IZMIR DEMOCRACY UNIVERSITY SOCIAL SCIENCES JOURNAL (IDUSoS)
Cilt/Volume:8 Sayı/Issue:2 Sayfa/Page:224-246

CLUSTERING OF POST-PANDEMIC UNEMPLOYMENT IN OECD COUNTRIES USING THE K-MEANS METHOD (2021–2023)

Öğr. Gör. Dr. Çetin GÖRÜR[1]

ABSTRACT

This study examines unemployment dynamics across OECD countries during 2021, 2022, and 2023 by applying the k-means clustering method. Annual unemployment rate data were drawn from the World Bank's World Development Indicators database and standardized before analysis. The number of clusters was identified using several diagnostic tools, including the silhouette score, the elbow method, and the gap statistic. Based on these results, countries were grouped according to similarities in unemployment rates. Thus allowing both cross-sectional and temporal comparisons

The analysis identifies three distinct clusters each year, which broadly mirror differences in economic development levels, labour market structures, and institutional capacities. Advanced economies such as Switzerland, Japan, Germany, and Norway consistently fall into the low-unemployment group, supported by diversified economic bases, strong vocational training systems, and effective active labour market policies. By contrast, Greece, Spain, and Türkiye remain in the high-unemployment group throughout the period, reflecting entrenched structural rigidities, skill mismatches, and a heavy reliance on tourism and other low value-added services. A subset of countries shifted between clusters over time, highlighting the role of post-pandemic recovery dynamics, inflationary pressures, and sector-specific growth patterns.

The findings indicate that unemployment disparities remain a persistent feature across OECD countries, influenced by both long-term structural factors and short-term economic fluctuations. By offering an updated post-pandemic perspective, this study contributes to the existing literature and provides insights relevant for policymaking. In countries with high unemployment, efforts should focus on diversifying economic sectors, enhancing vocational and technical education, and strengthening labour market institutions to facilitate job creation. Conversely, countries with consistently low unemployment should prioritize maintaining

[1] Van Yüzüncü Yıl Üniversitesi, Erciş Meslek Yüksekokulu, Muhasebe ve Vergi Bölümü, Muhasebe ve Vergi Uygulamaları Pr., e-posta: cetingorur@yyu.edu.tr, ORCID: 0000-0002-9556-5068

Makale Gönderim Tarihi (Manuscript Submit Date): 22.09.2025, Kabul Tarihi (Accepted Date): 09.10.2025, Basım Tarihi (Published Date): 31.12.2025

labour market resilience in the face of technological change and demographic shifts. Taken together, these results offer practical guidance for policymakers aiming to implement targeted, evidence-based strategies that address unemployment effectively within a rapidly evolving global economy.

Keywords: OECD, Post-Pandemic Unemployment, Clustering Analysis, Labour Market

Jel Codes: E24, E02, O11

OECD ÜLKELERİNDE PANDEMİ SONRASI İŞSİZLİĞİN K-MEANS YÖNTEMİYLE KÜMELENMESİ (2021–2023)

ÖZET

Bu çalışma, k-means kümeleme yöntemini uygulayarak 2021, 2022 ve 2023 yıllarında OECD ülkelerinde işsizlik dinamiklerini incelemektedir. Yıllık işsizlik oranı verileri, Dünya Bankası'nın Dünya Kalkınma Göstergeleri veritabanından alınmış ve analizden önce standartlaştırılmıştır. Kümelerin sayısı, siluet puanı, dirsek yöntemi ve boşluk istatistiği gibi çeşitli teşhis araçları kullanılarak belirlenmiştir. Bu sonuçlara dayanarak, ülkeler işsizlik oranlarındaki benzerliklere göre gruplandırılmış ve böylece hem kesitsel hem de zamansal karşılaştırmalar yapılabilmiştir.

Analiz, her yıl ekonomik gelişmişlik düzeyleri, işgücü piyasası yapıları ve kurumsal kapasitelerdeki farklılıkları genel olarak yansıtan üç farklı küme belirlemektedir. İsviçre, Japonya, Almanya ve Norveç gibi gelişmiş ekonomiler, çeşitlendirilmiş ekonomik temeller, güçlü mesleki eğitim sistemleri ve etkili aktif işgücü piyasası politikaları sayesinde sürekli olarak düşük işsizlik grubunda yer almaktadır. Buna karşılık, Yunanistan, İspanya ve Türkiye, köklü yapısal katılıklar, beceri uyumsuzlukları ve turizm ve diğer düşük katma değerli hizmetlere aşırı bağımlılık nedeniyle, dönem boyunca yüksek işsizlik grubunda kalmaktadır. Bazı ülkeler zaman içinde kümeler arasında geçiş yapmıştır, bu da pandemi sonrası toparlanma dinamiklerinin, enflasyonist baskıların ve sektöre özgü büyüme modellerinin rolünü vurgulamaktadır.

Bulgular, işsizlik eşitsizliklerinin, hem uzun vadeli yapısal faktörlerin hem de kısa vadeli ekonomik dalgalanmaların etkisiyle OECD ülkeleri genelinde kalıcı bir özellik olmaya devam ettiğini göstermektedir. Bu çalışma, güncellenmiş bir pandemi sonrası perspektif sunarak mevcut literatüre katkıda bulunmakta ve politika yapımına ilişkin önemli bilgiler sağlamaktadır. İşsizliğin yüksek olduğu ülkelerde, çabalar ekonomik sektörlerin çeşitlendirilmesi, mesleki ve teknik eğitimin geliştirilmesi ve istihdam yaratılmasını kolaylaştırmak için işgücü piyasası kurumlarının güçlendirilmesine odaklanmalıdır. Tersine, işsizliğin sürekli düşük olduğu ülkeler, teknolojik değişim ve demografik dönüşümler karşısında işgücü piyasasının dayanıklılığını korumaya öncelik vermelidir. Bu sonuçlar, hızla gelişen küresel ekonomide işsizliğin etkili biçimde ele alınabilmesi için politika yapıcılara kanıta dayalı stratejiler geliştirmede pratik bir rehber sunmaktadır.

Anahtar Kelimeler: OECD, Pandemi Sonrası İşsizlik, Kümelenme Analizi, İşgücü Piyasası

Jel Kod: E24, E02, O11

1. INTRODUCTION

Unemployment is not only a key indicator of macroeconomic performance but also a critical measure of social well-being, reflecting the extent to which an economy utilizes its labour resources effectively. High unemployment rates indicate underutilized labour potential, leading to a loss in output and a gap between actual and potential gross domestic product (GDP) (Blanchard & Johnson, 2013). Beyond its economic implications, unemployment carries substantial social costs, including increased poverty, income inequality, and social exclusion (OECD, 2023; Bell & Blanchflower, 2011).

The individual consequences of unemployment are equally severe. Joblessness can lead to long-term "scarring" effects, such as persistent wage penalties, diminished future job prospects, and deterioration in mental and physical health (Arulampalam, 2001; Paul & Moser, 2009). These effects tend to be more pronounced during economic downturns, where extended spells of unemployment exacerbate skill depreciation and hinder re-employment opportunities (Nichols et al., 2013).

Unemployment continues to be a central concern in economic policy because of its broad and multifaceted effects. Governments and international organizations closely monitor labour market developments, as changes in unemployment influence fiscal and monetary decisions, social welfare programs, and the implementation of active labour market measures (ILO, 2022; OECD, 2023). Analyses of unemployment patterns across countries or regions-often using clustering techniques-can shed light on structural differences, the effectiveness of policy interventions, and the resilience of labour markets to external shocks (Monfort et al., 2018; Yılmaz, 2022).

As a key macroeconomic indicator, the unemployment rate reflects both the state of the labour market and the wider trajectory of economic growth. Accordingly, researchers have examined unemployment from a variety of geographic, methodological, and socio-economic perspectives, frequently employing clustering methods to detect patterns and group regions or countries with similar characteristics.

At the regional level, cluster analysis has been widely applied to identify spatial disparities in unemployment and to guide policy decisions. For instance, Ardiansyah et al. (2024) studied open unemployment in West Java Province, Indonesia, during the COVID-19 pandemic using the k-means algorithm, with the silhouette method used to determine the most appropriate number of clusters.

At the national and cross-country level, other scholars have combined clustering with macroeconomic and labour market indicators. Ok (2022) conducted a fuzzy C-means cluster analysis on 18 unemployment-related indicators for 36 OECD and six non-OECD countries, comparing cluster structures across different numbers of groups using MATLAB and R, with validity indices applied to determine the optimal cluster count. Yılmaz (2022) used hierarchical and k-means clustering to group 35 OECD and EU countries according to seven labour force index variables, finding that Türkiye was placed in a cluster characterized by high workload and working hours flexibility but low autonomy and educational opportunities. Monfort et al. (2018) examined income inequality, redistribution, and unemployment convergence in the EU using cluster analysis, concluding that economic integration had not resulted in convergence to a single cluster, with traditional regional classifications remaining relevant.

The relationship between unemployment and other socio-economic factors has also been studied through clustering approaches. Aretz et al. (2024) analyzed GP density in Germany at the county level between 2015 and 2019 using Moran's I, LISA cluster analysis, and SLX models, finding that higher unemployment rates were associated with lower GP availability. Seppälä et al. (2024) clustered Finnish municipalities by socio-economic risk factors to test the inverse intervention law in child protection, identifying unemployment reduction as a driver in high-risk municipalities. Salimova et al. (2024) clustered regions in Russia's Volga Federal District by labour market and digital economy indicators, finding minimal impact of digitalization on employment levels. Zaharia (2024) examined Romanian counties, comparing 2012 and 2022 data, and found that educational infrastructure and teacher-student ratios had little effect on unemployment and school dropout rates, with economic and social development playing a greater role.

Other contributions have linked unemployment to sectoral and behavioral dynamics. Masserini et al. (2024) applied interrupted time series analysis to quarterly Italian data to assess the effects of the 2008 global financial crisis and the 2011 European sovereign debt crisis on unemployment trends, noting significant long-term increases and slow recovery. Moreover, Eygü (2018) examined the effect of inflation and foreign trade data on the unemployment rate using annual data from 1990 to 2017 in Türkiye through multiple regression analysis. The research findings indicate that there is a negative correlation between inflation and foreign trade data and the unemployment rate. Blasques et al. (2021) used a dynamic factor model and k-means clustering to study the relationship between unemployment and education participation in the Netherlands, revealing that part-time education enrollment correlated more strongly with unemployment than full-time programs. Engeloğlu and Yurdakul (2024) applied asymmetric causality tests followed by cluster analysis to consumer confidence data in 27 EU countries, identifying unemployment rate as one of the factors influencing consumer confidence and grouping countries with similar behavioral patterns.

The dynamics of unemployment during crisis periods have been analyzed in several studies. Mogos et al. (2024) examined unemployment in 27 EU countries during the COVID-19 pandemic using CRISP-DM methodology and EM and k-means clustering, tracking movements between clusters and suggesting coordinated strategies for countries with similar responses. Manea and Sorcaru (2021) discussed Romania's unemployment challenges during the COVID-19 and economic crises, highlighting structural mismatches between labour supply and demand and the impact of returning migrant workers. Poutanen et al. (2024) studied 643 Finnish participants in the Work Ability Programme, clustering them by main activity trajectories from 2005 to 2021, and identifying health and work ability disparities across groups, particularly for those transitioning from employment to unemployment. Also, Demir (2021) examined unemployment convergence among neighbouring Balkan countries from 1991 to 2020 using spatial and spatial panel econometric approaches. Moran's I and LM tests confirmed spatial dependence, and SAR, SEM, and random effects models were identified as appropriate. The results indicated significant spatial relationships, with unemployment levels in bordering countries influencing one another.

Finally, some studies have focused on specific labour market segments and demographic groups. Trentini (2024) investigated unemployment among older workers in 11 European countries between 2006 and 2019 using event history and sequence analysis, finding that low education increased the risk of unemployment and that Southern European countries had higher risks of long-term joblessness. Lietzmann and Hohmeyer (2024) examined non-standard employment (NSE) in Germany, clustering labour market trajectories of previously unemployed individuals from 2012 to 2015, and found that while NSE can offer employment opportunities, transitions to permanent jobs remain limited for many.

Accordingly, the primary objective of this study is to examine the changes in unemployment patterns across OECD countries over the period 2021–2023 using clustering analysis. By employing a set of relevant unemployment indicators, the research seeks to group countries with similar labour market characteristics and to track shifts in these groupings over time. This approach allows for the identification of structural changes in unemployment dynamics, the detection of emerging trends, and the assessment of convergence or divergence among member states in the post-pandemic economic environment. In doing so, the study contributes to the existing literature by offering a longitudinal and comparative perspective on unemployment clustering within the OECD, focusing on a recent and highly volatile period that has not yet been extensively analyzed. The findings are expected to provide both a deeper understanding of cross-country labour market developments and practical insights to inform evidence-based policy design aimed at enhancing employment outcomes within the OECD.

2. MATERIALS AND METHODS

This study analyses unemployment patterns in OECD countries for 2021-2023 using k-means clustering. Annual unemployment rates were obtained from the World Bank and standardised before analysis. The optimal number of clusters was identified using established diagnostic measures, and all computations were conducted in R to ensure reproducibility. All OECD member countries were included in the analysis in the study.

2.1. Data Sources and Variables

The present study utilizes annual national unemployment rates-as percentage values-from the World Bank's World Development Indicators database (World Bank, 2024). Specifically, unemployment rate data for OECD member countries covering 2021, 2022, and 2023 were downloaded directly from the World Bank's online portal. This approach ensures standardized, internationally comparable data across countries and time. The unemployment rate variable is defined as the share of the labour force that is without work but actively seeking employment and available to start work, which is the standard indicator used in cross-country macroeconomic research (World Bank, 2024).

2.2. Data Preparation

Once extracted, unemployment rates for each year were organized into a panel dataset with rows corresponding to countries and columns to years (2021, 2022, 2023). Prior to clustering, the data were inspected for missing values and outliers. Any missing observations were verified against the source database and confirmed to be unavailable, and countries with incomplete data for any of the three years were excluded from the final sample to ensure consistent clustering across years. No additional imputation was applied to preserve data integrity.

2.3. Clustering Method: K-Means Algorithm

To uncover patterns of similarity in unemployment dynamics, the study applies the k-means clustering algorithm, executed separately for each year as well as in a pooled three-year structure where appropriate (MacQueen, 1967). The k-means method partitions the set of n observations (countries) into k clusters $C = \{C_1, C_2, \dots, C_k\}$ with the goal of minimizing within-cluster sum of squared distances. The objective function minimized is:

$$\min_{C} \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2 \qquad (1)$$

here, x represents the unemployment rate of a given country for the relevant year(s), and μ_i denotes the centroid (mean value) of cluster C_i. The algorithm iteratively assigns each country to the nearest centroid and updates centroids until cluster memberships stabilize (Hartigan & Wong, 1979; İşleyen, 2021; Kanberoğlu et al., 2021).

Prior to clustering, unemployment values were standardized (transformed to z-scores) to ensure that clusters reflect relative deviations from mean performance rather than countries with high absolute rates dominating the grouping. Standardization is given by:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \qquad (2)$$

where x_{ij} is the unemployment rate for country i in year j, $\bar{x}_j$ is the mean across countries for that year, and s_j is the corresponding standard deviation (İşleyen, 2021; Kanberoğlu et al., 2021).

To determine the optimal number of clusters, three widely used diagnostic methods were applied: the silhouette score, the elbow method (based on within-cluster sum of squares), and the gap statistic (Tibshirani et al., 2001). The silhouette score evaluates both the cohesion within clusters and the separation between them, with higher average values indicating clearer and more distinct clusters. The elbow method assesses the reduction in within-cluster variance as the number of clusters (k) increases, identifying the point where adding additional clusters provides diminishing returns. The gap statistic, on the other hand, compares the observed clustering structure against a reference distribution, testing whether the identified clusters are meaningfully more structured than what might occur by chance. These diagnostics were calculated for k values from 1 to 10 to inform the selection of the most appropriate number of clusters.

3. ANALYSIS AND FINDINGS

This section reports the results of the clustering analysis of unemployment patterns across OECD countries for the years 2021, 2022, and 2023. To establish the optimal number of clusters, several diagnostic tools were employed, including silhouette scores, the elbow method, and the gap statistic. Based on these criteria, the analysis explored both two- and three-cluster solutions, which facilitated the identification of structural similarities among countries, the persistence of high- and low-unemployment groups, and transitional movements between clusters over time. Comparative evaluations across years provide insights into the stability or volatility of labour market conditions within the OECD, as well as the convergence or divergence trends among member states in the post-pandemic period. Next, Figure 1 presents the silhouette-based diagnostic used to determine the optimal number of clusters.

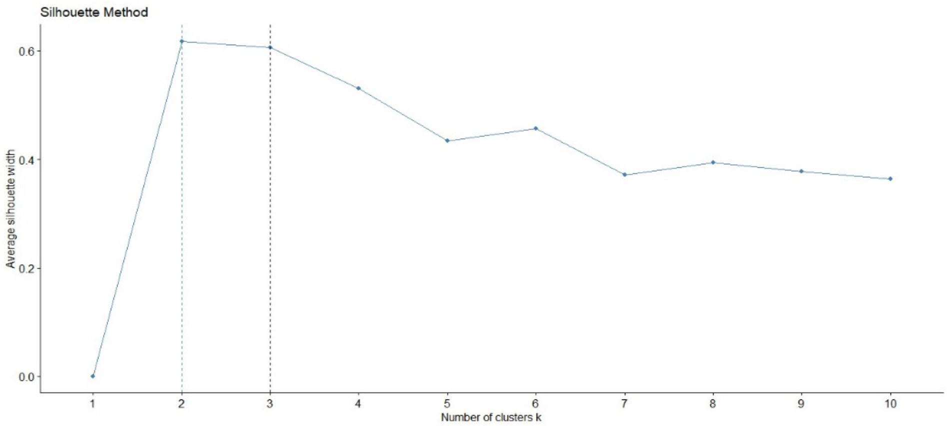

Figure 1: **Determining The Optimal Number of Clusters Using The Silhouette Method**

Figure 1 illustrates the silhouette analysis that you ran to evaluate clustering quality across candidate k values; the accompanying summary scores shown in Table 1 indicate that the silhouette score is highest at two clusters (0.617) and only marginally lower at three clusters (0.606), with a steady fall for larger k values. This pattern indicates that a two-cluster partition captures the most distinct separation in your unemployment indicators, while a three-cluster solution remains viable and produces only slightly less cohesion. In practical terms, the silhouette results suggest that the country observations form one relatively compact group and one other more separated group (or, for k=3, a compact low-group, a moderate group and a clearly higher-unemployment group), so both two- and three-cluster representations are defensible starting points for interpretation. The relatively rapid decline in silhouette values beyond k = 3 implies diminishing returns from further subdivision of the sample into many small clusters. Next, Figure 2 shows the elbow-method diagnostic for choosing k.

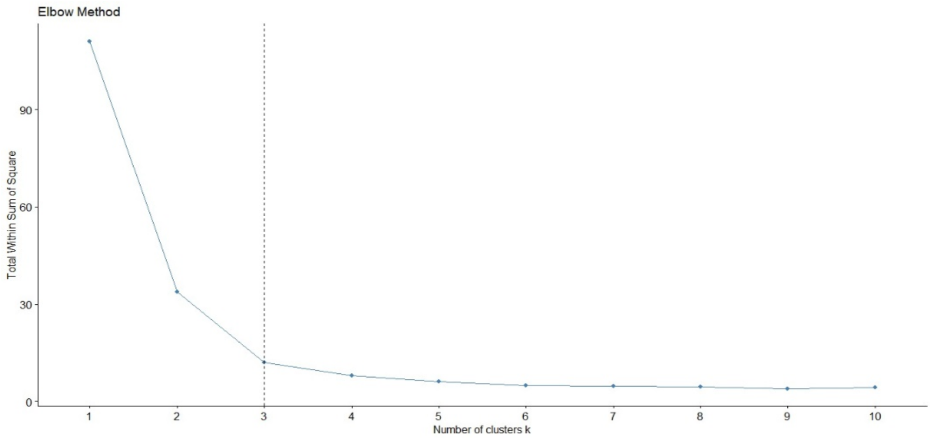

Figure 2: Determining The Optimal Number of Clusters Using The Elbow Method

Figure 2 (elbow plot) complements the silhouette evidence by showing how within-cluster dispersion changes with k; although the figure itself is qualitative, its function is to identify the point at which additional clusters produce only marginal gains in explained heterogeneity. Taken together with the silhouette results reported in Table 1, the elbow criterion appears to support a parsimonious solution (two or three clusters) because the reduction in within-cluster variance flattens after those values. In short, the elbow plot reinforces the conclusion that modeling the OECD sample with a very small number of clusters is appropriate: additional clusters beyond two or three would fragment groups without delivering substantially clearer separation. Next, Figure 3 reports the gap-statistic diagnostic for optimal k.

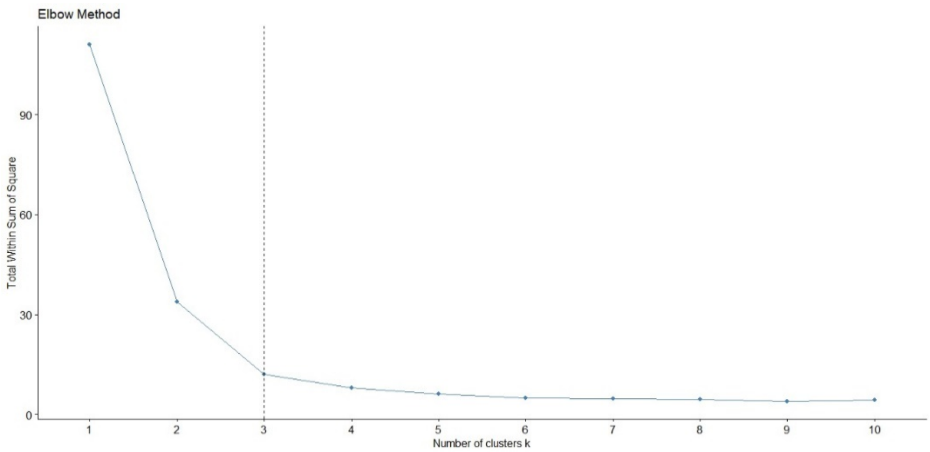

Figure 3: Determining The Optimal Number of Clusters Using The Gap Statistical Method

Figure 3 provides a gap-statistic assessment that-like silhouette and elbow-serves to identify the k that maximizes the difference between observed clustering and random reference distributions. While the gap plot is a separate criterion, its inclusion alongside silhouette and elbow gives the researcher multiple, confirming perspectives; in your case the combined diagnostics point toward two or three clusters as the most interpretable solutions. Using multiple diagnostics in this way strengthens confidence that observed groupings are not artefacts of a single index but reflect persistent structure in the unemployment indicators. Next, Table 1 tabulates silhouette (appropriateness) values for cluster numbers 1–10.

Table 1: Appropriateness Values of Cluster Numbers According to Silhouette Method Results

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10
0.000	0.617	0.606	0.530	0.434	0.457	0.371	0.393	0.377	0.363

Table 1 quantifies the silhouette results: k = 1 unsurprisingly yields 0.000, while k = 2 gives the highest score (0.617) and k = 3 is very close (0.606); thereafter silhouette values decline (k = 4: 0.530; k = 5: 0.434; and lower for larger k). This numeric pattern is informative for two reasons. First, a silhouette value above 0.5 for k = 2 and k = 3 indicates reasonably strong clustering structure for those solutions. Second, the relative proximity of the k = 2 and k = 3 scores suggests that both parsimonious options capture meaningful heterogeneity: k = 2 provides the strongest single split, while k = 3 subdivides that structure into an interpretable three-tier configuration (low, medium, high unemployment). Therefore, selection between two and three clusters should be guided by the substantive interpretability required by your research question: two clusters highlight a clear dichotomy in the OECD sample, while three clusters allow a middle group to be distinguished. Next, Table 2 reports average unemployment rates by year for the two-cluster solution.

Table 2: Average Unemployment Rates by Year for Two Groups

Clusters	2021	2022	2023
Cluster 1	5.661	9.854	3.912
Cluster 2	14.088	4.272	8.246

Table 2 highlights a notable temporal pattern in the two-cluster averages. In 2021, one cluster exhibits a high mean unemployment rate (14.088), while the other is comparatively low (5.661). In 2022, these average levels reverse (cluster means: 9.854 and 4.272), and in 2023, the original low/high relationship reappears (3.912 and 8.246). Two methodological considerations emerge from these figures. First, cluster labels (Cluster 1 vs. Cluster 2) are arbitrary and may change across years; hence, interpretation should focus on the composition of each group rather than the numerical label itself. Second, the year-to-year movement of average values indicates that country membership shifted substantially across years (or that the group labelled "Cluster 1" in 2022 is not the same group labelled "Cluster 1" in 2021). Interpreted substantively, the pattern suggests a marked reallocation of countries between the groups from 2021 to 2022 and further adjustments into 2023, which in turn reflect evolving labour-market conditions across OECD members during the pandemic recovery and its aftermath. These average changes require looking at membership tables to see which countries moved and to propose reasons for those moves. Next, Table 3 lists country membership for the two clusters in 2021.

Table 3: Distribution of OECD Countries for Two Clusters in 2021

Cluster 1	Cluster 2
United States, Canada, France, Australia, United Kingdom, Belgium, Germany, Italy, Netherlands, Sweden, Portugal, Norway, Switzerland, Iceland, Denmark, Ireland, Luxembourg, Japan, Korea Rep., Finland, Israel, Lithuania, Latvia, Austria, Estonia, New Zealand, Czechia, Hungary, Slovenia, Poland, Chile, Slovak Republic, Mexico.	Türkiye, Spain, Greece, Colombia, Costa Rica.

Table 3 shows that the lower-unemployment group in 2021 includes a broad set of advanced economies-United States, Canada, several Northern and Central European states, Japan, Korea, Israel, and a number of small advanced economies-while the high-average cluster in 2021 contains Türkiye, Spain, Greece, Colombia and Costa Rica. Given the 2021 averages in Table 2 (cluster with Türkiye et al. = 14.088), the 2021 partition therefore separates an overall low-unemployment core of mostly advanced industrialized economies from a smaller set of countries with substantially higher unemployment. The composition suggests that, at least in 2021, the clustering reflected a contrast between a large group of relatively low-unemployment OECD members and a small group that experienced particularly elevated unemployment rates. Possible explanations consistent with these membership lists (and rooted in the clustering outcome) include country-level exposures to specific pandemic-era shocks, structural labour-market rigidities, or shorter-run cyclical disturbances concentrated in the countries that form the high-average cluster in 2021; however, the precise causal mechanisms require country-level diagnostics beyond the clustering output. Next, Table 4 lists country membership for the two clusters in 2022.

Table 4: Distribution of OECD Countries for Two Clusters in 2022

Cluster 1	Cluster 2
Türkiye, France, Italy, Spain, Sweden, Greece, Colombia, Costa Rica, Chile.	United States, Canada, Australia, United Kingdom, Belgium, Germany, Netherlands, Portugal, Norway, Switzerland, Iceland, Denmark, Ireland, Luxembourg,

	Japan, Korea Rep., Finland, Israel, Lithuania, Latvia, Austria, Estonia, New Zealand, Czechia, Hungary, Slovenia, Poland, Slovak Republic, Mexico.

Table 4 reveals a substantial reconfiguration: in 2022 Türkiye, France, Italy, Spain, Sweden, Greece, Colombia, Costa Rica and Chile form one cluster, while the large set of advanced economies (United States, Canada, Australia, UK, Germany, Netherlands, Nordic countries, Japan, Korea, Israel, and many Central and Eastern European members) comprise the other. Compared with 2021, several countries that were previously grouped with the low-unemployment core (notably France, Italy and Sweden) now sit in the higher-average cluster for 2022 (which Table 2 shows had mean ~9.85). This shift indicates that 2022 saw an erosion of the earlier dichotomy: some countries experienced rising unemployment (or relative deterioration versus peers), moving them into the higher-average category. The pattern is consistent with heterogenous post-pandemic adjustment paths across OECD members-some countries quickly returned to low unemployment while others lagged. Again, the cluster labels are arbitrary; what matters is that a group containing Türkiye and several Southern European and Latin American countries registers a higher mean in 2022. Next, Table 5 lists country membership for the two clusters in 2023.

Table 5: Distribution of OECD Countries for Two Clusters in 2023

Cluster 1	Cluster 2
United States, Canada, Australia, United Kingdom, Belgium, Germany, Netherlands, Norway, Switzerland, Iceland, Denmark, Ireland, Luxembourg, Japan, Korea Rep., Israel, Austria, New Zealand, Czechia, Hungary, Slovenia, Poland, Slovak Republic, Mexico.	Türkiye, France, Italy, Spain, Sweden, Portugal, Greece, Finland, Lithuania, Colombia, Latvia, Estonia, Costa Rica, Chile.

Table 5 shows that in 2023 the high-average group includes Türkiye, France, Italy, Spain, Sweden, Portugal, Greece, Finland, Lithuania, Colombia, Latvia, Estonia, Costa Rica and Chile, and the other cluster contains the remaining advanced-economy group. Comparing 2023 to 2022, the high-unemployment cluster has broadened to include additional Northern and Baltic countries (Finland, Lithuania, Latvia, Estonia) in addition to the Southern European and Latin American economies that were already in the higher group. The 2023 two-cluster averages (Table 2) indicate the high cluster's mean is 8.246 while the other group's mean is 3.912, so the inter-group gap remains sizable. This evolution suggests that by 2023 several countries that had been in the low-unemployment core either experienced relative deterioration or were captured by a broader mid-to-high group; such movement implies either country-specific setbacks in labour markets or a relative improvement among the low cluster that widened the gap. The key substantive point is stability for a core set of low-unemployment OECD members versus persistent or resurfacing pressures in a heterogeneous group of Southern European, Baltic and some Latin American countries. Next, Figure 4 visualizes year-to-year group transitions for the two-cluster solution.

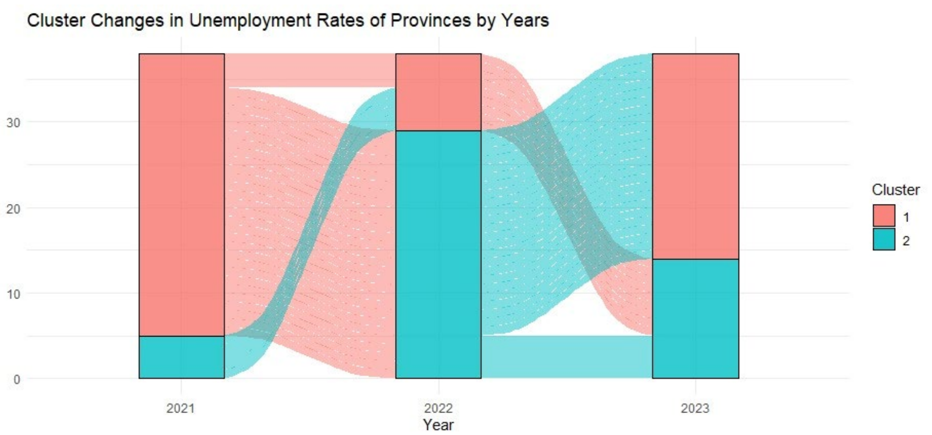

Figure 4: Unemployment Rates by Year in OECD Countries Using a Two-Group Classification by Year (2021–2023)

Figure 4 (the transition plot) graphically summarizes the flows noted above: a substantial core of advanced economies remains stably in the low-unemployment group across 2021–2023, while another set of countries-anchored by Türkiye, Spain and Greece-either remain persistently on the higher side or move in and out of the high-average group depending on the year. The figure helps to identify two types of dynamics: (1) persistent high-unemployment membership (countries that remain in the high group across all three years), and (2) transient movement (countries that switch clusters between years). Identifying which countries follow which pattern is important: persistent members of the high cluster indicate structural or long-run vulnerabilities in labour markets, whereas transients more likely reflect cyclical shocks, measurement timing, or short-term policy differences. In practice, the transition plot corroborates the table-based finding that a two-group representation captures both a stable low-unemployment core and a more variable high-unemployment set. Next, Table 6 reports average unemployment rates by year for the three-cluster solution.

Table 6: Average Unemployment Rates by Year for Three Groups

Clusters	2021	2022	2023
Cluster 1	7.325	3.610	3.421
Cluster 2	4.274	6.589	6.303

Cluster 3	14.088	11.536	9.921

Table 6 organizes means into three tiers: a moderate/upper tier (Cluster 1: 7.325 in 2021 → 3.610 in 2022 → 3.421 in 2023), a lower tier (Cluster 2: 4.274 → 6.589 → 6.303), and a clearly high tier (Cluster 3: 14.088 → 11.536 → 9.921). Interpreting these numbers, Cluster 3 is consistently the high-unemployment group across all three years but shows a declining average from 2021 to 2023 (from about 14.1 to about 9.9), which suggests partial improvement among the worst-performing countries during the sample window. Cluster 1's average falls markedly between 2021 and 2022 and remains low in 2023, while Cluster 2 moves from a low average in 2021 to a higher mid range in 2022–2023. The upshot is that the three-cluster partition reveals a more nuanced picture than the binary split: it separates the OECD sample into low, intermediate and high unemployment regimes, each exhibiting distinct temporal dynamics. Next, Table 7 lists country membership across three clusters in 2021.

Table 7: Distribution of OECD Countries Across Three Clusters in 2021

Cluster 1	Cluster 2	Cluster 3
Canada, France, Belgium, Italy, Sweden, Portugal, Iceland, Ireland, Finland, Lithuania, Latvia, Austria, Estonia, Chile, Slovak Republic.	United States, Australia, United Kingdom, Germany, Netherlands, Norway, Switzerland, Denmark, Luxembourg, Japan, Korea Rep., Israel, New Zealand, Czechia, Hungary, Slovenia, Poland, Mexico.	Türkiye, Spain, Greece, Colombia, Costa Rica.

Table 7 shows that in 2021 the highest-unemployment cluster (Cluster 3) again contains Türkiye, Spain, Greece, Colombia and Costa Rica-exactly the set that drove the high mean under the two-cluster solution-while Cluster 1 comprises a group of countries with intermediate means (Canada, France, Belgium, Italy, Sweden, Portugal, Iceland, Ireland, Finland, Lithuania, Latvia, Austria, Estonia, Chile, Slovak Republic) and Cluster 2 contains the low-unemployment core (United States, Australia, United Kingdom, Germany, Netherlands, Norway, Switzerland, Denmark, Luxembourg, Japan, Korea, Israel, New Zealand, Czechia, Hungary, Slovenia, Poland, Mexico). This three-fold partition confirms that the 2021 data are most naturally described as a tripartite separation: (a) a clearly low-unemployment block of advanced economies, (b) an intermediate block mostly composed of some Western and Southern European countries plus a few others, and (c) a small high-unemployment block including Türkiye and several Southern European/Latin American members. The three-cluster view therefore improves interpretability by isolating an intermediate set that the two-cluster dichotomy had grouped with the low-unemployment majority. Next, Table 8 lists country membership across three clusters in 2022.

Table 8: Distribution of OECD Countries Across Three Clusters in 2022

Cluster 1	Cluster 2	Cluster 3
United States, Australia, United Kingdom, Germany, Netherlands, Norway, Switzerland, Iceland, Denmark, Ireland, Luxembourg, Japan, Korea Rep., Israel, Austria, New Zealand, Czechia, Hungary, Slovenia, Poland, Mexico.	Canada, France, Belgium, Italy, Sweden, Portugal, Finland, Lithuania, Latvia, Estonia, Chile, Slovak Republic.	Türkiye, Spain, Greece, Colombia, Costa Rica,

Table 8 indicates a reallocation in 2022: the low-unemployment core now includes many of the same advanced economies (United States, Australia, UK, Germany, Netherlands, Nordic countries, Japan, Korea, Israel, Austria, New Zealand, Czechia, Hungary, Slovenia, Poland, Mexico), while an intermediate cluster collects Canada, France, Belgium, Italy, Sweden, Portugal, Finland, Baltic states, Chile and the Slovak Republic; and the high cluster continues to contain Türkiye, Spain, Greece, Colombia and Costa Rica. Compared with 2021, the intermediate cluster expanded and the core/low cluster remains large; the persistence of Türkiye, Spain and Greece in the high cluster across both 2021 and 2022 suggests enduring pressures in those countries' labour markets over this period, while the transfer of some countries into the intermediate group reflects relative deterioration versus the stable low-unemployment core. These membership movements illustrate that the three-cluster solution tracks not only absolute unemployment levels but also relative changes in how countries compare to their peers. Next, Table 9 lists country membership across three clusters in 2023.

Table 9: Distribution of OECD Countries Across Three Clusters in 2023

Cluster 1	Cluster 2	Cluster 3
United States,Australia, United Kingdom, Germany, Netherlands, Norway, Switzerland, Iceland, Ireland, Japan, Korea Rep., Israel, New Zealand, Czechia, Hungary, Slovenia, Poland, Mexico.	Canada, France, Belgium, Italy, Sweden, Portugal, Denmark, Luxembourg, Finland, Lithuania, Latvia, Austria, Estonia, Slovak Republic.	Türkiye, Spain, Greece, Colombia, Costa Rica, Chile.

Table 9 shows that the 2023 pattern is similar to 2022 with two important adjustments: Chile moves into the high cluster (joining Türkiye, Spain, Greece, Colombia, Costa Rica), and some countries that had been in the intermediate cluster remain there (Canada, France, Belgium, Italy, Sweden, Portugal, Denmark, Luxembourg, Finland, Baltic countries, Austria, Slovak Republic). The high cluster's average fell from 14.09 in 2021 to 9.92 in 2023 (Table 6), which indicates some improvement among the most affected countries, but their averages remain considerably higher than those of the other clusters. The presence of Chile in the 2023 high cluster suggests country-specific deterioration or stronger exposure to the factors that characterize the high group. Overall, the three-cluster narrative confirms a persistent high-unemployment group (largely Southern European and some Latin American members), an intermediate group with mixed outcomes, and a resilient low-unemployment core. Next, Figure 5 visualizes cluster transitions across the three-cluster solution for 2021–2023.

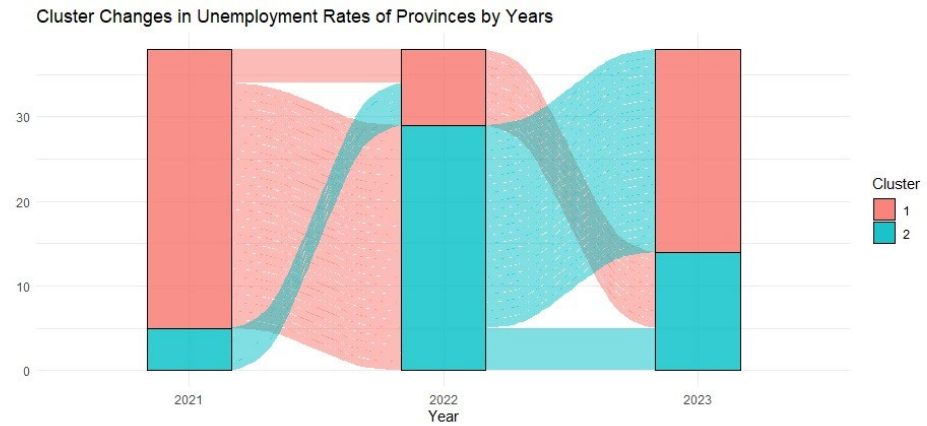

Figure 5: Unemployment Rates by year in OECD Countries Using a Three-Group Classification by Year (2021–2023)

Figure 5's transition visualization emphasizes the qualitative conclusions drawn from the tables: a stable low-unemployment core of advanced economies remains largely unchanged across years, while a smaller high-unemployment cluster persists (though with some shrinkage in mean values over time). The intermediate cluster acts as a buffer for countries that move between core and high groupings-France and Italy, for instance, are in the intermediate group rather than the low core across portions of the sample. The figure therefore underscores two types of heterogeneity: stable structural differences (countries that occupy the same tier in all years) and transitional movements (countries that move from intermediate to high or from low to intermediate). From a policy perspective, this visual message points to the need for targeted diagnostics: persistent members of the high cluster likely require structural or reform-oriented interventions, whereas countries that transition in and out may benefit from cyclical stabilization and short-term labour policies.

Across both the two- and three-cluster solutions, a consistent empirical pattern emerges. A core group of predominantly North American, Northern and Central European, and other advanced economies forms a stable low-unemployment cluster, while a smaller set-anchored by Türkiye and several Southern European and Latin American OECD members-constitutes a persistently higher-unemployment cluster. The primary distinction between the two- and three-cluster solutions lies in whether the intermediate group is modeled explicitly (k = 3) or absorbed into the low-unemployment cluster (k = 2). Year-by-year comparisons indicate a modest improvement in the highest-unemployment cluster's mean from 2021 to 2023, although substantial gaps between clusters persist.

The clustering results also provide clear guidance for subsequent analysis. Cluster membership identifies groups of countries with similar unemployment outcomes, and explaining why certain countries remain in the high cluster-or shift between clusters-requires examination of country-level factors, such as youth unemployment rates, sectoral employment shifts, short-term policy measures, recovery timing, or data measurement periods. The clustering thus offers a concise empirical typology that can guide these follow-up investigations: persistent membership in the high-unemployment cluster signals candidates for structural analysis, whereas transitional membership highlights sensitivity to short-term shocks or policy interventions.

DISCUSSION AND CONCLUSION

The clustering analysis of OECD countries based on unemployment rates for 2021, 2022, and 2023 reveals distinct groupings that broadly reflect differences in economic development levels and labour market structures. Developed economies, including Switzerland, Japan, and Germany, consistently formed clusters characterised by low unemployment rates, indicating resilient labour markets, diversified industrial bases, and effective employment policies. In contrast, countries with relatively higher unemployment rates-such as Greece, Spain, and Türkiye-clustered together, reflecting persistent structural challenges, sectoral imbalances, and weaker labour absorption capacities. Notably, several countries shifted between clusters over the three years, suggesting that macroeconomic shocks, post-pandemic recovery dynamics, and inflationary pressures have influenced their labour market positions.

The results of this study are broadly in line with previous research showing that labour market outcomes differ considerably across OECD countries, shaped by institutions, education systems, and the structure of production (Bell & Blanchflower, 2011; Monfort et al., 2018; Yılmaz, 2022). Similar to the arguments made by Monfort and colleagues (2018), the level of economic development emerges as a central factor in explaining cluster membership. Countries with higher income levels tend to have labour markets that are more resilient and less exposed to sudden fluctuations. On the other hand, the persistence of high-unemployment clusters confirms the findings of Yılmaz (2022), pointing to the difficulty that peripheral economies face in addressing structural unemployment, even when short-term recovery takes place.

What distinguishes this analysis from much of the earlier literature is its focus on the immediate aftermath of the COVID-19 pandemic. This was a period characterized by unstable growth, interruptions in global supply chains, and inflationary pressures. The variation between clusters can be understood through several factors. Economies with diversified industrial bases, a strong manufacturing sector, or a dynamic high-technology industry are better positioned to sustain employment and to soften the effects of economic shocks. Moreover, established labour market institutions-such as active labour market policies, collective bargaining mechanisms, and unemployment benefits-help to support employment stability. Robust systems of general and vocational education also play an important role by increasing workers' ability to adjust to technological change and re-enter employment more quickly.

In contrast, countries grouped in the high-unemployment cluster often share certain structural weaknesses. These include relatively low labour force participation, persistently high youth unemployment, regional imbalances, and a dependence on sectors that are particularly vulnerable to global shocks, such as tourism and agriculture. For these countries, policy measures should be directed at encouraging greater economic diversification, expanding and modernizing vocational and technical education, and ensuring that education is more closely aligned with the actual needs of the labour market. Complementary policies-such as targeted job-placement programs, wage subsidies for disadvantaged groups, or support for entrepreneurship-may also contribute to employment creation. In addition, promoting both regional and sectoral mobility of labour could reduce persistent inequalities.

For countries with low unemployment, the central challenge is different: maintaining labour market resilience in the face of technological advances and demographic shifts. Here, long-term investment in lifelong learning, skills development, and inclusive labour market policies becomes crucial.

In conclusion, narrowing unemployment disparities within the OECD requires a mix of macroeconomic stability, effective institutional frameworks, and carefully designed structural reforms. It is the combination of these elements, rather than reliance on a single policy, that appears most effective in addressing the complex nature of unemployment differences across countries.

Ethics committee approval

This article does not require ethics committee approval.

Conflict of interest statement

This article has no conflicts of interest with any individual or institution.

REFERENCES

Ardiansyah, M.F.H., Amany, N., Anugrah, C.I., & Syafitri, U.D. (2024). K-Means Clustering Application of Open Unemployment in 2020 Caused by COVID-19 in West Java Province. *ENTHUSIASTIC, International Journal of Applied Statistics and Data Science*, *4*(1), 1- 12.

Aretz, B., Frey, S., & Weltermann, B. (2024). Regional socioeconomic characteristics and density of general practitioners in Germany: A nationwide cross-sectional and longitudinal spatial analysis. *Public Health*, *236*, 338–346. https://doi.org/10.1016/j.puhe.2024.09.010

Arulampalam, W. (2001). Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal*, *111*(475), F585–F606. https://doi.org/10.1111/1468-0297.00664

Bell, D. N. F., & Blanchflower, D. G. (2011). Young people and the Great Recession. *Oxford Review of Economic Policy*, *27*(2), 241–267.

Blanchard, O., & Johnson, D. R. (2013). *Macroeconomics* (6th ed.). Pearson.

Blasques, F., Hoogerkamp, M.H., Koopman, S.J., & Werve, I. van de. (2021). Dynamic factor models with clustered loadings: Forecasting education flows using unemployment data. *International Journal of Forecasting*, *37*, 1426–1441. https://doi.org/10.1016/j.ijforecast.2021.01.026

Demir, Y. (2021). Balkan Ülkelerine Ait İşsizliğin Mekânsal Panel Ekonometri Yaklaşımı İle Analizi. *Balkan Sosyal Bilimler Dergisi, 10*(19) 26–35.

Engeloğlu, Ö., & Yurdakul, F. (2025). The Factors Causing Consumer Behavior: Asymmetric Causality And Cluster Analysis For EU Countries Through Consumer Confidence Index. *EGE AKADEMİK BAKIŞ/EGE ACADEMIC REVIEW*, *25*(1), 21-42. https://doi.org/10.21121/eab.20250102

Eygü, H. (2018). Enflasyon, İşsizlik Ve Dış Ticaret Arasındaki İlişkinin İncelenmesi: Türkiye Örneği (1990-2017). *Kastamonu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 20(2), 96-112. https://doi.org/iibfdkastamonu.408823

Hartigan, J. A., & Wong, M. A. (1979). A K-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *28*(1), 100–108. https://doi.org/10.2307/2346830

İşleyen, Ş. (2021). Clustering Analysis of Employment Sectors According To OECD Countries Using The K-Average Method. *International Journal of Contemporary Economics and Administrative Sciences*, *11*(1), 093–105. https://doi.org/10.5281/zenodo.5136506

International Labour Organization (ILO). (2022). *World employment and social outlook: Trends 2022*. https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@dgreports/@dcomm/@publ/documents/publication/wcms_834081.pdf

Kanberoglu, Z., Görür, Ç., Yıldırımçakar, İ., & Türkmenoğlu, M. (2021). Analysis of Covid-19 Pandemic Data According to the Countries Including in the Human Development Index by Using Clustering Analysis K-Average Method. *International Journal of Contemporary Economics and Administrative Sciences*, *11*(2), 454–468. https://doi.org/10.5281/zenodo.5831774

Lietzmann, T., & Hohmeyer, K. (2024). Unemployed and then? The role of non-standard employment in labour market trajectories after unemployment. *Int J Soc Welf*, *34*, e12698. https://doi.org/10.1111/ijsw.12698

MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations.* In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1, 281–297. University of California Press.

Manea, L.D., & Sorcaru, I.A. (2021). Cluster Analysis on the Unemployment Rate during Lockdown Period. *Annals of "Dunarea de Jos" University of Galati Fascicle I. Economics and Applied Informatics*, 31-36. https://doi.org/10.35219/eai15840409220

Masserini , L., Bini, M., Zeli A., & Forciniti, A. (2024). Measuring the impact of the 2008 and 2011 financial crises and the 2015 recovery on the unemployment rate in Italy.

Socio-Economic Planning Sciences, *95*, 102032. https://doi.org/10.1016/j.seps.2024.102032

Mogos, I.R., Dinu, M., Constantinescu, V.G., & Istrate, B. (2022). *Unemployment in European Union During the COVID-19 Pandemic. A Cluster Analysis. In: R. Pamfilie, V. Dinu, C. Vasiliu, D. Pleșea, L. Tăchiciu eds*. 8th BASIQ International Conference on New Trends in Sustainable Business and Consumption. Graz, Austria, 25-27 May 2022. Bucharest: ASE, 117-124. https://doi.org/10.24818/BASIQ/2022/08/014

Monfort, M., Cuestas, J. C., & Ordóñez, J. (2018). Real convergence in Europe: A cluster analysis. *Economic Modelling*, 689-694. https://doi.org/10.1016/j.econmod.2013.05.015

Monfort, M., Ordóñez, J., & Sala, H. (2018). Inequality and Unemployment Patterns in Europe: Does Integration Lead to (Real) Convergence? *Open Econ Rev*, *29*, 703–724. https://doi.org/10.1007/s11079-018-9488-x

Nichols, A., Mitchell, J., & Lindner, S. (2013). *Consequences of long-term unemploymen*t. Urban Institute. https://www.urban.org/sites/default/files/publication/23921/412887-Consequences-of-Long-Term-Unemployment.PDF

OECD (2023). *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*. OECD Publishing, Paris, https://doi.org/10.1787/08785bba-en

Ogryzek, M., & Jaskulski, M. (2025). Applying Methods of Exploratory Data Analysis and Methods of Modeling the Unemployment Rate in Spatial Terms in Poland*. Appl. Sci.*, *15*, 4136. https://doi.org/10.3390/app15084136

Ok, Y. (2022). Bulanık C - Ortalamalar İle Ülkelerin İşsizlik Göstergeleri Temelinde Kümelenmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (34), 507-512. https://doi.org/10.31590/ejosat.1083246

Paul, K. I., & Moser, K. (2009). Unemployment impairs mental health: Meta-analyses. *Journal of Vocational Behavior*, *74*(3), 264–282. https://doi.org/10.1016/j.jvb.2009.01.001

Poutanen, J., Gluschkoff , K., Kausto, J., & Joensuu, M. (2024). Main activity trajectory clusters of unemployed people with partial work ability and cluster features. *Scandinavian Journal of Public Health*, *52*, 918–926. https://doi.org/10.1177/14034948231210347

Salimova, G., Ableeva, A., Gusmanov, R., Sharafutdinov, A., & Nigmatullina, G. (2024). Employment in the Digital Economy Development: Regional Clustering. *Public Organization Review*, *24*, 141–160. https://doi.org/10.1007/s11115-023-00746-w

Seppälä, P., Zhu, N., Hietamäki, J., Häkkilä, L., Gawel, A., & Toikko, T. (2024). The threshold of child protection notifications is higher in municipalities with a high level of risk factors – Is this evidence of the inverse intervention law? *Child Abuse & Neglect*, 155,106963. https://doi.org/10.1016/j.chiabu.2024.106963

Tatarczak, A., & Boichuk, O. (2018). The multivariate techniques in evaluation of unemployment analysis of Polish regions. *Oeconomia Copernicana*, 9(3), 361–380. https://doi.org/10.24136/oc.2018.018

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 411–423. https://doi.org/10.1111/1467-9868.00293

Trentini, M. (2024). Labour market trajectories and unemployment of older workers in Europe after the Great Recession. *Sociology Compass*, *18*(5), 1-16. https://doi.org/10.1111/soc4.13215

World Bank. (2024). *World Development Indicators. Retrieved* (Access Date:12/06/2025), from https://data.worldbank.org/indicator/SL.UEM.TOTL.ZS

Yılmaz, T. (2022) . OECD Ülkelerinin İş Gücü Endeksi Açısından Kümeleme Analizi ile İncelenmesi. *Journal of International Management, Educational and Economics Perspectives*, *10*(1), 84–101.

Zaharia, M. (2024). The influence of pre-university education resources on school dropout and the unemployment rate in Romania is significant? *Journal of Research and Innovation for Sustainable Society (JRISS)*, *6*(2). https://doi.org/10.33727/JRISS.2024.2.22:202-214