MAPPING THE LANDSCAPE OF TRANSPORT INFRASTRUCTURE AND REGIONAL DEVELOPMENT: A COMPREHENSIVE BIBLIOMETRIC ANALYSIS

Thomas KRABOKOUKIS

Department of Planning and Regional Development, University of Thessaly, 382 21 Volos, Greece
tkrabokoukis@uth.gr

Serafeim POLYZOS

Department of Planning and Regional Development, University of Thessaly, 382 21 Volos, Greece
spolyzos@uth.gr

Dimitrios KANTIANIS

Department of Economic and Regional Development, Panteion University, 176 71 Athens, Greece
dkant@panteion.gr

Abstract

The presented herein bibliometric analysis delves into the intricate relationship between transport infrastructure and regional development, aiming to fill a gap in the existing literature. Utilizing a number of widely recognized bibliometric techniques, such as citation analysis, co-authorship analysis, co-occurrence analysis, and bibliographic coupling analysis, the study provides a holistic overview of key trends, research clusters, as well as methodological approaches in the field. Notably, the citation analysis reveals a diverse array of highly cited papers, ranging from green logistics to the impacts of COVID-19 on the global economy, underscoring the multidimensional nature of research in the domain. The co-authorship analysis illuminates complex collaborative networks among countries, with prominent clusters such as the European and the Asian-Pacific clusters, highlighting the dynamic and interconnected nature of international collaborations. The co-occurrence analysis identifies four thematic clusters, emphasizing the interconnectedness of infrastructure, regional development, economic growth, environmental considerations, and spatiotemporal dynamics. Methodological approaches, dominated by economic theory and quantitative methods, are also classified. Addressing limitations, including the potential bias introduced by specific keywords and database restrictions, the research provides valuable insights for researchers, practitioners, and policymakers, paving the way for further refinements in understanding the evolving landscape of transport infrastructures and regional development research.

Keywords: transport infrastructure; economic development; regional development; bibliometric analysis; cluster analysis; VOSviewer.

1. INTRODUCTION

In general, efficient transport systems play a key role in economic development both nationally and regionally. In many cases transport infrastructure is seen as an effective mean to boost the regional economy by improving transport links both within the target region and to or from other regions (He and Duchin, 2009; Venables et al. 2014). Transport infrastructure also plays a crucial role in supporting the urban economy and society, sometimes acting as catalysts for improving the standard of living in urban areas (Bronzini and Piselli, 2009; Kasikoen et al., 2019; Polyzos and Tsiotas, 2020).

In previous research work, the relationship between transport infrastructure and regional development was analyzed both from a theoretical perspective (Polyzos and Tsiotas, 2020) and a methodological approach (Polyzos and Tsiotas, 2023). In continuation of the previous work, in this article, an extensive bibliographic analysis is applied to identify trends and patterns in the impact of transport infrastructure on regional development.

Transportation stands as an indispensable infrastructure within contemporary society, exerting a profound influence on mobility, commercial activities, and the overall economic well-being of nations. Nevertheless, the exacerbating challenges posed by issues such as traffic congestion, soaring fuel expenses, and increased CO_2 emissions underscore the imperative need for immediate improvements in transportation safety and efficiency (Guerrero-Ibanez et al., 2015). In addition to these challenges, infrastructure plays a crucial role in elevating productivity, and enhancing the attractiveness of a region for the establishment of economic activities, thus influencing the sectoral development of the region (Polyzos, 2019; Krabokoukis and Polyzos, 2020; Krabokoukis and Polyzos, 2021).

Moreover, the impact of infrastructure extends to the reduction of production costs for businesses and the augmentation of their profits in correlation with the degree of infrastructure utilization as a facilitator in the production process. Specifically, superior infrastructure contributes to the concentration of businesses, economies of scale, and increased productivity levels (Polyzos, 2019). The multifaceted role of infrastructure in promoting regional and sectoral development underscores its significance as a critical driver of economic growth and prosperity.

Recognized as a fundamental instrument in the execution of economic and regional development policies, e.g., European Union's (EU) 2020 Strategy (Kantianis, 2021), infrastructure projects and particularly transportation initiatives, play a critical role in strategies directed at underdeveloped regions. The overarching goal of every regional development policy is to foster conditions conducive to heightened regional productivity and economic growth. Investments in projects and regional infrastructure emerge as indispensable means to achieve these developmental objectives. This recognition is particularly salient in

the context of less developed regions, where public infrastructure investments are acknowledged as pivotal tools for effective regional development (Polyzos, 2019). In such a framework, they can serve as instruments to address and mitigate the challenges inherent in the regional problem (Krabokoukis and Polyzos, 2020).

The synthesis of these perspectives sheds light on the critical nexus between transportation infrastructure, regional policies, and the challenges faced by underdeveloped regions, emphasizing the need for comprehensive and forward-looking strategies. The multifaceted relationship between transport infrastructure and regional economic growth unfolds through its influence on the geography of economic activities, urban land use patterns, and regional competitiveness (Venables et al., 2014; Tsiotas and Polyzos, 2018; Krabokoukis and Polyzos, 2022). Within this complexity of interactions, analyses become imperative to understand the varied impacts of transport infrastructure investments on regional development. Therefore, such analyses need to consider the relevant direct, indirect, and induced effects over time (Polyzos, 2019; Rosik and Wojcik, 2022). This comprehensive appreciation is essential for informing effective policy decision-making and facilitating sustainable development in underdeveloped regions.

Despite the recognition of these challenges, current literature lacks a thorough bibliometric analysis of today's state of research in transport infrastructure and regional development. The study aims to bridge this gap by conducting an extensive bibliometric analysis, offering an overview of authors, key trends, research groups and gaps in the field. These can be very important in identifying opportunities for further study, complementing previous studies that examined the relationship between transport infrastructure and regional development in a theoretical framework (Polyzos and Tsiotas, 2020) and from a methodological point of view (Polyzos and Tsiotas, 2023). This analysis will serve as a valuable resource for researchers, practitioners and policy makers seeking to address the complex issues at the intersection of transport and regional development. The study is based on a comprehensive sample of articles published by international journals and being indexed in the Scopus database. The main research questions are the following:

1. What are the key trends in research on transport infrastructures and regional development?

2. How do patterns of co-authorship among countries reflect collaborations in the examined field?

3. What are the main research clusters and thematic areas identified in the co-occurrence analysis of keywords, and how do these clusters enhance our understanding of the examined field?

4. What are the predominant methodological approaches highlighted in the field under investigation?

The research methods employed encompassed the application of several widely utilized bibliometric techniques, including citation analysis, co-authorship analysis, co-occurrence analysis, and bibliographic coupling analysis. Additionally, an analysis of methodological approaches was undertaken utilizing the Python programming language. It is necessary however, to acknowledge the inherent limitations of the bibliometric approach. Firstly, the scope of findings could be broadened by incorporating other well-established databases, such as the Web of Science (WoS) database. Secondly, the use of keywords might be puzzling (Seguí-Amortegui et al., 2019). Thirdly, potential limitations arise from the exclusion of non-English language articles. Lastly, the study's database was constrained to: i) articles as the document type, and ii) journals as the source type (Krabokoukis, 2023).

Next, the manuscript is structurally organized as follows. A comprehensive account of the research methods employed in the study, encompassing details on the sample, data collection, and analysis techniques, is presented in the second section. The subsequent section delves into the study's findings, emphasizing key discoveries and conclusions drawn from the analysis. Finally, the paper concludes by discussing the findings of the study, offering useful insights to assist researchers, practitioners, and policymakers actively engaged in the realm of the examined field.

2. METHDOLOGY AND DATA

In this study, the methodology and data analysis were executed using the VOSviewer tool and the Python programming language. The detailed procedures for bibliometric analysis are elaborated in Subsections from 2.1. to 2.6.

2.1. Dataset

The research encompasses the timeframe from 2010 to the 26th of November 2023. The database construction relied on the utilization of Scopus Keywords, and detailed information about the database is presented in Table 1. The decision to commence the analysis in 2010 was motivated by the aim to capture a thorough and contemporary understanding of the topic. This starting point is chosen for two primary reasons. Firstly, it marks a significant period post the global financial crisis, with subsequent years reflecting diverse economic recovery initiatives and policy changes. Secondly, advancements in technology impacting transportation infrastructure and regional development policies have been notably pronounced in recent years.

TABLE 1 – DATA RETRIEVAL CONSTRAINTS AND PARAMETERS FOR THE SCOPUS DATABASE

Data Information	Search Criteria and Results
Database:	Scopus
Search field:	Title, Abstract, Keywords
Keywords:	transport*, transportation infrastructure, transportation networks, regional transportation policies, transportation impact, regional transport systems, regional development, regional growth, regional economies, regional planning, regional connectivity, economic growth, spatial development, regional accessibility
Open access:	All
Years:	2010-2023
Author name:	All
Subject area:	Social sciences; Environmental science; Engineering; Energy; Economics, Econometrics and Finance; Business, Management and Accounting; Computer science; Earth and Planetary Science; Agricultural and Biological Sciences; Decision Sciences; Mathematics; Psychology; Neuroscience.
Publication stage:	Final
Document type:	Article
Source title:	All
Affiliation:	All
Funding sponsor:	All
Country:	All
Source type:	Journal
Language:	English
Search string:	(TITLE-ABS-KEY ("transport*" OR "transportation infrastructure" OR "transportation networks" OR "regional transportation policies" OR "transportation impact" OR "regional transport systems") AND TITLE-ABS-KEY ("regional development" OR "regional growth" OR "regional economies" OR "regional planning" OR "regional connectivity" OR "economic growth" OR "spatial development" OR "regional accessibility")) AND PUBYEAR > 2009 AND PUBYEAR < 2024 AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (SUBJAREA , "SOCI") OR LIMIT-TO (SUBJAREA , "ENVI") OR LIMIT-TO (SUBJAREA , "ENGI") OR LIMIT-TO (SUBJAREA , "ENER") OR LIMIT-TO (SUBJAREA , "ECON") OR LIMIT-TO (SUBJAREA , "BUSI") OR LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "EART") OR LIMIT-TO (SUBJAREA , "AGRI") OR LIMIT-TO (SUBJAREA , "DECI")) AND (LIMIT-TO (LANGUAGE , "English"))
Data extracted:	26 November 2023
Number of publications:	4,202

2.2 Citation Analysis

Citation analysis stands as a prevalent technique in bibliometric analysis. It involves evaluating the impact and influence of a specific paper or author in a particular field of study by quantifying the frequency with which the paper or author is referenced in other scholarly works (Seguí-Amortegui et al., 2019; Krabokoukis, 2023). This form of analysis provides a reliable measure of the reach and significance of scientific research.

2.3 Co-Authorship Analysis

Co-authorship analysis stands out as a widely acknowledged method for unveiling collaborative patterns among authors, institutions, and countries across various fields of study (Seguí-Amortegui et al., 2019). Beyond its foundational role in recognizing these collaborative trends, co-authorship analysis serves as a valuable instrument for illuminating complex networks of researchers. It aids in identifying influential figures, institutions, and emerging trends within the subject of inquiry (Krabokoukis, 2023; Krabokoukis and Polyzos, 2023).

2.4 Co-Occurence Analysis

Co-occurrence analysis proves to be a valuable method for discerning connections between keywords or concepts frequently found together in literature (Krabokoukis and Polyzos, 2023). The approach to author co-occurrence entails pinpointing the most commonly appearing keywords or concepts within a given set of papers (Krabokoukis, 2023). This method offers a comprehensive understanding of the interrelationships embedded in the scholarly discourse.

2.5 Bibliographic Coupling Analysis

Bibliographic coupling analysis serves to explore the interconnectivity and collaborative synergy among diverse scientific publications, relying on the commonality of their cited references (Krabokoukis and Polyzos, 2023). This analytical method operates as a valuable tool for clustering documents, significantly streamlining the dissemination of scientific knowledge, and improving the efficacy of document retrieval (Krabokoukis, 2023). Furthermore, it provides an extensive perspective on the theoretical foundation within the examined field.

2.6 Co-Citation Analysis

Co-citation analysis stands out as a vital method for the in-depth exploration of academic citations (Krabokoukis and Polyzos, 2023). This approach examines the frequency with which two documents are collectively cited by other scholarly works, serving as a robust indicator of their conceptual or thematic affinity. Its significance extends to identifying central themes and influential sources within a specific research domain, effectively mapping the intellectual landscape (Krabokoukis, 2023). Integrating co-citation analysis into the research framework empowers investigators to discern the subtle currents within the literature, unveiling the enduring impact of specific ideas and the intricate interplay of intellectual forces.

2.7 Methodology Analysis

The study additionally aims to elucidate the diverse methodological approaches employed in the international literature within the examined field. To achieve this objective, a categorization of specific primary methodological approaches was undertaken, and keywords were generated with the assistance of the ChatGPT tool (OpenAI, 2023). Python code was developed to search for these keywords within the abstracts of the database, subsequently classifying them accordingly. Table A in the Appendix presents the Python code that was implemented, and its flowchart presented in Figure 1. This comprehensive approach allows for the discernment of the evolution of methodological preferences and the emergence of innovative techniques, providing valuable insights into the scholarly landscape of the studied period.
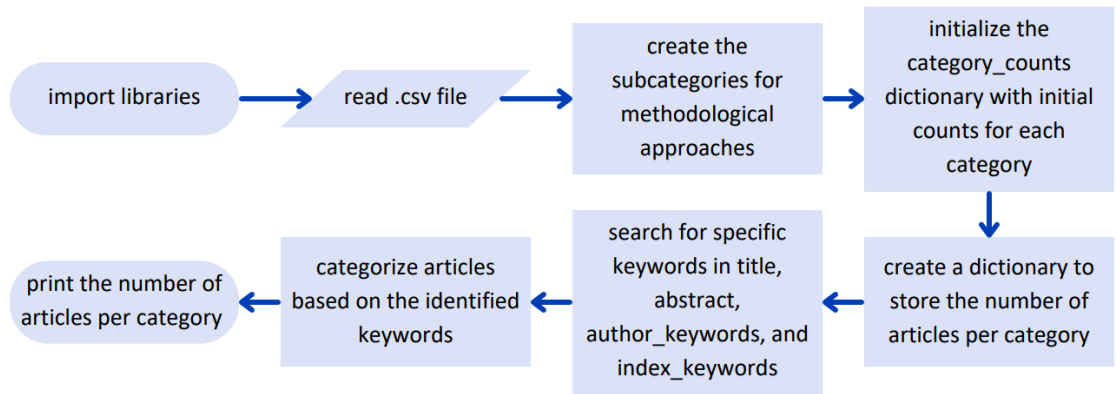
FIGURE 2 – FLOWCHART FROM THE PYTHON SCRIPT

3. RESULTS

3.1 Results of Citation Analysis

The results of the citation analysis are presented in Table 2, which outlines the top-10 documents based on their citation count as of November 26, 2023. Topping the list is the paper by Benjamin et al. (2016) titled "A North American hourly assimilation and model forecast cycle: The rapid refresh" with citation counts of 717. Following the studies by Dekker et al. (2012), Mathiesen et al. (2011), Stafford-Smith et al. (2017), and Ibn-Mohammed et al. (2021), each contributing significantly to the discourse with citation counts of 623, 559, 493, and 422, respectively. These papers span a range of topics, from green logistics and renewable energy systems to the impacts of COVID-19 on the global economy and ecosystems. The diversity of the top-cited documents underscores the multidimensional nature of research in the field, capturing key contributions that have garnered substantial attention within the scholarly community.

TABLE 2 – TOP 10 DOCUMENTS BY CITATIONS (UNTIL 26 NOVEMBER 2023)

No.	Ref.	Title	Citations
1	Benjamin et al., 2016	A North American hourly assimilation and model forecast cycle: The rapid refresh	717
2	Dekker et al., 2012	Operations Research for green logistics – An overview of aspects, issues, contributions and challenges	623
3	Mathiesen et al., 2011	100% Renewable energy systems, climate mitigation and economic growth	559
4	Stafford-Smith et al., 2017	Integration: the key to implementing the Sustainable Development Goals	493
5	Ibn-Mohammed et al., 2021	A critical review of the impacts of COVID-19 on the global economy and ecosystems and opportunities for circular economy strategies	422
6	Liu et al., 2015	High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China from 1990 to 2010	359
7	Guerrero-Ibanez et al., 2015	Integration challenges of intelligent transportation systems with connected vehicle, cloud computing, and Internet of Things technologies	351
8	Duncan et al., 2016	A space-based, high-resolution view of notable changes in urban NOx pollution around the world (2005–2014)	326
9	Umar et al., 2020	COP21 Roadmap: Do innovation, financial development, and transportation infrastructure matter for environmental sustainability in China?	312
10	Alemi et al., 2018	What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California	294

The findings from the analysis of Table 3, showcasing the top-10 authors and countries based on their publication count until November 26, 2023, reveal insights into the global landscape of research in the field. Leading the list of prolific authors is Bekun, F.V., with 11 publications. Following Lin, B., and Chen, Z., with 10 and 9 publications, respectively, signifying their noteworthy engagement in the field.

TABLE 3 – TOP 10 AUTHORS AND COUNTRIES BY PUBLICATIONS (UNTIL 26 NOVEMBER 2023)

No.	Author	Number of Documents	No.	Country	Number of Documents
1	Bekun, F.V.	11	1	China	1,146
2	Lin, B.	10	2	United States	708
3	Chen, Z.	9	3	United Kingdom	352
4	Khan, S.A.R.	9	4	Australia	200
5	Adedoyin, F.F.	8	5	India	188
6	Ewing, R.	8	6	Russian Federation	166
7	Fu, X.	8	7	Spain	148
8	Sharif, A.	8	8	Japan	147
9	Ahmed, Z.	7	9	Germany	137
10	Fernandes, E.	7	10	Italy	129

Shifting the focus to countries, China emerges as the dominant force, occupying the top spot with a remarkable 1,146 publications. The United States and the United Kingdom follow with substantial contributions of 708 and 352 publications, respectively. Australia, India, and the Russian Federation are included in the top 5, emphasizing their active involvement in the global research landscape.

3.2 Co-authorship Analysis

Figure 2 displays a co-authorship map of all countries in the Scopus dataset. The map is based on a threshold minimum number of documents of a country of 5, and 82 out of the 147 meet this threshold. The higher total link strengths are for China (647), United States (443), United Kingdom (386), Australia (181), Pakistan (173), Spain (126), France (117), Canada (116), Turkey (114), Malaysia (108), Netherlands (108), Germany (103), Hong Kong (101), and Italy (101). The total link strengths, as indicated by the numbers in parentheses, reflect the cumulative collaboration strength between countries in the co-authorship network. Higher values signify a greater volume of collaborative research output. Notably, countries like China, the United States, and the United Kingdom exhibit exceptionally high total link strengths, indicating prolific collaboration and substantial contributions to global research networks in the field. These countries, along with others listed, emerge as key contributors with extensive international research partnerships, underscoring their prominent roles in the collaborative landscape of the examined research domain.
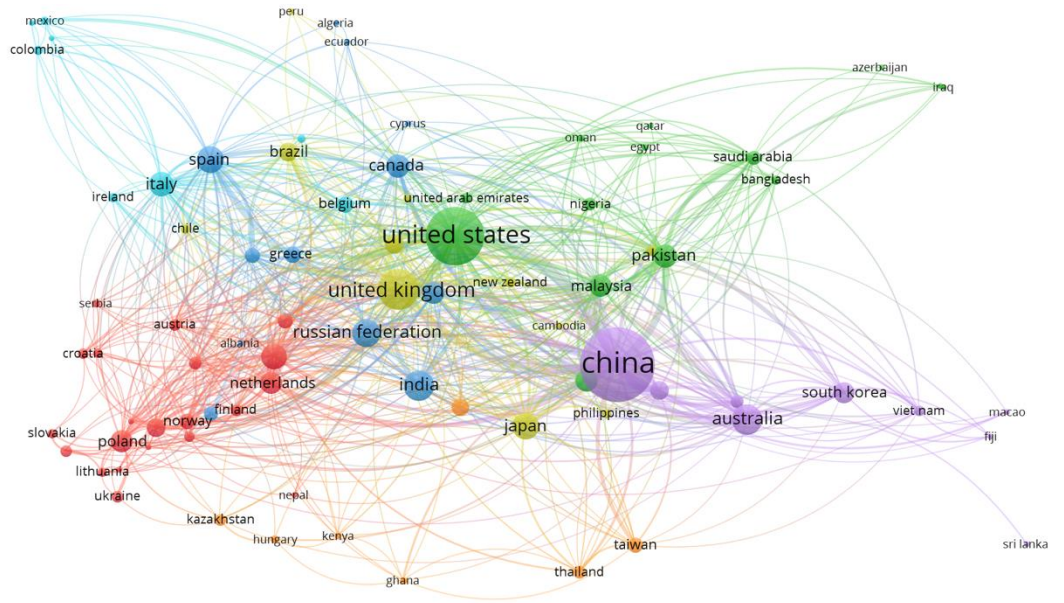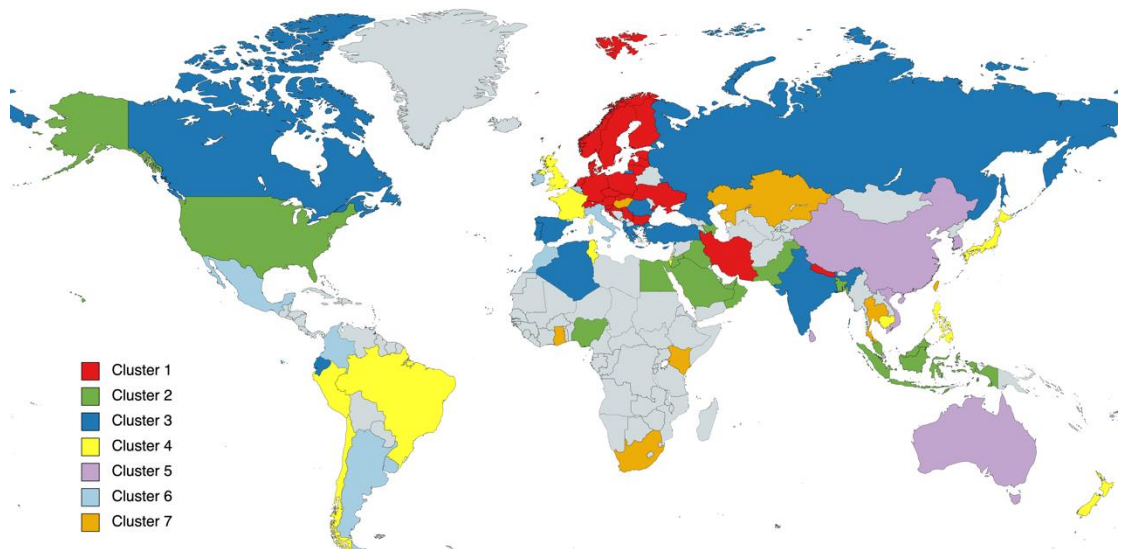
FIGURE 2 – CO-AUTHORSHIP NETWORK ANALYSIS OF COUNTRIES IN TRANSPORTATION AND REGIONAL DEVELOPMENT
RESEARCH

As shown in Figure 2 and Scheme 1, the largest cluster, represented in red (cluster 1), consists of 21
countries, followed by the green cluster (cluster 2) with 14 items, the blue cluster (cluster 3) with 12
countries, the yellow cluster (cluster 4) with 11 countries, the purple cluster (cluster 5) with 9 countries,
the light blue cluster (cluster 6) with 8 countries, and finally the orange cluster (cluster 7) with 7 countries.
It is crucial to search deeper into the clusters identified, as they offer valuable insights into the
collaborative networks among countries (Krabokoukis and Polyzos, 2023).

SCHEME 1 – ILLUSTRATION ON MAP OF THE CO-AUTHORSHIP NETWORK ANALYSIS OF COUNTRIES IN TRANSPORTATION
AND REGIONAL DEVELOPMENT RESEARCH

Cluster 1 (European cluster), represented in red, emerges as the most extensive cluster in terms of the number of included countries, encompassing a collaborative network of 21 nations. These countries, as follows (with the respective number of co-authored papers provided in parentheses), contribute significantly to a robust co-authorship network, signifying substantial collaboration in research on the analyzed field: Germany (137), Netherlands (101), Poland (100), Sweden (72), Iran (53), Norway (47), Switzerland (34), Finland (31), Austria (30), Czech Republic (30), Ukraine (30), Denmark (25), Croatia (23), Lithuania (22), Slovakia (20), Slovenia (18), Serbia (16), Latvia (7), Bulgaria (6), Estonia (6), and Nepal (6). Cluster 1 represents a noteworthy proportion of 14.19% of total analyzed papers, underscoring its prominence in shaping collaborative endeavors within the research domain.

Cluster 2 (Transatlantic cluster), represented in blue, stands out as the second most substantial cluster in terms of the number of papers with a remarkable 20.88% of total analyzed papers, and represents a diverse collaborative network among 14 countries. The countries contributing to this cluster, along with the respective number of papers (in parentheses), are: the United States (704), Indonesia (101), Malaysia (99), Pakistan (116), Saudi Arabia (46), Nigeria (39), United Arab Emirates (23), Bangladesh (20), Egypt (12), Oman (11), Qatar (8), Iraq (7), Azerbaijan (6), and Jordan (6). The substantial representation of the United States in this cluster, along with the active participation of countries from the Middle East and Southeast Asia, underscores the global nature of research collaboration in the analyzed field.

Cluster 3 (Pan-Continental cluster), distinguished by the blue color, is a collaborative network, involving 12 countries, that represents a notable 15.67% of total papers within the analyzed database. The countries, along with the respective number of papers (in parentheses), are: India (187), Canada (120), the Russian Federation (166), Spain (148), and Turkey (105), Greece (54), Portugal (52), Romania (40), Cyprus (9), Ecuador (7), Algeria (6), and Albania (5). India and Canada emerge as prominent contributors within this cluster, signifying their influential roles in fostering research collaboration. The inclusion of countries from Southern Europe, Eastern Europe, and North Africa enhances the diversity of perspectives and approaches within this network, contributing to the richness of the scholarly discourse.

Cluster 4 (Global cluster), depicted in yellow, constitutes a collaborative network that spans 11 countries, representing a 13.13% of total papers in the analyzed database, showcasing significant contributions to the research landscape in the field of transportation infrastructure and regional development. The countries, along with the respective number of papers (in parentheses), are: the United Kingdom (349), Japan (146), France (78), Brazil (73), New Zealand (23), Tunisia (23), Philippines (19), Chile (18), Israel (14), Cambodia (5), and Peru (5). Notably, the United Kingdom, France, and Japan emerged as key contributors. The inclusion of countries from South America, Asia, Europe, and Oceania further enhances the global perspective within Cluster 4.

Cluster 5 (Asian-Pacific cluster), represented in purple, emerges as the most productive collaborative network in terms of the number of articles and accounts for 27.63% of total papers in the analyzed database. It incorporates nine countries that have made substantial contributions to the field of transportation infrastructure and regional development. The countries, along with the number of papers (in parentheses), include: China (1,135), Australia (199), South Korea (98), Hong Kong (77), Singapore (35), Vietnam (25), Macao (6), Fiji (5), and Sri Lanka (5). Cluster 5 signifies its pivotal role in shaping the scholarly discourse within the specified domain. Notably, China emerged as a prominent contributor, showcasing the depth of its engagement in collaborative research efforts. The inclusion of countries from Oceania and Asia underscores the global nature of research within Cluster 5.

Cluster 6 (Latin-European cluster), represented in light blue, emerges as another productive collaborative network, accounting for a 4.67% of total articles in the analyzed database. This cluster comprises eight countries that have actively contributed to research in the field of transportation infrastructure and regional development. The countries, along with the corresponding number of papers (in parentheses), include Italy (128), Belgium (55), Ireland (22), Colombia (18), Mexico (17), Morocco (15), Uruguay (7), and Argentina (6). The collaborative efforts within Cluster 6 highlight a diverse geographical representation and collective contributions of these countries to advancing scholarly understanding in the specified domain.

Cluster 7 (Global South cluster), represented in orange, accounts for a 3.83% of total articles in the analyzed database and comprises seven countries actively engaged in collaborative research efforts. The countries, along with the corresponding number of papers (in parentheses), include: South Africa (66), Taiwan (51), Thailand (39), Kazakhstan (29), Hungary (17), Ghana (10), and Kenya (8). Despite being the smallest cluster in terms of article representation, Cluster 7 highlights the meaningful contributions and collaborative initiatives among these diverse countries, enriching the global perspectives within the specified research domain.

In conclusion, the co-authorship analysis reveals distinct collaborative patterns among countries in the field of transportation infrastructure and regional development. The variations among clusters underscore the dynamic and interconnected nature of international collaborations in advancing research within the specified field.

3.3 Results of Co-occurence Analysis

Figure 3 presents a comprehensive co-occurrence map of keywords within the Scopus dataset, revealing intricate patterns and associations among key terms. The map, established with a minimum threshold of 100 occurrences, highlights 62 keywords out of the extensive pool of 20,494. Notably, the keywords with the highest link strengths include "economic growth"; "China"; "economics"; "regional planning";

"economic development"; "sustainable development"; "carbon dioxide"; "carbon emission"; "transportation infrastructure"; and "economics and social effects", underlining their centrality in the scholarly discourse.

The co-occurrence analysis identifies four distinct clusters, each representing a thematic concentration within the field. **Cluster 1 (in red) encapsulates the mutual combination of transport infrastructure with economic prosperity**, especially within the framework of regional planning and development strategies. The presence of keywords like "high-speed train" and "road transport" underscores the pivotal role of transportation infrastructure in connecting various regions and promoting economic development. Additionally, terms such as "investment" and "regional development" highlight the importance of investments and regional planning in achieving economic goals. Therefore, this cluster is crucial for understanding how transportation infrastructure plays a critical role in promoting regional development and economic prosperity.

Cluster 2 (in green) is characterized by a focus on economic analysis, emissions control, and energy efficiency, illustrating the intersection of economic considerations and environmental impact in the realm of transportation. The inclusion of keywords like "climate change" and "greenhouse gases" underscores the cluster's commitment to exploring the environmental implications of transportation activities. Additionally, terms such as "investments" and "population statistics" suggest a comprehensive approach that considers economic, environmental, and societal factors in the analysis of transportation. Thus, this cluster provides valuable insights into the interdisciplinary nature of research focusing on the environmental aspects of transportation.
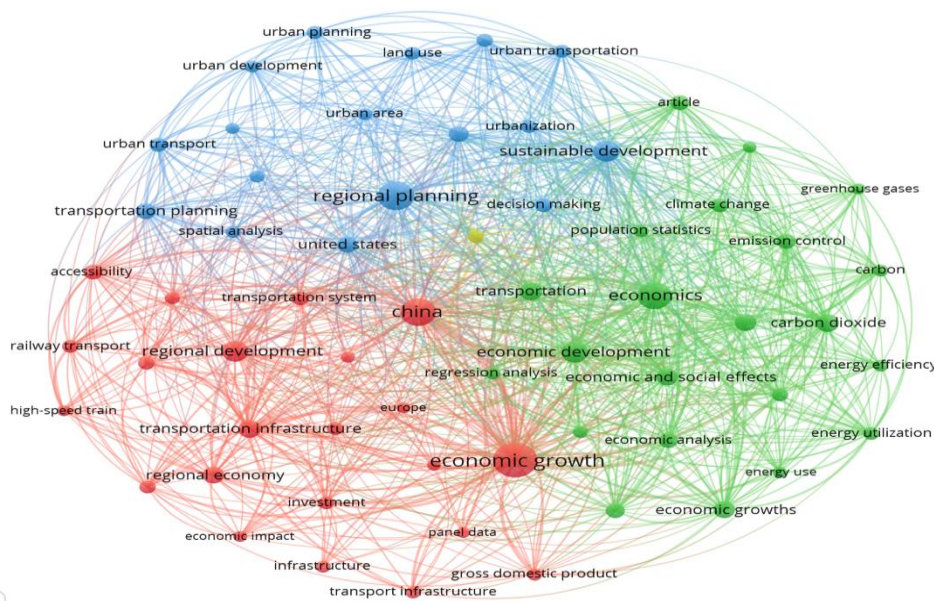

FIGURE 3 – CO-OCCURENCE ANALYSIS BY ALL KEYWORDS

TABLE 4 – KEYWORDS PER CLUSTER FROM THE CO-OCCURRENCE ANALYSIS BY ALL KEYWORDS

Cluster	Keywords
Cluster 1 (red)	Accessibility, air transportation, china, economic growth, economic impact, Europe, gross domestic production, high-speed train, infrastructure, investment, numerical model, panel data, railway transport, regional development, regional economy, road transport, transport infrastructure, transportation development, transportation economics, transportation infrastructure, transportation policy, transportation system
Cluster 2 (green)	Article, carbon, carbon dioxide, carbon emission, climate change, economic analysis, economic and social effects, economic development, economic growths, economics, emission control, energy efficiency, energy use, energy utilization, environmental economics, environmental impact, greenhouse gases, investments, population statistics, regression analysis, roads and streets, transportation.
Cluster 3 (blue)	Decision making, land use, metropolitan area, public transport, regional planning, spatial analysis, sustainability, sustainable development, transportation planning, united states, urban area, urban development, urban growth, urban planning, urban transport, urban transportation, urbanization
Cluster 4 (yellow)	Spatiotemporal analysis

Cluster 3 (in blue) places a significant emphasis on the decision-making process, spatial analysis, and the role of transportation in fostering urban and regional sustainability. The inclusion of keywords like "land use", "public transport", and "urbanization" suggests a comprehensive exploration of the interconnectedness between transportation and urban development. The focus on "sustainability" and "sustainable development" highlights the cluster's commitment to addressing the long-term viability and environmental impact of transportation within urban contexts. Additionally, terms such as "metropolitan area" and "united states" indicate a broad geographical scope, reinforcing the cluster's relevance to diverse urban settings. Overall, Cluster 3 contributes valuable insights into the multidimensional relationships between transportation, urban planning, and sustainable development.

Cluster 4 (in yellow), uniquely labeled "Spatiotemporal Analysis", spotlights the specific focus on the temporal and spatial aspects of transportation phenomena. This cluster is characterized by a singular keyword, "spatiotemporal analysis". While less populated, this cluster underscores the importance of considering the dynamic nature of transportation in scholarly research. By homing in on the interplay between time and space, this cluster contributes a perspective that complements the broader themes identified in the co-occurrence analysis.

Upon closer examination, key trends emerge as we delve deeper into the data. Notably, the pivotal keywords that wield substantial influence within the scholarly discourse include "economic growth" with 1062 occurrences and a total link strength of 3920, "China" with 692 occurrences and 3055 total link strength, "economics" with 648 occurrences and 2912 total link strength, "regional planning" with 709 occurrences and 2358 total link strength, and "economic development" with 402 occurrences and 1962 total link strength. This result aligns with previous bibliometric studies in the sustainability sector, where the economic dimension exhibited greater prominence compared to other sustainability dimensions

(Krabokoukis, 2023). The resonance of these key economic themes underscores their centrality in the broader discourse and emphasizes their enduring significance across diverse research domains.

In conclusion, the co-occurrence analysis offers valuable insights into the thematic concentrations and interconnections within the diverse landscape of transportation studies. It reveals a nuanced research domain, characterized by the identification of four distinct clusters. These clusters underscore the multidimensional nature of transportation research, spanning economic, environmental, regional development, urban planning, and spatiotemporal dimensions.

3.4 Results of Bibliometric Coupling Analysis

Figure 4 provides a detailed overview of the bibliographic coupling network analysis, revealing intricate connections among authors within the Scopus dataset. Constructed with a minimum threshold of five documents authored by an individual, the network identifies 51 authors from the initial 12,059 authors in the database, offering a visual representation of collaborative relationships and scholarly connections among them.
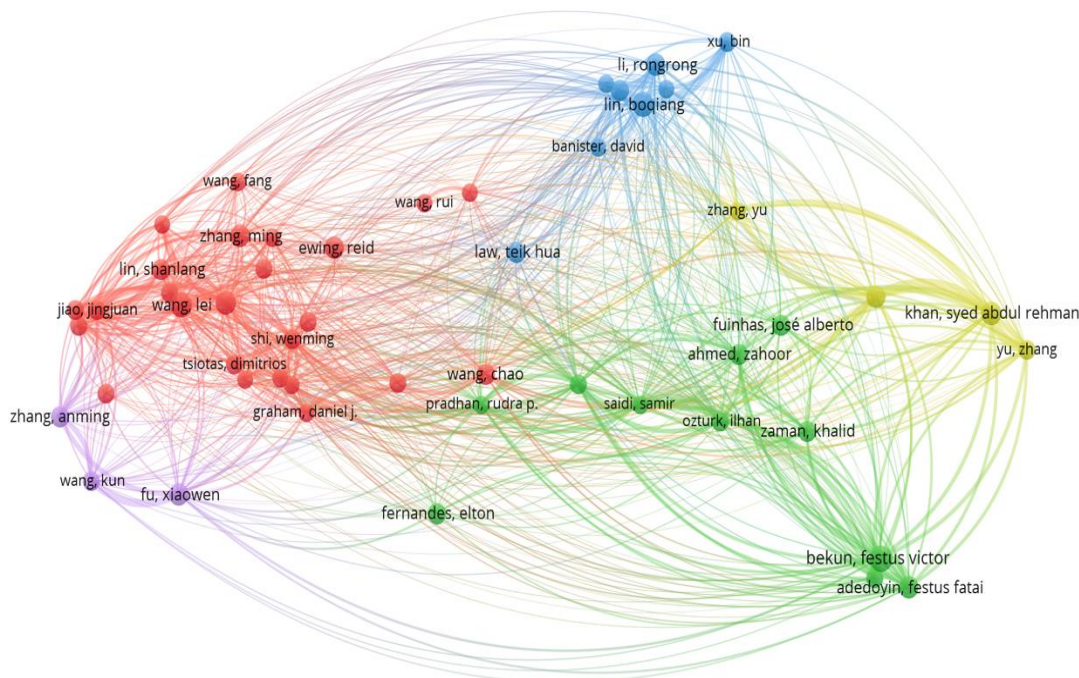

FIGURE 4 – BIBLIOMETRIC COUPLINK NETWORK ANALYSIS AND TOTAL LINK STRENGTH (IN PARENTHSESES) FROM THE
BIBLIOGRAPFIC COULING NETWORK ANALYSIS

A significant outcome of the bibliographic coupling analysis is the identification of authors with the highest total link strengths in the network, as arises from Table 5. Noteworthy contributors include Bekun, F. V. (2,192) with 11 papers, Khan, S. A. R. (1,901) with 9 papers, Adedoyin, F. F. (1,738) with 8 papers, Sharif,

A. (1,564) with 8 papers, and Saidi, S. (1,445) with 5 papers. These authors play pivotal roles within the network, signifying their substantial contributions and collaborations in the field of transport and regional development.

Furthermore, Figure 4 accentuates the largest cluster of connected authors and bigger total link strength, depicted in red and consisting of 25 items. This cluster represents a cohesive group of authors with shared research interests and collaborative endeavors. Other clusters, represented by green, blue, yellow, and purple nodes with 11, 8, 4, and 3 authors respectively, reflect various subgroups and themes within the broader field of transport and regional development research. Table 5 provides cluster-specific authors with corresponding total link strengths, further enhancing the understanding of collaborative patterns among authors.

TABLE 5 – IDENTIFYING THE INFLUENTIAL AUTHORS AND TOTAL LINK STRENTH (IN PARENTHSES) FROM THE BIBLIOGRAPHIC COUPLING NETWORK ANALYSIS

Cluster	Authors
Cluster 1 (red)	Chen, Z. (1,142), De Jong, M. (377), Ewing, R. (82), Graham, D. (510), Hayashi, Y. (164), Jiao, J. (1,239), Jin, F. (773), Kim, E. (385), Lee, P. T. (228), Lin, S. (770), Mi, J. (621), Nijkamp, P. (434), Shi, W. (947), Tsiotas, D. (116), Wang, C. (444), Wang, F. (220), Wang, J. (732), Wang, L. (939), Wang, P. (495), Wang, R. (87), Wang, W. (66), Wang, Y. (189), Witlox, F. (227), Zhang, F. (279), Zhang, M. (674)
Cluster 2 (green)	Adedoyin, F.F. (1,738), Ahmend, Z. (807), Balsalobre-Lorente, D. (1,321), Bekum, F.V. (2.192), Fernandez, E. (176), Fuinhas, J. A. (329), Ozturk, I. (921), Pradhan, R. (742), Saidi, S. (1,445), Shahbaz, M. (1,375), Zaman, K. (496)
Cluster 3 (blue)	Banister, D. (137), Law, T. H. (166), Li, R. (877), Lin, B. (997), Wang, L. (127), Wang, Q. (987), Wu, Y. (363), Xu, B. (738)
Cluster 4 (yellow)	Khan, S. A. R. (1,901), Sharif, A. (1,564), Yu, Z. (1,275), Zhang, Y. (746)

3.5 Results of Co-Citation Analysis

Figure 5 illustrates a co-citation map of authors cited in the Scopus dataset. The map is constructed based on a minimum threshold of 200 citations per author, with only 86 out of the 182,156 authors meeting this criterion. Notably, the authors with the highest total link strengths in the network are Wang, Y. (27,559), Zhang, Y. (27,243), Liu, Y. (22,525), Li, Y. (20,416), Wang, J. (19,715), as shown in Table 6. As depicted in Figure 5, the co-citation network is organized into four distinct clusters.

Cluster 1 represented in red, consists of 38 authors, followed by the green Cluster 2 with 27 authors, the blue Cluster 3 with 14 authors, and, finally, the yellow Cluster 4 with 7 authors. Table 6 provides an in-depth identification of collaborators within each cluster, along with their respective total link strengths obtained from the co-citation analysis. The identified clusters represent significant collaborative networks among authors, providing insights into the patterns of co-citation and the strength of relationships within the research community.
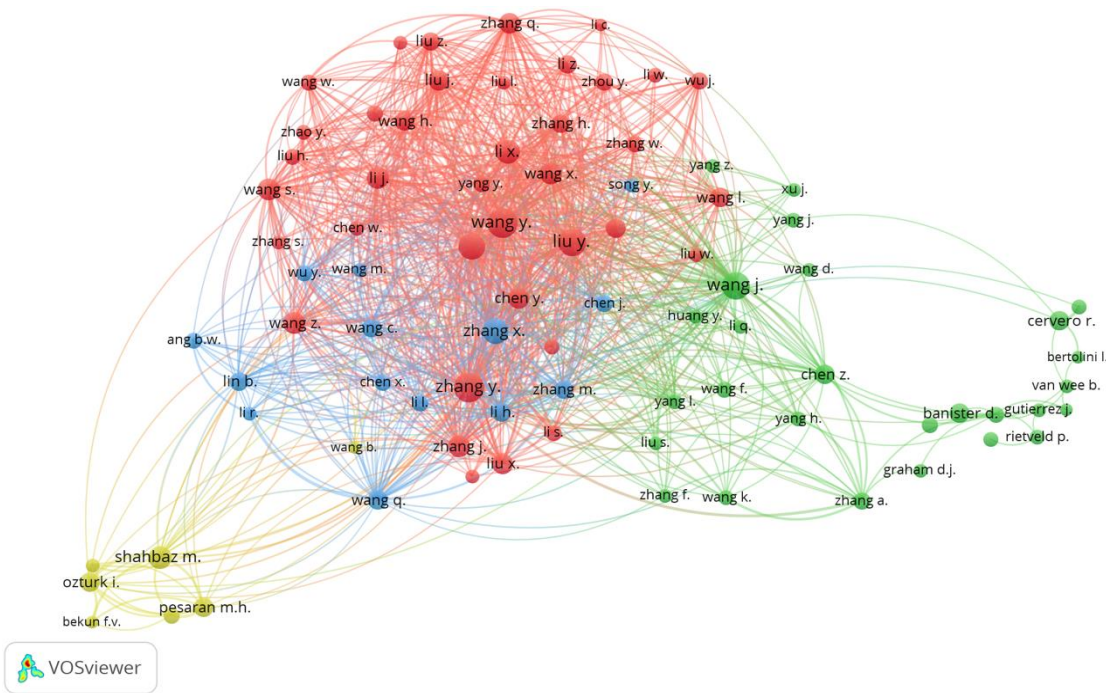
FIGURE 5 – CO-CITATION MAP OF AUTHORS IN TRANSPORT AND REGIONAL DEVELOPMENT RESEARCH

TABLE 6 – IDENTIFYING THE COLLABORATIVES AMONG THE AUTHORS AND THEIR TOTAL LINK STRENTH (IN PARENTHSES)
FROM THE CO-CITATION ANALYSIS

Cluster	Authors
Cluster 1 (red)	Wang, Y. (27,559), Zhang, Y. (27,243), Liu, Y. (22,525), Li, Y., (20,416), Li, X. (17,117), Wang, S. (15,884), Zhang, Q. (15,255), Wang, Z. (14,058), Li, J. (13,778), Liu, J. (13,412), Wang, X. (13,192), Zhang, J. (12,983), Liu, X. (12,275), Chen, Y. (12,028), Liu, Z. (11,428), Zhang, L. (11,306), Wang, L. (11,119), Zhang, H. (10,517), Wang, H. (10,469), Wang, W. (9,136), Zhou, Y. (9,190), Li, Z. (9,038), Wu, J. (8,741), Zhang, W. (8,033), Yang, Y. (7,848), Zhao, Y. (7,653) Liu, H. (7,637), Li, S. (7,422), Zhang, Z. (6,944), Liu, C. (6,437), Chen, W. (6,339), Chen, H. (6,244), Li, W. (6,187), Liu, L. (5,937), Yang, X. (5,839), Liu, W. (5,707), Li, C. (5,665), and Zhang, S. (5,082).
Cluster 2 (green)	Wang, J. (19,715), Chen, Z. (9,940), Yang, L. (8,291), Zhang, A. (7,684), Yang, H. (7,107), Wang, K. (6,807), Wang, F. (6,499), Zhang, F. (6,200), Liu, S. (5,718), Yang, J. (5,398), Huang, Y. (5,237), Vickerman, R. (5,048), Xu, J. (5,046), Li, Q. (5,037), Wang, D. (4,815), Yang, Z. (4,174), Cervero, R. (3,901), Banister, D. (3,778), Gutierrez, J. (3,057), Graham, D. J. (2,785), Rietveld, P. (2,258), Nijkamp, P. (2,209), Ewing, R. (2,069), Wegener, M. (1,928), Van Wee, B. (1,913), and Bertolini, L. (1,831).
Cluster 3 (blue)	Zhang, X. (18,235), Wang, Q. (12,130), Zhang, M. (10,888), Li, H. (10,072), Wang, C. (9,358), Lin, B. (9,408), Chen, J. (8,596), Li, L. (8,507), Wu, Y. (8,459), Chen, X. (7,361), Li, R. (7,196), Ang, B. W. (5,839), Song, Y. (5,671), and Wang, M. (5,026).
Cluster 4 (yellow)	Shahbaz, M. (9,762), Ozturk, I. (7,333), Pesaran, M. H. (6,194), Wang, B. (5,007), Bekun, F. V. (3,959), Shin, Y. (3,916), and Zaman, K. (3,631).

3.6 Results of Methodology Analysis

In this section, a bibliometric analysis is conducted to discern the key methodological approaches within the domain of transport infrastructures and regional development. The Python programming language was employed for this purpose. The outcomes of this analysis serve to underscore the prominence of specific methodologies, thereby offering valuable insights for both researchers and policymakers. A detailed script, created for this specific purpose, is presented in Table A of the Appendix. Some primary

methodological approaches have been identified, and corresponding keywords have been formulated for use in abstract searches. These searches aim to categorize articles under their respective methodological approaches. The specific methodologies and the corresponding number of articles for each are outlined in Table 7.

The outcomes delineated in Table 7, furnish a portrayal of the distribution of articles across diverse methodological frameworks. Among the total of 4,202 articles within the database, 2,398 were categorized under quantitative approaches, 155 under qualitative approaches, while 1,649 could not be classified based on the employed keywords in the analysis. Notably, articles falling under the "Economic Theory and Analysis" category amount to 1,531, constituting 60% of the categorized articles, predominantly centered on the subcategory "Economic Theory", totaling 1,404 articles. This underscores a predominant emphasis on economic frameworks, aligning with the corresponding conclusion drawn from Section 3.3, Results of Co-occurrence analysis. This concentration on economic perspectives highlights the significance attributed to economic factors in the planning and development of transportation infrastructure. By employing economic frameworks, researchers gain a comprehensive perspective on the intricate connections between transportation infrastructure and regional development. This approach provides a systematic lens through which the economic intricacies and implications embedded in the planning and execution of transportation projects can be analyzed. The advantages of this method lie in its ability to elucidate the economic drivers, impacts, and policy implications within the broader context of transport infrastructure, contributing to a more nuanced comprehension of the subject matter.

TABLE 7 – A QUANTITATIVE OVERVIEW OF THE METHODOLOGICAL APPROACHES IN RESEARCH ON TRANSPORT
INFRASTRUCTURES AND REGIONAL DEVELOPMENT

Main Category	Quantitative		Qualitative	
	Subcategory	Number of papers	Subcategory	Number of papers
Statistics and time series (541 papers)	Statistical modeling	386	Interviews	155
	Time series	100		
	Data analysis	55		
	Statistical tools	0		
Economic theory and analysis (1531 papers)	Economic theory	1,404		
	Investment analysis	8		
	Cost analysis	27		
	Economic policy	25		
	Network analysis	43		
	Supply chain analysis	24		
Geographic Information systems (326 papers)	Geographic systems	105		
	Spatial analysis	81		
	Demographic analysis	23		
	Machine learning	9		
	Scenario analysis	15		
	Environmental analysis	103		
Categorized papers	2,398		155	
Uncategorized papers	1,649			
Total papers	4,202			

Furthermore, the category "Statistics and Time Series" encompasses 541 papers, accounting for 21% of the articles, indicating a substantial emphasis on statistical and time-series methodologies. Through statistical modeling, researchers can quantitatively analyze and model complex relationships between transportation infrastructure and regional development. Time-series analysis allows the examination of trends and patterns over time, providing valuable insights into the dynamic nature of the studied phenomena. Additionally, data analysis serves as a robust tool for extracting meaningful information from datasets, contributing to the overall comprehensiveness of research. The strength of this methodology lies in its capacity to uncover temporal dynamics, model relationships, and derive empirical evidence, all being essential for a rigorous examination of the subjects under consideration.

The observation that the "Geographic Information Systems" category incorporates 326 papers (12% of the articles) implies a noteworthy concentration on GIS-related approaches. The Geographic Information Systems (GIS) methodology, encompassing Geographic Systems, Spatial Analysis, Demographic Analysis, Machine Learning, Scenario Analysis, and Environmental Analysis, provides a powerful lens through which one can gain deeper insights into the intricate dynamics of transportation infrastructure and regional development. By employing GIS, researchers can spatially analyze the geographical distribution of transportation networks, demographic patterns, and environmental factors. Spatial analysis further allows the exploration of the spatial relationships and patterns within our datasets, shedding light on the spatial dimensions of our research topics. The integration of machine learning enhances the ability to uncover complex patterns and trends within geographic datasets, while scenario analysis enables the assessment of potential future developments. Finally, environmental analysis within the GIS framework allows for a comprehensive understanding of the environmental implications of transportation and regional planning. Leveraging GIS methodologies enriches research with spatial intelligence, fostering a holistic comprehension of the multifaceted issues explored in this study.

Additionally, the "Interviews" category, containing 155 papers (7% of the articles), underscores the significance of qualitative research methods in contributing to the overall scholarly discourse. Through in-depth interviews, researchers gain a nuanced understanding of the perspectives, experiences, and insights of key stakeholders involved in the realm of transportation planning and development. The qualitative data derived from interviews offer a rich narrative, allowing the uncovering of contextual variations and unique challenges that may not be captured through quantitative methods alone. This approach provides a human-centric dimension to our exploration of transportation issues. Interviews contribute to the broader scholarly discourse by offering qualitative insights that complement and enrich the quantitative findings, ensuring a comprehensive and well-rounded understanding of the complex issues within our research domain.

These categorization insights shed light on the methodological orientations and thematic concentrations within the realm of transport infrastructures and regional development research. However, it is crucial to acknowledge the limitations of this analysis. The use of specific keywords, though indicative, might influence the outcomes, and alternative keywords could yield different results. To enhance the robustness of future studies, it is recommended to conduct a more in-depth analysis, both of the methodological approaches and the keywords employed. While this approach contributes to understanding the landscape of research in this domain, it is imperative to recognize that this analysis serves as a directional guide rather than an exhaustive exploration. As research evolves and new terminologies emerge, adapting the keyword selection becomes pivotal. Despite these limitations, this bibliometric analysis approach serves as an additional tool for researchers engaging in similar studies, offering a foundation for further refinement and improvement in future investigations.

4. CONCLUSIONS

The research embarked on an exploration of the landscape concerning transport infrastructure and regional development. The citation analysis unveiled a diverse range of highly cited papers, covering topics from green logistics to the impact of COVID-19 on the global economy. This diversity underscores the multidimensional nature of research in the field, capturing significant contributions that have garnered substantial attention within the scholarly community.

The co-authorship analysis illuminated intricate collaborative networks among countries. Clusters such as the European cluster and the Asian-Pacific cluster showcased the dynamic and interconnected nature of international collaborations, offering valuable insights into the global landscape of research in the field. The co-occurrence analysis identified four thematic clusters, emphasizing the interconnectedness of infrastructure, regional development, economic growth, environmental considerations, and spatiotemporal dynamics. These clusters provided an understanding of the multifaceted nature of studies in the examined field.

Additionally, an attempt was made to classify the articles under examination based on their methodological orientations. Economic theory and analysis approaches dominated the research landscape, with statistics models and geographic information systems also playing significant roles.

Addressing the limitations of this study, it is acknowledged that the use of specific keywords may have influenced outcomes, and alternative keywords could yield different results. Furthermore, the restriction to a particular database introduces bias. To enhance the robustness of future studies, a more in-depth analysis of both methodological approaches and keywords is recommended. This study serves as a

directional guide rather than an exhaustive exploration, recognizing the evolving nature of research and the emergence of new terminologies.

In conclusion, while this study provides a foundation for understanding the current state of research on transport infrastructure and regional development, there is a need for further refinement and improvement. As the field evolves, continuous adaptation of research methodologies and exploration of emerging themes will be pivotal for advancing scholarly discourse and addressing the complex challenges in this domain.

REFERENCES

Alemi, F., Circella, G., Handy, S., & Mokhtarian, P. (2018). What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. *Travel Behaviour and Society*, 13, 88–104. https://doi.org/10.1016/j.tbs.2018.06.002

Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., Olson, J. B., James, E. P., Dowell, D. C., Grell, G. A., Lin, H., Peckham, S. E., Smith, T. L., Moninger, W. R., Kenyon, J. S., & Manikin, G. S. (2016). A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review*, 144(4), 1669–1694. https://doi.org/10.1175/MWR-D-15-0242.1

Bronzini, R., Piselli, P. (2009). Determinants of long-run regional productivity: the role of R&D, human capital and public infrastructure. *Regional Science and Urban Economics,* vol. 39(2), pp. 187-19.

Dekker, R., Bloemhof, J., & Mallidis, I. (2012). Operations Research for green logistics – An overview of aspects, issues, contributions and challenges. *European Journal of Operational Research*, 219(3), 671–679. https://doi.org/10.1016/j.ejor. 2011.11.010

Duncan, B. N., Lamsal, L. N., Thompson, A. M., Yoshida, Y., Lu, Z., Streets, D. G., Hurwitz, M. M., & Pickering, K. E. (2016). A space-based, high-resolution view of notable changes in urban NO x pollution around the world (2005–2014). *Journal of Geophysical Research: Atmospheres*, 121(2), 976–996. https://doi.org/10.1002/ 2015JD024121

Guerrero-Ibanez, J. A., Zeadally, S., Contreras-Castillo, J. (2015). Integration challenges of intelligent transportation systems with connected vehicle, cloud computing, and internet of things technologies. *IEEE Wireless Communications*, 22(6), 122–128. https://doi.org/10.1109/MWC.2015.7368833

He L., Duchin F. (2009). Regional Development in China: Interregional Transportation Infrastructure and Regional Comparative Advantage. *Economic Systems Research*, vol. 21(1), pp. 3–22.

Ibn-Mohammed, T., Mustapha, K. B., Godsell, J., Adamu, Z., Babatunde, K. A., Akintade, D. D., Acquaye, A., Fujii, H., Ndiaye, M. M., Yamoah, F. A., & Koh, S. C. L. (2021). A critical analysis of the impacts of COVID-19 on the global economy and ecosystems and opportunities for circular economy strategies. *Resources, Conservation and Recycling*, 164, 105169. https://doi.org/10.1016/ j.resconrec. 2020.105169

Kantianis, D. D. (2021). Stochastic Socio-economic Ex-ante Evaluation of Road Transport Infrastructure Projects, In: Horobet A., Belascu L., Polychronidou P., Karasavvoglou A. (eds) Global, Regional and Local Perspectives on the Economies of Southeastern Europe, *Springer Proceedings in Business and Economics,* Springer, Cham. https://doi.org/10.1007/978-3-030-57953-1_20

Kasikoen K. M., Suprajaka S., Martini E. (2019). Impacts of inter-urban transportation railway to regional development (Case study: Sukaraja District - Bogor Regency - West Java Province). *IOP Conf. Series: Earth and Environmental Science*, 340, 012028, IOP Publishing, doi:10.1088/1755-1315/340/1/012028.

Krabokoukis, T. (2023). Exploring the State of Research on Tourism Sustainability: A Bibliometric Analysis in the Post-COVID Era. *Highlights of Sustainability*, 2(2), 50–61. https://doi.org/10.54175/hsustain2020005

Krabokoukis, T., Polyzos, S. (2022). Drawing an indicator of tourism competitiveness and examining its relationship with tourism seasonality for the Greek prefectures. *Regional Science Inquiry*, XIV(2), 55–70.

Krabokoukis, T., Polyzos, S. (2020). Spatial inequalities of tourist activity in Greece: A Shift-Share analysis. *Anatolia*, 32(1), 93–105. https://doi.org/10.1080/13032917. 2020.1851271

Krabokoukis, T., Polyzos, S. (2021). An Investigation of Factors Determining the Tourism Attractiveness of Greece's Prefectures. *Journal of the Knowledge Economy*, 12, 1997–2015. https://doi.org/10.1007/s13132-020-00704-7

Krabokoukis, T., Polyzos, S. (2023). A Bibliometric Analysis of Integrating Tourism Development into Urban Planning. Sustainability, 15(20). https://doi.org/10.3390/ su152014886

Liu, F., Zhang, Q., Tong, D., Zheng, B., Li, M., Huo, H., He, K. B. (2015). High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China from 1990 to 2010. *Atmospheric Chemistry and Physics*, 15(23), 13299–13317. https://doi.org/10.5194/acp-15-13299-2015

Mathiesen, B. V., Lund, H., Karlsson, K. (2011). 100% Renewable energy systems, climate mitigation and economic growth. *Applied Energy*, 88(2), 488–501. https://doi.org/10.1016/j.apenergy.2010.03.001

OpenAI. (2023). ChatGPT [Large language model]. https://chat.openai.com

Polyzos, S. (2019). *Regional Development*, Athens. Greece: Kritiki Publishing.

Polyzos, S., Tsiotas, D. (2020). The contribution of transport infrastructures to the economic and regional development: A review of the conceptual framework. *Theoretical and Empirical Researches in Urban Management*, 15(1).

Polyzos, S., Tsiotas, D. (2023). Interregional transport infrastructures and regional development: A methodological approach. *Theoretical and Empirical Researches in Urban Management*, 18(2).

Rosik, P., Wójcik, J. (2022). Transport Infrastructure and Regional Development: A Survey of Literature on Wider Economic and Spatial Impacts, *Sustainability*, MDPI, vol. 15(1), pages 1-19, December.

Seguí-Amortegui, Clemente-Almendros, Medina, & Grueso. (2019). Sustainability and Competitiveness in the Tourism Industry and Tourist Destinations: A Bibliometric Study. *Sustainability*, 11(22), 6351. https://doi.org/10.3390/su11226351

Stafford-Smith, M., Griggs, D., Gaffney, O., Ullah, F., Reyers, B., Kanie, N., Stigson, B., Shrivastava, P., Leach, M., O'Connell, D. (2017). Integration: The key to implementing the Sustainable Development Goals. *Sustainability Science*, 12(6), 911–919. https://doi.org/10.1007/s11625-016-0383-3

Tsiotas, D., Polyzos, S. (2018). The Complexity in the Study of Spatial Networks: An Epistemological Approach. *Networks and Spatial Economics*, 18(1), 1–32. https://doi.org/10.1007/s11067-017-9354-1

Umar, M., Ji, X., Kirikkaleli, D., Xu, Q. (2020). COP21 Roadmap: Do innovation, financial development, and transportation infrastructure matter for environmental sustainability in China? *Journal of Environmental Management*, 271, 111026. https://doi.org/10.1016/j.jenvman.2020.111026

Venables, A., Laird, J. J., Overman, H. G. (2014). *Transport investment and economic performance: Implications for project appraisal.* Paper commissioned by UK Department for Transport, 1-79.

APPENDIX

TABLE A – PYTHON CODE: CATEGORIZATION OF RESEARCH ARTICLES IN THE FIELD OF TRANSPORT AND REGIONAL
DEVELOPMENT BASED ON METHODOLOGICAL APPROACHES: A KEYWORD ANALYSIS.

```
import pandas as pd
import numpy as np  # Importing numpy for NaN checking

# Read the CSV file from the desktop
df = pd.read_csv('~/Desktop/scopus.csv')

# Create a dictionary for categories
categories = {
'Scenario analysis': [
  "scenario analysis", "spatial scenario", "future land use", "scenarios assessment",  "spatial decision support systems",
"geospatial scenario planning techniques",  "urban scenarios", "scenario planning methods", "scenario development
techniques", "scenario evaluation tools", "scenario forecasting methods", "scenario impact analysis software",  "scenario
assessment techniques", "scenario building methods", "scenario-based planning software",  "spatial scenario design
techniques", "scenario analysis approaches", "spatial scenario modeling software"],
'Machine learning': [
  "machine learning in geography", "machine learning", "geospatial machine learning techniques", "spatial pattern recognition
methods", "spatial clustering algorithms", "machine learning in remote sensing", "geospatial deep learning models",
"supervised learning methods", "unsupervised learning algorithms", "deep learning models", "neural networks analysis",
"machine learning applications in geography", "geospatial machine learning applications", "geospatial data analysis with
machine learning", "machine learning for spatial analysis", "geospatial data mining", "machine learning for geographic data",
"geospatial predictive modeling", "geospatial feature engineering", "geospatial machine learning tools", "geospatial deep
learning frameworks", "geospatial neural networks", "spatial pattern recognition software", "geospatial unsupervised learning",
"geospatial deep learning applications", "machine learning for remote sensing", "geospatial machine learning libraries",
"machine learning for geospatial data analysis", "geospatial machine learning models", "machine learning for spatial data",
"geospatial data classification", "geospatial data regression analysis", "geospatial machine learning frameworks", "machine
learning for geospatial applications", "geospatial data interpretation with machine learning", "machine learning in geographic
information systems", "geospatial machine learning algorithms"],
'Demographic analysis': [
  "demography analysis", "population dynamics", "migration analysis", "demographic modeling", "population projections",
"population health analysis", "population statistics", "demographic trends", "population distribution analysis", "demographic
data interpretation", "demographic data modeling", "demographic research", "demographic data collection", "demographic data
management", "population growth analysis", "demographic data collection methods", "demographic data management
software", "demographic data analysis techniques", "population dynamics modeling", "migration data analysis", "demographic
data trends", "population distribution modeling", "population health trends", "demographic data research", "population growth
trends", "population statistics analysis", "demographic research methods", "demographic data modeling techniques"],
'Spatial analysis': [
  "spatial analysis", "spatial distribution", "spatiotemporal analysis", "spatial autocorrelation", "spatial econometrics", "spatial
pattern analysis", "spatial input output analysis", "spatial data modeling", "spatial data visualization", "spatial data
interpretation", "spatial data processing", "spatial data representation", "spatial data summarization", "geospatial analysis
techniques", "geospatial data management", "spatial analysis software", "geospatial data analysis tools", "spatial analysis
methods", "spatial distribution techniques", "spatiotemporal analysis methods", "spatial autocorrelation techniques", "spatial
econometrics methods", "spatial pattern recognition", "spatial input output analysis methods", "geospatial modeling
techniques", "spatial data visualization methods", "spatial data interpretation techniques", "spatial data processing methods",
"spatial data representation software", "spatial data summarization techniques", "geospatial analysis approaches", "geospatial
data management software", "spatial analysis software tools", "spatial distribution software", "spatiotemporal analysis
```

software", "spatial autocorrelation software", "spatial econometrics software", "spatial pattern analysis software", "spatial input output analysis software", "geospatial modeling software", "spatial data visualization tools", "spatial data interpretation software", "spatial data processing tools", "spatial data representation methods", "spatial data summarization software", "geospatial analysis software tools", "geospatial data management tools"],

'Geographic systems': [
    "geographic information systems", "geospatial technology", "spatial analysis", "geospatial analysis", "geospatial data", "geospatial mapping", "geospatial databases", "spatial modeling", "geographic data", "spatial data management", "GIS techniques", "spatial decision support", "geospatial software", "geospatial data processing", "geographic data visualization", "geospatial technology applications", "cartography", "geospatial intelligence", "geospatial data analysis", "geospatial data interpretation", "geospatial data representation", "geospatial data summarization", "geospatial analysis techniques", "geospatial data management", "remote sensing", "GPS", "geospatial data collection", "geospatial data collection methods", "geospatial analysis software", "geospatial data analysis tools", "geospatial information systems", "geospatial data systems", "geospatial information technology", "geographic data analysis", "geospatial data processing tools", "geospatial mapping software", "GIS applications", "geospatial data modeling", "geospatial database management", "cartographic techniques", "geospatial data visualization software", "geospatial mapping techniques", "GIS data analysis", "geospatial technology solutions", "geospatial analysis tools", "geospatial data interpretation methods", "geospatial data representation software", "geographic data management tools", "remote sensing technology", "GPS systems", "geospatial data collection software", "geospatial data analysis methods", "geospatial software applications", "geographic information technology", "spatial data analysis", "geospatial data processing techniques", "geospatial mapping tools", "GIS data management", "geospatial modeling software", "geospatial database systems", "cartographic software", "geospatial data visualization tools", "geospatial mapping solutions", "GIS data processing", "geospatial technology platforms", "geospatial analysis methods", "geospatial data interpretation software", "geospatial data representation methods", "geographic data management software", "remote sensing tools", "GPS technology", "geospatial data collection techniques", "geospatial data analysis software", "geospatial information management", "geographic data systems", "spatial data processing", "geospatial data modeling software", "geospatial database technology", "cartographic tools", "geospatial data visualization techniques", "geospatial mapping applications", "GIS data visualization", "geospatial technology tools", "geospatial analysis applications", "geospatial data interpretation techniques", "geospatial data representation tools", "geographic data management solutions", "remote sensing applications", "GPS software", "geospatial data collection platforms", "geospatial data analysis techniques"],

'Supply chain analysis': [
    "supply chain analysis", "supply chain modeling", "logistics analysis", "transportation analysis", "retail location analysis", "supply chain management", "supply chain logistics", "supply chain efficiency", "supply chain risk analysis", "supply chain optimization", "logistics management", "distribution network analysis", "supply chain planning", "inventory management", "supply chain data analysis", "supply chain modeling software"],

'Network analysis': [
    "network analysis in geography", "network analysis", "spatial network analysis methods", "transportation network analysis tools", "network optimization techniques", "network modeling software", "social network analysis methods", "network data analysis software", "network connectivity analysis methods", "network structure analysis tools", "graph theory analysis techniques", "network performance analysis software", "spatial network modeling techniques", "transportation network modeling software", "network reliability analysis methods", "network routing analysis tools", "geospatial network analysis applications", "geospatial network modeling software"],

'Economic policy': [
    "monetary policy", "fiscal policy", "economic regulations", "economic policy evaluation", "economic policy formulation", "economic policy impact", "monetary policy analysis", "fiscal policy planning", "economic policy effectiveness", "policy implementation", "policy evaluation methods", "economic reform policies", "policy impact assessment", "policy formulation techniques", "economic policy recommendations", "economic policy evaluation models", "policy impact assessment software", "policy formulation software"],

'Cost analysis': [
    "cost-effectiveness analysis", "cost control", "cost estimation", "cost efficiency", "cost-benefit assessment", "cost minimization", "cost optimization", "cost management", "cost allocation", "cost benefit analysis", "cost analysis methods", "cost-effectiveness evaluation", "cost estimation techniques", "cost control strategies", "cost optimization models", "cost estimation software", "cost accounting", "cost modeling", "cost projection", "cost reduction", "cost variance analysis", "cost estimation models", "cost management techniques", "cost reduction strategies", "cost projection methods", "cost accounting software"],

'Investment analysis': [
    "investment analysis", "investment strategies", "investment appraisal", "investment valuation", "investment risk", "investment management", "investment evaluation", "investment modeling", "investment decision-making", "investment economics", "investment planning", "investment trends", "investment performance", "investment opportunities", "investment return", "investment management strategies", "investment risk assessment", "portfolio analysis", "capital budgeting", "investment assessment", "investment modeling techniques", "financial analysis", "investment decision support", "investment software", "investment performance evaluation", "investment economics models", "investment valuation methods", "investment management techniques", "foreign direct investments"],

'Economic theory': [

"economic theory", "economics", "microeconomics", "macroeconomics", "economic models", "economic growth theory", "economic equilibrium", "demand and supply", "economic cycles", "cost-benefit analysis", "economic policy analysis", "economic forecasting", "economic development", "economic modeling", "economic theories", "economic principles", "economic research", "economic policy formulation", "economic policy impact", "Kuznets", "monetary economics", "microeconomic theory", "macroeconomic theory", "economic research methods", "econometrics", "economic modeling techniques", "economic analysis", "economic principles and concepts", "econometric models", "economic equilibrium theory", "economic policy modeling", "economic data analysis", "economic statistics", "economic policy evaluation", "economic impact assessment", "economic theory applications"],

'Statistical tools': [

"statistical software", "statistical packages", "data analysis software", "data visualization tools", "data analysis tools", "data modeling software", "statistical programming", "data analytics software", "statistical computing", "data visualization software", "data mining tools", "statistical methods software", "data reporting tools", "statistical software tools", "statistical software packages", "data management software", "statistical modeling software", "statistical analysis software", "data exploration software", "data cleansing software", "data transformation software", "data profiling software", "data validation software", "data integration software", "data aggregation software", "data classification software", "data clustering software", "data regression software", "data feature engineering software", "data quality assessment software", "data mining software", "business intelligence software", "business analytics software", "data warehousing software", "dashboard software", "reporting software", "ETL (Extract, Transform, Load) software", "data governance software", "data quality software", "data cleaning software", "data enrichment software", "data integration software", "data warehouse software", "data extraction software", "data loading software", "data mapping software", "data modeling tools", "data exploration tools", "data aggregation tools", "data validation tools", "data classification tools", "data clustering tools", "data feature engineering tools", "data quality assessment tools", "data cleaning tools", "data transformation tools", "data enrichment tools", "data integration tools", "data warehouse tools", "data extraction tools", "data loading tools", "data mapping tools", "data governance tools", "data quality tools", "business intelligence tools", "business analytics tools", "data warehousing tools", "dashboard tools", "reporting tools", "ETL tools"],

'Data analysis': [

"data analysis", "quantitative analysis", "data visualization", "data preprocessing", "data cleansing", "data transformation", "data exploration", "data mining", "data interpretation", "data analytics", "data management", "data processing", "data patterns", "data modeling", "data representation", "data summarization", "data analysis techniques", "data profiling", "data summarization methods", "data quality assessment", "data validation", "data exploration tools", "data modeling techniques", "data visualization software", "data analysis software", "data preprocessing software", "data analytics software", "data management software", "data mining tools", "data analytics tools", "data management tools", "data transformation techniques", "data cleansing methods", "data quality analysis", "data integration", "data aggregation", "data classification", "data clustering", "data correlation analysis", "data regression analysis", "data feature engineering", "data outlier detection", "data pattern recognition", "data visualization techniques", "data summarization software"],

'Time series': [

"time series", "time-series", "autoregressive integrated moving average", "ARIMA" "seasonal decomposition", "stationarity", "time series modeling", "time series data", "time series analysis methods", "forecasting methods", "trend analysis", "exponential smoothing", "stationary time series", "non-stationary time series", "seasonal patterns", "time series data preprocessing", "time series forecasting software", "time series data visualization", "time series data interpretation", "spectral analysis", "wavelet analysis", "GARCH models", "state space models", "time series econometrics", "time series modeling software"],

'Statistical modeling': [

"statistical models", "statistical modeling techniques", "linear regression", "multivariate analysis", "time series forecasting", "hypothesis testing", "regression analysis", "time series analysis", "statistical analysis", "statistical software", "data modeling", "statistical techniques", "data analysis", "statistical methods", "statistical tests", "probability distribution", "data regression", "statistical inference", "predictive modeling", "factor analysis", "principal component analysis", "ARIMA time series analysis", "variance analysis", "statistical classification", "data quality analysis", "pattern extraction", "handbook of statistical analysis"],

'Interviews': [

"interviews", "qualitative interviews", "semi-structured interviews", "researcher-conducted interviews", "qualitative interview methodology", "in-depth interviews", "open-ended interviews", "narrative analysis", "thematic analysis", "qualitative research methods", "qualitative data analysis", "qualitative data collection", "qualitative data interpretation", "qualitative research design", "coding strategies", "content analysis", "researcher reflexivity", "participant observation", "case study analysis", "case recording", "qualitative case analysis", "comparative study", "cross-case analysis", "single case study", "qualitative data coding", "qualitative data saturation", "qualitative data validity", "qualitative data reliability", "interview transcription", "qualitative research ethics", "qualitative research sampling", "qualitative data triangulation", "qualitative data trustworthiness", "qualitative data transferability", "qualitative research reporting", "qualitative data software", "qualitative data management", "qualitative data analysis techniques"],

'Environmental analysis': [

"atmospheric transport analysis", "climate models", "coastal zone analysis", "environmental impact assessment", "environmental modeling", "environmental planning", "environmental data analysis", "environmental data interpretation", "environmental data management", "climate change analysis", "atmospheric data analysis", "environmental policy analysis", "environmental sustainability analysis", "environmental impact modeling", "environmental risk assessment", "environmental data collection methods", "environmental data management software", "environmental impact assessment methods", "climate

```
change modeling", "atmospheric transport modeling", "coastal zone modeling", "environmental impact assessment tools",
"environmental modeling techniques", "environmental planning methods", "climate change assessment", "atmospheric data
interpretation", "environmental policy evaluation", "environmental sustainability assessment", "environmental risk modeling",
"environmental data collection techniques", "environmental data management techniques"],
}
# Creating a dictionary to store the number of articles per category
category_counts = {category: 0 for category in categories}
category_counts['Other'] = 0  # Adding the 'Other' category
# Categorizing the articles
for idx, row in df.iterrows():
    title = str(row['Title']).lower()
    abstract = str(row['Abstract']).lower()
    author_keywords = str(row['Author Keywords']).lower()
    index_keywords = str(row['Index Keywords']).lower()
    article_category = "Other"  # Default category
    for category, keywords in categories.items():
        for keyword in keywords:
            if keyword in title or keyword in abstract or keyword in author_keywords or keyword in index_keywords:
                article_category = category
                break
    category_counts[article_category] += 1
# Printing the number of articles per category
for category, count in category_counts.items():
print(f'{category}: {count}')
```