CLUSTER ANALYSIS OF THE EU REGIONAL COMPETITIVENESS INDEX OF NUTS-2 REGIONS

Pavol ORŠANSKÝ

Abstract

This research investigates the complex dynamics of regional development within the European Union by performing a cluster analysis of the EU Regional Competitiveness Index (RCI 2.0) across 234 NUTS-2 regions. The central issue addressed is the "Capital City Bias" and the challenge of balancing industrial productivity with the quality of life for residents. Furthermore, the study explores the "middle-income trap," a problematic state where regions transitioning through developmental stages may face a policy vacuum if basic infrastructure is neglected before innovation ecosystems are fully mature. The primary objective is to identify hidden patterns and specific similarities within regional groupings to move beyond simple rankings and better understand the unique developmental needs of different clusters. To achieve this, the study utilizes the k-means++ clustering algorithm, an advanced iteration of Lloyd's algorithm that employs a heuristic for more effective centroid seeding to improve both running time and solution quality. The research focuses on the three core sub-indices of the RCI: Basic (including institutions and infrastructure), Efficiency (labor market and higher education), and Innovation (technological readiness and business sophistication). To determine the optimal number of clusters for each sub-index, the Calinski-Harabasz criterion (variance ratio criterion) is applied, ensuring that the resulting data partitions are both dense and well-separated. Furthermore, Non-negative Matrix Factorization (NNMF) is employed as a sophisticated visualization tool, allowing for the transformation of multidimensional regional data into a two-dimensional plane while preserving essential Euclidean norms. The results demonstrate a persistent geographical divide in Europe, characterized by a stark "elitism" in capital cities compared to their stagnating peripheries, providing critical insights for the tailoring of future Cohesion Policies.

Key words:
cluster analysis, Regional Competitiveness Index, k-means++ clustering

JEL Classification M12, M54, O32 https://doi.org/10.52665/ser20250207

INTRODUCTION

Since 2010, the EU Regional Competitiveness Index (RCI) has been measuring the major factors of competitiveness for all the NUTS-2 level regions across the European Union. The Index measures, with a rich set of indicators, the ability of a region to offer an attractive environment for firms and residents to live and work. Since the 2022 edition of the RCI uses an updated methodological framework, to facilitate comparison over time. In addition, starting from the original data used in 2016 and 2019, the scores have been re-calculated using the new methodology, labelled as RCI 2.0, 2016 edition, and RCI 2.0, 2019 edition. The resulting rankings do not replace the RCI rankings published in 2016 and 2019, produced with the old methodology. The RCI is composed of three sub-indices: **Basic**, **Efficiency** and **Innovation**, and of 11 pillars that describe the different aspects of competitiveness.

The **Basic sub-index** refers to the key basic drivers of all types of economies. It identifies the main issues that are necessary to develop regional competitiveness and includes five pillars: (1) *The Institutions*, (2) *The Macroeconomic Stability*, (3) *The Infrastructures*, (4) *The Health* and (5) *The Basic Education*. The **Efficiency sub-index** includes three pillars: (6) *Higher education, training and lifelong learning*, (7) *Labor market efficiency* and (8) *Market size*. Lastly, the **Innovation sub-index** includes the three pillars that are the drivers of improvement at the most advanced stage of economic development: (9) *Technological readiness*, (10) *Business sophistication* and (11) *Innovation*. The final

RCI 2.0 is weighted arithmetic mean of these three sub-indices, which are weighed differently per development stage (gross domestic product (GDP) per head in purchasing power standards (PPS) expressed as an index with the EU-27 average set to 100), as shown in Table 1. For more details of each pillar or others information about the methodology see (Dijkstra 2023).

Table 1: Table of sub-indexes weights of the RCI

Stage of Development	Sub-index weight		
	Basic	Efficiency	Innovation
GDP index[1] < 75	30%	50%	20%
GDP index[1] ∈ [75,100]	25%	50%	25%
GDP index[1] > 75	20%	50%	30%

Source: author's processing

[1] GDP/ head (PPS), Index EU-27 = 100.

In our work we try to find some specific similarities in each type of sub-index which are other than those in other groups. In other words, we do **cluster analysis** of every sub-index in relation to NUTS-2 regions of the EU. Cluster analysis involves applying clustering algorithms with the goal of finding hidden patterns or groupings in a data set. It is therefore used frequently in exploration data analysis but is also used for anomaly detection and preprocessing for supervised learning. Clustering algorithms form groupings in such a way that data within a group (or cluster) has a higher measure of similarity than data in any other cluster. Various similarity measures can be used, including Euclidean, probabilistic, cosine distance, and correlation. Most unsupervised learning methods are a form of cluster analysis. Clustering algorithms fall into two broad groups: (1) *Hard clustering*, where each data point belongs to only one cluster, such as the popular k-means method and (2) *Soft clustering*, where each data point can belong to more than one cluster, such as in Gaussian mixture models. Examples include phonemes in speech, which can be modeled as a combination of multiple base sounds, and genes that can be involved in multiple biological processes. We use **k-means clustering**, or Lloyd's algorithm (Lloyd 1982), which is an iterative, data-partitioning algorithm that assigns n observations to exactly one of k clusters defined by centroids, where k is chosen before the algorithm starts. We use an improved version of this algorithm called the **k-means++ algorithm**. The k-means++ algorithm uses a heuristic to find centroid seeds for k-means clustering. According to Arthur and Vassilvitskii (Arthur and Vassilvitskii 2007), k-means++ improves the running time of Lloyd's algorithm, and the quality of the final solution.

1 LITERATURE OVERVIEW

The literature on the EU Regional Competitiveness Index (RCI) reveals a central "problematic": the challenge of reconciling administrative boundaries with functional economic realities while balancing social well-being against industrial productivity. Academic debate in this area is primarily structured around three core tensions.

A recurring theme in the literature is the dual nature of regional competitiveness. While traditional indices (like the WEF's Global Competitiveness Index) focus on business productivity, the RCI problem lies in its attempt to measure a region's attractiveness for both firms and residents (Annoni & Dijkstra, 2019). This creates a theoretical friction: policies that benefit firms (e.g., lower corporate taxes or flexible labor markets) may sometimes conflict with the "quality of life" metrics (e.g., high social protection and environmental standards) that make a region attractive to residents.

Scholars frequently highlight the "Modifiable Areal Unit Problem" (MAUP) as a significant hurdle in RCI research. The index utilizes NUTS 2 administrative regions, which are often criticized for being "artificial" constructions that do not reflect actual labor markets or commuting patterns. Literature points out that this can lead to the "Capital City Bias", where a capital's high performance masks deep-seated stagnation in its immediate rural periphery, complicating the delivery of effective Cohesion Policy.

The RCI employs a unique methodology where pillars are weighed differently based on a region's stage of development (GDP per capita). The problematic identified here is the potential for a "middle-income trap." Literature (Dijkstra et al., 2023) suggests that as regions transition from "Basic" to "Efficiency" and "Innovation" stages, the shift in priorities can lead to a policy vacuum where basic infrastructure is neglected before innovation ecosystems are fully mature.

2 METHODOLOGY

The main method used in our work is cluster analysis which refers to a family of algorithms and tasks rather than one specific algorithm. It can be achieved by various algorithms that differ significantly in their understanding of what constitutes a cluster and how to efficiently find them. Popular notions of clusters include groups with small distances between cluster members, dense areas of the data space, intervals or particular statistical distributions. Clustering can therefore be formulated as a multi-objective optimization problem. The appropriate clustering algorithm and parameter settings (including parameters such as the distance function to use, a density threshold or the number of expected clusters) depend on the individual data set and intended use of the results. It is an iterative process of knowledge discovery or interactive multi-objective optimization that involves trial and failure. There is a common denominator: a group of data objects, which is one of the reasons why there are so many clustering algorithms.

k-means clustering is a method of vector quantization, originally from signal processing, that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells (partition of a plane into regions close to each of a given set of objects). k-means clustering minimizes within-cluster variances (squared Euclidean distances), but not regular Euclidean distances, which would be the more difficult Weber problem: the mean optimizes squared errors, whereas only the geometric median minimizes Euclidean distances. For instance, better Euclidean solutions can be found using k-medians and k-medoids. The problem is computationally difficult (nondeterministic polynomial - hard); however, efficient heuristic algorithms converge quickly to a local optimum.

Given a set of observations $(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n)$, where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k ($\leq n$) sets $\boldsymbol{S} = (S_1, S_2, \dots, S_k)$ so as to minimize the within-cluster sum of squares (WCSS) (i.e. variance). Formally, the objective is to find

$$\operatorname{argmin}_{\boldsymbol{S}} \sum_{i=1}^{k} \sum_{\boldsymbol{x} \in S_i} \|\boldsymbol{x} - \boldsymbol{C}_i\|^2 = \operatorname{argmin}_{\boldsymbol{S}} \sum_{i=1}^{k} |S_i| \operatorname{Var} S_i \quad (1)$$

where $\|\dots\|$ is the L^2 norm (Euclidean distance) between the two vectors and $\boldsymbol{C}_i$ is the mean (also called centroid) of points in S_i, i.e.

$$\boldsymbol{C}_i = \frac{1}{|S_i|} \sum_{\boldsymbol{x} \in S_i} \boldsymbol{x}, \quad (2)$$

where $|S_i|$ is the size of S_i. This is equivalent to minimizing the pairwise squared deviations of points in the same cluster

$$\operatorname{argmin}_{\boldsymbol{S}} \sum_{\boldsymbol{x}, \boldsymbol{y} \in S_i} \|\boldsymbol{x} - \boldsymbol{y}\|^2 \quad (3)$$

The equivalence can be deduced from identity

$$|S_i| \sum_{\boldsymbol{x} \in S_i} \|\boldsymbol{x} - \boldsymbol{C}_i\|^2 = \frac{1}{2} \sum_{\boldsymbol{x}, \boldsymbol{y} \in S_i} \|\boldsymbol{x} - \boldsymbol{y}\|^2 \quad (4)$$

Since the total variance is constant, this is equivalent to maximizing the sum of squared deviations between points in different clusters (between-cluster sum of squares, BCSS) (Kriegel 2017).

k-means clustering, or Lloyd’s algorithm, is an iterative, data-partitioning algorithm that assigns n observations to exactly one of k clusters defined by centroids, where k is chosen before the algorithm starts. The algorithm proceeds as follows:

- Choose k initial cluster centers (centroid). For example, choose k observations at random or use the k-means ++ algorithm for cluster center initialization (the default).

- Compute point-to-cluster-centroid distances of all observations to each centroid
- There are two ways to proceed: (1) *Batch update* - assign each observation to the cluster with the closest centroid, (2) *Online update* - individually assign observations to a different centroid if the reassignment decreases the sum of the within-cluster, sum-of-squares point-to-cluster-centroid distances.
- Compute the average of the observations in each cluster to obtain k new centroid locations.
- Repeat steps 2 through 4 until cluster assignments do not change, or the maximum number of iterations is reached.

k-means++ improves the running time of Lloyd's algorithm, and the quality of the final solution. The k-means++ algorithm chooses seeds as follows, assuming the number of clusters is k.

- Select an observation uniformly at random from the data set, x. The chosen observation is the first centroid and is denoted C_1.

- Compute distances from each observation to C_1. Denote the distance between C_1 and the observation m as $d(x_m, C_j)$.
- Select the next centroid, C_2 at random from x with probability

$$P = \frac{d^2(x_m, C_1)}{\sum_{j=1}^{n} d^2(x_j, C_1)}. \quad (5)$$

- To choose center j, we compute the distances from each observation to each centroid and assign each observation to its closest centroid. For each $m = 1, \dots, n$ and $p = 1, \dots, j-1$, select centroid j at random from x with probability

$$P = \frac{d^2(x_m, C_p)}{\sum_{x_h \in C_p} d^2(x_h, C_p)}, \quad (6)$$

where C_p is the set of all observations closest to centroid C_p and x_m belongs to C_p.

- Repeat step 4 until k centroids are chosen.

The algorithms use a two-phase iterative algorithm to minimize the sum of point-to-centroid distances, summed over all k clusters. I. This first phase uses batch updates, where each iteration consists of reassigning points to their nearest cluster centroid, all at once, followed by recalculation of cluster centroids. This phase occasionally does not converge with a solution that is a local minimum. That is, a partition of the data where moving any single point to a different cluster increases the total sum of distances. This is more likely for small data sets. The batch phase is fast, but potentially only approximates a solution as a starting point for the second phase. This second phase uses online updates, where points are individually reassigned if doing so reduces the sum of distances, and cluster centroids are recomputed after each reassignment. Each iteration during this phase consists of one passing through all the points. This phase converges to a local minimum, although there might be other local minimums with lower total sum of distances. In general, finding the global minimum is solved by an exhaustive choice of starting points, but using several replicates with random starting points typically results in a solution that is a global minimum

An important task in clustering is the correct determination of the number of clusters. This ensures that the data is properly and efficiently divided. The correct choice of k is often ambiguous, with interpretations depending on the shape and scale of the distribution of points in a data set and the desired clustering resolution of the user. In addition, increasing k without penalty will always reduce the amount of error in the resulting clustering, to the extreme case of zero error if each data point is considered its own cluster (i.e., when k equals the number of data points, n). Intuitively then, the optimal choice of k will strike a balance between maximum compression of the data using a single cluster, and maximum accuracy by assigning each data point to its own cluster. If an appropriate value of k is not apparent from prior knowledge of the properties of the data set, it must be chosen somehow. There are several categories of methods for making this decision. We use the Caliński-Harabasz criterion. This criterion is sometimes called the variance ratio criterion

(VRC). The Caliński-Harabasz index is defined as

$$VRC_k = \frac{SS_B}{SS_W} \cdot \frac{n-k}{k-1}, \quad (7)$$

where SS_B is the overall between-cluster variance, SS_W is the overall within-cluster variance k is the number of clusters, and n is the number of observations. The overall between-cluster variance SS_B is defined as

$$SS_B = \sum_{i=1}^{k} n_i \|m_i - m\|^2, \quad (8)$$

where k is the number of clusters, n_i is the number of observations in i-th cluster, m_i is the centroid of i-th cluster, m is the overall mean of the sample data. The overall within-cluster variance SS_W is defined

$$SS_W = \sum_{i=1}^{k} \sum_{x \in C_i} \|x - m_i\|^2, \quad (9)$$

where k is the number of clusters, x is a data point, C_i is the i-th cluster, m_i is the centroid of i-th cluster. Well-defined clusters have a large between-cluster variance (SS_B) and a small within-cluster variance (SS_W). The larger the VRC_k ratio, the better the data partition. To determine the optimal number of clusters, maximize VRC_k with respect to k. The optimal number of clusters corresponds to the solution with the highest Caliński-Harabasz index value (Caliński & Harabasz 1974).

Non-negative matrix factorization (NNMF), also non-negative matrix approximation is a group of algorithms in multivariate analysis and linear algebra where a matrix V is factorized into (usually) two matrices W and H, with the property that all three matrices have no negative elements. This non-negativity makes the resulting matrices easier to inspect. The factorization uses an iterative algorithm starting with random initial values for W and H. Because the root means square residual D might have local minima, repeated factorizations might yield different W and H. Sometimes the algorithm converges to a solution of lower rank than k, which can indicate that the result is not optimal. More detailed information can be seen in (Michael 2007).

3 FINDINGS

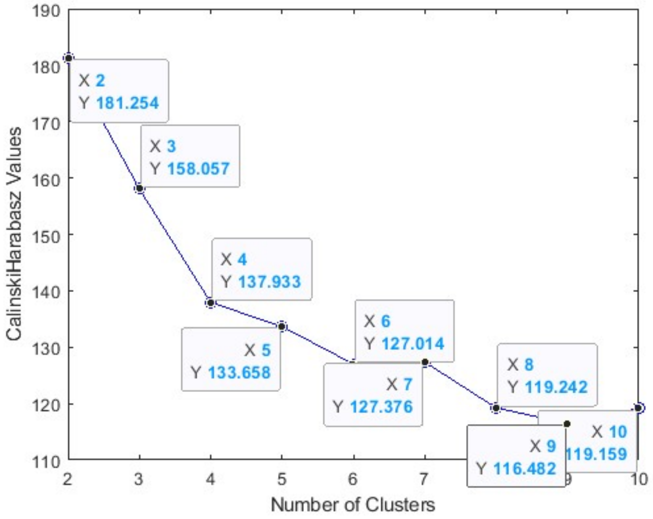
The results of the cluster analysis are divided into three subcategories according to individual subindexes, i.e. Basic, Efficiency and Innovation according to NUTS-2 regions.

Basic sub-index cluster analysis of EU countries

The Basic sub-index includes five pillars: the institutions, the macroeconomic stability, the infrastructures, the health and the basic education. That means we must analyse 234 vectors (number NUTS-2 regions) of dimension 5 (number of Basic sub-index pillars). First, we must find the best number of clusters for this analysis. Using MatLab function *evalclusters*, which creates a clustering evaluation object containing data used to evaluate the optimal number of data clusters, we find that the best number of clusters is 2 (see Figure 1). Higher value of Caliński-Harabasz index means the clusters are dense and well separated, although there is no "acceptable" cut-off value. We need to choose that solution which gives a peak or at least an abrupt elbow on the line plot of Caliński-Harabasz indices. The choice of only two clusters is also suitable considering the dendrogram (see Figure 2) where we can see intermixture of clusters. In the next step (using MatLab function *kmeans*), we sorted the data into these two clusters and found their centroids.

The centroids characterizing the values of institutions, macroeconomic stability, infrastructures, health and basic education are this four points $C_1 = [68.88, 81.26, 62.51, 87.93, 85.16]$ and $C_2 = [134.50, 120.62, 107.96, 105.71, 111.45]$. For better further visualization, we chose Nonnegative matrix factorization (NNMF) to display the 5-dimensional space in the plane (see Figure 3.). We see at least a separation of clusters, but a clear division into two groups with better and worse ratings is evident. This division is also visually obvious at map of EU (Figure 4), where we can see division of the NUTS-2 regions of EU with dividing Europe by diagonal running from south-west to north-east. Except for the region of Central Bohemia and Prague itself, the entire eastern block is part of the cluster together with Greece and Italy, part of the Iberian Peninsula.

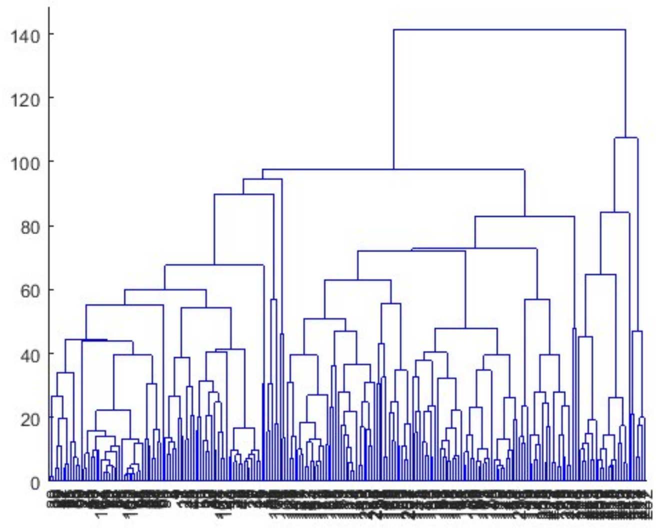
Figure 1: Line-plot of Caliński-Harabasz values vs number of clusters for the Basic sub-index dataset of NUTS-2 regions



1 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

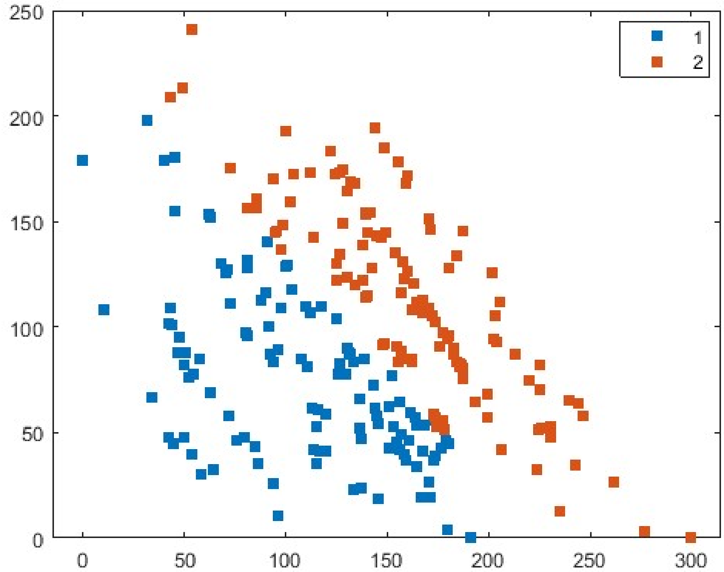
2

3 Figure 2: Dendrogram for the Basic sub-index dataset of NUTS-2 regions



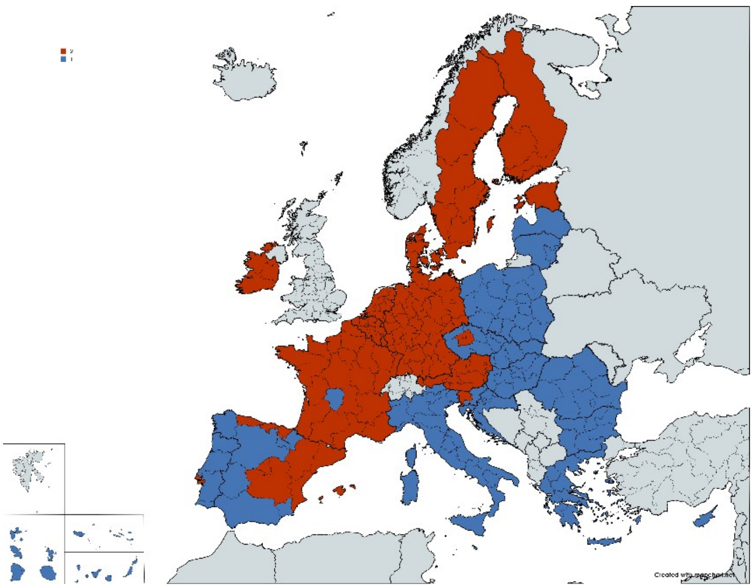
4 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

Figure 3: NNMF visualization for the Basic sub-index dataset of NUTS-2 regions



5 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

Figure 4: The NUTS-2 maps of the Basic sub-index dataset for two clusters

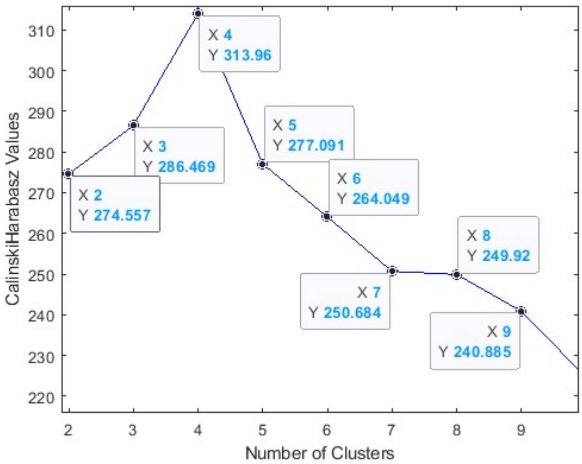


6 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

Efficiency sub-index cluster analysis of EU countries

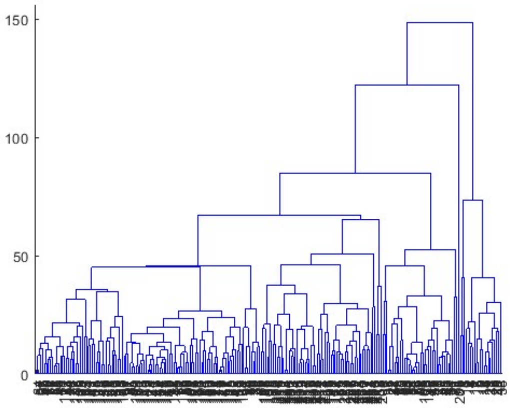
The Efficiency sub-index includes three pillars: higher education, training and lifelong learning, the labor market efficiency and the market size. So, we have 234 vectors of dimension 3. If we evaluate the number of clusters, we obtain the number of 4 clusters (see Figure 5 also Figure 6).

Figure 5: Line-plot of Caliński-Harabasz values vs number of clusters for the Efficiency sub-index dataset of NUTS-2 regions



7 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

Figure 6: Dendrogram for the Efficiency sub-index dataset of NUTS-2 regions

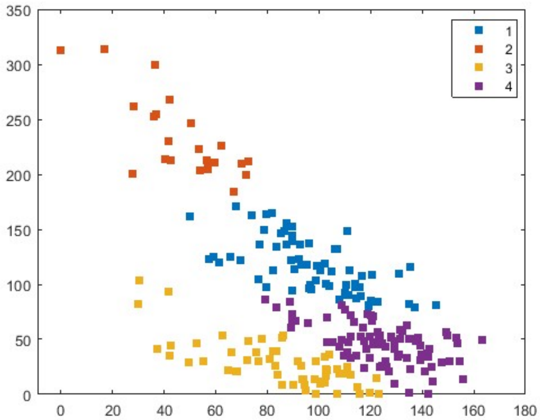


8 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

The centroids of these four clusters, which represent lifelong learning, the labor market and the market size, are $C_1 = [107.95, 111.89, 111.54]$, $C_2 = [114.03, 116.10, 207.33]$, $C_3 = [69.33, 70.17, 33.60]$ and

$C_4 = [103.99, 103.14, 54.23]$. If we transform the obtained clusters into two-dimensional vectors, the visualization can be seen on the NNMF visualization (Figure 7).

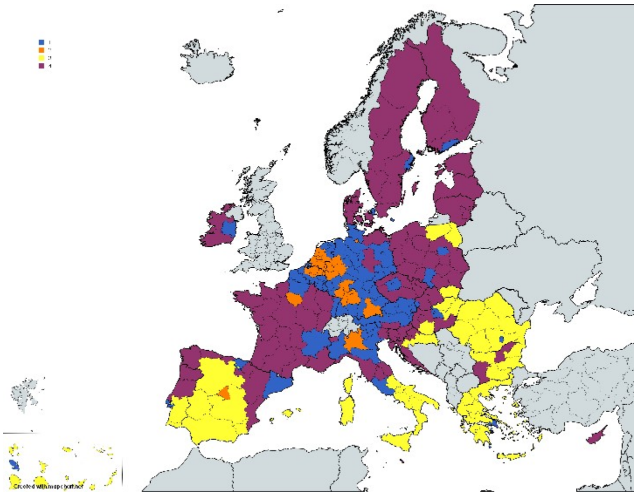
Figure 7: NNMF visualization for the Efficiency sub-index dataset of NUTS-2 regions



9 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)).*

After drawing the clusters on the geographical map, the categorization according to the regional development in the north-south and east-west direction is obvious. And there is also a strong elitism around the developed capital cities and their agglomerations, or cen-ters such as the Ruhr, northern Italy, or the BENELUX countries (Figure 8).

Figure 8: The NUTS-2 maps of the Efficiency sub-index dataset for four clusters

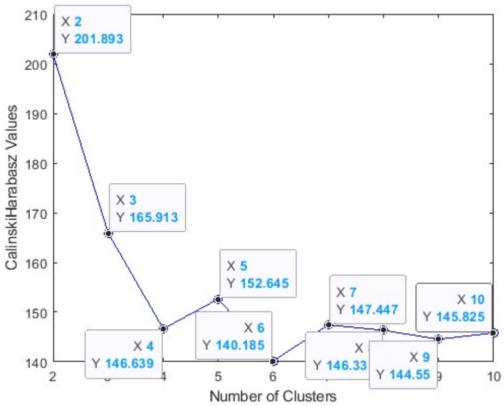


10 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

Innovation sub-index cluster analysis of EU countries

Innovation sub-index includes the three pillars that are the drivers of improvement at the most advanced stage of economic development: technological readiness, business sophistication and innovation. That means we have also 234 vectors of dimension 3. If we draw the dependence between the Caliński-Harabasz values and the number of clusters, we see that the highest value is for the basic division into two clusters, but considering the elbow rule, we can also choose the number of clusters 4 (see Figure 9).

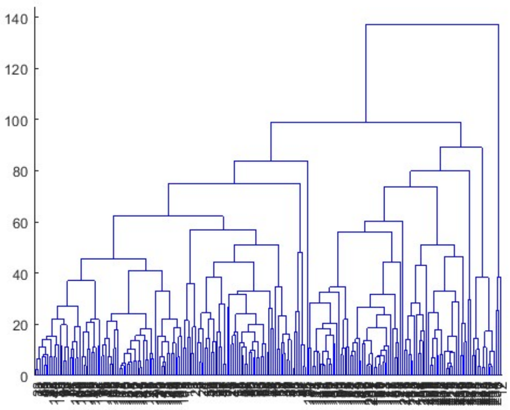
Figure 9: Line-plot of Caliński-Harabasz values vs number of clusters for the Innovation sub-index dataset of NUTS-2 regions



11 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)).*

From the dendrogram also can be also seen that the best way to analyze this dataset is for number of clusters 2 or 4 (see Figure 10).

Figure 10: Dendrogram for the Innovation sub-index dataset of NUTS-2 regions

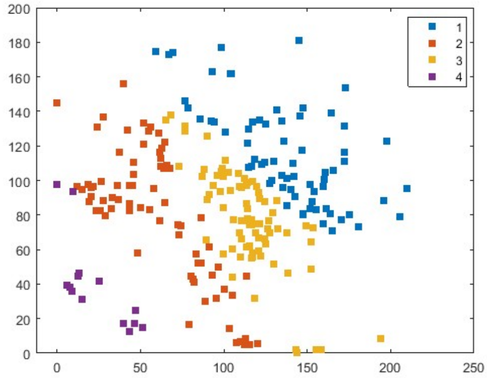


12 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

Using the MatLab function *kmeans* we get four clusters with their centroids (representing technological readiness, business sophistication and innovation) $C_1 = [130.11,\ 127.75, 131.84]$

$C_2 = [88.18, 66.31, 60.00]$, $C_3 = [93.65, 105.73, 97.14]$ and $C_4 = [35.89, 20.95, 39.03]$. Using the Nonnegative matrix factorization, we can visualize the resulted clustering in two-dimensional way (Figure 11). And now it is clear that the choice of four clusters is better way than only dividing the dataset into two clusters.

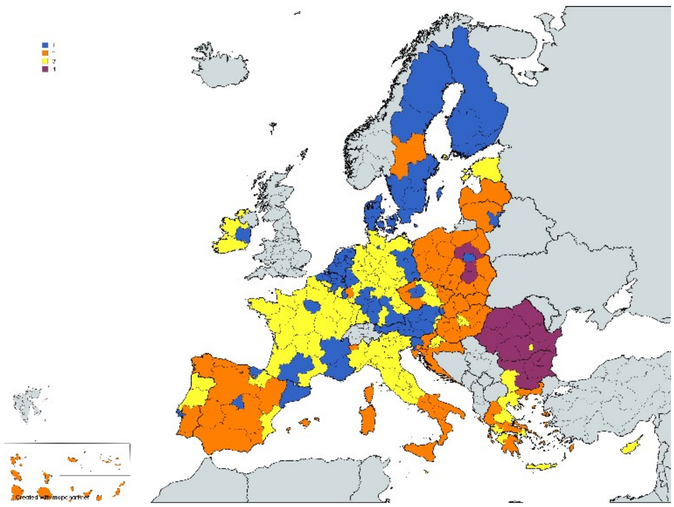
Figure 11: NNMF visualization for the Innovation sub-index dataset of NUTS-2 regions



13 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

The map of NUTS-2 region according to our obtained clustering indexation for Innovation pillars is very similar to the previous one for Efficiency pillars. It means there can been seen all the stereotypes about the rate of innovativeness, e.g. more innovativeness western Europe in contrast to the east Europe (Romania, Bulgaria and easter regions of Poland), typical centers of innovations as (BENELUX, Ruhr, the capitals with their surroundings). Also, the difference between the south and north of Europe can be seen.

Figure 12: The NUTS-2 maps of the Innovations sub-index dataset for four clusters



14 *Source: author's processing from data of the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO)*

4 DISCUSSION

In our article, we analyzed individual assessments of the level of competitiveness of EU regions (NUTS-2) according to three basic sub-indexes (Basic, Efficiency, Innovation). As a starting point, we use the EU Regional Competitiveness Index (RCI), whose individual components (evaluation indices) we evaluated using cluster analysis. We com-pared the obtained categorizations with the regional characteristics of the given regional territorial unit.

The cluster analysis was conducted within the MATLAB environment, utilizing the advanced **k-means++** algorithm. In addition to standard clustering techniques, a notable feature of this study is the application of Non-negative Matrix Factorization (NNMF) for data visualization. This method facilitates the transformation of multidimensional matrices into a two-dimensional space while preserving the structural relationships between entities, thereby significantly enhancing the clarity of the results.

CONCLUSION

This study successfully applied advanced cluster analysis to the EU Regional Competitiveness Index (RCI 2.0) to identify patterns of economic development across 234 NUTS-2 regions. By employing the k-means++ algorithm and validating results through the Caliński-Harabasz criterion, the research moved beyond simple rankings to reveal distinct regional groupings based on the three core sub-indices: Basic, Efficiency, and Innovation.

The analysis of the Basic sub-index revealed a fundamental geographical divide in Europe, separating more developed regions from an "eastern bloc" that includes Greece, Italy, and parts of the Iberian Peninsula. In contrast, the Efficiency and Innovation sub-indices highlighted a more complex four-cluster structure. These findings underscore a significant "elitism" surrounding capital cities and major industrial hubs like the Ruhr and BENELUX countries, which consistently outperform their peripheries.

Methodologically, the use of Non-negative Matrix Factorization (NNMF) proved to be a highly effective tool for visualizing multidimensional competitiveness data in a two-dimensional space while preserving essential Euclidean norms. Ultimately, these results confirm that regional competitiveness in the EU remains characterized by persistent north-south and east-west disparities, as well as a stark contrast between innovative urban centers and stagnant rural regions. These insights are critical for tailoring future Cohesion Policies to the specific developmental needs of each identified regional cluster.

ACKNOWLEDGMENTS

This work was financially supported by the Scientific Grant Agency (VEGA), which is a grant agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic [grant VEGA, Reg. No. 1/0448/24] “Study of key determinants of human capital and economic growth in the conditions of digital economy development”.

REFERENCES

Aiginger, K. (2006). Competitiveness: From a dangerous obsession to a welfare creating ability with positive externalities. Journal of Industry, Competition and Trade, 6(2), 161–177. https://doi.org/10.1007/s10842-006-9475-6

Annoni, P., & Dijkstra, L. (2010). EU Regional Competitiveness Index 2010. Publications Office of the European Union. https://doi.org/10.2776/17706

Annoni, P., & Dijkstra, L. (2019). The EU Regional Competitiveness Index 2019. European Commission, Directorate-General for Regional and Urban Policy.

Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99–120.

Bristow, G. (2010). Critical reflections on regional competitiveness: Theory, policy, practice. Routledge.

Charron, N., Dijkstra, L., & Lapuente, V. (2014). Regional governance matters: Quality of government within European Union Member

States. Regional Studies, 48(1), 68–90. https://doi.org/10.1080/00343404.2013.770141
Dijkstra, L., Annoni, P., & Kozovska, K. (2023). EU Regional Competitiveness Index 2.0 - 2022 edition (Working Paper 02/2023). European Commission.
Iammarino, S., Rodríguez-Pose, A., & Storper, M. (2019). Regional inequality in Europe: Evidence, theory and policy implications. Journal of Economic Geography, 19(2), 273–298. https://doi.org/10.1093/jeg/lby021
Krugman, P. (1994). Competitiveness: a dangerous obsession. Foreign Affairs, 73(2), 28–44.
Maza, A., & Hierro, M. (2024). Regional competitiveness in the EU: a spatial perspective. Papers in Regional Science, 103(1), 100005. https://doi.org/10.1016/j.pirs.2024.100005
Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what this means for the future of the EU). Cambridge Journal of Regions, Economy and Society, 11(1), 189–209. https://doi.org/10.1093/cjres/rsx024
Dijkstra, L., Papadimitriou, E., Martinez, B. C., Dominicis, L., & Kovacic, M. (2023). EU regional competitiveness index 2.0 (Working Paper WP 01/2023). Publications Office of the European Union. https://doi.org/10.2776/46106.

Contact:

Pavol Oršanský, Mgr. PhD.,
Department of Economics and Economy,
Faculty of Socio-Economic Relations,
Alexander Dubček University of Trenčín in Trenčín
Študentská 3
911 50, Trenčín
e-mail: pavol.orsansky@tnuni.sk

Lloyd, S. Least squares quantization in PCM, in IEEE Transactions on Information Theory, vol. 28, no. 2, pp. 129-137, March 1982, https://doi.org/10.1109/TIT.1982.1056489.
Arthur, D., Vassilvitskii S. (2007) K-means++: The Advantages of Careful Seeding. SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. 2007, pp. 1027–1035. https://doi.org/10.1145/1283383.1283494.
Kriegel, H. P., Schubert, E. & Zimek, A. The (black) art of runtime evaluation: Are we comparing algorithms or implementations? Knowl Inf Syst 52, 341–378 (2017). https://doi.org/10.1007/s10115-016-1004-2
Caliński, T., Harabasz, J. (1974). a dendrite method for cluster analysis. Communications in Statistics, 3(1), 1–27. https://doi.org/10.1080/03610927408827101.
Michael, W. B., Browne, B., Langville, A. N., Pauca, V. P., Plemmons, R. J. (2007). Algorithms and Applications for Approximate Nonnegative Matrix Factorization. Computational Statistics & Data Analysis 52, no. 1 (September 2007), pp.155–73. https://doi.org/10.1016/j.csda.2006.11.006.