PRETHODNO PRIOPĆENJE PRELIMINARY COMMUNICATION
UDK: 656.7:519.23(4)
JEL classification: C38; R41; Z30
DOI: https://doi.org/10.22598/at/2025.37.2.213

Ivan JAJIĆ *
Katarina ĆURKO **
Jasmina PIVAR ***

KLASTERIRANJE EUROPSKIH ZEMALJA PREMA OPORAVKU PUTNIČKOG ZRAČNOG PROMETA NAKON PANDEMIJE COVID–19 (2018.–2023.): EKSPLORATIVNA ANALIZA

CLUSTERING EUROPEAN COUNTRIES BY POST-COVID-19 AIR-PASSENGER RECOVERY (2018–2023): AN EXPLORATORY ANALYSIS

SAŽETAK: Autori rada analiziraju Eurostatove podatke o putničkom zračnom prometu u 29 europskih zemalja u godinama 2018. (prije COVID-a), 2020. (COVID-19) i 2023. (oporavak). Primjenom metode klasteriranja k-srednje vrijednosti na standardiziranim varijablama izdvajaju se heterogeni obrasci oporavka koji grupiraju zemlje u tri klastera. Odabir broja klastera temelji se na metodi lakta, Silhouette pristupu i Calinski–Harabaszovu indeksu. Rezultati pokazuju da je oporavak do 2023. godine ostao neujednačen. Analiziraju se svi klasteri, uz postotne promjene i provjere robusnosti. Raspravljaju se implikacije i politike u zračnom prometu i turizmu (planiranje kapaciteta, razvoj ruta, ciljane mjere) uz naglašenu ulogu informacijskih tehnologija u pospješivanju analitičkih procesa.

KLJUČNE RIJEČI: k-srednja vrijednost, statistička analiza, vizualizacija, COVID-19, međunarodni turistički tokovi

ABSTRACT: The authors analyze Eurostat's air-passenger data for 29 European countries during three reference years – 2018 (pre-COVID), 2020 (COVID-19), and 2023 (recovery). By applying the k-means clustering method on standardized variables heterogenous recovery patterns are classified. Model selection is supported by the Elbow method, Silhouette approach, and Calinski–Harabasz index. The findings reveal heterogeneous recovery by 2023. All clusters are reported in detail including percentage changes and robustness checks. Managerial and policy implications for air transport and tourism (capacity planning, route development, targeted incentives) and information technologies' contribution to enhancing analytical processes and supporting data-driven decision-making are discussed.

KEY WORDS: k-means, statistical analysis, visualization, COVID-19, international tourist flows

* Ivan Jajić, PhD, Assistant, University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia, e-mail: ijajic@net.efzg.hr, ORCID: 0000-0002-5081-3531

** Full Professor Katarina Ćurko, PhD, University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia, e-mail: kcurko@net.efzg.hr, ORCID: 0000-0002-2489-6791

*** Assistant Professor Jasmina Pivar, PhD, University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia, e-mail: jpivar@net.efzg.hr, ORCID: 0000-0002-9223-4239

1. UVOD

Ograničenja putovanja uvedena radi provedbe javnozdravstvenih politika za sprječavanje širenja bolesti COVID-19 prouzročila su, kao kolateralna posljedica, krizu u brojnim industrijama. Posebno je bio pogođen zračni prijevoz putnika, što je snažno utjecalo na zrakoplovnu industriju. Države su zatvarale granice, uvodile mjere *lockdowna*, a putnici nisu mogli ili nisu željeli putovati. U prvim mjesecima pandemije (siječanj–svibanj 2020.) mreže međunarodnih i domaćih zračnih luka bile su pogođene različitim intenzitetom. COVID-19 mnogo je snažnije utjecao na međunarodne nego na domaće letove. Intraeuropski zračni promet gotovo je stao u prvim tjednima i mjesecima zatvaranja, uz izrazito ograničenu povezivost u odnosu na predpandemijsko razdoblje. U Kini je zabilježen pad prometa u veljači, dok su se u SAD-u dogodile promjene, ali blažeg intenziteta nego u Europi (Sun *et al*., 2020). Nakon početnog vala 2020. godine, novi sojevi virusa te politika nulte tolerancije na COVID-19 uzrokovale su dodatne poremećaje u zračnom prometu u pojedinim zemljama. Primjerice, dok se broj putnika u Europi i SAD-u oporavljao, u Kini je u 2021. i 2022. godini ponovno zabilježen pad (To i Lee, 2024).

Pandemija je u Europi izazvala izraženije i heterogenije poremećaje u zračnom prometu. Kako su zemlje bilježile različite dinamike pada i oporavka, sve je nužnije sustavno proučiti obrasce oporavka kako bi se poduprla akademska istraživanja i buduće javne politike. Polazeći od razlike između pandemije COVID-19 i oporavka europskog putničkog zračnog prometa, postavljeno je sljedeće istraživačko pitanje: IP – Kako su se europske zemlje grupirale prema obrascima post-COVID oporavka u putničkom zračnom prometu? Radi odgovora na istraživačko pitanje, cilj istraživanja je dvostruk: *(i) identificirati skupine zemalja sa sličnim*

1. INTRODUCTION

The constraints on traveling due to public health policies for preventing the spread of COVID-19 infection led to crises, as collateral damage, for many industries. Passenger air travel was particularly affected, which significantly impacted the airline industry. Countries closed their borders, set lockdown measures, and passengers were unable or reluctant to travel. During the earlier months of the COVID-19 pandemic (January-May 2020) worldwide, international and domestic airport networks were affected differently. COVID-19 affected international flights much more intensively than domestic flights. Intra-European air traffic stopped during the first weeks and months of the lockdown with highly limited connectivity compared to the situation before the pandemic. China' air traffic recorded a decline in February and the USA implemented changes, but they were less severe than in Europe (Sun *et al*., 2020). After the initial wave in 2020, new strains of COVID-19 and zero-tolerance policy on COVID-19 caused additional disruptions in air traffic of some countries. For example, while the number of passengers in Europe and the USA was recovering, China saw a fresh drop in 2021 and 2022 (To and Lee, 2024).

The COVID-19 pandemic caused more heterogeneous and considerable disruptions to air travel in Europe. As countries experienced different dynamics of declines and recoveries, it is becoming more and more necessary to study systematically the different recovery patterns to support both academic studies and future public policy. Starting from the relationship between the COVID-19 pandemic and the recovery of Europe's air passenger traffic, the following research question was formulated: RQ — How did European countries cluster according to their post-COVID recovery patterns in air passenger traffic? To answer the research question, *obrascima oporavka primjenom metode k-srednje vrijednosti na standardiziranim pokazateljima zračnog prometa* te *(ii) interpretirati implikacije tih klastera za razvoj turizma, upravljanje prometom i oblikovanje politika*. Sukladno tome, formuliraju se hipoteze istraživanja:

H1: Europske zemlje moguće je podijeliti u najmanje dva statistički značajna klastera na temelju predpandemijskih i postpandemijskih podataka o zračnom prometu.

H2: Veličina oporavka između 2020. i 2023. godine značajno se razlikuje među klasterima.

H3: U većini zemalja putnički promet iz 2023. godine nije u potpunosti nadmašio razine iz 2018. godine.

Otkrivanjem heterogenosti obrazaca oporavka ovaj rad pridonosi istraživanjima u zračnom prometu i turizmu primjenom informacijskih tehnologija i podatkovno-analitičkih metoda na interdisciplinaran način. Rad ističe skupine zemalja koje pokazuju veću otpornost, te one koje zaostaju u oporavku, nudeći primjenjive spoznaje donositeljima politika, zrakoplovnim prijevoznicima i menadžerima u turizmu. Te su spoznaje relevantne za upravljanje krizama, planiranje kapaciteta i ciljane mjere potpore u širem sustavu turizma i prometa.

Zračna povezanost oblikuje tokove turista, sezonalnost i konkurentnost odredišta (ACI Europe, 2023; ACI Europe & SEO Amsterdam Economics, 2024), stoga neujednačen post-COVID oporavak ima izravne implikacije na obnovu potražnje, održavanje međunarodnih linija i jačanje regionalne otpornosti (ACI Europe, 2023; Frontier Economics, 2023). Klasifikacijom europskih zemalja u klastere oporavka, ovaj rad nudi sažetu tipologiju koja turističkim dionicima može pomoći u prilagodbi intervencija prema skupinama zemalja. Primjerice, destinacije u sporije oporavljajućim klasterima mogu zahtijevati ciljane marketinške aktivnosti i potporne politike za obnovu potražnje, dok će

the authors set a two-fold research goal: *i) to identify groups of countries with comparable recovery patterns using k-means clustering on standardized air-traffic indicators*, and (*ii) to interpret the implications of these clusters for tourism development, transport management, and policy design*. Therefore, the following research hypotheses were formulated:

H1: The European countries can be divided into at least two statistically meaningful clusters based on their pre- and post-COVID air-traffic data.

H2: The magnitude of the recovery between 2020 and 2023 differs significantly across clusters.

H3: In most countries, 2023 passenger volumes did not fully surpass the 2018 levels.

By uncovering heterogeneity in recovery patterns, this paper contributes to both aviation and tourism research by applying information technology and data-analytic methods in an interdisciplinary way. It highlights which country groups exhibit resilience or lag in recovery providing actionable insights for policymakers, airlines, and tourism managers, which are relevant for crisis management, capacity planning, and targeted support measures in the wider tourism and transport sectors.

Air connectivity shapes tourism flows, seasonality, and destination competitiveness (ACI Europe, 2023; ACI Europe & SEO Amsterdam Economics, 2024). Hence, uneven post-COVID recovery has direct implications for rebuilding demand, sustaining international routes, and enhancing regional resilience (ACI Europe, 2023; Frontier Economics 2023). By classifying European countries into recovery clusters, this paper provides a concise typology that can support tourism stakeholders' efforts in tailoring interventions by country groups. Thus, the destinations in slower-recovering clusters may require targeted marketing campaigns and support policies to rebuild demand, while those in faster-recovering clusters may need destinacije u brže oporavljajućim klasterima zahtijevati upravljanje kapacitetima i strategije diverzifikacije za održiv rast.

Podaci o zračnom prometu za 29 europskih zemalja preuzeti su iz javno dostupne baze Eurostat. Odabrane su godine ključne za pregled utjecaja pandemije na vrijednosti zračnog prometa: predpandemijska 2018., pandemijska 2020. i postpandemijska 2023. godina. Zbog složenosti i varijabilnosti oporavka među državama, klaster analiza predstavlja prikladnu metodu za ispitivanje procesa oporavka. Ona omogućuje identifikaciju skupina zemalja koje su se oporavljale na sličan način ili sličnim putanjama na temelju usporedive kvantitativne perspektive zračnog prometa. U odnosu na konvencionalnije regresijske okvire, klasteriranje je prikladnije za eksplorativne svrhe jer identificira strukturne odnose unutar promatranog razdoblja, umjesto da polazi od unaprijed zadanih odnosa koji su nužni u regresiji. Stoga je klasteriranje dobro prilagođeno eksplorativnim istraživanjima u postpandemijskim uvjetima.

Analizom vrijednosti zračnog prometa prije, tijekom i nakon pandemije, rad pruža uvid u poremećaje u prometu i obrasce oporavka. Nadalje, uključivanje R koda povećava transparentnost, ponovljivost i metodološku jasnoću u istraživanjima temeljenima na podacima; kod je prikazan u prilogu rada. Primjenom metoda klasteriranja na postpandemijske podatke o zračnom prometu, rad klasificira zemlje prema obrascima oporavka te pridonosi novim empirijskim dokazima o otpornosti europskog zračnog prometa. Nalazi imaju implikacije za prometnu politiku, planiranje oporavka nakon pandemije te šira istraživanja turizma i mobilnosti.

Struktura rada je sljedeća: nakon uvoda, pregled literature donosi pozadinu primjene klasteriranja u analizi zračnog prometa. Treće poglavlje opisuje metodologiju, dok četvrto poglavlje prikazuje empirijske rezultate. Završno poglavlje sažima ključne spoznaje, ograničenja i prijedloge za buduća istraživanja.

capacity management and diversification strategies for sustainable growth.

Air traffic data from 29 European countries were collected from the Eurostat database. The data for the crucial years for determining the impact of COVID-19 pandemic on the air traffic values were selected: pre-pandemic 2018, pandemic 2020, and post-pandemic 2023. Due to the complexity and variation of recovery from the pandemic in the air traffic context across different countries, cluster analysis provides a useful method of examining the recovery process. It allows identifying clusters of countries that recovered in a similar manner or trajectory based on comparative quantitative air traffic perspective. Clustering is more suitable for exploratory purposes than more conventional regression frameworks as it allows identifying structural relationships in the specified period as against pre-defined relationships required in regression. Hence, clustering is well-suited for exploratory research under post-pandemic conditions.

By examining air traffic values before, during, and after the COVID-19 pandemic, this paper will provide insights into air traffic disruptions and recovery patterns. Furthermore, embedding R code in this paper will potentially enhance transparency, replicability, and methodological clarity in data-driven research. The code is shown in the Appendix. Utilizing clustering methods to analyze post-pandemic air traffic data, this paper classifies countries in line with recovery patterns and contributes new empirical evidence on the resilience of European air transport. The findings have implications for transport policy, pandemic recovery planning, and wider tourism and mobility investigations.

The article is structured as follows: after Introduction, Literature Review presents the background on the use of clustering in air traffic analyses. The third section describes methodology, while the fourth one provides empirical findings. The final section summarizes the most important research findings, limitations, and suggestions for future research.

2. PREGLED LITERATURE

Industrija zračnog prometa i zrakoplovnih prijevoznika posljednjih je godina doživjela izrazite promjene. Pandemija bolesti COVID-19 snažno je pogodila zrakoplovnu industriju (Mohd Kamal *et al*., 2024). Nadalje, različita ograničenja putovanja u pojedinim zemljama promijenila su percepciju zračnog prijevoza i trajno utjecala na navike potrošača (Mohd Kamal *et al*., 2024). U sektoru zračnih prijevoznika nužan je podatkovno utemeljen pristup, osobito u upravljanju tokovima zračnog prometa. Taj je aspekt presudan u pristupu specifičnoj zračnoj luci kako bi se osigurali sigurni dolasci zrakoplova, ali i smanjila zagušenost u samoj luci radi učinkovitijeg upravljanja slotovima i flotom (Di Luigi *et al*., 2022; Lui *et al*., 2020).

Podatkovno utemeljeni pristupi koji se oslanjaju na metode klasteriranja koriste se za detaljniju analizu zračnog prometa. Klasteriranje je primijenjeno za uvid u intraeuropske putanje letova tijekom jedne ljetne sezone (Lui *et al*., 2020; Bolić *et al*., 2022). Analiza je pokazala jasnu povezanost klastera putanja s tipovima zrakoplova i operativnim troškovima leta. Slično tome, pristup temeljen na klasteriranju primijenjen je za analizu putanja u zračnoj luci Toulouse–Blagnac radi boljeg razumijevanja prometa i rasporeda slijetanja (Olive i Morio, 2019). Nadalje, analiza podataka o kretanju, potpomognuta klasteriranjem obrazaca osjetljivim na relevantnost koje uzima u obzir samo određene putanje, demonstrirana je na tri studije slučaja s realnim podacima zračnog prometa (Andrienko *et al*., 2018). Na temelju eigengap-kriterija korišten je automatski hijerarhijski algoritam klasteriranja za izvedbu kategorija mreže zračnog prometa u terminalnom području; algoritam je obradio 404 leta pravcem sjever–jug i izdvojio tri kategorije (Zhang *et al*., 2020).

Novija literatura u sve većoj mjeri razmatra učinke pandemije COVID-19 na zračni promet. Suau-Sanchez *et al*. (2020) proveli

2. LITERATURE REVIEW

The air traffic and airline industry have seen drastic changes in the recent years. The COVID-19 pandemic severely affected the airline industry (Mohd Kamal *et al*., 2024). Furthermore, various countries' travel restrictions changed the perception of air travel and impacted consumers' habits permanently (Mohd Kamal *et al*., 2024). The airline carrier sector needs a data-driven approach, especially in air traffic flow management. This aspect is vital in approaching a specific airport to ensure safe and sound arrivals of aircraft as well as to decrease the congestion in the airport itself for better slot and aircraft management (Di Luigi *et al*., 2022; Lui *et al*., 2020).

Data-driven approaches using clustering methods are used more detailed analyses of air traffic. Clustering was applied to gain insights into intra-European flight trajectories during one summer season (Lui *et al*., 2020; Bolić *et al*., 2022). The analysis showed a clear relation between trajectory clusters with aircraft types and operational flight costs. Similarly, a clustering-based approach was applied to analyze trajectories in the Toulouse-Blagnac airport for a better understanding of traffic and landing schedules (Olive and Morio, 2019). Then, movement data analysis supported by relevance-aware pattern clustering that considers only certain trajectories was demonstrated on three case studies with real air traffic data (Andrienko *et al*., 2018). An eigengap-based automatic hierarchical clustering algorithm was used to derive categories of the air traffic network in the terminal area. It analyzed the data on 404 north-south-oriented aircraft flight trajectories and derived three categories (Zhang *et al*., 2020).

Recent literature has increasingly examined the effects of COVID-19 on air transport. (Suau-Sanchez *et al*., 2020) conducted a preliminary evaluation of the impacts of COVID-19 on air transport and discussed the su preliminarnu ocjenu utjecaja pandemije na zračni promet, raspravili strukturne ranjivosti industrije i zaključili da bi pandemija mogla izazvati dugoročne promjene u zrakoplovnim operacijama diljem svijeta. Xiaoqian *et al*. (2021) na sličan su način sveobuhvatno ispitali poremećaje uzrokovane pandemijom COVID-19 te primijenili mrežni pristup mobilnosti kako bi izolirali uočljive promjene u tokovima zračnog prometa i operativnoj otpornosti. Jangik i Rafferty (2021) istražili su obrasce oporavka u zrakoplovnom sektoru, osobito u regiji Azija–Pacifik, gdje su uočili da su putanje oporavka i politike, promatrane kroz njihovu učinkovitost, ne samo značajne nego i izrazito varijabilne unutar regije.

Istodobno, istraživanja u turizmu ističu da je zračna povezanost temelj protoka turista i konkurentnosti destinacija (ACI Europe, 2023). Neujednačen oporavak zračnog prometa diljem Europe odražava i šire nalaze kako je postpandemijski povratak bio izrazito asimetričan, pri čemu su neka tržišta znatno zaostajala za drugima (Frontier Economics, 2023). Unatoč studijama koje se bave poremećajima i izazovima oporavka, većina se usredotočuje ili na široke regionalne obrasce (npr. kontinentalne učinke) ili na strategije pojedinih zrakoplovnih prijevoznika. Ovaj rad zauzima drugačiji pristup primjenom metoda klasteriranja radi sustavne klasifikacije obrazaca oporavka zračnog prometa među državama članicama Europe, nudeći detaljniju perspektivu na razini pojedine zemlje.

Metode razdvajanja govornih segmenata pilota u različite klastere uz pomoć aglomerativnog hijerarhijskog klasteriranja prikazane su i vrednovane na dva javno dostupna skupa podataka: korpusu ATCO2 i LDC-ATCC (Khalil *et al*., 2023). Osim toga, i slikovni podaci o zračnom prometu mogu se analizirati klasteriranjem. Primjerice, *Competitive Learning Riemannian Quantization* primijenjen je kao metoda klasteriranja na stvarnim podacima radi definiranja homogenih zona zračnog prostora (Le Brigant i Puechmorel, 2019). Ta-

industry's structural vulnerabilities to conclude that the pandemic might provoke long-term changes in aviation operations around the globe. Xiaoqian *et al*. (2021) similarly examined COVID-19-derived disruptions comprehensively and mobilized a mobility network approach to help isolate observable changes in air traffic flows and operational resilience. Jangik and Rafferty (2021) explored recovery patterns in the aviation sector, specifically in the Asia-Pacific region, where they observed that recovery trajectories and polices, regarding their effectiveness, were not only significant but also underpinned extensive variation across the region.

Concurrently, tourism research highlights that air connectivity is fundamental for tourist flows and destination competitiveness (ACI Europe, 2023). The uneven recovery of air transport across Europe also reflects broader analyses showing that the post-pandemic rebound was highly asymmetric, with some markets lagging significantly behind others (Frontier Economics, 2023). Despite previous studies addressing pandemic-related disruptions and recovery challenges, most focused either on broad regional patterns (e.g., continent-wide impacts) or on the strategies of individual airlines. This paper takes a different approach by applying clustering methods to systematically classify air traffic recovery patterns across European member states, offering a more detailed, country-level perspective.

Methods for separating the speech parts of pilots into different clusters based on the speaker's voice using agglomerative hierarchical clustering have been presented and evaluated on two publicly available datasets: the ATCO2 corpus and the Linguistic Data Consortium Air Traffic Control Corpus (LDC-ATCC) (Khalil *et al*., 2023). Furthermore, air traffic image data can also be analyzed by using clustering algorithms. For instance, Competitive Learning Riemannian Quantization was used as a clustering method on real data to get homogenous zones of airspace (Le Brigant and Puechmorel, 2019).

kođer, pristup k-srednjih vrijednosti korišten je za identifikaciju malih bespilotnih letjelica na temelju grafičkih reprezentacija radiofrekvencijskih signala, što može poslužiti kao temelj za sustave ranog upozoravanja (Swinney *et al*., 2022). U kontekstu zrakoplovstva, odabir kontrolora zračnog prometa ključan je element upravljanja sigurnošću. Intrinzični čimbenici kontrolora mogu utjecati na vizualnu pažnju pa je važno razviti sustavan pristup procjeni tih čimbenika kako bi se razlikovale razine izvedbe vizualne pažnje. Klasteriranje je pritom korišteno kao dio sustavne procjene intrinzičnih obilježja polaznika studija za kontrolore zračnog prometa te je pomoglo razlikovati četiri razine kvaliteta: neadekvatnu, umjerenu, dobru i izvrsnu (Zhang *et al*., 2018). U ovome se radu problem podsektorizacije zračnog prostora modelira kao višekriterijski problem klasteriranja u složenim mrežama, koji se zatim rješava primjenom metode minimalne omeđujuće geometrije za dizajn konveksnih i kompaktnih granica. Slijedom toga, točke i rute u indijskom zračnom prostoru predstavljene su kao mrežni graf, a prometna opterećenja nasumično su dodijeljena čvorovima kako bi se vodio dekompozicijski diskretni algoritam rojnih čestica. Ovakav je pristup generirao klastere koji predstavljaju podkategorije zračnog prostora i pokazao da mreže zračnog prometa imaju hijerarhijsku strukturu (Chandra *et al*., 2024).

Podaci o zračnom prometu, posebice obujam prometa, mogu se koristiti i za analizu makroekonomskog razvoja. Klasteriranje k-srednjih vrijednosti primijenjeno je na grupiranje mjesečnih volumena zračnog prometa tijekom 168 mjeseci u pet klastera, koji su potom poslužili kao podloga za daljnju analizu makroekonomskih čimbenika (Chen *et al*., 2020).

Neki su se autori bavili unaprjeđenjem tehnika temeljenih na klasteriranju za otkrivanje obrazaca putanja zračnog prometa. Primjerice, komprimirani podaci o putanjama analizirani su brzom metodom klasteriranja temeljenom na gustoćama (DPCA) radi poboljšanja točnosti predviđanja obra-

Additionally, the k-means clustering approach was used to identify small unmanned aerial systems based on graphical signal representations. This can provide a basis for creating early warning systems (Swinney *et al*., 2022). In aviation, selecting air traffic controllers is key to aviation safety management. The intrinsic factors of the controllers can influence visual attention performance, so it is important to develop a systematic assessment approach of intrinsic factors to distinguish the levels of visual attention performance. Furthermore, clustering was used as part of the systematic assessment of intrinsic qualities of air traffic control students, which facilitated distinguishing four levels of intrinsic qualities: inadequate, moderate, good, and excellent (Zhang *et al*., 2018). This paper models the airspace sub-sectorization problem as a multi-objective complex network clustering problem, which is then solved by applying the minimum bounding geometry method to design convex and compact boundaries. Subsequently, Indian airspace waypoints and routes are presented as a network graph, and traffic loads were randomly allotted to the vertices to guide the decomposition-based discrete particle swarm optimization algorithm. This approach generated clusters representing sub-categories of airspace, which showed that air transport networks have a hierarchical structure (Chandra *et al*., 2024).

Air traffic data, particularly air traffic volume, can be used for analysing macroeconomic development. The K-means clustering method was used to group air-traffic volume data during 168 months into five clusters, which then served as the basis for further analysis of macroeconomic factors (Chen *et al*., 2020).

Some authors dealt with the improvement of clustering-based techniques for mining air-traffic trajectory patterns. For example, the compressed trajectory data is analyzed using a fast-clustering algorithm based on density peaks (DPCA) to improve the accuracy of pattern prediction (Tang *et al*., 2022). The approach that consists of coarse clusterzaca (Tang *et al*., 2022). Pristup koji se sastoji od klasteriranja, detekcije iznimaka i konstrukcije agregiranih ruta razvijen je za identifikaciju i procjenu često korištenih putanja letova (Bombelli *et al*., 2019). Hijerarhijsko klasteriranje, particijsko klasteriranje i K-Shape klasteriranje testirani su kako bi se odredila najprikladnija metoda za analizu vremenskih nizova zračnog prometa u Maleziji; najboljom se pokazala particijska metoda koja koristi dinamičko poravnanje u vremenu (DTW) (Ghani *et al*., 2020).

Iako je više radova primijenilo metode klasteriranja za istraživanje obrazaca zračnog prometa (Bombelli *et al*., 2019; Chen *et al*., 2020), većina se studija usredotočila na optimizaciju putanja, sektorizaciju ili modernizaciju upravljanja zračnim prometom (ATM). Vrlo je malo pokušaja da se sveobuhvatno ispita obujam putničkog zračnog prometa u odnosu na pandemiju COVID-19. Konkretno, ne postoji sustavna analiza klasteriranja europskih zemalja korištenjem višegodišnjih podataka o prometu prije, tijekom i nakon pandemije, uz oslanjanje na javno dostupne skupove podataka poput Eurostatova avia_paocc. Ovo istraživanje popunjava važan znanstveni jaz identificiranjem razlika u obrascima oporavka putničkog prometa na makrorazini primjenom tehnika klasteriranja. Zatim, kombiniranje alata R i JASP povećava transparentnost, ponovljivost i metodološku jasnoću u ovom podatkovno utemeljenom istraživanju.

3. METODOLOGIJA

3.1. Podaci o zračnom prometu

Za ovo istraživanje autori su koristili otvorene podatke iz baze Eurostata (skup podataka: Air passenger transport between reporting countries, šifra: avia_paocc, pristupljeno 2025). Navigacija u Eurostatu: Detailed datasets – Overview of air passenger transport by country and airports (avia_pao) – Air passenger transport between reporting countries (avia_paocc). Iz ovoga su skupa po-

ing, fine clustering, outlier detection, and aggregate route construction was developed for the identification and approximation of well-traveled flight trajectories (Bombelli *et al*., 2019). Hierarchical clustering, partitional clustering, and K-Shape clustering were tested to find the best method for the analysis of time series air-traffic data over Malaysia. Partitional clustering algorithm using the Dynamic Time Warping proved to be the best (Ghani *et al*., 2020).

While several research papers applied clustering techniques for exploring air traffic patterns (Bombeni *et al*., 2019; Chen *et al*., 2020), most studies focused on trajectory optimization, sectorization, or modernization of air traffic management (ATM). There have been very few attempts to comprehensively examine passenger air traffic volume concerning the COVID-19 pandemic. Precisely, there is no systematic analysis of clustering European countries using multi-year passenger traffic based on the traffic before, during, and after, using a publicly available dataset, such as Eurostat's avia_paocc. This research scientifically contributes to an important gap in identifying differences in recovery patterns for passenger traffic concerning COVID-19 at the macro level using clustering techniques. Furthermore, combining R and JASP in this paper will enhance transparency, replicability, and methodological clarity in this data-driven research.

3. METHODOLOGY

3.1. Air-traffic data

The research used open data from the Eurostat database (dataset: Air passenger transport between reporting countries, code: avia_paocc, accessed 2025). Data navigation tree at Eurostat is Detailed datasets – Overview of air passenger transport by country and airports (avia_pao) – Air passenger transport between reporting countries (avia_paocc). The dataset yielded information on 29 dataka izdvojene informacije za 29 europskih zemalja s potpunim zapisima za tri referentne godine: 2018. (predpandemijska osnovica), 2020. (pandemijska godina) i 2023. (godina postpandemijskog oporavka). Tako pripremljeni skup sadrži tri varijable: godina, broj putnika i zemlja. Varijabla „godina“ ima tri modaliteta (2018., 2020., 2023.), „broj putnika“ je numerička diskretna varijabla koja označuje ukupne vrijednosti, a „zemlja“ je nominalna varijabla s 29 modaliteta, kao što je prikazano u Tablici 1. Ove varijable omogućuju konstruiranje obilježja koja zahvaćaju kako razinu putničkog prometa tako i relativne promjene tijekom tri odabrane godine. Umjesto modeliranja dugoročnih trendova, analiza se usredotočuje na obrasce poremećaja i oporavka povezane s pandemijom COVID-19. Metodološki pristup temelji se na klasičnom statističkom klasteriranju (k-srednja vrijednost), etabliranoj eksplorativnoj tehnici odabranoj zbog interpretativnosti i transparentnosti.

European countries with complete records for three reference years: 2018 as the pre-pandemic baseline, 2020 as the pandemic year, and 2023 as the post-pandemic recovery year. The prepared dataset, therefore, includes three variables: year, number of passengers, and country. The ‘year’ has three modalities (2018, 2020, 2023), the ‘number of passengers’ is a discrete variable indicating total counts, and the ‘country’ is a nominal variable with 29 modalities as shown in Table 1. These variables allow construction of features that capture both the level of passenger traffic and the relative changes across the three selected years. Rather than modelling long-term trends, the analysis focuses on patterns of disruption and recovery patterns associated with the COVID-19 pandemic. The methodological approach is based on classical statistical clustering (k-means), established exploratory data analysis technique selected for interpretability and transparency.

Tablica 1: Varijable korištene za klaster analizu

Varijabla	Tip varijable	Modaliteti / Min-Max
Godina	Ordinalna	2018., 2020., 2023.
Broj putnika	Numerička, diskretna	[287 787; 236 451 496]
Zemlja	Nominalna	Austrija, Belgija, Bugarska, Hrvatska, Cipar, Češka, Danska, Estonija, Finska, Francuska, Njemačka, Grčka, Mađarska, Island, Irska, Italija, Latvija, Litva, Luksemburg, Malta, Nizozemska, Norveška, Poljska, Portugal, Rumunjska, Slovačka, Slovenija, Španjolska, Švedska

Izvor: rad autora

Table 1: Variables used in the research for clustering

Variable	Variable type	Modalities / Min-Max
Year	Ordinal	2018, 2020, 2023
Number of passengers	Numerical, discrete	[287787; 236451496]
Country	Nominal	Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, and Sweden

Source: Authors’ elaboration

3.2. Metodološki okvir — k-srednja vrijednost klasteriranje

Metodološki okvir kombinira postupke klasteriranja s višestrukim indeksima validnosti. Algoritam k-srednje vrijednosti korišten je za particioniranje standardiziranih obilježja u k klastere, dok je t-SNE primijenjen isključivo radi vizualizacije dobivenih skupina (Roweis *et al*., 2000). Stabilnost modela procijenjena je metodom lakta, Silhouette koeficijentom i Calinski–Harabasz indeksom. Ova kombinacija uravnotežuje jednostavnost i tumačivost te osigurava da analiza zadrži robusnost u postavci s relativno malim brojem jedinica (small-N) i niskom dimenzionalnošću (Akogul i Erisoglu, 2017; Ikotun *et al*., 2023). Postupak klasteriranja k-srednje vrijednosti minimizira unutar-klastersku varijancu, opisanu funkcijom kvadratne pogreške:

$$J = \sum_{j=1}^{k} \sum_{i \in C_j} \left\| x_i - \mu_j \right\|^2 \qquad (1)$$

gdje *k* označava broj klastera; i indeksira opažanja (ovdje: zemlje); C_j je skup opažanja u klasteru j; x_i je vektor obilježja za opažanje i; µ_j je centroid klastera j. Funkcija J predstavlja zbroj kvadrata euklidskih udaljenosti unutar klastera koji algoritam k-srednje vrijednosti minimizira.

Optimalan broj klastera određen je metodom lakta, koja identificira točku nakon koje daljnje povećavanje k ne donosi bitno smanjenje unutar-klasterske varijance. Radi interpretacije, klasteri i zemlje vizualizirani su t-SNE-om (Roweis *et al*., 2000). Valjanost rješenja dodatno je procijenjena pomoću više indeksa. Uz metodu lakta, razmatrani su Akaikeov informacijski kriterij (AIC) i Bayesov informacijski kriterij (BIC), pri čemu niže vrijednosti označavaju bolji kompromis između prilagođenosti i ekonomičnosti. Silhouette pristup (raspon −1 do 1; više je bolje) koji odražava veću koheziju i odvojenost također je primijenjen za procjenu konzistentnosti unutar klastera. Dodatno su

3.2. Methodological framework - k-means clustering

The methodological framework combines clustering procedures with multiple validation indices. K-means was employed to partition standardized features into k clusters, while t-SNE was used exclusively for visualization of the resulting groups (Roweis *et al*., 2000). Model stability was evaluated using the Elbow method, the Silhouette coefficient, and the Calinski–Harabasz index. This combination balances simplicity and interpretability, ensuring that the analysis remains robust in a relatively small-N, low-dimensional setting (Akogul and Erisoglu, 2017; Ikotun *et al*., 2023). The K-means clustering procedure is to minimize intra-cluster variance, described by the squared error function:

$$J = \sum_{j=1}^{k} \sum_{i \in C_j} \left\| x_i - \mu_j \right\|^2 \qquad (1)$$

where *k* is the number of clusters, *i* indices an observation (here: country), *Cj* is the set of observations assigned to cluster *j*, *xi* is the feature vector for observation *i*, and *µj* (or *cj*) is the centroid of cluster *j.* The function *J* represents the within-cluster sum of squared Euclidean distances minimized by the k-means algorithm.

The optimal number of clusters was determined using the Elbow method, which identifies the point at which further increases in *k* do not substantially reduce within-cluster variance. To aid interpretation, clusters and participating countries were visualized using t-SNE (Roweis *et al*., 2000,). The validity of the clustering solution was further assessed through multiple indices. Besides the Elbow method, also considered were the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), where lower values indicate a better balance between model fit and parsimony. The Silhouette approach (range –1 to 1; higher is better) reflecting primijenjeni Dunnov indeks (više je bolje), Calinski–Harabasz indeks (vrijednosti > 100 upućuju na dobro razlučive klastere), Pearsonov γ (raspon −1 do 1, s višim vrijednostima koje odražavaju jaču unutarnju konzistentnost) i entropija (0 = savršeno klasteriranje; > 0 = viša nesigurnost) (Hendrawati *et al*., 2019; Ikotun *et al*., 2025). Skupno, ti pokazatelji čine robustan okvir za vrednovanje stabilnosti klaster rješenja.

Empirijska implementacija oslonila se na kombinaciju R i JASP, dva otvorena statistička okruženja. R je korišten za pripremu podataka, izračun postotnih promjena i grafičke prikaze, a JASP je poslužio za deskriptivnu statistiku, metodu lakta i evaluaciju klastera (Akogul i Erisoglu, 2017; Ikotun *et al*., 2023). Inicijalni skup podataka (široki format: zemlje × godine) preoblikovan je u dugi format u R-u radi lakše manipulacije. Brojevi putnika za 2018., 2020. i 2023. godinu poslužili su kao primarna obilježja. Uz to su izračunate postotne promjene za razdoblja između 2018. i 2020., 2020. i 2023. te 2018. i 2023. godine kako bi se obuhvatili šok i dinamika oporavka. Prije klasteriranja, sva obilježja standardizirana su radi usporedivosti.

Kombiniranjem apsolutnih razina i stopa promjene, metodologija nadilazi statičko klasteriranje obujma te uključuje usporedna ponašanja oporavka. Ovi relativni indikatori (u nastavku: postotne promjene klastera) kvantificiraju poremećaj i oporavak u razdobljima između 2018. i 2020., 2020. i 2023. te 2018. i 2023. godine. Takav pristup omogućuje identifikaciju heterogenih obrazaca otpornosti među europskim zemljama. Na kraju, klasteriranje putničkih podataka, potvrđeno višestrukim indeksima, daje nove empirijske dokaze o postpandemijskom oporavku zračnog prometa. Ne samo da primjenjuje utvrđene statističke tehnike, već ih i operacionalizira unutar okvira specifičnog za pojedinu domenu, doprinoseći tako empirijskim spoznajama i praktičnim implikacijama politike u okviru prometa i turizma u Europi.

higher cohesion and separation was also applied to evaluate consistency within clusters. Additional indices included the Dunn index (higher is better), the Calinski–Harabasz index (values >100 indicate distinct clusters), Pearson's γ (range 1 to 1, with higher values reflecting stronger internal consistency), and entropy values (0 indicates perfect clustering, >0 higher uncertainty) (Hendrawati *et al*., 2019; Ikotun *et al*., 2025). Together, these measures provided a robust framework for evaluating the stability of the cluster solution.

The empirical implementation relied on a combination of R and JASP, two open-source statistical platforms. R was used for data wrangling, calculation of percentage changes, and graphical presentation, while JASP provided a transparent interface for descriptive statistics, Elbow analysis, and cluster evaluation (Akogul and Erisoglu, 2017; Ikotun *et al*., 2023). The initial dataset, organized in wide format (countries × years), was reshaped into long format in R for easier manipulation. Passenger numbers for 2018, 2020, and 2023 served as the primary features. In addition, percentage changes between 2018 and 2020, 2020 and 2023, and 2018 and 2023 were computed to capture both the disruptive shock and subsequent recovery dynamics. These features were standardized before clustering to ensure comparability across variables.

By combining absolute levels with change-rate measures, the methodology moves beyond static clustering of volumes and incorporates comparative recovery behaviors. These relative indicators, hereafter referred to as cluster percentage changes, capture disruption and recovery between 2018 and 2020, 2020 and 2023, and 2018 and 2023. This approach enables the quantitative identification of heterogeneous resilience patterns among European countries. Ultimately, passenger data based clustering, validated by multiple indices, provides new empirical evidence on post-pandemic air traffic recovery. It not only applies the established statistical techniques but also operational-

4. REZULTATI

Ovaj odjeljak prikazuje rezultate deskriptivne statistike, metode lakta i grafičke rezultate klastera t-SNE-a. Nadalje, uključeni su rezultati valjanosti i evaluacijskih metrika klasteriranja s jasno definiranim vremenskim komponentama te njihovim grafičkim prikazima. Na kraju se varijacije godišnjih promjena prikazuju i grafički i numerički, uz statističku ocjenu.

Deskriptivna statistika prikazana je za podatke o zračnom prometu po državama i godinama (Tablica 2). Radi usporedivosti, broj država zadržan je na 29 tijekom cijelog trogodišnjeg razdoblja jer su sve imale raspoložive podatke. Podaci su prikazani u logaritmiranom obliku kako bi se olakšala interpretacija. Koeficijent asimetričnosti upućuje na približnu simetričnost uz male pozitivne i negativne odmake, koji su zbog svoje veličine zanemarivi. S druge strane, negativna spljoštenost ukazuje na „lakše repove" u odnosu na normalnu raspodjelu, tj. manju zastupljenost ekstremnih vrijednosti. Ipak, procjene su osjetljive na veličinu uzorka te ih treba tumačiti oprezno. Shapiro–Wilkov test normalnosti nije značajan, što upućuje na to da se hipoteza o normalnosti raspodjele ne odbacuje.

izes them within a domain-specific framework, thus contributing to the empirical literature and to practical policy implications in the European transport and tourism sectors.

4. EMPIRICAL FINDINGS

This section includes the results of descriptive statistics, the Elbow method, and t-SNE cluster graphical results. Furthermore, the results of cluster validity and evaluation metrics with a defined time component and their graphical representation will be included. Lastly, yearly change variations will be imposed through both graphical and numerical values with statistical evaluation.

Descriptive statistics is shown for air traffic data for countries per year (Table 2). For easier analysis, the number of countries remained at 29 for the whole three-year period because they had data available for the given period. The data is shown in the log format to ease the data representation. Skewness coefficient shows that the data is symmetric with small differences in positive and negative values, but they are disregarded due to their small size. Conversely, negative kurtosis indicates lighter tails relative to normal distribution, i.e. lower presence of extreme values. Nevertheless, the estimates are sample-size sensitive and should be interpreted cautiously. The Shapiro-Wilk normality test is non-significant, which indicates normal data distribution.

Tablica 2: Deskriptivna statistika zračnog prometa u promatranim godinama

Metrika/Godine	2018.	2020.	2023.
Validnost	29	29	29
Nedostaje	0	0	0
Medijan	7,348	6,821	7,361
Srednja vrijednost	7.341	6,756	7,327
Std. devijacija	0,566	0,607	0,579
Koeficijent varijacije	0,077	0,090	0,079
Asimetrija	0,023	-0,124	-0,054
Std. pogreška asimetrije	0,434	0,434	0,434
Spljoštenost	-0,577	-0,509	-0,449
Std. pogreška spljoštenosti	0,845	0,845	0,845
Shapiro-Wilk	0,974	0,975	0,982
P-vrijednost Shapiro-Wilk	0,669	0,687	0,888
Minimum	6,258	5,459	6,103
Maksimum	8,348	7,764	8,374

Izvor: rad autora

Table 2: Descriptive statistics (air traffic data) for the investigated years

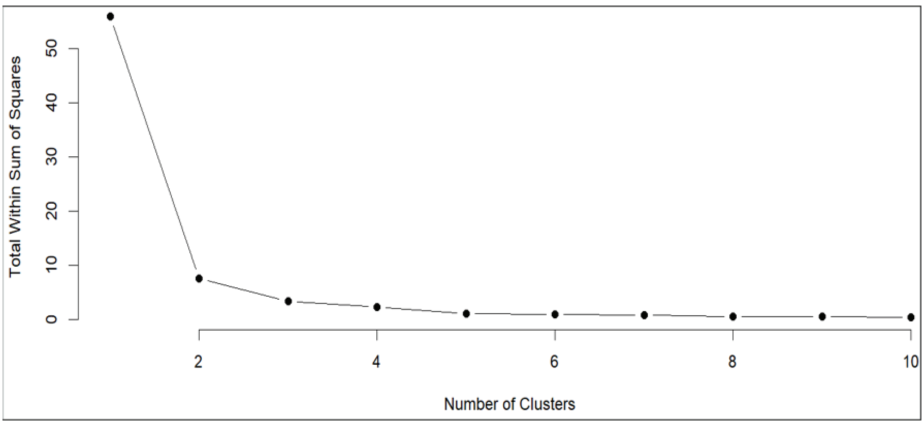
Metrics/Years	2018	2020	2023
Valid	29	29	29
Missing	0	0	0
Median	7.348	6.821	7.361
Mean	7.341	6.756	7.327
Std. Deviation	0.566	0.607	0.579
Coefficient of variation	0.077	0.090	0.079
Skewness	0.023	-0.124	-0.054
Std. Error of Skewness	0.434	0.434	0.434
Kurtosis	-0.577	-0.509	-0.449
Std. Error of Kurtosis	0.845	0.845	0.845
Shapiro-Wilk	0.974	0.975	0.982
P-value of Shapiro-Wilk	0.669	0.687	0.888
Minimum	6.258	5.459	6.103
Maximum	8.348	7.764	8.374

Source: Authors' elaboration

Slika 1 prikazuje krivulju metode lakta za istraživanje klastera. Krivulja upućuje na to da se smanjenje unutar-klasterske varijance usporava kod rješenja s tri klastera. Isti je zaključak dobiven za sve tri promatrane godine.

Figure 1 shows the Elbow method curve for research clusters. The curve shows that the decrease of intra-cluster variance slows down at the three-cluster solution. The same conclusion was yielded for all three research years.

Slika 1: Metoda lakta za istraživane klastere / Figure 1: The Elbow method for research clusters



Izvor: izrada autora / Source: Authors'elaboration

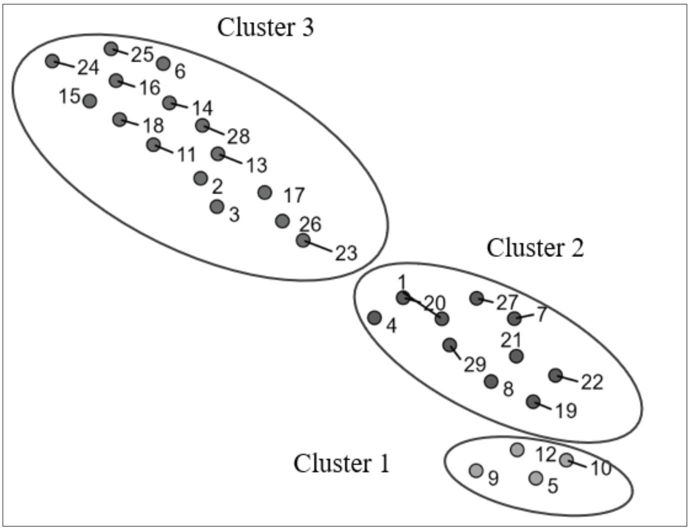
Dijagram t-SNE prikazuje položaj zemalja i pripadnost klasterima za tri izdvojena podatkovna klastera. Metodologija pojednostavljuje optimizaciju i poboljšava vizualnu interpretaciju (Roweis *et al*., 2000). Vrijednosti su prikazane na Slici 2. Zemlje su tijekom sve tri godine ostale u istom klasteru, što znači da ih je COVID-19 mogao promijeniti po poretku unutar klastera, ali ne i po pripadnosti skupini.

Klaster 1: (zemlje br. 5, 9, 10, 12): Njemačka (DE), Francuska (FR), Španjolska (ES), Italija (IT). Klaster 2; (zemlje br. 1, 4, 7, 8, 19, 20, 21, 22, 27, 29): Nizozemska (NL), Belgija (BE), Poljska (PL), Austrija (AT), Grčka (GR), Portugal (PT), Švedska (SE), Norveška (NO), Finska (FI), Danska (DK). Klaster 3 (zemlje br. 2, 3, 6, 11, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26, 28): Hrvatska (HR), Slovenija (SI), Slovačka (SK), Mađarska (HU), Bugarska (BG), Rumunjska (RO), Latvija (LV), Litva (LT), Estonija (EE), Cipar (CY), Malta (MT), Luksemburg (LU), Češka (CZ), Irska (IE), Island (IS).

The t-SNE chart denotes the countries' position and cluster representation given the three data clusters. This methodology simplifies optimization and improves visual representation (Roweis *et al*., 2000). The values can be seen in Figure 2. The countries remained in the same cluster across the three years, meaning that COVID-19 did not change their group but might change their order within the cluster.

Cluster 1: (countries no. 5, 9, 10, 12) includes Germany (DE), France (FR), Spain (ES), and Italy (IT). Cluster 2: (countries no. 1, 4, 7, 8, 19, 20, 21, 22, 27, 29) brings together the Netherlands (NL), Belgium (BE), Poland (PL), Austria (AT), Greece (GR), Portugal (PT), Sweden (SE), Norway (NO), Finland (FI), and Denmark (DK). Cluster 3 (countries no. 2, 3, 6, 11, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26, 28) consists of Croatia (HR), Slovenia (SI), Slovakia (SK), Hungary (HU), Bulgaria (BG), Romania (RO), Latvia (LV), Lithuania (LT), Estonia (EE), Cyprus (CY), Malta (MT), Luxembourg (LU), Czechia (CZ), Ireland (IE), and Iceland (IS).

Slika 2: t-SNE prikaz klastera za istraživane klastere / Figure 2: t-SNE Cluster plot for researched clusters



Izvor: izrada autora / Source: Authors' elaboration

Pokazatelji valjanosti klasteriranja K-srednje vrijednosti prikazani su u Tablici 3. Koeficijent determinacije je visok, iznosi 94 %, što znači da model adekvatno objašnjava varijacije zavisne varijable. Metrike uspješnosti modela, poput Akaikeovog informacijskog kriterija (AIC), Bayesovog informacijskog kriterija (BIC) i vrijednosti Silhouette pristupa, jasno su vidljive (Hendrawati *et al*., 2019). Vrijednosti AIC-a (22,970) i BIC-a (35,280) odražavaju kompromis između prilagođenosti modela i njegove složenosti, pri čemu nešto viša vrijednost BIC-a upućuje na konzervativniju kaznu za složenost modela. Silhouette koeficijent od 0,650 (na ljestvici na kojoj se vrijednosti iznad 0,5 smatraju umjereno snažnima, a vrijednosti iznad 0,7 snažnima) ukazuje na umjereno dobru razdvojenost među klasterima.

K-means clustering validity indicators are presented in Table 3. The coefficient of determination is high, standing at 94 %, meaning that the model properly explains the variations in the dependent variable. Model performance metrics such as the Akaike information criterion (AIC), Bayesian information criterion (BIC), and the Silhouette approach values are visible (Hendrawati *et al*., 2020). The AIC (22.970) and BIC (35.280) values reflect the trade-off between model fit and complexity, with slightly higher BIC values suggesting a conservative penalty for model complexity. The Silhouette score of 0.650 (on scale where values above 0.5 are considered moderately strong and values above 0.7 are considered strong) indicates moderately good separation among the clusters.

Tablica 3: Sažetak modela – k-srednja vrijednost klasteriranje

Klaster	N	R^2	AIC	BIC	Silhouette
3	29	0,941	22,970	35,280	0,650

Izvor: rad autora

Table 3: Model Summary - K-means clustering

Clusters	N	R^2	AIC	BIC	Silhouette
3	29	0.941	22.970	35.280	0.650

Source: Authors' elaboration

Tablica 4 prikazuje metrike uspješnosti klastera za svaku godinu. Vrijednost Silhouette koeficijenta najviša je za Klaster 3, što znači da je on najreprezentativniji od tri klastera. Klaster 3 najhomogenija je i najoštrije razgraničena skupina jer ostvaruje najvišu Silhouette vrijednost (0,776) te najniži zbroj kvadrata unutar klastera (0,444). Klaster 1 također pokazuje relativno dobru razdvojenost (Silhouette = 0,673), ali je najmanje brojan (4 države). Klaster 2 slabije je kohezivan (Silhouette = 0,464) i ima veći zbroj kvadrata (1,813), ali obuhvaća samo 10 drža-

Table 4 presents cluster performance metrics for each year. The Silhouette score value is the highest for Cluster 3, meaning it is the most representative of the three. Cluster 3 is the most homogeneous and well-discriminated group because it presents the highest Silhouette value (0.776) and the lowest within-cluster sum of squares value (0.444). Cluster 1 also has quite good separation (0.673 of Silhouette), but it is far less populated (4 countries). Cluster 2 is less cohesively bonded (Silhouette score 0.464) and has a larger sum of squares (1.813), but only va, što upućuje na određeni stupanj unutarnje disperzije.

Metrike u Tablici 5 potvrđuju da sam model ima zadovoljavajuću kvalitetu. Dunnov indeks kvantificira odnos razdvojenosti i kompaktosti klastera (zabilježena vrijednost 0,154) te ukazuje na znatno preklapanje klastera i samo umjerenu kompaktnost. Stoga opažena vrijednost od 0,154 sugerira da postoji određeni stupanj preklapanja i da bi klasteri mogli biti jasnije razgraničeni. Calinski-Harabaszov indeks, s vrijednošću 206,748, implicira dobru razdvojenost i vrlo kvalitetno klasteriranje. Umjerena klasterska struktura vidljiva je i iz Pearsonova γ koeficijenta od 0,565. Konačno, entropija od 0,981 upućuje na visok stupanj miješanja među klasterima, tj. da, iako je klasteriranje valjano, i dalje postoji određena nesigurnost u pogledu točne pripadnosti pojedinih promatranih jedinica klasterima. Ukupni rezultati potvrđuju solidne metrike uspješnosti modela.

Postotne promjene po klasterima prikazane su u Tablici 6, pri čemu je 2020. godina uzeta kao referentna. Klaster 1 (n = 4), koji okuplja najveća zrakoplovna tržišta (DE, FR, ES, IT), zabilježio je prosječan pad od –72,5 % između 2018. i 2020. godine, nakon čega je uslijedio oporavak od +261 % u razdoblju od 2020. do 2023 godine. Unatoč tom oporavku, obujam prometa u 2023. godini i dalje je bio nešto ispod razine iz 2018. godine. Klaster 2 (n = 10), koji obuhvaća srednje velika tržišta, pokazao je sličnu kontrakciju od –71,3 % 2020. godine te naknadni oporavak od +253 % do 2023. godine, pri čemu su razine prometa gotovo dosegnule predpandemijske vrijednosti. Klaster 3 (n = 15), koji predstavlja manja tržišta, doživio je najizrazitiji pad (–75,1 %), ali i najsnažniji oporavak (+301 %), a razine prometa u 2023. premašile su one iz 2018. godine.

Zajedno promatrani, ovi rezultati potvrđuju izražene razlike među klasterima. Iako su sva tri klastera snažno pogođena pandemijom COVID-19, najveća tržišta (Klaster 1) i srednje velika tržišta (Klaster 2) pokazala su sporiji i djelomičan oporavak, dok su

contains 10 countries, which suggests some internal dispersion.

The metrics in Table 5 confirm that the model itself is of satisfactory quality. The Dunn index quantifies separation versus compactness (0.154 observed value) with an indication of substantial cluster overlap and only moderate compactness. Therefore, the observed value of 0.154 shows that there is some degree of overlap and that the clusters could be more distinct. The Calinski-Harabasz index, with the value of 206.748, implies good separation and excellent clustering. A moderate clustering structure is noticeable by Pearson's coefficient γ of 0.565. Lastly, the entropy of 0.981 suggests a high degree of cluster mixing, meaning that although valid, some uncertainty in exact cluster assignments remains. Therefore, the overall results confirm fair model performance metrics.

Cluster percentage changes are presented in Table 6, whereby 2020 was used as the benchmark year. Cluster 1 (n = 4) that groups the largest aviation markets (DE, FR, ES, IT), experienced an average decline of –72.5 % between 2018 and 2020, which was followed by a rebound of +261 % from 2020 to 2023. Despite this recovery, the air traffic volumes in 2023 remained slightly below the 2018 levels. Cluster 2 (n = 10), comprising medium-sized markets, showed a similar contraction of –71.3 % in 2020 and a subsequent recovery of +253 % by 2023, reaching near pre-pandemic traffic levels. Cluster 3 (n = 15), representing smaller markets, underwent the steepest decline (–75.1 %) as well as the strongest rebound (+301 %) with the air traffic volumes in 2023 surpassed those from 2018.

Cumulatively, these results confirm marked cross-cluster contrasts. While all three groups were severely affected by the COVID-19 pandemic, the largest markets (Cluster 1) and medium-sized markets (Cluster 2) showed slower and only partial recovery, whereas smaller markets in Cluster 3, al-manja tržišta u Klasteru 3, unatoč najjačem početnom padu, ostvarila najbrži oporavak te čak premašila predpandemijske razine broja putnika.

Tablica 4: Metrike klastera

Klaster	1	2	3
Veličina	4	10	15
Heterogenost unutar klastera	0,546	0,365	0,089
Zbroj kvadrata	2,712	1,813	0,444
Silhouette	0,673	0,464	0,776
Centar 2018.	2,298	-0,035	-0,589
Centar 2020.	2,281	0,002	-0,610
Centar 2023.	2,289	-0,020	-0,597

Izvor: rad autora

Tablica 5: Metrike modela

Metrike	Vrijednost
Maksimum dijametar	1,889
Minimum separacija	0,290
Pearson's γ	0,565
Dunnov indeks	0,154
Entropija	0,981
Calinski-Harabasz indeks	206,748

Izvor: rad autora

Tablica 6: Klaster promjene (%)

Klaster	Prosj_ promjena_ 2018_2020	Prosj_ promjena_ 2020_2023
1	-72,50 %	261 %
2	-71,30 %	253 %
3	-75,10 %	301 %

Izvor: rad autora

Nakon prikaza ukupnih postotnih promjena za sve klastere, autori se detaljnije usmjeravaju na Klaster 3, budući da je ova skupina zabilježila i najizraženiju kontrakciju i najsnažniji oporavak. Dok su Klaster 1 i

though having experienced the largest initial decline, rebounded fastest and even exceeded pre-pandemic passenger volumes.

Table 4: Cluster performance metrics

Cluster	1	2	3
Size	4	10	15
Explained proportion within-cluster heterogeneity	0.546	0.365	0.089
Within sum of squares	2.712	1.813	0.444
Silhouette score	0.673	0.464	0.776
Center 2018	2.298	-0.035	-0.589
Center 2020	2.281	0.002	-0.610
Center 2023	2.289	-0.020	-0.597

Source: Authors' elaboration

Table 5: Model performance metrics

Metrics	Value
Maximum diameter	1.889
Minimum separation	0.290
Pearson's γ	0.565
Dunn index	0.154
Entropy	0.981
Calinski-Harabasz index	206.748

Source: Authors' elaboration

Table 6: Cluster changes (%)

Cluster	Avg_change_ 2018_2020	Avg_change_ 2020_2023
1	-72,50 %	261 %
2	-71,30 %	253 %
3	-75,10 %	301 %

Source: Authors' elaboration

Following the summary of overall percentage shifts across all clusters, the paper focuses on examining Cluster 3 in more detail as this group exhibited both the sharpest contraction and the strongest rebound. While Clusters 1 and 2 adhered to more moderate

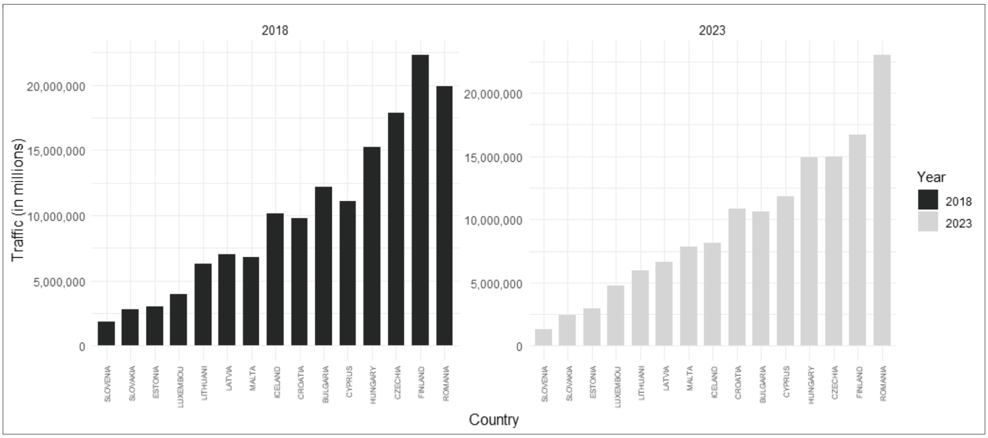
Klaster 2 slijedili umjerenije obrasce, ostajući do 2023. godine bliže predpandemijskim razinama, zemlje u Klasteru 3 pokazale su izrazito heterogene obrasce oporavka, što opravdava dodatnu analizu. Stoga Slike 3–5 i Tablice 7 i 8 prikazuju dinamiku unutar Klastera 3, ističući poremećaje i oporavak specifičan za pojedine zemlje.

Slika 3 prikazuje vizualni prikaz zračnog prometa za Klaster 3 u 2018. i 2023. godini po zemljama. Vidljive su vrijednosti oporavka, pri čemu su neke zemlje zabilježile daljnji rast, dok su druge i dalje ispod početnih razina prije pandemije COVID-19. Općenito gledano, razine zračnog prometa su se oporavile ili su na putu potpunog oporavka.

patterns remaining closer to pre-pandemic levels by 2023, Cluster 3 countries displayed highly heterogeneous recovery patterns, which warranting further examination. Figures 3–5 and Tables 7 and 8, therefore, illustrate the dynamics within Cluster 3, highlighting country-specific disruptions and rebounds.

Figure 3 shows the visual presentation of air traffic for Cluster 3 in the 2018–2023 period countries. There are evident recovery values visible with some countries experiencing further growth, while others are still below their initial values before the COVID-19 pandemic. Overall, the air traffic values have recovered or are advancing toward complete recovery.

Slika 3: Usporedba Klastera 3 između 2018. i 2023. godine po državama / Figure 3: Cluster 3 comparison between 2018 and 2023 by countries



Izvor: autori, obrada u programu R / Source: Authors' elaboration using R

Oporavak zemalja Klastera 3 između 2020. i 2023. godine djeluje snažno u većini promatranih slučajeva, kako je prikazano u Tablici 7. U prosjeku je zabilježen pad vrijednosti zračnog prometa od 80 % u 2020. godini, pri čemu su među najteže pogođenima bile Slovenija (–84 %), Slovačka (–82 %) i Island (–85 %). Kada se u analizu uključi 2023. godina, vidljiv je izrazit obrazac oporavka, osobito u Hrvat-

The recovery of Cluster 3 countries between 2020 and 2023 appears strong across most of the observed countries as shown in Table 7. On average, in 2020 there was an 80 % drop air traffic values whereby Slovenia (-84 %), Slovakia (-82 %), and Iceland (-85 %) were among the hardest affected. If 2023 is factored into analysis, a significant recovery pattern is evident – especially across Croaskoj (+455 %), na Cipru (+408 %) i na Islandu (+429 %). Nadalje, neke zemlje, poput Luksemburga (+20 %), Malte (+15 %) i Rumunjske (+16 %), čak su nadmašile razine iz 2018. godine, što upućuje na potpuni oporavak i rast.

tia (+455 %), Cyprus (+408 %), and Iceland (+429 %). Furthermore, some countries, such as Luxembourg (+20 %), Malta (+15 %), and Romania (+16 %), even surpassed their 2018 levels, indicating full recovery and growth.

Tablica 7: Promjene u Klasteru 3 (broj putnika i postotak)

Država	2018.	2020.	2023.	Promjena prometa 2018./2020.	Promjena prometa 2020./2023.	Promjena prometa 2018./2023.
Bugarska	12.181.375	3.738.156	10.583.590	-69 %	183 %	-13 %
Hrvatska	9.798.678	1.958.355	10.868.005	-80 %	455 %	11 %
Cipar	11.095.888	2.327.823	11.828.636	-79 %	408 %	7 %
Češka	17.893.941	3.834.479	15.002.832	-79 %	291 %	-16 %
Estonija	2.995.830	858.165	2.946.323	-71 %	243 %	-2 %
Finska	22.268.840	5.477.611	16.673.588	-75 %	204 %	-25 %
Mađarska	15.212.355	3.965.443	14.921.822	-74 %	276 %	-2 %
Island	10.167.367	1.531.223	8.101.686	-85 %	429 %	-20 %
Latvija	7.039.116	1.995.459	6.620.657	-72 %	232 %	-6 %
Litva	6.259.643	1.809.106	5.994.734	-71 %	231 %	-4 %
Luksemburg	3.988.804	1.426.183	4.793.798	-64 %	236 %	20 %
Malta	6.805.643	1.752.445	7.815.593	-74 %	346 %	15 %
Rumunjska	19.870.674	6.626.452	22.981.524	-67 %	247 %	16 %
Slovačka	2.814.330	501.750	2.435.876	-82 %	385 %	-13 %
Slovenija	1.810.567	287.787	1.268.352	-84 %	341 %	-30 %

Izvor: rad autora

Table 7: Cluster 3 changes (passenger number and percentage)

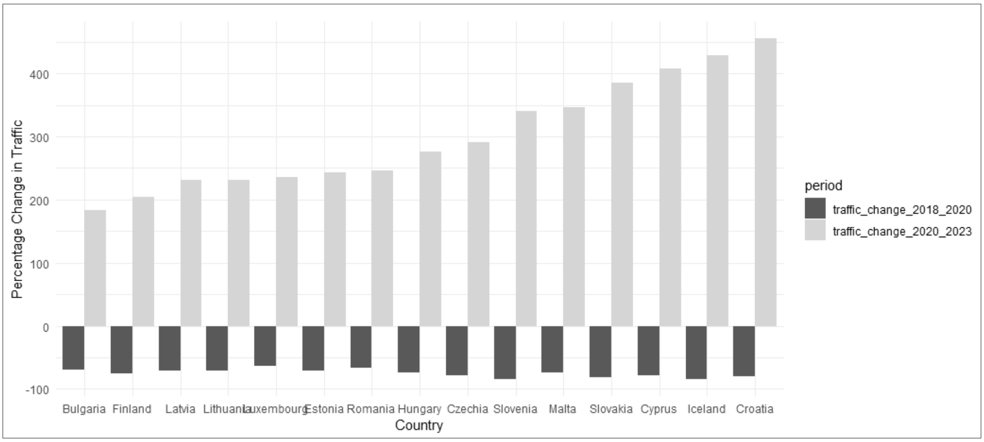
Country	2018	2020	2023	Traffic change 2018/2020	Traffic change 2020/2023	Traffic change 2018/2023
Bulgaria	12,181,375	3,738,156	10,583,590	-69 %	183 %	-13 %
Croatia	9,798,678	1,958,355	10,868,005	-80 %	455 %	11 %
Cyprus	11,095,888	2,327,823	11,828,636	-79 %	408 %	7 %
Czechia	17,893,941	3,834,479	15,002,832	-79 %	291 %	-16 %
Estonia	2,995,830	858,165	2,946,323	-71 %	243 %	-2 %
Finland	22,268,840	5,477,611	16,673,588	-75 %	204 %	-25 %
Hungary	15,212,355	3,965,443	14,921,822	-74 %	276 %	-2 %
Iceland	10,167,367	1,531,223	8,101,686	-85 %	429 %	-20 %
Latvia	7,039,116	1,995,459	6,620,657	-72 %	232 %	-6 %
Lithuania	6,259,643	1,809,106	5,994,734	-71 %	231 %	-4 %
Luxembourg	3,988,804	1,426,183	4,793,798	-64 %	236 %	20 %
Malta	6,805,643	1,752,445	7,815,593	-74 %	346 %	15 %
Romania	19,870,674	6,626,452	22,981,524	-67 %	247 %	16 %
Slovakia	2,814,330	501,750	2,435,876	-82 %	385 %	-13 %
Slovenia	1,810,567	287,787	1,268,352	-84 %	341 %	-30 %

Source: Authors' elaboration

Detaljnija analiza Klastera 3 prikazana je na Slici 4. Vidljivo je da su sve zemlje zabilježile negativne vrijednosti u usporedbi 2018. i 2020. godine, dok se u 2023. godini tri zemlje izdvajaju s najvišim vrijednostima rasta. Riječ je o Hrvatskoj, Islandu i Cipru. Grafički prikaz izrađen je u programskom okruženju R, pri čemu je konstruiran stupčasti dijagram za razdoblja između 2018. i 2020. te između 2020. i 2023. godine.

Further detailed analysis of Cluster 3 is visible in Figure 4. It is evident that all countries yielded negative values when comparing 2018 and 2020, while in 2023 three countries exhibit the most substantial growth values. These countries are Croatia, Iceland, and Cyprus. The graphic representation has been elaborated in R software to create a bar plot for the periods 2018-2020 and 2020-2023.

Slika 4: Promjena prometa u Klasteru 3 (2018.–2020. i 2020.–2023. godina) / Figure 4: Traffic change in Cluster 3 (2018-2020 and 2020-2023)

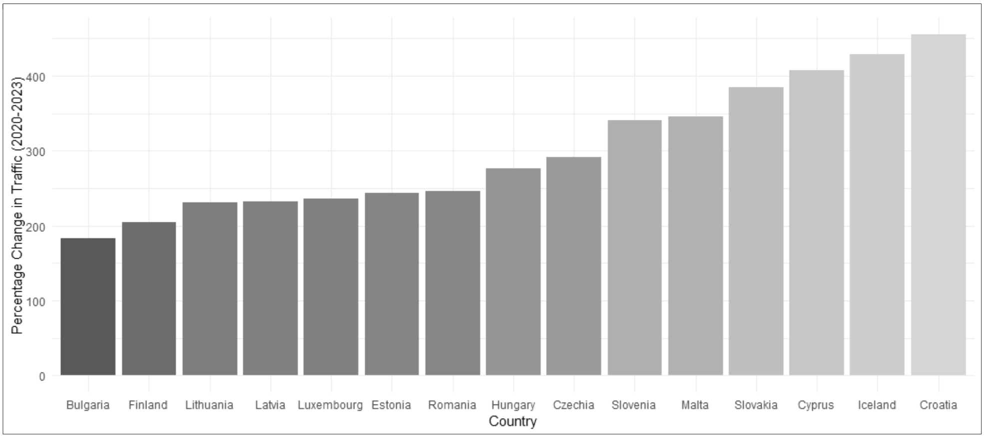


Izvor: autori, obrada u programu R / Source: Authors' elaboration using R

Slika 5 prikazuje koliko su se pojedine zemlje nosile s predpandemijskim i postpandemijskim godinama u odnosu na COVID-19. Rezultati su mješoviti te upućuju na to da neke zemlje u Klasteru 3 još uvijek nisu dosegnule ili nadmašile razine iz 2018. godine. Nadalje, kako bi se mogle uočiti faze oporavka uz pomoć vizualizacije u programu R prikazane na Slici 5, izrađeni su grafički prikazi za Klaster 3 za 2018. i 2023. godinu.

Figure 5 shows how well countries coped with the pre-COVID-19 and post-COVID-19 pandemic years. The results are mixed and indicate that some countries in Cluster 3 have not yet reached or surpassed the values recorded in 2018. Furthermore, in order to facilitate the visualization of recovery phases using R software in Figure 5, Cluster 3 plots for both 2018 and 2023 were created.

Slika 5: Promjena prometa u razdoblju od 2018. do 2023. godine / Figure 5: Traffic change in period between 2018 and 2023



Izvor: autori, obrada u programu R / Source: Authors' elaboration using R

Kako bi se opravdale prikazane promjene, provedena je statistička analiza. Primijenjeni su zavisni t-testovi za usporedbu srednjih vrijednosti zračnog prometa između 2018. i 2020. te 2018. i 2023. godine s df = 28 za N = 29 uparenih parova država. Za usporedbu 2020. i 2023. godine pretpostavke o jednakosti varijanci nisu bile zadovoljene pa je primijenjen Welchov t-test za nejednake varijance, pri čemu je dobiven ne-cjelobrojni stupanj slobode df = 16,541 (Tablica 8). Nalazi upućuju na statistički značajne razlike (p < 0,01) između 2018. i 2020. te između 2020. i 2023. godine, dok usporedba između 2018. i 2023. godine ne pokazuje značajnost. Svi testovi provedeni su u programskom okruženju R, a kod je dostupan u prilogu rada.

To justify the changes shown, statistical analysis was performed. Paired t-tests were implemented to compare mean air traffic values between 2018 vs 2020 and 2018 vs 2023, with df = 28 for N = 29 matched country pairs. When comparing 2020 and 2023, the assumptions of equal variances were not met, and therefore Welch's unequal-variance t-test was applied, resulting in a non-integer df = 16.541 (see Table 8). The results indicate statistically significant differences (p < 0.01) between 2018 and 2020, and 2020 and 2023, while the comparison between 2018 and 2023 does not indicate a statistically significant difference. All tests were conducted in R, and the code is available in the Appendix.

Tablica 8: Statistička analiza Klastera 3 primjenom zavisnog t-testa usporedbe 2018. i 2023., 2018. i 2020. te 2020. i 2023. godine

Razdoblje	t-vrijednost	Stupnjevi slobode	p-vrijednost	Srednja razlika	95 %-tni interval pouzdanosti (donja granica)	9 5%-tni interval pouzdanosti (gornja granica)
2018. u odnosu na 2023.	0,3877	28	0,7012	642.838,3	-2.753.607	-2.753.607 do 4.039.283
2018. u odnosu na 2020.	40,061	28	0,000413	-3,4E+07	-33.831.394	51.129.990 do -16.532.798
2020. u odnosu na 2023.	-42,912	16,541	0,000523	-6.982.106	-6.982.106	-10.423.677 do -3.542.533

Izvor: Obrada korištenjem R koda

Table 8: Cluster 3 statistical analysis using paired t-test across 2018 vs. 2023, 2018 vs. 2020, and 2020 vs. 2023

Period	t-Value	Degrees of freedom	p-Value	Mean difference	95 % confidence interval (lower limit)	95 % confidence interval (upper limit)
2018 vs. 2023	0.3877	28	0.7012	642,838.3	-2,753,607 to 4,039,283	-2,753,607 to 4,039,283
2018 vs. 2020	40.061	28	0.0004134	-33,831,394	-51,129,990 to -16,532,798	51,129,990 to -16,532,798
2020 vs. 2023	-42.912	16.541	0.0005231	-6,982,106	-10,423,677 to -3,542,533	-10,423,677 do -3,542,533

Source: Authors' elaboration using R

Statistička analiza pruža čvrste dokaze o heterogenom utjecaju pandemije COVID-19 na europski zračni promet. U 2020. godini, broj putnika u zemljama Klastera 3, koje se pretežito sastoje od malih i srednje velikih tržišta, smanjio se u prosjeku za 72 %. Ova snažna kontrakcija odražava kombinirani učinak međunarodnih ograničenja putovanja, nacionalnih *lockdowna* i urušavanja potražnje za međunarodnom mobilnošću (Maneenop i Suntichai, 2023). Značajnost tog pada potvrđena je zavisnim t-testovima,

The statistical analysis provides robust evidence of the heterogeneous impact of the COVID-19 pandemic on European air traffic. In 2020, passenger volumes in Cluster 3 countries, consisting mainly of small and medium-sized markets, declined on average by 72 %. This sharp contraction reflected the combined effects of international travel restrictions, national lockdowns, and collapsing demand for international mobility (Maneenop and Suntichai, 2023). The significance of this decline is confirmed by pri čemu su p-vrijednosti ispod praga od 1 %. Međutim, do 2023. godine ista skupina zemalja bilježi snažan oporavak, s prosječnim povećanjem od 9,5 milijuna putnika u odnosu na 2020. godinu. Ovaj oporavak statistički je značajan na razini od 1 % (t = –42,912; p = 0,0005), a 95 %-tni interval pouzdanosti (–10.423.677 do –3.542.533) potvrđuje robusnost rezultata. Nasuprot tomu, razlika između 2018. i 2023. godine nije statistički značajna, što upućuje na to da se europski zračni promet u velikoj mjeri oporavio, ali još uvijek nije premašio predpandemijske razine. Ovi nalazi u velikoj su mjeri u skladu s ranijim istraživanjima koja pokazuju neujednačen i nedovršen oporavak globalnog zrakoplovstva (Sun *et al*., 2021; To i Lee, 2024).

Iz perspektive turizma ovi obrasci oporavka naglašavaju različite izazove i prilike među skupinama zemalja. Prethodna istraživanja pokazuju da zračna povezanost izravno oblikuje turističke tokove, konkurentnost destinacija i otpornost regionalnih turističkih sustava (ACI Europe, 2023; Tang *et al*., 2022). Klaster 1, koji uključuje neka od najvećih europskih tržišta, i dalje je ispod obujma iz 2018. godine te se suočava s produljenim razdobljem obnove potražnje. To je u skladu s dokazima o usporenom oporavku velikih tržišnih čvorišta zbog njihove ovisnosti o međukontinentalnim tokovima (Frontier Economics, 2023). Zemlje Klastera 2, sa srednje snažnim obrascima oporavka, približile su se predpandemijskim razinama. Za te destinacije ključni je prioritet javnih politika očuvanje održivosti linija te jačanje konkurentnosti u sve nestabilnijem turističkom okruženju. Suprotno tome, zemlje Klastera 3, iako najteže pogođene 2020., do 2023. godine ostvarile su najsnažniji oporavak. Ova dinamika potvrđuje ranije nalaze da manja i perifernija tržišta, premda osjetljivija na vanjske šokove, mogu pokazati brz oporavak nakon ukidanja ograničenja (Maneenop *et al*., 2023). Za te zemlje izazov više nije poticanje potražnje, nego osiguravanje održivog rasta i upravljanje sezonskom vo-

paired t-tests, with p-values below the 1 % threshold. By 2023, however, the same group of countries exhibited a strong rebound, with an average increase of 9.5 million passengers compared to 2020. This recovery is statistically significant at the 1 % level (t = –42.912, p = 0.0005), and the 95 % confidence interval (–10,423,677 to –3,542,533) confirms the robustness of the results. In contrast, the difference between 2018 and 2023 was not statistically significant, which suggests that, while European air traffic has largely recovered, it has not yet exceeded pre-pandemic volumes. These findings are broadly consistent with previous studies documenting uneven and incomplete recovery in global aviation (Sun *et al*., 2021; To and Lee, 2024).

From the tourism perspective, these recovery patterns highlight differentiated challenges and opportunities across country groups. Prior research shows that air connectivity directly shapes tourist flows, destination competitiveness, and the resilience of regional tourism systems (ACI Europe, 2023; Tang *et al*., 2022). Cluster 1, which includes some of the largest European markets, remains below 2018 volumes and faces prolonged demand rebuilding. This aligns with evidence of delayed recovery in large hub markets due to their dependence on long-haul flows (Frontier Economics, 2023). Cluster 2 countries, with moderate recovery patterns, are close to restoring pre-pandemic levels. For these destinations, the key policy priority is ensuring route viability and sustaining competitiveness in an increasingly volatile tourism environment. By contrast, Cluster 3 countries, although the most severely affected in 2020, experienced the strongest rebound by 2023. This dynamic confirms earlier observations that smaller and peripheral markets, while more vulnerable to external shocks, can display rapid rebounds once restrictions are lifted (Maneenop *et al*., 2023). For these countries, the challenge is no longer demand stimulation but ensuring sustainable growth and addressing seasonal volatillatilnošću. Nadalje, hipoteza H1 je potvrđena, budući da je analiza potvrdila prisutnost triju statistički značajnih klastera. Hipoteza H2 također je potvrđena, s obzirom na značajne razlike u intenzitetu oporavka između 2020. i 2023. godine. Hipoteza H3, međutim, samo je djelomično potvrđena: iako su pojedine zemlje, poput Luksemburga, Malte i Rumunjske, premašile obujme iz 2018. godine, većina zemalja u Klasterima 1 i 2 u 2023. godini ostaje ispod predpandemijskih razina. Ukupni rezultati potvrđuju heterogenu prirodu europskog oporavka zračnog prometa te naglašavaju važnost ciljanih politika.

Konačno, osim sadržajnog doprinosa području zrakoplovstva i turizma, rad doprinosi i razvoju informacijskih znanosti. Primjenom metode k-srednje vrijednosti, validacijom rezultata pomoću više pokazatelja te korištenjem ponovljivih analitičkih postupaka u programima R i JASP, analiza pokazuje kako se otvoreni podaci i statističke metode mogu operacionalizirati u interdisciplinarnom kontekstu. Metodološki doprinos pokazuje kako informacijske tehnologije i pristupi utemeljeni na podacima mogu poduprijeti donošenje odluka utemeljenih na dokazima u turizmu, prometu i kriznom menadžmentu, osnažujući vezu između primijenjenih društvenih znanosti i informacijskih znanosti.

5. ZAKLJUČAK

Autori su u ovom radu ispitali utjecaj pandemije COVID-19 na europski putnički zračni promet primjenom metode k-srednje vrijednosti na podatke Eurostata za godine 2018., 2020. i 2023. godine. Mjere uspješnosti modela potvrdile su primjerenost dobivenog rješenja klasteriranja, a na istraživačko pitanje – Kako su se europske zemlje grupirale prema obrascima post-COVID oporavka u putničkom zračnom prometu? – odgovoreno je identifikacijom triju jasno razlučivih klastera. Klaster 3 pokazao je najvišu unutarnju homogenost (91,1 %), što je opravdalo njegovu detaljniju analizu. Unatoč snažnom

ity. Furthermore, hypothesis H1 is supported, as the analysis confirmed the presence of three statistically meaningful clusters. Hypothesis H2 is also supported, given the significant variation in recovery magnitudes between 2020 and 2023. Hypothesis H3, however, is only partially supported, while a few countries, such as Luxembourg, Malta, and Romania, surpassed their 2018 volumes, most countries in Clusters 1 and 2 remained below pre-pandemic levels in 2023. Overall outcomes confirm the heterogeneous nature of European air-traffic recovery and emphasize the importance of targeted policy approaches.

Finally, beyond its substantive contribution to aviation and tourism, the paper also advances the field of information science. By applying k-means clustering, validating results with multiple indices, and employing reproducible workflows in R and JASP, the analysis demonstrates how open data and statistical methods can be operationalized in an interdisciplinary context. This methodological contribution illustrates how information technologies and data-driven approaches can support evidence-based decision-making in tourism, transport, and crisis management, strengthening the bridge between applied social sciences and information science.

5. CONCLUSION

The authors of this paper examined the impact of the COVID-19 pandemic on European air passenger traffic by applying k-means clustering to Eurostat data for the years 2018, 2020, and 2023. The model performance metrics confirmed the adequacy of the clustering solution, and the research question: How did European countries cluster according to their post-COVID recovery patterns in air passenger traffic? – was addressed through the identification of three distinct clusters. Cluster 3 displayed the highest internal homogeneity (91.1 %), which justified its further detailed analysis. Despite padu 2020. godine, izražen oporavak do 2023. godine upućuje na to da su mjere poput kampanja cijepljenja, ublažavanja ograničenja putovanja i obnove potražnje bile učinkovite u vraćanju zračne povezanosti. Statistički testovi potvrdili su značajne razlike između 2020. i 2023. godine, dok između 2018. i 2023. godine nije utvrđena značajna razlika, što upućuje na zaključak da se europski zračni promet uvelike vratio na predpandemijske razine, ali ih još nije trajno nadmašio.

Rezultati potvrđuju hipoteze H1 i H2, potvrđujući postojanje statistički značajnih klastera i razlika u putanjama oporavka, dok je hipoteza H3 djelomično potvrđena jer većina zemalja u Klasterima 1 i 2 još nije premašila obujme iz 2018. godine. Rezultati doprinose teorijskoj raspravi o otpornosti i oporavku u zrakoplovstvu nudeći izvornu, podatkovno utemeljenu klasifikaciju post-COVID obrazaca među europskim zemljama.

Iz perspektive turizma, rezultati naglašavaju važnost zračne povezanosti za konkurentnost destinacija te neujednačenu brzinu oporavka potražnje u Europi. Zemlje u sporije oporavljajućim klasterima trebat će ciljane marketinške aktivnosti i potporne javne politike za obnovu međunarodnih turističkih tokova, dok se zemlje u brže oporavljajućim klasterima suočavaju s izazovom upravljanja održivim rastom i sezonskom volatilnošću. Ove spoznaje nude praktične smjernice donositeljima politika, zrakoplovnim prijevoznicima i menadžerima u turizmu u oblikovanju diferenciranih strategija oporavka i jačanju pripravnosti za buduće krize.

Osim doprinosa turizmu i zrakoplovstvu, rad pridonosi i informacijskim znanostima demonstrirajući primjenu tehnika klasteriranja, otvorenih podataka i ponovljivih istraživačkih tokova u R-u i JASP-u na domenski specifičan problem. Ovo interdisciplinarno gledište pokazuje kako podatkovno-analitički pristupi mogu povezati prometne studije, istraživanja u turizmu i informacijske znanosti, podupirući donošenje odluka utemeljenih na dokazima u više sektora.

the severe contraction in 2020, the strong rebound observed by 2023 indicates that measures such as vaccination campaigns, the easing of travel restrictions, and renewed demand were effective in restoring air connectivity. Statistical tests confirmed significant differences between 2020 and 2023, while no significant differences were found between 2018 and 2023, suggesting that European air traffic has largely regained its pre-pandemic levels without yet surpassing them.

The findings support H1 and H2, confirming the existence of statistically meaningful clusters and significant differences in recovery trajectories, while H3 is only partially supported, as most countries in Clusters 1 and 2 have not exceeded their 2018 volumes. These results contribute to the theoretical debate on resilience and recovery in aviation by providing an original, data-driven classification of European countries' post-COVID patterns.

From a tourism perspective, the results underline the importance of air connectivity for destination competitiveness and the uneven speed of demand recovery across Europe. Countries in slower-recovering clusters will need targeted marketing efforts and policy support to rebuild international tourist flows, while those in faster-recovering clusters face the challenge of managing sustainable growth and seasonal volatility. These insights offer practical guidance for policymakers, airlines, and tourism managers in shaping differentiated recovery strategies and strengthening preparedness for future crises.

Beyond tourism and aviation, the paper also contributes to information science by demonstrating the application of clustering techniques, open data, and reproducible workflows in R and JASP to a domain-specific problem. This interdisciplinary perspective illustrates how data-analytic approaches can bridge transport studies, tourism research, and information science, supporting evidence-based decision-making across sectors.

Rad, naravno, ima ograničenja. Usmjeren je na europske zemlje za koje su bili dostupni konzistentni podaci, bez eksplicitnog uvažavanja gospodarske i regulatorne heterogenosti koja također može oblikovati oporavak. Strategije zračnih prijevoznika, cijene goriva i drugi vanjski šokovi bili su isključeni. Buduća istraživanja trebala bi proširiti analizu na druge svjetske regije, integrirati dodatne eksplanatorne varijable te ugraditi biheviorаlne dimenzije poput promjena u poslovnim putovanjima i turističkoj potražnji. Takvi bi napori produbili razumijevanje postpandemijske otpornosti zračnog prometa i pružili dodatne smjernice za prometne i turističke politike.

The paper is not without limitations. The focus was on European countries for which consistent data were available, without accounting for economic or regulatory heterogeneity that may also shape recovery. Airline strategies, fuel prices, and other external shocks were excluded. Future research should extend the analysis to other world regions, integrate additional explanatory variables, and incorporate behavioral dimensions such as shifts in business travel and tourism demand. Such efforts would broaden the understanding of post-pandemic aviation resilience and provide further guidance for both transport and tourism policy.

LITERATURA – REFERENCES

ACI EUROPE & SEO Amsterdam Economics (2024) *Benefits of airports & air connectivity: The economic and social impact of European airports and air connectivity*. Brussels: ACI EUROPE. Available at: https://www.aci-europe.org/downloads/resources/ACI%20EUROPE%20SYNOPSIS%20-%20BENEFITS%20OF%20AIRPORTS%20%20AIR%20CONNECTIVITY%20digital.pdf (Accessed: 4 November 2025)

ACI EUROPE (2023) *Airport industry connectivity report 2023*. Brussels: ACI EUROPE. Available at: https://www.aci-europe.org/downloads/resources/ACI%20EUROPE%20Airport%20Industry%20Connectivity%20Report%202023.pdf (Accessed: 4 November 2025)

Akogul, S. & Erisoglu, M. (2017) An approach for determining the number of clusters in a model-based cluster analysis. *Entropy, 19*(9), 452. DOI: https://doi.org/10.3390/e19090452

Andrienko, G., Andrienko, N., Fuchs, G. & Garcia, J. M. (2018) Clustering trajectories by relevant parts for air traffic analysis. *IEEE Transactions on Visualization and Computer Graphics, 24*(1), 34-44. DOI: https://doi.org/10.1109/tvcg.2017.2744322

Bolić, T., Castelli, L., De Lorenzo, A. & Vascotto, F. (2022) Trajectory clustering for air traffic categorization. *Aerospace, 9*(5), 227. DOI: https://doi.org/10.3390/aerospace9050227

Bombelli, A., Segarra-Torné, E., Trumbauer, K. D. & Mease, K. D. (2019) Improved clustering for route-based Eulerian air traffic modeling. *Journal of Guidance, Control, and Dynamics, 42*(5). DOI: https://doi.org/10.2514/1.G003939

Chandra, S., Hazra, A., Verma, A. & Sooraj, K. P. (2024) On the possibilities of efficient air traffic monitoring through complex network clustering-based airspace sub-sectorization: A multi-objective discrete particle swarm optimization approach. *Journal of the Transportation Research Board, 3*, 100036. DOI: https://doi.org/10.1177/03611981241263829

Chen, J.-H., Wei, H.-H., Chen, C.-L., Wei, H.-Y., Chen, Y.-P. & Ye, Z. (2020) A practical approach to determining critical macroeconomic factors in air-traffic volume based on K-means clustering and decision-tree classification. *Journal of Air Transport Management, 82*, p. 101743. DOI: https://doi.org/10.1016/j.jairtraman.2019.101743

Di Luigi, G., Lulli, G. & Lancia, C. (2022) Data-driven optimization for Air Traffic Flow Management with trajectory preferences. *arXiv:2211.06526v1* (Cornell University). DOI: https://doi.org/10.48550/arXiv.2211.06526

Eurostat (2025) *Air transport measurement – passengers and freight*. Eurostat. Available at: DOI: https://doi.org/10.2908/AVIA_PAOCC (Accessed: 4 November 2025)

Frontier Economics (2023) *Uneven post-pandemic take-off for Europe's airlines. London: Frontier Economics*. Available at: https://www.frontier-economics.com/uk/en/news-and-insights/articles/article-i20675-uneven-post-pandemic-take-off-for-europe-s-airlines/(Accessed: 4 November 2025).

Ghani, W., Nursyirwan, I. F., Ali, M. N. & Jusoh, W. N. W. (2020) Comparison of time series clustering approaches on air traffic congestion pattern. *International Journal of Advanced Science and Technology, 29*(6), 1837-1847.

Hendrawati, T., Wigena, A. H., Sumertajaya, I. M. & Sartono, B. (2019) Performance evaluation of AIC and BIC in time series clustering with Piccolo method. Proceedings of the EAI International Conference on … (conference paper). DOI: https://doi.org/10.4108/eai.2-8-2019.2290340

Ikotun, A. M., Habyarimana, F. & Ezugwu, A. E. (2025) Cluster validity indices for automatic clustering: A comprehensive review. *Heliyon, 11*(2), e41953. DOI: https://doi.org/10.1016/j.heliyon.2025.e41953

Ikotun, M., Ezugwu, A. E., Abualigah, L., Abuhaija, B. & Heming, J. (2023) K-means Clustering Algorithms: A comprehensive review, variants analysis, and advances in the era of Big Data. *Information Sciences, 622*, 178-210. DOI: https://doi.org/10.1016/j.ins.2022.11.139

Jangik, J. & Rafferty, P. (2021) How The Speed Limit Policy Affects Travel Speed?: Quasi-experimental Approach. *Transport Policy 103*(January), 2-10. DOI: https://doi.org/10.1016/j.tranpol.2021.01.003.

Khalil, D., Prasad, A., Motlicek, P., Zuluaga-Gomez, J., Nigmatulina, I., Madikeri, S. & Schuepbach, C. (2023) An automatic speaker clustering pipeline for the Air Traffic Communication domain. *Aerospace, 10*(10), p. 876. DOI: https://doi.org/10.3390/aerospace10100876

Le Brigant, A. & Puechmorel, S. (2019) Quantization and clustering on Riemannian manifolds with an application to air traffic analysis. *Journal of Multivariate Analysis, 173*, 685-703. DOI: https://doi.org/10.1016/j.jmva.2019.05.008

Li, J.-Q., Zhang, H.-Y., Zhang, Y. & Liu, H.-T. (2018) Systematic assessment of intrinsic factors influencing visual attention performances in air traffic control via clustering algorithm and statistical inference. *PLOS ONE, 13*(10). DOI: https://doi.org/10.1371/journal.pone.0205334

Lui, G. N., Klein, T. & Liem, R. P. (2020) Data-driven approach for aircraft arrival flow investigation at terminal maneuvering area. *AIAA Aviation 2020 Forum*. DOI: https://doi.org/10.2514/6.2020-2869

Maneenop, S. & Suntichai, K. (2023) Comparative Evidence on the Impact of the COVID-19 Outbreak and Vaccine Discovery on the Global Transportation Industry. *Transportation Research Interdisciplinary Perspectives, 17*(January), 100741. DOI: https://doi.org/10.1016/j.trip.2022.100741.

Mohd Kamal, S. B., Yunos, M., Akasah Ismail, M. & Lahap, J. (2024) Post-pandemic aviation travel: The alteration of travel behaviour. *International Journal of Academic Research in Business and Social Sciences, 14*(11). DOI: https://doi.org/10.6007/ijarbss/v14-i11/23445

Olive, X. & Morio, J. (2019) Trajectory clustering of air traffic flows around airports. *Aerospace Science and Technology, 84*, 776-781. DOI: https://doi.org/10.1016/j.ast.2018.11.031

Roweis, S. T. & Lawrence, K. S. (2000) Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science, 290*(5500), 2323-2326. DOI: https://doi.org/10.1126/science.290.5500.2323

Suau-Sanchez, P., Voltes-Dorta, A. & Cugueró-Escofet, N. (2020) An Early Assessment of the Impact of COVID-19 on Air Transport: Just Another Crisis or the End of Aviation as We Know It?. *Journal of Transport Geography, 86*(June), 102749. DOI: https://doi.org/10.1016/j.jtrangeo.2020.102749.

Sun, X., Wandelt, S. & Zhang, A. (2020) How did covid-19 impact air transportation? A first peek through the lens of complex networks. *Journal of Air Transport Management, 89*, 101928. DOI: https://doi.org/10.1016/j.jairtraman.2020.101928

Sun, X., Wandelt, S., Zheng, C. & Zhang, A. (2021) COVID-19 Pandemic and Air Transportation: Successfully Navigating the Paper Hurricane. *Journal of Air Transport Management, 94*(April), 102062. DOI: https://doi.org/10.1016/j.jairtraman.2021.102062.

Swinney, J. & Woods, J. C. (2022) K-means clustering approach to UAS classification via graphical signal representation of radio frequency signals for Air Traffic Early warning. *IEEE Transactions on Intelligent Transportation Systems, 23*(12), 24957-24965. DOI: https://doi.org/10.1109/tits.2022.3202011

Tang, X., Liu, Y. & Chen, K. (2022) Air traffic trajectory operation mode mining based on clustering. *Applied Sciences, 12*(12), 5944. DOI: https://doi.org/10.3390/app12125944

To, W. M. & Lee, P. K. C. (2024) Modeling of the COVID-19 impact on air passenger traffic in the US, European countries, and China. *Journal of Air Transport Management, 115*, 102556. 2024. DOI: https://doi.org/10.1016/j.jairtraman.2024.102556

Zhang, Z., Zhang, A., Sun, C., Guan, J. & Huang, X. (2020) The reliability analysis of air traffic network based on trajectory clustering of terminal area. *IEEE Access, 8*, 75035-75042. DOI: https://doi.org/10.1109/access.2020.2988586

Dodatak: R kod

Sljedeći isječci ilustriraju glavni analitički tijek rada: podaci su transformirani u „long" format, izračunate su postotne promjene, klasteri su vizualizirani, a zavisni t-testovi primijenjeni su za usporedbe uparenih zemalja (df = N−1). Cjeloviti R skripti dostupni su na zahtjev ili kao dopunski materijal.

Appendix A: R code

The excerpts below illustrate the main workflow: data were reshaped into long format, percentage changes were computed, clusters were visualized, and paired t-tests were applied for matched country comparisons (df = N−1). Full R scripts are available upon request or as supplementary material.

```
# --- Data reshaping: from wide (countries × years) to long format ---
wide_data <- wide_data %>%
  tibble::rownames_to_column(var = "country")
long_data <- wide_data %>%
  pivot_longer(cols = c(`2018`, `2020`, `2023`),
               names_to = "year", values_to = "traffic") %>%
  mutate(year = as.integer(year))
# --- Percentage change calculations (example for 2018–2020 and 2020–2023) ---
changes <- long_data %>%
  pivot_wider(names_from = year, values_from = traffic) %>%
  mutate(
    pct_18_20 = (`2020` - `2018`) / `2018` * 100,
    pct_20_23 = (`2023` - `2020`) / `2020` * 100
 )
# --- Visualization: bar plot for traffic changes in Cluster 3 countries ---
ggplot(cluster3_long_changes, aes(x = country, y = pct_change, fill = period)) +
  geom_col(position = "dodge") +
  labs(title = "Traffic change for Cluster 3 countries",
       x = "Country", y = "Change (%)") +
  theme_minimal()
# --- Example of paired t-test (2018 vs. 2023, matched by country) ---
t_test_result <- t.test(wide_data$`2018`, wide_data$`2023`, paired = TRUE)
print(t_test_result)
```

Primljeno: 29. svibnja 2025. / Submitted: 29 May 2025
Prihvaćeno: 3. studenoga 2025. / Accepted: 3 November 2025