences between the cases in the different clusters. The observed significance levels were not corrected for this and therefore cannot be interpreted as tests for the hypothesis that the means of the clusters are equal.</td></tr>
</table>

Source: Unukić, 2024

The analysis of the cluster analysis results in Table 3 shows that three clusters have emerged, which differ significantly in terms of their number. Clusters 1, 2, and 3 contain 22, 411, and 194 regions, respectively.

Table 3 Number of clusters

<table>
<tr><th colspan="3">Number of cases in each cluster</th></tr>
<tr><td rowspan="3">Cluster</td><td>1</td><td>22</td></tr>
<tr><td>2</td><td>411</td></tr>
<tr><td>3</td><td>194</td></tr>
<tr><td colspan="2">Valid</td><td>627</td></tr>
<tr><td colspan="2">Missing</td><td>.000</td></tr>
</table>

Source: Unukić, 2024

The post-hoc tests are carried out according to the calculated clusters and can be seen in Table 4 Since the assumption of homogeneity of variance of the observed variables was violated ($p < 0.05$), the authors use the Games-Howell test, which shows the differences between the clusters obtained on the basis of the observed variables.

Table 4 Games-Howell test

Multiple Comparisons							
Games-Howell							
Dependent Variable	(I) QCL_1 Cluster Number of Case	(J) QCL_1 Cluster Number of Case	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Zgusnas Zscore (Population Density)	1	2	2.379*	.078	.000	2.184	2.575
		3	2.289*	.078	.000	2.093	2.486
	2	1	-2.379*	.078	.000	-2.575	-2.184
		3	-.089*	.010	.000	-.113	-.066
	3	1	-2.289*	.078	.000	-2.486	-2.093
		2	.089*	.010	.000	.066	.113
ZdBDPpc Zscore (deflated GDP per capita)	1	2	1.798*	.095	.000	1.559	2.035
		3	.894*	.101	.000	.644	1.144
	2	1	-1.798*	.095	.000	-2.035	-1.559
		3	-.903*	.043	.000	-1.005	-.802
	3	1	-.894*	.101	.000	-1.144	-.644
		2	.903*	.043	.000	.802	1.005
ZdBI1 Zscore (deflated gross domestic expenditure on research and experimental development)	1	2	.726*	.128	.000	.402	1.049
		3	-.347*	.137	.044	-.686	-.008
	2	1	-.726*	.128	.000	-1.049	-.402
		3	-1.073*	.049	.000	-1.189	-.956
	3	1	.347*	.137	.044	.008	.681832
		2	1.073*	.049	.000	.956	1.189
* The mean difference is significant at the 0.05 level.							

Source: Unukić, 2024

Looking at the characteristics of the clusters obtained, i.e., their average values in the observation period, they can be assigned characteristic names following the example of Martin (2004). Cluster 1 is characterised above all by the highest average GDP per capita (approx. 19,058 euros/person) and the highest average population density (1,447.7 persons/km2). For the reasons mentioned above, Cluster 1 will be referred to as the hub of knowledge in the remainder of the text.

Using the same variables to determine the name of the cluster and the region groups of Martin (2004), Cluster 2 is characterised by the lowest GDP per capita (7,833.86 euros/person) and the lowest population density (83.19 persons/km2), so that it can be classified into the group of regions as production locations and Cluster 2 bears this name in the text that follows.

As Cluster 3 has a relatively high GDP per capita (13,468.79 euros/person) and a relatively low population density, it can be classified into the group of regions with growing yields according to Martin (2004) and will bear this name in the text that follows.

Cluster 1, regions as hubs of knowledge, comprises 22 observation units that refer to the regions of two large cities, the Croatian capital Zagreb (HR05) and the Romanian capital Bucharest (RO32). Throughout the observation period, these two capitals belong to Cluster 1, which is not unusual as capitals usually have the highest GDP per capita and the highest population density.

Cluster 2, regions as production locations, contains the largest number of observation units, namely 411. This cluster is characterised by the lowest GDP per capita, the highest foreign direct investment per capita, the longest duration of compulsory education and the lowest population density. In 2011, Cluster 2 comprised 43 regions, in 2012 - 41 regions, in 2013 - 43 regions, in 2014 - 40 regions, in 2015 - 36 regions, in 2016 - 37 regions, in 2017 - 36 regions, in 2018, 2019 and 2020 - 34 regions, and in 2021 - 33 regions. It is evident that there have been changes in the regions during the observation period, but only between Cluster 2 (regions as production locations) and Cluster 3 (regions with growing yields), as Cluster 1 (regions as hubs of knowledge) has an unchanged number of observed units throughout the observation period.

Cluster 3 (regions with growing yields) comprises 194 observation units and is characterised by average GDP per capita values, lower population density, the largest share of employees in the manufacturing sector and the highest availability of the Internet in the regions. This cluster comprised 12 regions in 2011, 14 regions in 2012, 12 regions in 2013, 15 regions in 2014, 19 regions in 2015, 18 regions in 2016, 19 regions in 2017, 21 regions in 2018, 2019 and 2020, and 22 regions in 2021.

Table 5 shows the transition matrix of the clusters for the first and last year of observation.

Table 5 Transition matrix[1] (2011–2021)

Transition from → to	Hubs of Knowledge	Production Locations	Growing Yields	Total
Hubs of Knowledge	2 (Stable)	0	0	2
Production Locations	0	27 (Stable)	6	33
Growing Yields	0	7	15 (Stable)	22
Total	2	34	21	57

Source: Authors, based on data available in Unukić, 2024

Looking at the stability of the clusters, it can be seen that the hubs of the knowledge cluster, represented by Zagreb and Bucharest, have remained completely stable and have maintained their high GDP per capita and population density. Most of the regions in the manufacturing cluster have also been able to maintain their position, which is due to a consistent economic structure based on cost advantages and export specialisation. Similarly, the growing yields cluster was also stable, with most regions continuing to show moderate GDP growth and low population density.

Mobility between clusters was observed mainly between production locations and growing yields, with some regions moving upwards due to improved competitiveness and growth momentum. Conversely, regions that moved from growing yields to production locations may indicate stagnation or lower productivity. The shifts between hubs of knowledge and other clusters were negligible, underlining their consolidated position.

The hubs of knowledge cluster shows remarkable stability, which underlines its established role as an economic leader. Although the production locations cluster is largely stable, it shows moderate transitions to the growing yields cluster, reflecting the evolving economic structures. The growing yields cluster shows both upward and downward mobility, indicating that it is more dynamic but also more susceptible to economic change.

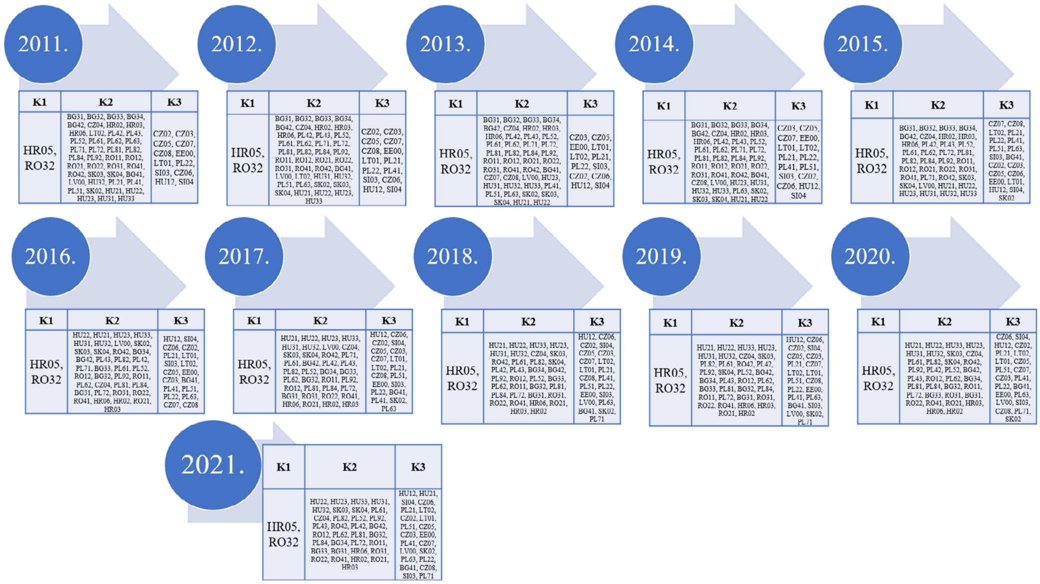
Policy implications according to transition matrix include:

- To maintain their leadership role, knowledge centres need continuous investment in innovation and infrastructure.
- For production locations, policy should focus on structural change, such as promoting R&D and workforce development.
- In the growing yields cluster, downward movers need targeted interventions to counter stagnation, while upward movers would benefit from strategic investments in productivity-enhancing sectors to consolidate their progress.

The following Figures 3 and 4 provide an overview of the changes in the clusters during the observation period.

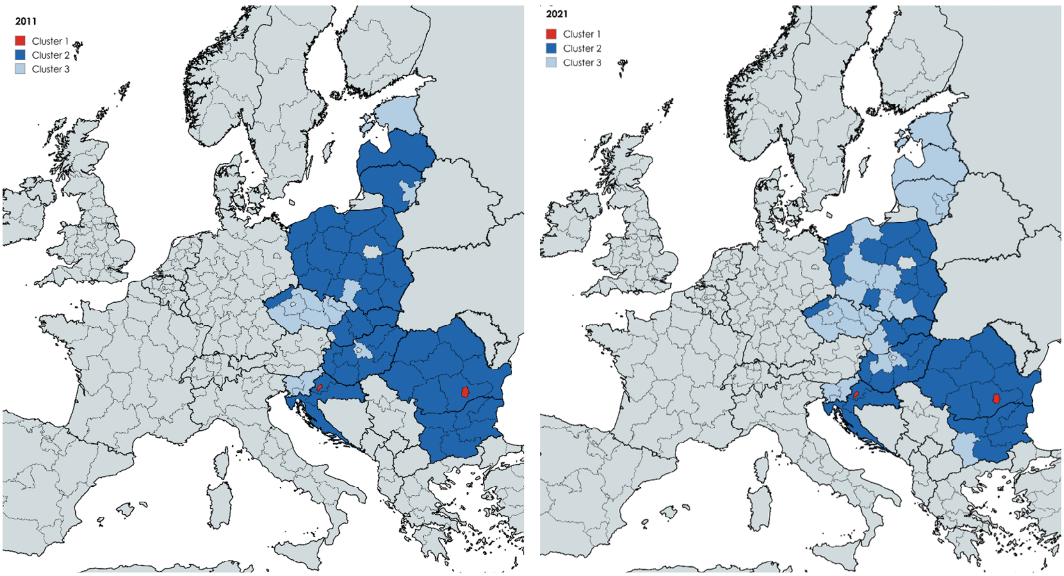
1 Diagonal cells (e.g., Knowledge Hubs → Knowledge Hubs) represent regions that remained in the same cluster throughout the observed period; off-diagonal cells (e.g., Production Locations → Regions with Growing Yields) represent regions that transitioned between clusters.

Figure 3 Clusters throughout the observed period



Source: Unukić, 2024

Figure 4 Maps of clusters in 2011 and 2021



Source: Authors, based on data available in Unukić, 2024

4. Discussion: Challenges and perspectives

The empirical findings support Martin's (2004) conceptual framework within the CEE context, confirming the presence of regional differentiation that corresponds to the model's theoretical categories. Nonetheless, several challenges remain. The ongoing divergence in competitiveness levels among post-transition CEE regions raises questions about the long-term effectiveness of the EU cohesion policies in promoting convergence.

While traditional indicators such as GDP per capita and R&D intensity remain relevant, emerging evidence (Kouskoura et al., 2024) shows that sustainability and digital capacity are increasingly shaping regional competitiveness outcomes. Similarly, recent spatial clustering analyses (Sánchez & Cuadrado-Roura, 2024) indicate that stronger sectoral specialisation contributes to regional resilience—an aspect not fully captured by the current quantitative framework.

Future research should therefore combine statistical clustering techniques with spatial econometric and network-based approaches to better encompass qualitative dimensions such as governance efficiency, institutional quality, and innovation networks.

5. Conclusion

This study is an exploratory effort to empirically test Martin's (2004) Regional Competitiveness Hat model in the context of post-transition EU member states of Central and Eastern Europe. Although the analysis provides meaningful insights, it is preliminary in scope and calls for further empirical investigations using broader datasets and expanded indicator frameworks.

According to the research results, the first cluster includes the capital cities Zagreb (H505) and Bucharest (RO32). This cluster is characterised by a high GDP per capita and a high population density, which places it in the group of regions as hubs of knowledge, according to Martin's (2004) classification. The second cluster is marked by a low GDP per capita and a low population density, placing it within the category of production locations. In contrast, the third cluster shows an upward trend in GDP per capita while maintaining a low population density, features that align with Martin's definition of regions with growing yields.

In this study, the Hungarian region of Nyugat-Dunántúl falls into the second cluster, just as it did in Martin's original research. The region is characterised by a relatively low GDP per capita but stands out for its strong productivity and high levels of foreign investment. The k-means clustering results align well with Martin's regional competitiveness hat model, supporting the existence of the three regional types (knowledge hubs, production locations, and regions with growing yields), as outlined in his framework. This alignment not only reinforces the validity of Martin's model, but also extends its relevance to the NUTS 2 regions of Central and Eastern Europe during the 2011-2021 period. In doing so, the findings demonstrate that the model holds up well in the context of post-transition EU countries, where economic and institutional transformations have been particularly significant. In the hypothetical case of contradictory results, these findings would indicate either methodological limitations or the existence of unique regional dynamics that Martin's framework does not fully capture, thus providing opportunities to refine the theoretical model. The novelty of this research lies in the application of Martin's model to CEE regions–a context largely unexplored in previous studies. Moreover, the real value of this study comes from combining the model with advanced methods like k-means clustering, which adds a strong, data-driven foundation. By doing so, the research offers practical insights into key factors that influence regional competitiveness, such as GDP per capita, population density, and R&D spending. These findings help identify distinct regional patterns and support the central hypothesis of the paper.

The findings of this study are highly relevant for shaping regional development policies within the European Union, especially in Central and Eastern Europe (CEE). By categorising regions into three aforementioned distinct groups, the study offers a clear and practical framework for crafting tailored economic strategies. This classification helps policymakers recognise the unique strengths and challenges of each region. For instance, knowledge hubs could benefit from increased investment in education, research, and innovation, while manufacturing-based regions (production locations) may need better infrastructure or reforms in labour market policies.

The study also identifies key drivers of regional competitiveness, such as GDP per capita, population density, and spending on research and development. These insights can guide the allocation of EU cohesion funds and national resources, helping ensure that financial support goes where it can have the greatest impact, whether to boost high-potential areas or support those in need.

Importantly, the research sheds light on the economic transitions that CEE countries have undergone in recent decades. This makes it particularly valuable for understanding how competitiveness is evolving in the region. Moreover, it offers lessons for EU candidate countries, providing a useful model for their integration and development strategies.

Overall, by providing solid empirical data, this study supports the EU's cohesion policy and underscores where efforts to reduce regional disparities should be focused. This alignment with EU objectives makes development initiatives more effective in promoting balanced growth and economic convergence across the Union.

Despite valuable contributions of the study, several methodological limitations should be acknowledged. The clustering approach was guided by theoretical reasoning rather than formal validation indices, such as the elbow or silhouette method. The analysis was based on a limited set of variables (GDP per capita, population density, and R&D expenditure), which, although central to regional competitiveness, do not fully capture institutional quality, infrastructure, or human capital. These factors may have affected the precision and explanatory depth of the clusters. Future research should therefore expand the set of indicators and include complementary validation techniques to strengthen methodological reliability and provide more nuanced insights into regional dynamics.

Looking ahead, the study also lays a foundation for future policy-making. It calls for expanding the research to include more regions, particularly EU candidate countries, and suggests incorporating new indicators, such as digitalisation and green economy metrics. Doing so would offer deeper insights into emerging competitiveness trends and help policymakers stay ahead of evolving regional challenges.

To build a more complete picture, future research should broaden its scope to cover the entire European economic landscape. This would allow for a richer understanding of regional dynamics across varying levels of development and adaptation, and ultimately support more inclusive and forward-thinking regional strategies.

Building on the findings presented here, several open questions warrant further exploration. To what extent might digitalisation and the green transition reshape the spatial patterns of competitiveness across CEE? How do institutional quality and governance capacity interact with economic variables to influence cluster stability? Furthermore, can the inclusion of non-economic factors (such as social capital or innovation culture) enhance the explanatory power of Martin's framework? Addressing these questions would contribute to a more comprehensive understanding of regional dynamics and support further refinement of competitiveness theory.

References

1. Chrobocińska, K. (2023). Competitiveness of Regions in Selected Countries of Central and Eastern Europe. *Central European Economic Journal*, *10*(57), 392-402. https://doi.org/10.2478/ceej-2023-0022
2. Devčić, K., Tonković Pražić, I. & Župan, Ž. (2012). Klaster analiza: primjena u marketinškim istraživanjima. *Zbornik radova Međimurskog veleučilišta u* Čakovcu, *3*(1), 15-22.
3. Dijkstra, L., Papadimitriou, E., Martinez, B. C., De Dominicisi, L. & Kovacic, M. (2023). *EU Regional Competitiveness Index 2.0: 2022 edition*. Publications Office of the European Union.
4. Eva, M., Cehan, A., Corodescu-Roșca, E. & Bourdin, S. (2022). Spatial patterns of regional inequalities: Empirical evidence from a large panel of countries. *Applied Geography*, 140, 102638. https://doi.org/10.1016/j.apgeog.2022.102638
5. Everitt, B. S., Landau, S., Leese, M. & Stahl, D. (2011). *Cluster analysis.* John Wiley & Sons. https://doi.org/10.1002/9780470977811
6. Ferrarini, F., Muzzioli, S. & De Baets, B. (2024). A TOPSIS analysis of regional competitiveness at European level. *Competitiveness Review*, *34*(7), 52-72. https://doi.org/10.1108/CR-01-2024-0005
7. Gamidullaeva, L., Korostyshevskaya, E., Myamlin, A. & Podkorytova, O. (2022). Exploring Regional Industrial Growth: Does Specialization Explain It?. *Economies*, *10*(7). https://doi.org/10.3390/economies10070172

8. Giełczewski, M., Piniewski, M. & Domański, P. D. (2022). Mixed statistical and data mining analysis of river flow and catchment properties at regional scale. *Stochastic Environmental Research and Risk Assessment*, *36*(9), 2861-2882. https://doi.org/10.1007/s00477-022-02169-3
9. Grassia, M. G., Marino, M., Mazza, R., Misuraca, M., Zavarrone, E. & Friel, M. (2024). Regional competitiveness: A structural-based topic analysis on recent literature. *Social Indicators Research*, *173*(1), 83-108. https://doi.org/10.1007/s11205-022-02951-4
10. Kouskoura, A., Kalliontzi, E., Skalkos, D. & Bakouros, I. (2024). Assessing the key factors measuring regional competitiveness. *Sustainability*, 16(6), 2574. https://doi.org/10.3390/su16062574
11. Landau, S. & Everitt, B. S. (2003). *A handbook of statistical analyses using SPSS.* Chapman and Hall/CRC. https://doi.org/10.1201/9780203009765
12. Madhulatha, T. S. (2012). An overview on clustering methods. *arXiv preprint arXiv:1205.1117.* https://doi.org/10.9790/3021-0204719725
13. Martin, R. L. (2004). *A study on the factors of regional competitiveness. A draft final report for the European Commission Directorate-General Regional Policy*. Cambridge: University of Cambridge.
14. Melecký, L. & Staníčková, M. (2014). NUTS 2 regions classification: comparison of cluster analysis and DEA method. In *17th International Colloquium on Regional Sciences. Conference Proceedings* (pp. 45-53). https://doi.org/10.5817/CZ.MUNI.P210-6840-2014-4
15. Nowak, P. (2011). *Nowy paradygmat rozwoju regionalnego na przykładzie wybranych regionów Europy Zachodnie* [Doctoral dissertation, Poznań University of Economics and Business]. Poznań University of Economics and Business.
16. Porter, M. E. (2012). The Economic Performance of Regions. In Martin, R. et al. (Eds.), *Regional Competitiveness* (pp. 131-160). Routledge.
17. Psycharis, Y., Kallioras, D. & Pantazis, P. (2020). Regional Inequalities in Central and Eastern European Countries: The Role of Capital Regions and Metropolitan Areas. *Economic Development and Financial Markets: Latest Research and Policy Insights from Central and Southeastern Europe*, 3-20. https://doi.org/10.1007/978-3-030-32426-1_1
18. Rahmat, S. & Sen, J. (2021). A multi-model approach to assess the relative weights and sensitivities of the factors of regional competitiveness. *Journal of Urban & Regional Analysis*, *13*(1). https://doi.org/10.37043/JURA.2021.13.1.3
19. Ručinska, S. I. L. V. I. A. & Ručinsky, R. (2007). Factors of regional competitiveness. In *2nd Central European Conference in Regional Science–CERS* (pp. 902-911).
20. Sánchez, A. M. & Cuadrado-Roura, J. R. (2024). Sectoral–regional resilience and productive specialization: a comparison among the last three crises. *The Annals of Regional Science*, *73*(1), 61-86. https://doi.org/10.1007/s00168-024-01261-6
21. Shvindina, H. (2022). Economic competitiveness: An overview of multilevel concept. *Responsible Consumption and Production*, 160-172. Springer. https://doi.org/10.1007/978-3-319-95726-5_63
22. Snieška, V. & Bruneckienė, J. (2009). Measurement of Lithuanian regions by regional competitiveness index. *Engineering Economics*, *61*(1).
23. Unukić, I. (2024). *Izvori regionalne ekonomske transformacije posttranzicijskih zemalja EU* [Doctoral dissertation, Josip Juraj Strossmayer University of Osijek]. Josip Juraj Strossmayer University of Osijek.
24. Wieland, T. (2019). REAT: A Regional Economic Analysis Toolbox for R. *Region*, *6*(3). https://doi.org/10.18335/region.v6i3.267