Convergence or divergence? Trends in the digitalisation index cluster over the years

Üzeyir Fidan
Usak University,
Department of Computer Programming,
Türkiye
Email: uzeyir.fidan@usak.edu.tr

This study examines the digitalisation levels of European countries between 2017 and 2022 using the Digital Economy and Society Index (DESI). In this study, four main dimensions of the DESI were considered. Initially, hierarchical and non-hierarchical clustering analysis methods were used to group countries according to similar digitalisation levels. Then, the stability of clusters and transitions between clusters were determined by applying association rule analysis. Findings show that European countries are clustered at five levels of digitalisation. Most countries have remained in the same cluster over the years, while some countries have moved to neighbouring clusters. Moreover, common digitalisation characteristics of the countries in each cluster are revealed. The results of this study will guide countries in evaluating their digitalisation levels and devising strategies. This study is expected to contribute to policymakers' decision-making processes in the field of digital transformation.

Keywords:
digitalisation,
data analytics,
DESI dimensions,
European Union,
cluster analysis

Online first publication date: 7 November 2024

Introduction

Digitalisation is a phenomenon that deeply affects the social and economic structures of societies. Digitalisation refers to the structural transformation of society and the economy through the spread of information and communication technologies (Tran et al. 2023, Matthess–Kunkel 2020, Liu 2022). Various indices have been developed to measure the dimensions of digitalisation and differences in digitalisation levels between countries (Grigorescu et al. 2021). These indices include several indicators,

from digital infrastructure to the skill levels of people. They are also used to compare the use of digital technology in public services. These indices combine indicators covering different dimensions of digitalisation, from infrastructure to skills, and from technology use to economic and social impacts (Evangelista et al. 2014, Česnauskė 2019, Borowiecki et al. 2021). Digitalisation indices provide guidance to policymakers and explain countries' digital profiles, allowing international organisations to monitor and evaluate countries.

The Digital Economy and Society Index (DESI) developed by the European Commission is a multi-dimensional index created to comparatively monitor and evaluate the digitalisation levels of member states as part of the EU's digital single market strategy (EC 2022). Thanks to the DESI, countries can identify their digital competitiveness using objective criteria and determine policy priorities. The index also serves as a tool for member states to develop competitive and integrated digital markets across the EU.

This study focuses on the trends in DESI scores over time. The main research question is whether there is convergence or divergence between countries in terms of their level of digitalisation. In this regard, answers to the following research questions are sought:

RQ1: Which clusters do countries fall into within the digitalisation index over the years and what are the common characteristics of the countries in these clusters?

RQ2: Do the differences between countries with high and low levels of digitalisation decrease or increase over time? Is there transitivity between clusters?

RQ3: How stable have the country clusters been in terms of the digitalisation index over the years?

The changes in the index data in recent years are analysed, and convergence/ divergence trends between countries are quantitatively evaluated. Based on the findings of this study, the validity of the 'digital divide' hypothesis and the impact of digitalisation on global inequality are discussed.

Existing studies in the field of digitalisation have generally focused on the relationships between digitalisation indices and various socioeconomic indicators (Bodor et al. 2014). However, previous research has paid little attention to the temporal evolution of digitalisation indices. This study aims to fill an important gap in the literature by examining the course of digitalisation indices over the years using cluster analysis. By applying cluster analysis separately for each year, this study examines how the cluster structures formed by countries according to their digitalisation levels have changed over time. Findings, such as changes in the number of clusters, mergers or divisions of clusters and countries' transitions between clusters, provide important clues and help explain convergence or divergence trends among countries in the field of digitalisation. Thus, this study presents a large picture that guides policymakers. The main motivation of this study is to contribute to the debate on convergence/divergence between countries in the field of digitalisation by

using innovative quantitative methods such as cluster analysis. It is hoped that the findings will shape global policies to reduce the digital divide. The contribution of this study to the literature is that it examines the course of differences between countries in the field of digitalisation over time using comprehensive data.

The remainder of this paper is organised as follows. In the next section, a literature review and previous studies on the subject are summarised, and the contribution of this study is stated. In the methodology section, the dataset and statistical techniques used in the analysis are explained in detail. The findings section presents the results of the empirical analyses, which are interpreted in tables and figures. In the discussion and conclusion section, the empirical findings are evaluated by considering the literature, the results of the study are summarised and policy recommendations are made.

Conceptual framework and related works

Analysing the digitalisation levels of countries is vital for increasing their global competitiveness and welfare levels (Szeles–Simionescu 2020, Mărginean–Orăştean 2017). In this study, the digitalisation levels of countries are examined through cluster and association analyses using the DESI. The DESI is a multi-dimensional Digital Economy and Society Index developed by the European Commission. It consists of four main dimensions: human capital, connectivity, digital technology integration and digital public services. The DESI is an important tool for comprehensively assessing countries' digital competitiveness.

In previous studies, DESI data were analysed using various statistical techniques, such as regression, correlation and factor analyses. In their study, Bruno et al. (2023) developed a DESI-based tool with fewer requirements. This guided digital transformation by evaluating the digital divide between countries. Stavytskyy et al. (2019) determined the effects of sub-dimensions by applying panel regression to DESI data. They also criticised the calculation methodology of the DESI and claimed that current trends could not be calculated from historical data. Bánhidi et al. (2020) multi-dimensionally examined the usability of the DESI and evaluated its results in comparison with cluster analysis. Thus, they compared the cluster analysis and DESI results. Olczyk–Kuc-Czarnecka (2022) adopted an approach based on sensitivity analysis and examined whether weight selection of the dimensions included in the DESI index was beneficial. This study concludes that excluding Internet data and digital public services from the DESI does not significantly change the explanatory power of the index. This suggests that the existing components of the DESI and the weighting methodology are generally explanatory and accurate. Note that the authors state that it may be useful to optimise the index components for different countries and periods. In conclusion, sensitivity analyses are important for the construction of composite indices, such as the DESI.

Nagy–Somosi (2022) investigated how digital transformation in the economy and society affects social innovation capacity. The findings of their study demonstrate that the integration of digital technology plays a critical role in digital transformation. Zdanowska et al. (2020) analysed innovation capacity and its impact on city development. Tóth–Nagy (2023) analysed the socio-digital structure of European countries and revealed important inequalities. Because these inequalities were discovered, digital skills, economics, and performance data were also studied. The most important finding was that household income is a determining factor. In another study on inequalities, particularly the slower development of index scores in Eastern European countries and its reasons were investigated (Tőkés 2022, Müller-Frączek 2019). Imran et al. (2022) investigated the direct effect of the DESI dimensions on sustainable development goal index (SGDI) using panel regression modelling. Russo (2020) explained the development and applications of the DESI and stated that it is an important starting point for local development projects defined by the European 2020 strategy. Kovács et al. (2022) determined that convergence between European states could be detected using the annual DESI database. Bakumenko–Minina (2020) identified two groups of countries based on their level of digitalisation and concluded that indicators are decisive for the digital development of economies and societies.

These studies demonstrate that when examined in depth through statistical analyses, the multi-dimensional structure of the DESI can serve as a guide for policymaking. In this study, DESI data are analysed using cluster analysis method. Cluster analysis is effective for revealing the structure of a dataset by grouping observations with similar characteristics in the same cluster. In addition, modelling the association analysis in an integrated manner with clustering analysis yields important results. This study is one of the first studies to analyse DESI data using cluster analysis based on years and to reveal transitivity among clusters. The results obtained are considered important for guiding countries' digitalisation strategies and developing relevant policy recommendations.

Data, ethics and methods

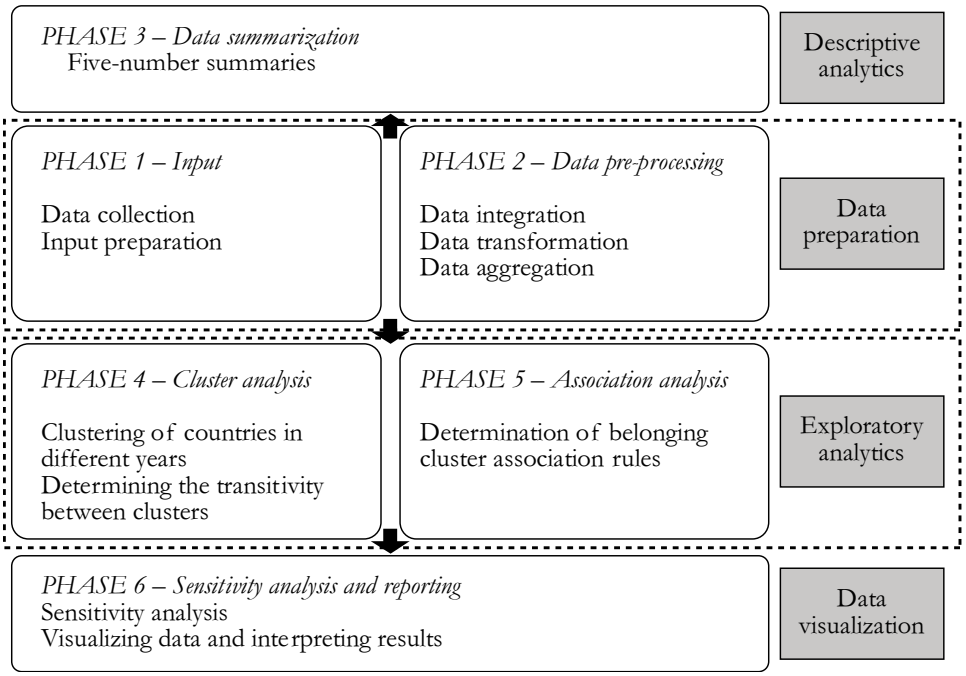
In the latest European Commission report, the digitalisation levels of 27 EU countries were analysed using data from 2017 to 2022. Four main dimensions are considered in the analysis: human capital, connectivity, digital technology integration and digital public services. Data on each dimension are presented separately by country. In addition, the DESI was calculated by taking an equally weighted average of the scores given to countries for each dimension. Thus, the overall digitalisation level of countries can be expressed by a single index value. Thanks to the multi-dimensional analysis presented in this report, countries' digitalisation performances have been comprehensively evaluated.

In this study, primary data collection methods, such as questionnaires, interviews and observations, were not employed. Because the data were secondary data obtained from reports published by the European Commission, no ethics committee approval was required.

In this study, descriptive and exploratory data analysis methods were applied using the DESI 2017–2022 data published by the European Commission. The dataset was analysed with the help of descriptive statistics and summary tables. The European countries were then grouped according to similar digitalisation characteristics by clustering analysis using data mining techniques. In addition, association rule analysis was used to explore the relationships among variables. The research model illustrating these analytical processes is presented as a whole in Figure 1. Thus, trends in digitalisation in Europe have been examined in depth by combining different analysis techniques on a multi-dimensional dataset.

Figure 1

Research methodology



In the first phase, the data to be used were determined and prepared for analysis. These data were obtained for different years.

In the second phase, data pre-processing was performed using data integration, transformation, reduction, discretisation and collection. In the data integration process, separate tables, variable names and scales for 2017–2022 were standardised and combined into a single dataset. In the data transformation process, normalisation

processes were applied to the analyses. Finally, a single index value was calculated for the four sub-dimensions with data collection (the results of the data collection process were compared with the total DESI scores). As a result, the multi-dimensional dataset was transformed into a suitable and consistent form for the analyses.

In the third phase, descriptive analyses were performed on the dataset using descriptive statistics. In this context, descriptive tables containing measures of central tendency and the dispersion of variables were prepared. In addition, a correlation matrix showing the relationships between variables was created. Thanks to these descriptive statistics, the first impressions of the data were gained, and this information helped form the basis for the analyses. These analyses also played an important role in understanding the structure and characteristics of the dataset.

After the preliminary data analyses, the exploratory data analysis was initiated to answer the research questions. In the fourth phase, dendrogram visualization was used to determine the appropriate number of clusters for clustering analysis. Using the dendrogram graph, the natural clustering tendencies of the data were observed and the optimal number of clusters was determined. A two-step clustering analysis was then performed according to the determined number of clusters. Using the two-step method, countries were subjected to preliminary clustering based on their similar characteristics and then final clusters were formed. To define the structural characteristics of the clusters obtained through clustering analysis, descriptive statistics of the variables were calculated for the observations belonging to each cluster. In this way, the common digitalisation characteristics of countries that form different clusters can be identified. The central tendency, diffusion and shape criteria of each cluster were analysed and interpreted. Thus, the differences and similarities between clusters were statistically revealed. Cluster-specific descriptive statistics are very useful for clearly profiling each group.

After the clustering analysis, in the fifth phase, the cluster memberships of the countries over the years were analysed. To begin, we determined in which cluster each country was located in each year of the study. Then, using the a priori algorithm, association rules between cluster labels and years were extracted. Through association analysis, it was determined which cluster memberships were frequently observed, and thus, the stability of the clusters was evaluated (Yüncü−Fidan 2020). The fact that countries remained consistent cluster members over time indicates that the cluster is stable. The obtained results allow the clustering performance to be analysed in temporal dimension.

In the sixth phase, sensitivity analyses were conducted to evaluate possible biases arising from the equal weighting of the DESI scores. Initially, different weights were determined for the sub-dimensions using the entropy method, and the DESI scores were recalculated. The new clustering results were interpreted by comparing them to the original results. In addition, the changes in the cluster structures that emerged when the number of clusters was increased were analysed. The consistency and robustness of the findings obtained via these sensitivity analyses were tested. Finally,

the results are summarised and interpreted using visualization techniques. The culmination, summary and results of the techniques and tests used in this study are clearly presented.

Results

Detailed descriptive statistics for the variables in the dataset were calculated annually, and the characteristics of the data were defined and interpreted. Descriptive statistics, such as measures of central tendency (mean and median) and dispersion (standard deviation, interquartile range, minimum and maximum), were obtained for the variables. Thus, the data were summarised with basic statistical properties of each variable, and the information hidden in the raw data was revealed with the help of descriptive statistics. By interpreting the results obtained, an idea about the structure and content of the data was obtained, and important points such as trends, variations, patterns and relationships in the data were analysed (Table 1).

Table 1

Descriptive statistics

Denomination	D1	D2	D3	D4	D1	D2	D3	D4
	2017				2018			
Max	16.165	9.864	8.814	16.782	16.479	10.238	9.710	17.231
Q_3	12.554	7.933	7.117	14.786	12.369	8.430	7.448	15.910
Median	10.266	6.519	5.700	11.600	10.312	6.952	6.221	12.680
Q_1	9.116	5.527	4.794	9.229	9.429	5.999	5.466	10.181
Min	6.973	3.168	2.530	1.853	6.869	3.351	2.853	2.451
Mean	10.840	6.694	5.731	11.807	10.914	7.086	6.316	12.631
Standard deviation	2.231	1.686	1.821	3.538	2.287	1.696	1.903	3.546
	2019				2020			
Max	16.445	10.973	10.732	18.138	16.872	12.005	12.250	19.410
Q_3	13.145	9.015	8.462	16.998	12.539	10.488	9.684	18.019
Median	10.711	7.936	6.825	13.550	10.926	9.040	7.525	14.577
Q_1	9.615	6.941	5.743	10.801	9.880	8.004	6.085	11.979
Min	6.993	4.072	3.018	2.958	7.133	4.786	3.279	3.724
Mean	11.206	8.046	6.917	13.466	11.458	9.132	7.632	14.488
Standard deviation	2.298	1.727	2.111	3.618	2.339	1.862	2.375	3.712
	2021				2022			
Max	17.632	18.029	13.353	20.631	17.848	19.272	14.772	21.841
Q_3	12.696	12.186	10.657	19.240	14.149	15.307	10.831	20.767
Median	11.115	11.069	8.433	15.790	11.437	14.401	9.183	16.838
Q_1	10.210	9.481	6.552	13.099	10.441	12.519	6.958	14.350
Min	7.517	7.779	3.522	4.541	7.729	9.957	3.788	5.261
Mean	11.673	11.282	8.435	15.689	12.037	14.263	9.241	16.830
Standard deviation	2.379	2.301	2.679	3.830	2.495	2.073	2.932	4.005

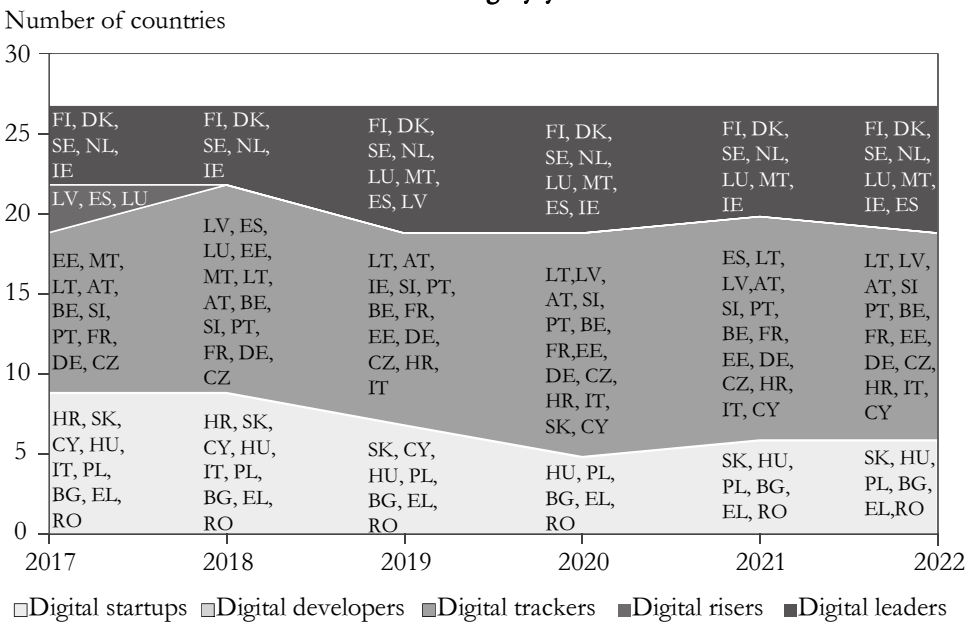
Because the DESI scores were calculated as equally weighted sums of four sub-dimensions (human capital, connectivity, digital technology integration and digital public services), they are not included separately in the tables.

RQ1: Cluster analysis – identifying clusters

Six hierarchical clustering analyses were performed on the dataset. For each year, the optimal number of clusters was determined by examining the dendrogram graphs created using the Ward method in detail. According to the number of clusters, the two-step clustering algorithm was applied to the dataset (Figure 2).

Figure 2

Clustering by years



Notes: AT – Austria, BE – Belgium, BG – Bulgaria, CY – Cyprus, CZ – Czech Republic, DE – Germany, DK – Denmark, EE – Estonia, EL – Greece, ES – Spain, FI – Finland, FR – France, HR – Croatia, HU – Hungary, IE – Ireland, IT – Italy, LT – Lithuania, LU – Luxembourg, LV – Latvia, MT – Malta, NL – Netherland, PL – Poland, PT – Portugal, RO – Romania, SE – Sweden, SI – Slovenia, SK – Slovakia.

After applying the algorithm, the membership of the countries to the clusters and the characteristics of each cluster were reported in detail, and sub-groups of the countries in the dataset with similar characteristics were successfully identified. Using the DESI, the digitalisation levels of the countries were generally classified into five different clusters using cluster analysis. These clusters are named 'digital startups', 'digital developers', 'digital trackers', 'digital risers' and 'digital leaders', and they rank the level of digitalisation from low to high. The digital startups cluster represents the

countries with the lowest level of digitalisation, whereas the digital leaders cluster represents the countries with the highest level of digitalisation. According to the DESI, the digitalisation levels of countries are comprehensively clustered and the qualitative characteristics of each cluster in terms of digitalisation are summarised through nomenclature. An annual analysis revealed that no country was included in the digital developers cluster.

RQ2: Cluster analysis – boundaries of clusters

Table 2 presents the mean values (cluster centres) and standard deviations for each dimension of the clusters formed using the four dimensions of the DESI. To evaluate the quality of the clustering results, the silhouette coefficient of concordance was calculated. The silhouette coefficient is calculated as the ratio of the proximity of points within a cluster to the distance of points in different clusters. Therefore, the silhouette coefficient was used to quantitatively evaluate the homogeneity of the clusters and their level of differentiation from each other. High silhouette values indicate high internal homogeneity and separation.

Table 2

Silhouette measure of cohesion and separation

Years	Cluster	Human capital		Connectivity		Integration of digital technology	
		mean	standard deviation	mean	standard deviation	mean	standard deviation
2017	1	8.325	0.906	5.645	1.494	3.732	1.146
	2	10.497	1.014	6.043	0.977	5.981	0.776
	3	11.672	1.008	8.361	1.363	5.396	1.452
	4	14.214	1.150	7.797	1.771	8.365	0.386
2018	1	8.897	1.418	5.833	1.485	4.582	1.351
	2	10.949	1.112	7.522	1.320	6.458	1.062
	3	14.456	1.190	8.207	1.830	9.068	0.454
2019	1	8.799	1.075	6.696	1.670	4.547	1.367
	2	11.155	1.534	7.599	0.914	7.268	1.021
	3	13.387	1.921	9.897	1.139	8.463	2.179
2020	1	8.506	1.017	8.209	2.275	4.195	0.721
	2	10.915	1.018	8.503	1.363	7.446	1.140
	3	14.253	1.434	10.808	1.361	10.106	1.656
2021	1	9.207	1.312	9.902	1.401	4.871	1.106
	2	11.173	1.096	10.668	1.620	8.391	1.099
	3	14.788	1.582	13.694	2.459	11.579	1.798
2022	1	9.302	1.219	12.894	1.019	5.400	1.342
	2	11.399	1.035	13.834	1.895	9.194	1.276
	3	15.125	1.467	15.986	1.928	12.198	2.283

(Table continues on the next page.)

(Continued.)

Years	Cluster	Digital public services		Silhouette measure of cohesion and separation
		mean	standard deviation	−1.0 −0.5 0.0 0.5 1.0
2017	1	7.554	2.787	
	2	11.318	1.644	
	3	15.526	0.891	
	4	15.018	0.652	
2018	1	8.697	2.600	
	2	13.958	1.896	
	3	16.265	0.629	
2019	1	9.082	2.990	
	2	13.508	1.812	
	3	17.237	0.562	
2020	1	9.321	3.483	
	2	14.131	2.028	
	3	18.342	0.568	
2021	1	10.652	3.342	
	2	15.874	2.178	
	3	19.635	0.613	
2022	1	11.562	3.468	
	2	16.673	1.987	
	3	21.037	0.409	

RQ3: Association analysis – stability of clusters

By analysing the number of clusters formed by years and the number of countries in each cluster, important inferences can be made regarding the balance of clusters, dominant clusters and transitions between clusters (Figure 3). For this purpose, an a priori analysis was conducted, and all rules were analysed in detail. To reveal cluster stability, the obtained rules and their lift, leverage and conviction values are presented in Table 3.

Table 3

Association rules based on clustering analysis results

		Confidence	Lift	Leverage	Conviction	n
Year 2017 = digital startups	all other years = digital startups	1.0	4.5	0.14	3.89	5/5
Year 2017 = digital trackers	all other years = digital trackers	0.9	2.7	0.21	3.33	10/9
Year 2017 = digital leaders	all other years = digital leaders	1.0	5.4	0.12	3.26	4/4

In this analysis, to obtain all rules, dominant rules were determined from all rules with a support value of 0.1. Rules with high confidence values are expected to exhibit

stability in clusters. At the same time, lift values greater than 1 indicate a strong relationship between related elements. Thus, cluster stability and transition patterns were revealed.

Figure 3

Generalised cluster structure

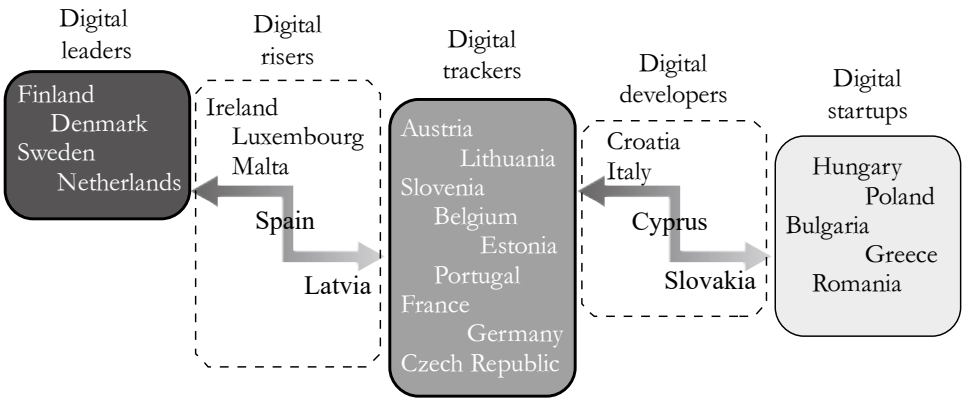


Figure 3 shows the generalised cluster structure obtained by applying the clustering analysis and association rule to the 6-year dataset. The evaluation of the clustering results shows that the three main clusters, namely, digital startups, digital trackers and digital leaders, consistently consist of the same countries over 6 years. These three main clusters represent countries with significant differences in digitalisation levels. The digital risers and digital developers clusters are two clusters with high inter-cluster mobility. Because of the clustering analysis, five different clusters were formed. The digital leaders cluster consists of countries that are pioneers and leaders in digitalisation. The digital risers cluster consists of countries that have made rapid progress in this field. The digital trackers cluster consists of countries that follow leaders but cannot catch up with them, and the digital developers cluster consists of countries that develop digitalisation infrastructure. Finally, the digital startups cluster includes developing countries that are just starting digitalisation. The names given to clusters were determined according to countries' digitalisation performance and levels. As shown in Figure 3, countries in these two clusters tend to transition to neighbouring clusters in some years. Thus, when the 6-year dataset is evaluated in its entirety, both stable main clusters and sub-clusters with high inter-cluster mobility can be distinguished.

Ireland, Luxembourg and Malta are close to the digital leaders cluster in terms of the DESI. However, Latvia performs close to the digital followers cluster. Spain is a country with high stability in this sub-cluster and is positioned at the centre of the cluster. Croatia and Italy are close to the digital followers cluster, whereas Slovakia is close to the digital starters cluster. Cyprus is located in the centre of the digital developers cluster. These countries' transitional positions provide important clues to

the clusters that they may evolve towards in the future. In summary, the transitivity relations between clusters and potential development directions of countries in the field of digitalisation can be exposed. To understand deep relationships, Table 4 was created.

Table 4

Lower–upper cluster distances and rate of change

Code[a)]	Human capital			Connectivity			Integration of digital technology		
	d_1	d_2	r (%)	d_1	d_2	r (%)	d_1	d_2	r (%)
Digital leaders									
FI	–	4.224	2.012	–	0.456	16.976	–	5.233	11.448
DK	–	1.173	2.054	–	4.591	17.025	–	4.959	10.516
NL	–	2.158	2.847	–	2.844	15.961	–	3.478	9.825
SE	–	1.871	2.797	–	0.382	10.895	–	4.521	10.788
Digital risers									
IE	0.320	3.809	1.818	1.365	1.891	25.915	3.256	1.390	6.811
MT	1.831	2.297	3.232	3.500	0.244	15.612	2.054	2.591	11.130
ES	3.150	0.979	1.639	0.678	3.934	17.267	4.452	0.194	9.972
LU	1.537	2.591	2.491	1.924	1.332	8.601	5.347	0.701	9.590
LV	4.946	0.818	1.873	4.231	0.975	5.680	7.630	2.984	15.534
Digital trackers									
EE	0.136	2.594	0.073	3.569	2.505	15.001	0.420	0.327	11.858
AT	0.885	1.845	0.300	0.565	0.499	22.173	0.253	1.001	11.190
SI	2.560	0.170	1.838	0.293	1.357	15.257	0.421	1.169	9.141
FR	1.156	1.574	1.919	1.365	2.429	22.443	1.561	0.813	9.488
DE	2.382	0.348	1.842	2.149	3.213	20.367	0.580	0.168	11.224
LT	3.009	0.278	3.094	2.343	1.279	13.659	0.227	0.521	8.537
PT	2.138	0.592	3.399	1.785	0.721	14.018	0.141	0.607	7.264
BE	1.451	1.279	3.151	4.725	3.661	10.554	2.451	3.199	9.466
CZ	2.226	0.504	2.318	1.509	0.445	19.378	1.079	0.331	8.316
Digital developers									
IT	2.710	0.186	2.206	1.813	2.325	27.531	0.745	5.096	12.889
CY	1.411	1.485	1.369	1.200	1.713	29.906	0.604	3.748	8.596
HR	1.105	4.001	2.481	1.480	0.967	20.201	0.258	4.094	12.004
SK	0.820	2.076	3.251	1.038	0.526	14.267	2.482	1.869	7.789
Digital startups									
HU	1.281	–	1.107	0.783	–	17.952	3.394	–	10.934
PL	1.636	–	2.846	1.989	–	17.845	3.071	–	12.575
EL	0.860	–	2.135	1.223	–	33.340	2.134	–	10.136
BG	2.746	–	1.726	0.942	–	18.540	4.908	–	8.971
RO	3.164	–	2.105	0.189	–	11.929	5.003	–	7.360

(Table continues on the next page.)

(Continued.)

Code[a)]	Digital public services			DESI total		
	d_1	d_2	r (%)	d_1	d_2	r (%)
			Digital leaders			
FI	–	1.093	6.431	–	11.007	7.792
DK	–	0.019	6.574	–	10.743	8.406
NL	–	0.299	6.996	–	8.778	8.168
SE	–	0.143	6.868	–	6.632	7.370
			Digital risers			
IE	0.203	3.016	8.028	5.143	10.106	8.727
MT	0.387	3.605	6.235	6.999	8.250	7.893
ES	0.185	3.034	7.210	7.109	8.140	8.463
LU	0.222	2.997	5.912	9.030	6.219	6.080
LV	1.363	1.856	6.188	18.170	2.922	5.867
			Digital trackers			
EE	2.047	8.947	0.990	2.079	9.363	5.266
AT	2.718	4.182	7.601	3.915	7.526	8.556
SI	3.374	3.526	8.430	5.221	6.221	8.407
FR	3.910	2.990	8.217	5.262	6.180	9.575
DE	4.896	2.004	7.347	5.708	5.734	9.629
LT	0.298	6.602	7.224	5.877	5.565	7.676
PT	3.771	3.129	6.516	7.835	3.607	7.438
BE	4.560	2.340	7.168	8.284	3.158	7.103
CZ	4.634	2.266	9.237	9.448	1.994	9.099
			Digital developers			
IT	3.226	3.346	9.023	3.379	10.953	11.930
CY	3.466	3.105	9.293	4.280	10.052	10.798
HR	4.454	2.118	8.953	5.086	9.246	9.446
SK	4.847	1.725	7.470	9.187	5.145	7.861
			Digital startups			
HU	0.502	–	8.549	3.390	–	9.168
PL	0.093	–	10.195	6.602	–	10.221
EL	4.001	–	10.431	8.218	–	11.881
BG	0.874	–	9.206	9.470	–	9.573
RO	8.587	–	23.328	16.565	–	9.548

a) For country codes see Figure 2.

Notes: the variables in Table 4 show the following information: d_1 is the distance of the country to supercluster centre, d_2 is the distance of the country to the sub-cluster centre and r (%) is the average rate of change between 2017 and 2022.

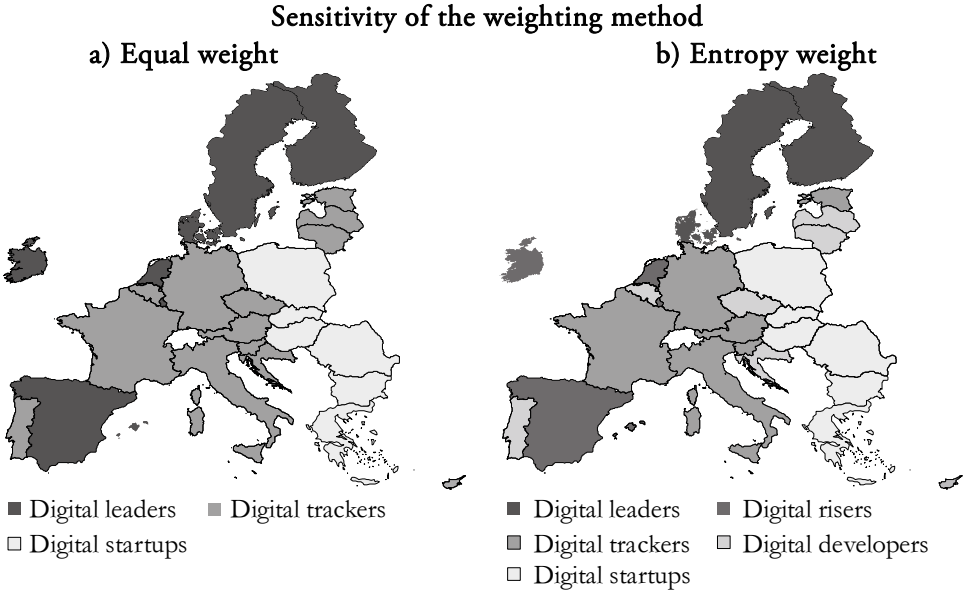
For countries to improve their digitalisation performance, the four main dimensions of the DESI (connectivity, human capital, internet usage and digital technology integration) need to be examined in detail. Each country must balance its current position and move to higher clusters. In this context, in line with the targets set to improve countries' positions, countries' distance to the upper cluster centre and their distance to the lower clusters are considered comparatively. Thus, as progress is

made in DESI indicators, the transition probabilities of countries between clusters can be modelled. A low value of d_1, which indicates the country's distance to a supercluster centre, indicates that the country is relatively close to a supercluster. Similarly, the low value of d_2, which indicates the distance to the sub-cluster centre, indicates that the country has close similarities with a sub-cluster. In addition, for each country, the r value obtained from the average of the annual percentage changes in the relevant sub-dimension scores between 2017 and 2022 is used to model the tendency to move to a supercluster. The value of r indicates the level of willingness of a country to move to a supercluster.

Sensitivity analysis

The data and data calculation methodology used herein were based on the assumption of equal weighting of the DESI scores. In this section, we discuss possible changes when using the more objective entropy technique. To test these differences, the weights of the DESI dimensions were calculated from the 2022 data. According to the calculated weights, human capital = 0.183, connectivity = 0.091, digital technology integration = 0.451 and digital public services = 0.275. According to the index scores calculated with equal weights, the impact of human capital and connectivity dimensions decreased slightly, whereas the impact of digital technology integration and digital public services dimensions increased. The 2022 clustering results obtained using the equal-weight and entropy methods are presented in Figure 4. Based on the results, we conclude that different weighting methods may be effective relative to the results depending on the nature of the dataset.

Figure 4



Conclusions

In this study, the digitalisation levels of the countries were examined by cluster analysis using the DESI data. Due to the multi-year dataset, the clustering structure over time and the stability of the clusters were evaluated.

According to these findings, three dominant clusters formed between 2018 and 2022, while four clusters formed only in 2017. The annual analysis revealed that the main determinant of the cluster structure is the characteristics of inequalities. The findings obtained in this study, which are based on DESI data for the period under review, show that similar conclusions have been reached in other academic studies on the subject. Accordingly, it can be concluded that the structural characteristics of regional inequalities effectively determine the basic dynamics of clusters formed during the relevant period.

In the analyses conducted to examine inter-cluster transitivity, the descriptive statistics of each cluster were used first. The findings demonstrate that cluster centres differ significantly from each other in terms of each dimension. In addition, the silhouette analysis results indicate that clustering quality is acceptable. With association rules analysis, the most frequently observed rules were identified, and the stability of the clusters was evaluated. In addition to the stable clusters, two new clusters that can be labelled as transition clusters were identified. On a yearly basis, time-dependent changes in these transition clusters reveal the transitivity between clusters. These findings contribute to a better understanding of the dynamics of regional developmental disparities and provide important implications that can guide policymakers.

Sensitivity analysis revealed that the weighting method used in the DESI calculation had a determinant effect on the obtained findings. The current equal weighting approach does not adequately reflect the internal dynamics of the dataset, and it can be improved. In future studies, determining the weights of the DESI components using data-driven statistical techniques will provide more consistent and reliable results. Revising the weighting methodology will contribute to a more accurate measurement of countries' digitalisation performance.

Countries that comprise the digital leaders group performed above 50% on average in all DESI dimensions. Improvement in all dimensions is the common denominator of this group. In particular, approximately 84% achievement in the digital public services dimension is the overall group success. The most prominent element distinguishing this group from the others is the digital technology integration dimension. Although countries in the digital leaders group are leading in this dimension compared with those in other groups, this dimension still constitutes the weakest link of the group. Therefore, despite the digital leaders group's advanced position in digital transformation, companies have deficiencies that must be overcome in terms of technology integration. The performance of the countries that formed the digital risers group proved that they were in a transition cluster. In terms of the digital public services dimension, the digital risers group is quite similar to the

digital leaders group and differs markedly from the digital trackers group. However, this relationship is reversed in the digital technology integration dimension. Therefore, the most prominent common characteristic of the digital risers group is the positive results that it has achieved in e-government applications despite its low performance in digital technology integration. The general profile of the group reflects the characteristics of a cluster in transition. Countries in the digital trackers group performed in a way that reflected the overall EU average. They were significantly behind the digital leaders group and were removed from all dimensions. They can target higher groups by increasing their steady rise, especially in terms of the connectivity and digital technology integration dimensions. Countries in this cluster resemble upper-level digital public services groups but are closer to the lower clusters in terms of the other three dimensions. Although the digital developers group outperformed the higher level group in the connectivity dimension, it was below the EU average in the digital public services dimension. The most prominent common characteristic of this group is inadequate e-government applications. Digital developers are a transition group that can make substantial contributions to their digitalisation performance by improving a single dimension compared with other clusters. Focusing especially on the digital public services dimension, which is a weak link, can quickly increase the overall digitalisation index upwards. In contrast to countries in the digital leaders group, those in the digital startups group performed below 50% on average in all DESI dimensions. On the one hand, this situation indicates a high potential for improvement in all dimensions; on the other hand, it also harbours the risks of falling out of the digital transformation race. When the 6-year period was analysed, it was observed that the average rates of change achieved in the connectivity dimension stood out and investments in this area increased. However, inadequacies in the human capital and digital technology integration dimensions are among the main reasons for the overall low digitalisation performance of this group.

As a result, the countries were grouped into five different clusters according to the DESI. Among these clusters, digital leaders, digital followers and digital beginners were found to consist of the same countries over 6 years. However, digital ascenders and digital developers are more transient, with some countries moving to neighbouring clusters over time.

Managerial implications

The most critical implication of the study for policymakers is derived from the combined evaluation of Table 4 and Figure 3. First, the rapid progress made in the connectivity and digital technology integration dimensions of the DESI over the 6-year period analysed is remarkable. These sub-dimensions share the common characteristic of being based on economic fundamentals. The emphasis on infrastructure investments by European Commission President Ursula von der Leyen in her 2020 State of the Union speech also supports this trend (EC 2022). In

conclusion, the progress achieved in the economically strategic areas of connectivity and digital technology integration is in line with the policy priorities shaped at EU level and is expected to continue to increase in the coming period.

The digital public services dimension of the DESI focuses on the dissemination of e-government applications. Although all member countries are making intensive efforts in this area, they are not represented sufficiently in the index scores. Digital technologies increasingly place new demands and expectations on the public sector. Realising the full potential of these technologies poses a significant challenge for public institutions. Some member states allocate more than half of their digital budgets to the digitalisation of public services (e.g. Malta, Lithuania, Finland and Croatia). However, the existing shortcomings limit the impact of this intensive effort.

The human capital dimension of the DESI is addressed in the context of equipping citizens with digital skills. As emphasised in the European Commission's DESI 2020 report, the EU aims for at least 80% of citizens to acquire basic digital skills by 2030 and to close the digital divide between men and women (EC 2022). However, most member countries are still far from these targets. Countries that want to stay ahead in the digital transformation race can turn human capital development into an opportunity. Among the dimensions of the DESI, steps to be taken towards human capital have the highest potential to bring countries to the next level.

The results of this study are important for revealing the current level of digitalisation and the potential development direction of each country. Considering these findings, countries can determine their digitalisation policies and targets and take the necessary steps to solidify them. The results of this study can be useful for decision-makers and relevant institutions as follows:

- Countries can objectively evaluate their current situation in digitalisation by observing their own digitalisation levels and in which clusters they are located.
- As cluster information is obtained, countries can position themselves compared to other countries and set targets to reach their digitisation goals.
- Countries that want to increase their level of digitalisation can observe and take note of countries in the closest and upper cluster as role models.
- Countries in transitional clusters can identify the clusters to which they can evolve and adapt their strategies.
- National and international institutions can identify countries to support in their digitalisation journey.
- Technology companies can gain important insights into the geographies in which they should invest.

Limitations and future studies

This study has some limitations. Only DESI data were used in this study. Given the multi-dimensional nature of digitalisation, combining different indices and indicators may yield more comprehensive results. In addition, the analyses were limited to a 6-year dataset. To identify long-term trends, longer term data are needed.

In future studies, analyses can be repeated using longer term datasets and different digitalisation indicators. In addition, testing different clustering algorithms and analysing the causality relations between countries can expand the scope of this study. Investigating the real-life implications of the findings obtained through case studies and modelling the diffusion of digitalisation policies across countries are topics for future studies.

REFERENCES

BAKUMENKO, L. P.–MININA, E. A. (2020): International index of digital economy and society (I-DESI): trends in the development of digital technologies *Statistics and Economics* 17 (2): 40–54. https://doi.org/10.21686/2500-3925-2020-2-40-54

BÁNHIDI, Z.–DOBOS, I.–NEMESLAKI, A. (2020): What the overall digital economy and society index reveals: A statistical analysis of the DESI EU28 dimensions *Regional Statistics* 10 (2): 42–52. https://doi.org/10.15196/RS100209

BODOR, Á.–GRÜNHUT, Z.–HORECZKI, R. (2014): Socio-cultural cleavages in Europe *Regional Statistics* 4 (2): 106–125. https://doi.org/10.15196/RS04207

BOROWIECKI, R.–SIUTA-TOKARSKA, B.–MAROŃ, J.–SUDER, M.–THIER, A.–ŻMIJA, K. (2021): Developing digital economy and society in the light of the issue of digital convergence of the markets in the European Union countries *Energies* 14 (9): 2717. https://doi.org/10.3390/en14092717

BRUNO, G.–DIGLIO, A.–PICCOLO, C.–PIPICELLI, E. (2023): A reduced composite indicator for digital divide measurement at the regional level: an application to the digital economy and society index (DESI) *Technological Forecasting and Social Change* 190: 122461. https://doi.org/10.1016/j.techfore.2023.122461

ČESNAUSKĖ, J. (2019): Digital economy and society: Baltic states in the EU context *Economics and Culture* 16 (1): 80–90. https://doi.org/10.2478/jec-2019-0009

EVANGELISTA, R.–GUERRIERI, P.–MELICIANI, V. (2014): The economic impact of digital technologies in Europe *Economics of Innovation and New Technology* 23 (8): 802–824. https://doi.org/10.1080/10438599.2014.918438

GRIGORESCU, A.–PELINESCU, E.–ION, A. E.–DUTCAS, M. F. (2021): Human capital in digital economy: An empirical analysis of Central and Eastern European countries from the European Union *Sustainability* 13 (4): 2020. https://doi.org/10.3390/su13042020

IMRAN, M.–LIU, X.–WANG, R.–SAUD, S.–ZHAO, Y.–KHAN, M. J. (2022): The influence of digital economy and society index on sustainable development indicators: the case of European Union *Sustainability* 14 (18): 11130. https://doi.org/10.3390/su141811130

KOVÁCS, T. Z.–BITTNER, B.–HUZSVAI, L.–NÁBRÁDI, A. (2022): Convergence and the Matthew effect in the European Union based on the DESI index *Mathematics* 10 (4): 613. https://doi.org/10.3390/math10040613

LIU, T. C. (2022): Digital policy in european countries from the perspective of the digital economy and society Index *Policy & Internet* 14 (1): 202–218. https://doi.org/10.1002/poi3.274

MĂRGINEAN, S.–ORĂȘTEAN, R. (2017): Measuring the digital economy: European Union countries in global rankings *Revista Economică* 69 (5): 73–80.

MATTHESS, M.–KUNKEL, S. (2020): Structural change and digitalization in developing countries: conceptually linking the two transformations *Technology in Society* 63: 101428. https://doi.org/10.1016/j.techsoc.2020.101428

MÜLLER-FRACZEK, I. (2019): Dynamic measurement of complex phenomena in assesing the Europe 2020 strategy effects *Regional Statistics* 9 (1): 32–53. https://doi.org/10.15196/RS090107

NAGY, S.–SOMOSI, M. V. (2022): The relationship between social innovation and digital economy and society *Regional Statistics* 12 (2): 3–29. https://doi.org/10.15196/RS120202

OLCZYK, M.–KUC-CZARNECKA, M. (2022): Digital transformation and economic growth-DESI improvement and implementation *Technological and Economic Development of Economy* 28: 775–803. https://doi.org/10.3846/tede.2022.16766

RUSSO, V. (2020): Digital Economy and Society Index (DESI). European guidelines and empirical applications on the territory. In: SARASOLA SÁNCHEZ-SERRANO, J.–MATURO, F.–HOŠKOVÁ-MAYEROVÁ, Š. (eds.): *Qualitative and quantitative models in socio-economic systems and social work. Studies in systems, decision and control* 208: 427–442., Springer, Cham. https://doi.org/10.1007/978-3-030-18593-0_31

STAVYTSKYY, A.–KHARLAMOVA, G.–STOICA, E. A. (2019): The analysis of the digital economy and society index in the EU *TalTech Journal of European Studies* 9 (3): 245–261. https://doi.org/10.1515/bjes-2019-0032

SZELES, M. R.–SIMIONESCU, M. (2020): Regional patterns and drivers of the EU digital economy *Social Indicators Research* 150 (1): 95–119. https://doi.org/10.1007/s11205-020-02287-x

TŐKÉS, G. E. (2022): Global digital divide: the case of the east EU countries. A comparative analysis of the east EU countries based on their digital economy and society index (2018–2022) *ME. DOK Média-Történet-Kommunikáció* 17 (4): 5–30.

TÓTH, B.–NAGY, G. (2023): Digital skill-based centre-periphery differences in Europe, 2019 *Regional Statistics* 13 (4): 583–616. https://doi.org/10.15196/RS130401

TRAN, T. L.–HERDON, M.–PHAN, T. D.–NGUYEN, T. M. (2023): Digital skill types and economic performance in the EU27 region, 2020–2021 *Regional Statistics* 13 (3): 536–558 https://doi.org/10.15196/RS130307

YÜNCÜ, V.–FIDAN, Ü. (2020): Çalışan sessizliği ve algılanan örgütsel destek arasındaki ilişkinin veri madenciliği ile incelenmesi *Acta Infologica* 4 (2): 123–137. https://doi.org/10.26650/acin.801398

ZDANOWSKA, N.–ROZENBLAT, C.–PUMAIN, D. (2020): Evolution of urban hierarchies under globalisation in Western and Eastern Europe *Regional Statistics* 10 (2): 3–26. https://doi.org/10.15196/RS100202

INTERNET SOURCE

EUROPEAN COMMISSION [EC] (2022): *Digital economy and society index report (DESI) 2022.* https://digital-strategy.ec.europa.eu/en/policies/desi (downloaded: February 2024)