

Review

A Bibliometric Analysis of Integrating Tourism Development into Urban Planning

Thomas Krabokoukis *  and Serafeim Polyzos

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

* Correspondence: tkrabokoukis@uth.gr

Abstract: This study conducts a literature analysis on integrating tourism development into urban planning to understand the current state of research in the field and identify potential areas for future study. The analysis, based on 11,134 articles published between 2010 and 2023, in the Scopus database, aims to identify emerging research themes, emphasize interdisciplinary collaboration, and reveal global research networks. Utilizing citation analysis, co-authorship analysis, co-occurrence analysis, bibliographic coupling analysis, and co-citation analysis with the use of VOSviewer (version 1.6.19) and Python programming language, the study sheds light on key trends in the integration of tourism development into urban planning. The findings provide a comprehensive understanding of the field's current landscape, highlighting the importance of interdisciplinary cooperation and global research networks. This study offers valuable insights for researchers, practitioners, and policymakers, facilitating the advancement of sustainable urban tourism practices and guiding future research endeavors in integrating tourism development into urban planning.

Keywords: tourism development; bibliometric analysis; literature analysis; VOSviewer; keyword analysis; content analysis; urban planning



Citation: Krabokoukis, T.; Polyzos, S. A Bibliometric Analysis of Integrating Tourism Development into Urban Planning. *Sustainability* **2023**, *15*, 14886. <https://doi.org/10.3390/su152014886>

Academic Editor: J. Andres Coca-Stefaniak

Received: 3 October 2023

Revised: 12 October 2023

Accepted: 14 October 2023

Published: 15 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Tourism and urban planning are two critical areas that have a significant impact on sustainability [1–7], and competition between cities [7–9] is a modern phenomenon that characterizes the functioning of small and large cities. Urban areas are often the primary destinations for tourists, and while tourism development can bring positive economic benefits, it can also have negative effects on the environment and local communities [7].

In the realm of research focusing on sustainability within the fields of tourism and urban planning, a plethora of valuable insights have emerged from recent studies. These studies delve into various facets, including the concept of “over tourism” and its intricate relationship with sustainability [3], the interwoven relationship between tourism, sustainability, and competitiveness [9], and the ever-evolving domain of digital transformation in tourism [10]. They also cast their gaze on the post-COVID era, acknowledging the profound impacts of the pandemic on the tourism sector [2] and on the realm of corporate social responsibility (CSR) within the tourism sector [4]. Additionally, they explore the role of tourism management and marketing in achieving sustainability [5]. These studies highlight the necessity for a holistic approach when deciphering the intricacies of sustainability within the context of tourism and urban planning [6], presenting a multifaceted perspective.

Despite the recognition of these challenges, the existing body of literature lacks a comprehensive analysis of the current state of research in sustainable tourism and urban planning. In this context, our research aims to bridge this gap by conducting an extensive bibliometric analysis of the available literature. We intend to offer a systematic overview of key trends, authors, and research clusters in the field, thus contributing to a more profound understanding of sustainable tourism and urban development. This analysis will serve as a

valuable resource for researchers, practitioners, and policymakers seeking to address the complex issues at the intersection of tourism and urban planning.

In order to grasp the current state of research and identify potential areas for further investigation, it is critical to review the relevant literature [9]. The purpose is to determine where research gaps and opportunities currently exist [2]. This could contribute to the development of tourism practices that minimize negative impacts and maximize positive outcomes. The objective of this paper is to conduct a bibliometric analysis of the literature in this area to identify key trends and gaps.

The study will be based on a comprehensive sample of articles published in international journals, in the Scopus database, and the analysis will be conducted using a variety of tools and techniques, including co-authorship analysis, keyword analysis, and content analysis. The research questions of the study are the following:

1. What are the key trends and themes in tourism and urban planning research, as indicated by the top-cited documents and authors in the Scopus dataset?
2. How do co-authorship patterns among countries reflect collaborations in the field of tourism and urban planning, and what are the implications for international research networks and partnerships?
3. What are the major research clusters and thematic areas in the co-occurrence analysis of keywords, and how do these clusters contribute to our understanding of sustainable tourism development, environmental conservation, and urban growth?
4. How does co-citation analysis shed light on the most influential authors in the field, and what does it tell us about the intellectual foundations and knowledge dissemination in tourism and urban planning research?

The main goals of conducting a bibliometric analysis are to identify gaps and trends in a particular field of research [2–6]. These research questions are, therefore, important for providing a comprehensive analysis of the state of research on sustainable tourism and identifying opportunities for further study. This information can help guide future research efforts and support the development of tourism practices that are more resilient in the face of crises. The research methods used in this study involved the application of bibliometric techniques, specifically citation analysis, co-authorship analysis, co-occurrence analysis, and content analysis. These methods were selected for their ability to provide a comprehensive overview of the literature, including its major themes, authors, and publications [2–6]. However, it is important to note that the bibliometric approach has certain limitations. First, findings could be extended by using other well-known databases [2], such as Web of Science (WoS). Second, the use of keywords might be puzzling [9]. Third is the exclusion of non-English language articles [2] and fourth is the restriction of the database used in this study to articles as the document type and journal as the source type [2,9].

The paper is organized as follows. First, an overview of the literature in the field will be provided. Second, the research methods used in the study, including the sample, data collection, and analysis techniques, will be described. Third, the results of the study, including the main findings and conclusions, will be presented. Finally, the implications of the study for researchers, practitioners, and policymakers working in the field of sustainable tourism will be discussed.

2. Methodology and Data

The present study employed the VOSviewer and the Python language to carry out the methodology and data analysis. The bibliometric analysis was conducted as detailed in Sections 2.1–2.6.

2.1. Dataset

Keywords were then employed in Scopus to extract the database used in the study, which covers the period from 2010 to 11 September 2023. Table 1 illustrates the specific details of the database used. The decision to begin the analysis in 2010 was prompted by the recognition of significant shifts and developments in the field of sustainable tourism and

urban planning during the past decade [7,10]. The emergence of new technologies [11–13], evolving environmental concerns [2,9], and changing urban dynamics [7] have all contributed to a rapidly transforming landscape within this domain. Consequently, it becomes imperative to assess recent research trends and identify the evolving challenges and opportunities faced by researchers, practitioners, and policymakers. The present study aims to bridge this gap by examining these recent developments, providing a comprehensive analysis of the state of research, and shedding light on the emerging trends and areas that warrant further exploration.

Table 1. Data retrieval constraints and parameters for the Scopus database.

Database:	Scopus
Search field:	Title, Abstract, Keywords
Keywords:	Sustainable tourism development, urban planning, urban competitiveness, smart cities, community-based, tourism impacts on tourism areas, policy and governance for sustainable tourism, stakeholder and engagement in urban tourism
Open access:	All
Years:	2010–2024
Author name:	All
Subject area:	All
Publication stage:	All
Document type:	Article
Source title:	All
Affiliation:	All
Funding sponsor:	All
Country:	All
Source type:	Journal
Language:	English
Search string:	(TITLE-ABS-KEY (sustainable AND tourism AND development) OR TITLE-ABS-KEY (urban AND planning AND tourism) OR TITLE-ABS-KEY (urban AND competitiveness AND in AND tourism) OR TITLE-ABS-KEY (smart AND cities AND tourism) OR TITLE-ABS-KEY (community-based AND tourism) OR TITLE-ABS-KEY (tourism AND impacts AND on AND urban AND areas) OR TITLE-ABS-KEY (policy AND governance AND for AND sustainable AND tourism) OR TITLE-ABS-KEY (stakeholder AND engagement AND in AND urban AND tourism)) AND PUBYEAR > 2009 AND PUBYEAR < 2025 AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English")) AND (LIMIT-TO (SRCTYPE, "j"))
Data extracted:	11 September 2023
Number of publications:	11,134

2.2. Citation Analysis

A common method in bibliometric analysis is citation analysis [2–6,9,10,14]. The impact and influence of a particular paper or author within a given field of study are assessed by calculating the frequency with which a single paper or author is mentioned within other publications. This analysis is also regarded as an accurate way to gauge the effect and scope of scientific research.

2.3. Co-Authorship Analysis

In bibliometric research, co-authorship analysis is a widely recognized method for uncovering collaboration trends among authors, institutions, and countries across diverse fields of study [2–6,9,10,14]. Beyond its fundamental role in identifying these collaborative patterns, co-authorship analysis serves as a valuable tool for shedding light on intricate networks of researchers, pinpointing influential figures, institutions, and emerging trends within the subject of inquiry.

2.4. Co-Occurrence Analysis

It is possible to comprehend the connections between keywords or concepts that commonly appear together in the literature by using co-occurrence analysis, a method that

is extensively employed in bibliometric research [2–6,9,10,14]. The author co-occurrence approach involves identifying the most frequently occurring keywords or concepts within a set of papers.

2.5. Bibliographic Coupling Analysis

Bibliographic coupling analysis is a method for probing the interconnectedness and synergy among various scientific publications, all contingent on their shared references [14]. The analysis functions as a tool for document grouping, greatly facilitating the provision of scientific information and enhancing the efficiency of document retrieval. Additionally, it offers an expansive viewpoint on the theoretical basis of the area under examination.

2.6. Co-Citation Analysis

Co-citation analysis emerges as another crucial method for delving into academic citations [2–6,9,10,14]. This technique explores how frequently two documents are jointly cited by other scholarly works, serving as a powerful indicator of their conceptual or thematic affinity. It extends its utility to pinpointing central themes and influential sources within a specific research domain, effectively mapping the intellectual landscape. Incorporating co-citation analysis into the research framework empowers investigators to discern the hidden currents of literature review, revealing the timeless influence of certain ideas and the interplay of intellectual forces.

3. Results

3.1. Results of Citation Analysis

Table 2 provides the structure of citations in the relevant research area by presenting the top 10 documents with the highest number of citations until 11 September 2023. The table highlights the authors, title, and the number of citations for each document. The most cited article is “Integrating local and scientific knowledge for environmental management” by Raymond et al. (2010). These documents represent significant contributions to the field of environmental management, smart and connected communities, sustainable tourism development, and other related topics.

Table 2. Top 10 documents by citations (until 11 September 2023).

No.	Ref.	Title	Citations
1	[15]	Integrating local and scientific knowledge for environmental management	725
2	[13]	Internet of Things and Big Data Analytics for Smart and Connected Communities	640
3	[16]	UTCI-Why another thermal index?	607
4	[17]	Influence analysis of community resident support for sustainable tourism development	570
5	[18]	Green, circular, bio economy: A comparative analysis of sustainability avenues	534
6	[19]	Is overtourism overused? Understanding the impact of tourism in a city context	415
7	[20]	Conceptual foundations for understanding smart tourism ecosystems	413
8	[21]	A critical review of the impacts of COVID-19 on the global economy and ecosystems and opportunities for circular economy strategies	390
9	[22]	The Private Sector and the SDGs: The Need to Move Beyond ‘Business as Usual’	381
10	[23]	Residents’ support for tourism development: The role of residents’ place image and perceived tourism impacts	376

Table 3 provides insights into the top 10 authors, countries, and journals by the number of publications until 11 September 2023. These tables provide a comprehensive overview of the most influential documents, authors, countries, and journals in the field, offering valuable insights into the landscape of academic research and publication.

Table 3. Top 10 authors, countries, and journals by publications (until 11 September 2023).

No.	Author	Number of Documents	No.	Country	Number of Documents	No.	Journal	Number of Documents
1	Giampiccoli, A.	49	1	China	1939	1	Sustainability	1347
2	Hall, C.M.	34	2	USA	928	2	Journal of sustainable tourism	505
3	Mtapuri, O.	30	3	Spain	922	3	African journal of hospitality tourism and leisure	185
4	Castanho, R.A.	24	4	United Kingdom	773	4	Wit transactions on ecology and the environment	157
5	Saarinen, J.	22	5	Australia	649	5	Worldwide hospitality and tourism themes	150
6	Han, H.	20	6	Italy	615	6	Goejournal of tourism and geosites	145
7	Boley, B.B.	19	7	South Africa	463	7	Tourism management	145
8	Couto, G.	18	8	Malaysia	412	8	Current issues in tourism	134
9	Law, R.	18	9	Indonesia	395	9	Land	133
10	Gössling, S.	17	10	Portugal	376	10	Journal of cleaner production	123

3.2. Results of Co-Authorship Analysis

3.2.1. Co-Authorship Analysis Using VOSviewer

Figure 1 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 10, and 95 out of the 205 meet this threshold. The higher total link strengths are for China (989), the United Kingdom (889), the USA (812), Australia (575), Spain (568), Italy (406), South Africa (384), Portugal (309), and Germany (300). The eight clusters identified in the map are described as follows. As shown in Figure 1, the largest cluster, represented in red (cluster 1), consists of 26 countries, followed by the green cluster (cluster 2) with 13 items, the blue cluster (cluster 3) with 13 countries, the yellow cluster (cluster 4) with 12 countries, the purple cluster (cluster 5) with 10 countries, the light blue cluster (cluster 6) with 10 countries, the orange cluster (cluster 7) with 8 countries, and finally, the brown cluster (cluster 8) with 3 countries. This analysis' results are also depicted in Scheme 1.

In addition to the co-authorship analysis presented in Figure 1, it is crucial to delve deeper into the clusters identified, as they offer valuable insights into the collaborative networks among countries within the Scopus dataset.

Cluster 1, represented in red, is the largest cluster, consisting of 26 countries (Albania, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Germany, Estonia, Greece, Hungary, Italy, Iraq, Israel, Latvia, Kazakhstan, Lithuania, Montenegro, Romania, Poland, Russia, Serbia, Slovakia, Slovenia, Ukraine), accounting for 20.02% of the total documents. These countries exhibit a robust co-authorship network, indicating extensive collaboration in research related to the analyzed field. Germany leads with 280 papers, followed by Romania and Poland with 279 each. The strong presence of European countries underscores their active engagement in scholarly activities and collaborative endeavors.

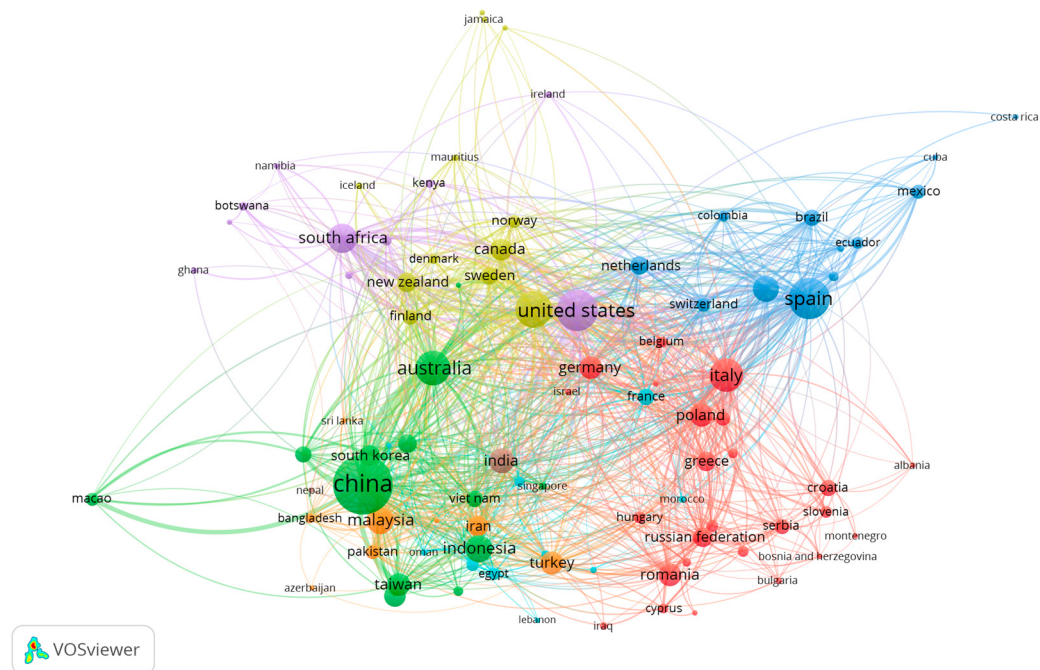
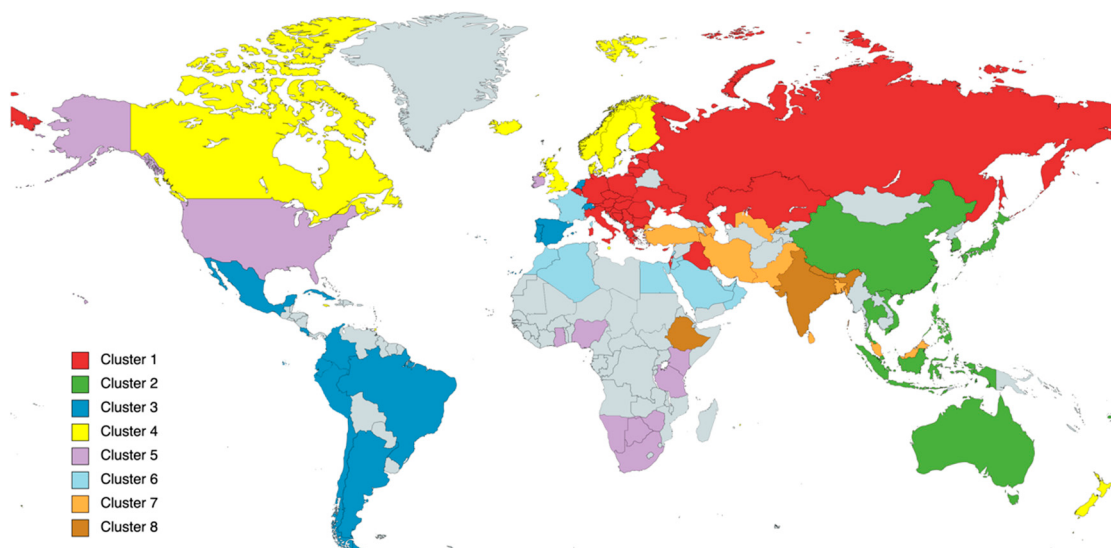


Figure 1. Co-Authorship network analysis of countries in tourism and urban planning research.



Scheme 1. Co-Authorship network analysis of countries in tourism and urban planning research.

The green cluster, cluster 2, includes 13 countries (Australia, China, Fiji, Hong Kong, Indonesia, Japan, Philippines, Singapore, South Korea, Taiwan, Thailand, Vietnam, and Macao), representing 29.20% of the total documents. Japan has the most documents in this cluster. It signifies robust research collaboration among Asian and Oceanian nations, with a particular focus on technology, innovation, and environmental studies. The substantial total link strengths underscore the significance of international cooperation in addressing global challenges.

The blue cluster, cluster 3, features 13 countries (Argentina, Brazil, Chile, Colombia, Costa Rica, Cuba, Mexico, Ecuador, Netherlands, Peru, Spain, Portugal, Switzerland), accounting for 14.22% of the total documents. Spain has the most documents in this cluster. This cluster highlights transcontinental collaboration, showcasing interconnected research efforts between Europe and Latin America. It indicates a shared interest in research

topics transcending geographical boundaries, potentially related to tourism, economics, or sustainable development.

In the yellow cluster, cluster 4, there are 12 countries (Denmark, Canada, Iceland, Finland, Jamaica, Malta, Mauritius, New Zealand, Norway, Sweden, Trinidad and Tobago, the United Kingdom) representing 11.91% of the total documents. The United Kingdom has the most documents in this cluster. These nations exhibit a robust co-authorship network, indicating close research ties. This cluster likely reflects the well-established research infrastructure and academic partnerships in these regions.

The purple cluster, cluster 5, consists of 10 countries (Botswana, Ghana, Ireland, Kenya, Namibia, South Africa, Nigeria, Tanzania, the United States, Zimbabwe), representing 10.88% of the total documents. The United States has the most documents in this cluster. It suggests collaborative research efforts between African and American institutions, with topics potentially encompassing biodiversity, conservation, and sustainable development in these regions.

The light blue cluster, cluster 6, includes 10 countries (Egypt, Algeria, France, Jordan, Morocco, Lebanon, Oman, Qatar, Saudi Arabia, United Arab Emirates), representing 3.55% of the total documents. France has the most documents in this cluster. This cluster underscores the collaboration among countries in the Arab world and North Africa, highlighting shared research interests and scholarly exchanges.

The orange cluster, cluster 7, is composed of eight countries (Azerbaijan, Iran, Bangladesh, Pakistan, Malaysia, Sri Lanka, Turkey, Uzbekistan), representing the 7.40% of the total documents. Malaysia has the most documents in this cluster. This cluster signifies the growing research contributions from these regions and their active involvement in the global academic landscape.

The final brown cluster, cluster 8, consists of three countries (India, Ethiopia, Nepal), representing 2.81% of the total documents. India has the most documents in this cluster. This cluster reflects research ties within South Asia and East Africa, suggesting increased engagement in collaborative research initiatives.

These clusters shed light on the global nature of research collaboration within the field, emphasizing the diverse geographical origins of scholarly contributions and the importance of international partnerships in advancing knowledge and addressing complex challenges.

3.2.2. Deeper Co-Authorship Analysis Using Python

Except for results from the VOSviewer, a Python script was developed to conduct an analysis and extract meaningful keywords per cluster, as defined previously through co-authorship analysis in VOSviewer. Initially, considering that the Scopus database includes an “Affiliations” column containing comprehensive information (such as Department, University, Address, etc.) rather than just the country, the clusters, each representing a group of countries, were defined. Subsequently, a function was implemented to extract country information from the “Affiliations” column. This function iterated through the “Affiliations”, matching them with predefined lists of countries in each cluster. If a match was found, the country was assigned. Otherwise, it was labeled as “Unknown”. Subsequently, this function was applied to create a new “Country” column in the dataset, facilitating the organization of articles by cluster based on their country affiliations.

Subsequently, a keyword extraction mechanism was implemented for the articles within each cluster. The CountVectorizer from the Scikit-learn library was employed to convert the article text into numerical vectors and calculate word frequencies. To ensure the meaningfulness of the extracted keywords, non-noun words were filtered out using WordNet, a lexical database for the English language. For each article, its title, abstract, author keywords, and index keywords were concatenated into a single text corpus. This corpus was processed, with any missing values replaced by empty strings, and the CountVectorizer was applied to generate vectors.

Upon obtaining the word frequencies, the results were filtered to include only nouns through WordNet’s part-of-speech tagging. This step enhanced the extraction of context-

tually relevant keywords. Ultimately, the top noun keywords, typically up to 50, were selected for each article within a cluster.

The script then presented the results, showcasing the top keywords for each cluster. These keywords were arranged in descending order of frequency and limited to the most significant nouns. This approach was designed to furnish a concise and meaningful set of keywords for each cluster of articles, facilitating a more profound comprehension and categorization of research topics.

In conclusion, this Python script, entitled “Automated Categorization and Keyword Extraction of Academic Articles Based on Author Affiliations”, is a valuable tool for researchers aiming to analyze and categorize academic articles based on author affiliations, as well as extract pertinent keywords. The code can be found in the Appendix A, and the results are shown in Table 4. Additionally, the code’s flowchart is presented in Figure 2.

Table 4. Top keywords per cluster as extracted from Python script.

Cluster	Cluster Title	Extracted Keywords
Cluster 1	European perspective	tourism, development, study, management, tourist, sustainability, analysis, local, research, area, areas, heritage, planning, destination, social, results, approach, paper, use, ecotourism, impact, natural, model, tourists, data, authors, water, region, landscape, land, community, environment, activities, energy, policy, case, city
Cluster 2	Asia–Pacific	tourism, development, study, china, analysis, management, tourist, research, model, sustainability, community, ecotourism, industry, results, area, social, local, environment, data, tourists, carbon, planning, areas, heritage, impact, system, conservation, resources, destination, paper, city, factors, water, land, use, value, method, quality, approach, ecosystem
Cluster 3	Spanish-speaking regions	tourism, development, management, sustainability, tourist, analysis, study, social, spain, water, areas, destination, local, research, planning, area, heritage, community, results, impact, model, approach, use, can, paper, ecotourism, change, natural, destinations, conservation, policy, case, city, activities, ecosystem, growth, data, authors
Cluster 4	Regions known for environmental consciousness	tourism, development, sustainability, management, study, local, community, an, social, planning, research, ecotourism, nature, conservation, tourist, change, use, policy, destination, climate, paper, analysis, areas, approach, land, industry, taylor, communities, tourists, results, group, impact, was, impacts, national, area, case, limited
Cluster 5	African context	tourism, development, conservation, community, management, ecotourism, wildlife, study, local, natural, resource, botswana, communities, policy, sustainability, approach, benefits, areas, data, national, use, nature, energy, tourist, resources, analysis, social, area, climate, park, change, paper, planning, ecosystem, hunting, namibia, using, activities, africa
Cluster 6	Middle East and North Africa	tourism, development, heritage, study, analysis, management, area, sustainability, egypt, tourist, research, planning, areas, social, local, ecotourism, land, approach, case, community, site, natural, model, sites, water, energy, paper, growth, use, cities, industry, results, conservation, policy, change, system, algeria, city
Cluster 7	Eurasian and South Asian Group	tourism, development, study, malaysia, ecotourism, management, sustainability, iran, analysis, social, using, model, local, results, area, heritage, planning, water, community, growth, research, tourist, conservation, carbon, data, use, areas, land, approach, tourists, energy, environment, method, system, decision, quality, natural, industry, impact, criteria
Cluster 8	South Asian	tourism, development, study, india, ecotourism, conservation, analysis, management, local, sustainability, land, community, research, area, data, water, tourist, natural, use, model, using, heritage, tourists, areas, nature, assessment, change, impact, environment, quality, approach, planning, region, city, livelihood, potential, activities, resources, ecosystem

In Cluster 1 (European perspective), the research primarily focuses on various facets of tourism development integrated into urban planning. Topics such as sustainability, local management, and heritage preservation emerge as key themes. Researchers delve into the sustainable management of tourist areas, with a keen interest in ecological impacts and social factors.

In Cluster 2 (Asia–Pacific), research emphasizes aspects such as community involvement, ecotourism, and sustainability. It explores the complex dynamics of tourism management in these regions, with a particular focus on carbon footprint reduction, industry growth, and conservation efforts.

Cluster 3 (Spanish-speaking regions) sheds light on sustainable tourism management, social impacts, and water resource utilization. Researchers within this cluster closely examine the relationship between tourism and local communities, advocating for responsible planning. Key research areas encompass destination management, cultural heritage preservation, and environmental policy.

In Cluster 4 (regions known for environmental consciousness), research focuses on the intricate balance between tourism development and conservation, on nature conservation, climate change adaptation, and local community engagement. Notable topics include the impacts of tourism on natural resources, policy frameworks, and climate-conscious planning.

In Cluster 5 (African context), research is deeply rooted in wildlife conservation, natural resource management, and community benefits. It examines the positive and negative impacts of tourism on local communities and wildlife habitats. Key themes include ecotourism, energy sustainability, and the economic potential of wildlife-related activities.

In Cluster 6 (Middle East and North Africa), research explores heritage tourism, natural resource management, and community involvement. Notable topics include site-specific conservation, cultural heritage preservation, and sustainable urban planning. It provides valuable insights into the preservation of historical and natural assets in the context of tourism development.

In Cluster 7 (Eurasian and South Asian Group), research emphasizes the utilization of resources, carbon footprint reduction, and environmental impact assessment. Key areas of interest encompass water resource management, community growth, and sustainable tourism criteria. It offers valuable insights into sustainable practices and decision frameworks.

In Cluster 8 (South Asian), research examines the ecological and economic dimensions of tourism development. Topics of interest include natural resource assessment, environmental impact, and potential livelihood opportunities. It sheds light on the intricate relationship between tourism, conservation, and community well-being.

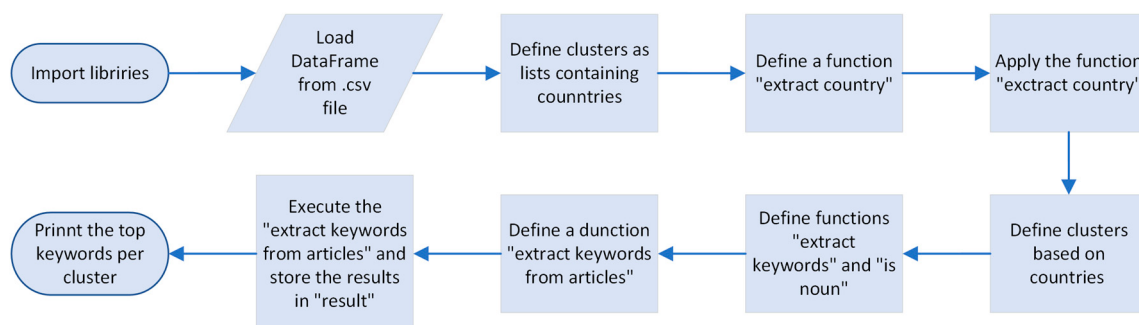


Figure 2. Flowchart from the Python script “Automated Categorization and Keyword Extraction of Academic Articles Based on Author Affiliations”.

3.3. Results of Co-Occurrence Analysis

Figure 3 displays a co-occurrence map of all keywords in the Scopus dataset. The map is based on a threshold minimum number of occurrences of 200, and only 53 out of the 34,327 keywords meet this threshold. The higher total link strengths are for sustainable development (10,965), tourism development (9473), ecotourism (8229), tourism (8178), sustainability (6596), tourist destination (4636), China (4172), sustainable development (3634), and tourism management (3599). The three clusters identified in the map are described as follows.

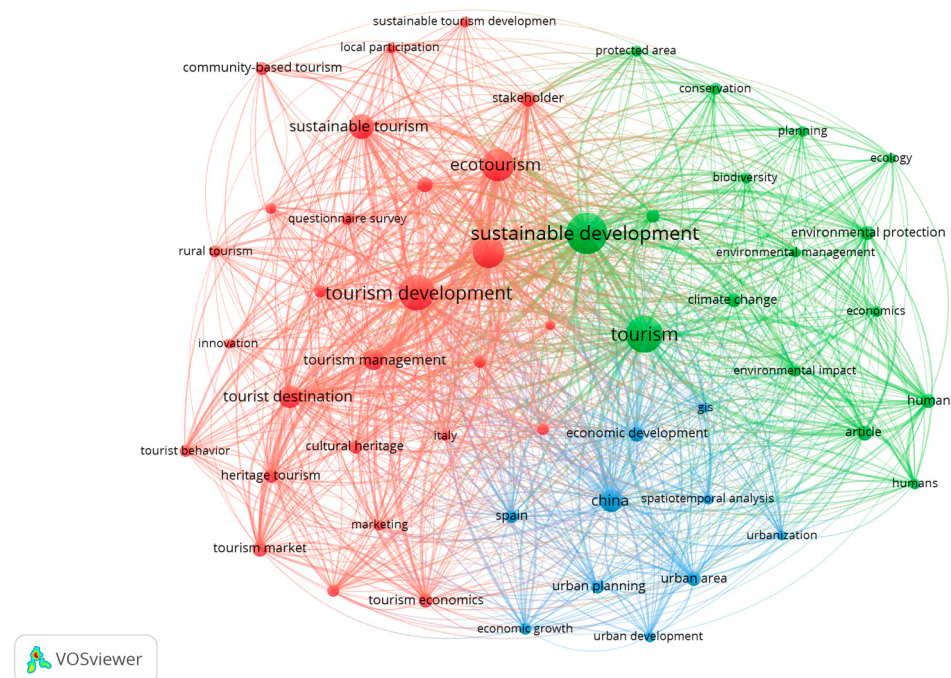


Figure 3. Co-occurrence analysis by all keywords.

Utilizing these keywords, Python was employed within the Jupyter Notebook of Anaconda to classify the articles in the database into their respective clusters. The code, entitled “Automated Article Categorization Script based on Keywords”, can be found in the Appendix B. Additionally, the code’s flowchart is presented in Figure 4.

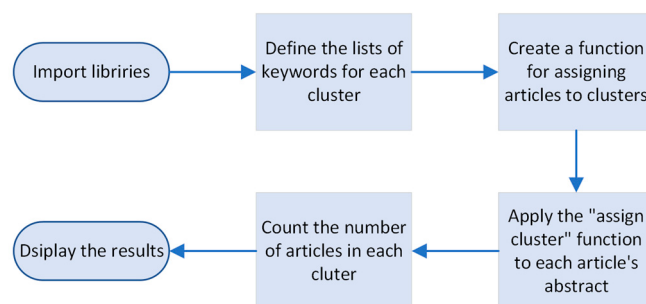


Figure 4. Flowchart from the Python script “Automated Article Categorization Script based on Keywords”.

Table 5 provides an overview of the distribution of research papers among the identified clusters. It is evident that the research predominantly focuses on the first two clusters. The Python code defines a script for assigning articles to different clusters based on the presence of specific keywords within their abstracts. First, the code defines lists of keywords associated with each cluster. These keywords represent the main themes or topics related to each cluster. Then, it defines a Python function called `assign_cluster(article_text)` that takes the text of an article’s abstract as input and assigns it to one of the predefined clusters based on the presence of specific keywords. If any of the keywords in the article text match those in the predefined lists for a cluster, the article is assigned to that cluster. If none of the keywords match, the article is labeled as “Uncategorized”.

Table 5. Distribution of research papers among clusters.

Cluster	Number of Papers
1	4640
2	6157
3	67
Uncategorized	270
Total	11,134

Next, the code applies the `assign_cluster` function to each article’s abstract in a dataset referred to as data. The code counts the number of articles assigned to each cluster and stores these counts in a variable called `cluster_counts`. Finally, the code displays the results by printing out the number of articles assigned to each cluster along with the cluster’s name.

Cluster 1 (red color) is entitled “Sustainable Tourism Development and Governance Approaches”. This cluster of keywords centers around the sustainable development of tourism in urban and rural areas, with a particular focus on governance and community involvement. “Community-based tourism”, “sustainable tourism”, and “sustainable tourism development” form the core themes of this cluster, highlighting the importance of fostering tourism that benefits both the environment and local communities. The inclusion of “COVID-19” indicates a contemporary concern, emphasizing the need for resilient tourism practices in the face of global challenges [2]. “Local participation” and “stakeholder” underscore the significance of involving local communities and various stakeholders in decision-making processes. “Innovation” and “strategic approach” highlight the evolving nature of tourism management, while “marketing” and “perception” delve into the ways in which tourism destinations are promoted and perceived. Together, these keywords suggest that this cluster explores sustainable tourism development through a governance lens, with an emphasis on community involvement, strategic planning, and adaptability.

Cluster 2 (green color) is entitled “Environmental Conservation and Sustainable Tourism”. This cluster is centered around the intersection of environmental conservation and sustainable tourism. Keywords like “biodiversity”, “climate change”, “conservation”, and “environmental impact” emphasize the importance of mitigating the negative effects of tourism on the environment. “Sustainable development” and “tourism” highlight the overarching theme of balancing tourism activities with environmental preservation. “Protected area” signifies the significance of safeguarding natural habitats and landscapes in the context of tourism. The presence of “decision making” and “planning” keywords underscores the necessity for informed choices and strategic planning when it comes to tourism development in ecologically sensitive regions. Overall, this cluster explores the challenges and opportunities for harmonizing tourism with environmental conservation, with a focus on sustainable practices, policy decisions, and ecological preservation.

Cluster 3 (blue color) is entitled “Urban Development and Economic Growth”. Cluster 3 primarily revolves around urban development and economic growth, particularly in the context of China and Spain. “Economic development” and “economic growth” are central themes, indicating a focus on the economic implications of tourism in urban areas. “Urban planning” and “urbanization” highlight the role of planning and development strategies in shaping urban landscapes for tourism. “Spatiotemporal analysis” and “GIS” (Geographic Information Systems) suggest a data-driven approach to understanding the spatial and temporal aspects of urban tourism development. “China” and “Spain” are specific geographical contexts where urban tourism and development are of particular interest. This cluster explores how tourism can drive economic growth and development in urban regions, with an emphasis on geographic and economic analyses.

3.4. Results of Bibliographic Coupling Analysis

Figure 5 presents a bibliographic coupling network analysis that highlights the interconnections among authors within the Scopus dataset. To create this network, a minimum threshold of four documents authored by an individual was applied, leading to the identifi-

cation of 30 authors out of the initial 10,624 authors in the dataset. Notably, this network showcases the collaborative relationships and scholarly connections among these authors. It is worth mentioning that while 30 authors are included in the network, not all of them are directly connected to each other, which suggests diverse research interests and collaboration patterns within this group. The figure, however, primarily focuses on the largest cluster of connected authors, which includes 26 individuals.

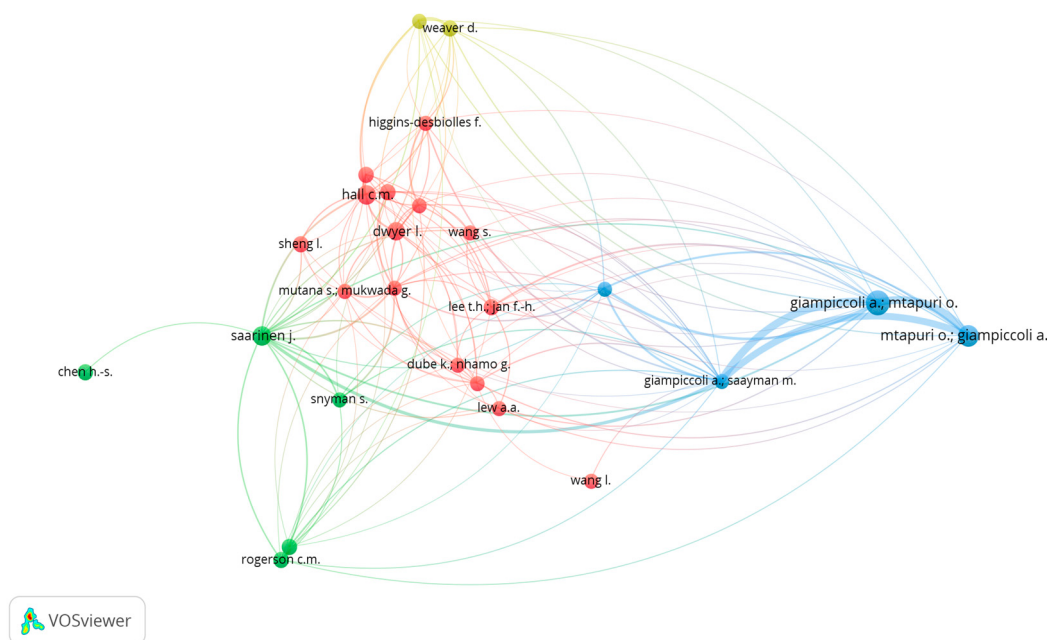


Figure 5. Bibliographic coupling network analysis of influential authors in tourism and urban planning research.

Key insights from this bibliographic coupling analysis include the identification of authors with the highest total link strengths in the network, such as Giampiccoli, A. (326), Mtapuri, O. (326), Giampiccoli, A., Saayman, M. (317), Mtapuri, O., Giampiccoli, A. (273), Saanen, J. (97), and Ramaano, A.I. (61). These authors are pivotal within the network, signifying their significant contributions and collaborations in the field of tourism and urban planning research.

Furthermore, the figure showcases the largest cluster of connected authors, depicted in red, which consists of 15 items. This cluster represents a cohesive group of authors with shared research interests and collaborative endeavors. Additionally, there are smaller clusters represented by the green, blue, and yellow nodes, which consist of five, four, and two items, respectively. These clusters reflect the various subgroups and themes within the broader field of tourism and urban planning research, highlighting the multifaceted nature of the scholarly community in this domain.

3.5. Results of Co-Citation Analysis

Figure 6 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 392 out of the 370,130 authors meeting this criterion. Notably, the authors with the highest total link strengths in the network are Hall, C.M. (110,643), Gossling, S. (70,263), Gursoy, D. (61,075), Wang, Y. (54,898), Zhang, Y. (53,606), Zhang, J. (52,378), Liu, Y. (50,613), Nunkoo, R. (49,566), and Li, Y. (49,144). As depicted in Figure 4, the largest cluster, represented in red, consists of 163 authors, followed by the green cluster with 123 authors, the blue cluster with 79 authors, the yellow cluster with 12 authors, the purple cluster with 9 authors, and finally, the blue cluster with 8 authors.

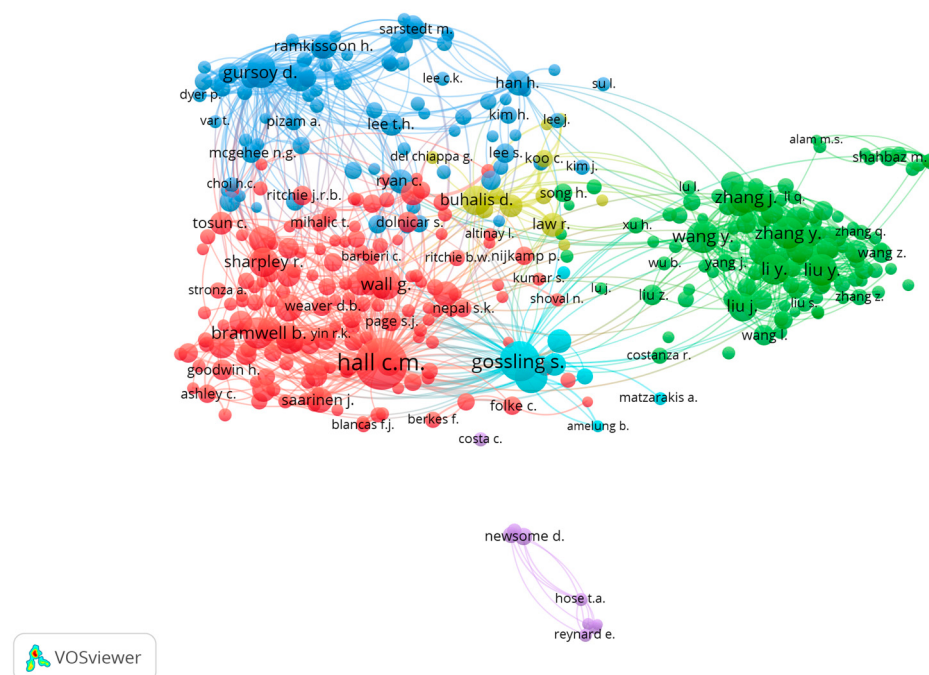


Figure 6. Bibliographic coupling network analysis of influential authors in tourism and urban planning research.

4. Conclusions

In this comprehensive bibliometric analysis within the fields of sustainable tourism and urban planning, we aimed to unearth trends, gaps, and influential factors, exploring the deeper implications of our findings. The study unveiled several pivotal discoveries.

The examination of top-cited documents and authors within the Scopus dataset brought to light significant trends and themes in tourism and urban planning research. Notable works, such as “Integrating local and scientific knowledge for environmental management” and “Internet of Things and Big Data Analytics for Smart and Connected Communities”, underscored the profound importance of interdisciplinary approaches and the integration of emerging technologies within these domains.

Through co-authorship analysis conducted with VOSviewer, we brought out the mosaic of tourism development integrated into urban planning. This endeavor identified eight distinct clusters, emphasizing the global reach of research collaboration and the remarkable diversity in the geographical origins of scholarly contributions. Moreover, the Python script developed for keyword extraction enriched our understanding of research topics within each cluster, providing a comprehensive array of meaningful keywords. These revelations lay a foundation for future research directions and foster interdisciplinary collaborations within this dynamic field.

Co-occurrence analysis of keywords offered profound insights into thematic areas permeating the literature. Two major clusters emerged: “Sustainable Tourism Development and Governance Approaches” and “Environmental Conservation and Sustainable Tourism”. These clusters underscored the significance of governance, community involvement, and environmental conservation in the realm of sustainable tourism practices. They also emphasized the necessity to strike a harmonious balance between economic growth and ecological preservation.

Delving into co-citation analysis revealed the most influential authors in the field, including luminaries such as Hall, C.M., Gossling, S., and Gursoy, D., among others. These authors have made seminal contributions to the literature, effectively shaping the intellectual foundations and facilitating the dissemination of knowledge within the spheres of tourism and urban planning research.

The implications derived from this study extend to researchers, practitioners, and policymakers, each holding a distinct role in translating these discoveries into real-world impact. Researchers can use our findings as a launching pad for their future inquiries, establishing a robust foundation for deeper investigations. This study fervently advocates the value of interdisciplinary approaches and the seamless integration of cutting-edge technologies within their research endeavors.

Practitioners, whose mission is to bring sustainable initiatives to life, can glean vital insights from our analysis. Moving beyond the theoretical realm, they can embrace the pivotal role of governance approaches in achieving the delicate equilibrium between economic growth and environmental conservation. Furthermore, this study underscores the transformative role of community engagement in molding the landscape of sustainable urban tourism.

Policymakers, as the architects of decision making, are urged to draw extensively from the knowledge generated by our analysis. The theme of international collaboration is underscored, illuminated by the global network unveiled in the co-authorship analysis. Additionally, this study stresses the importance of policies that not only foster tourism, urban development, and economic growth but also prioritize environmental conservation and community well-being.

However, it is essential to acknowledge the limitations of our analysis. Our bibliometric approach focused on English-language articles within the Scopus database, potentially omitting valuable non-English literature. Additionally, the analysis centered on articles as the document type, excluding other publication forms from consideration.

5. Future Direction

Given the increasing importance of technology and data analytics in tourism, further investigation into smart tourism ecosystems is warranted. This may involve exploring the integration of IoT, big data, and AI to enhance tourist experiences and resource management. Additionally, the COVID-19 pandemic underscored the vulnerability of tourism to global crises. Future research can delve deeper into strategies for building resilience in the tourism industry, including crisis management, risk assessment, and adaptive strategies. Lastly, with growing environmental concerns, the concept of a circular economy is gaining traction. Future studies can explore circular economy strategies in the context of sustainable tourism, investigating how tourism activities can minimize waste and promote sustainability. In such a context, future research can benefit from cross-disciplinary collaborations between tourism experts, urban planners, environmental scientists, and technology specialists. This interdisciplinary synergy can lead to innovative solutions and a more comprehensive understanding of complex issues.

Additionally, to gain deeper insights into sustainable tourism and urban planning, researchers can conduct comparative studies across different regions and countries. Comparing strategies, policies, and outcomes in diverse contexts can reveal best practices and highlight areas where specific interventions are needed. In this framework, future research can investigate effective methods for involving local communities, businesses, and governments in sustainable tourism initiatives. Understanding the dynamics of stakeholder engagement can lead to more successful and inclusive tourism practices.

Author Contributions: Conceptualization, T.K.; methodology, T.K.; software, T.K.; validation, S.P. and T.K.; formal analysis, T.K.; investigation, T.K.; resources, T.K.; data curation, T.K.; writing—original draft preparation, T.K.; writing—review and editing, T.K.; visualization, T.K. and S.P.; supervision, S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Python code. Automated categorization and keyword extraction of academic articles based on author affiliations.

```

import pandas as pd
import numpy as np
from sklearn.feature_extraction.text import CountVectorizer
import nltk
from nltk.corpus import wordnet
import nltk
nltk.download('wordnet')

# Define clusters as lists
cluster1 = ["Albania", "Austria", "Belgium", "Bosnia_and_Herzegovina", "Bulgaria", "Croatia",
"Cyprus", "Czechia", "Germany", "Estonia", "Greece", "Hungary", "Italy", "Iraq", "Israel", "Latvia",
"Kazakhstan", "Lithuania", "Montenegro", "Romania", "Poland", "Russia", "Serbia", "Slovakia",
"Slovenia", "Ukraine"]
cluster2 = ["Australia", "China", "Fiji", "Hong_Kong", "Indonesia", "Japan", "Philippines",
"Singapore", "South_Korea", "Taiwan", "Thailand", "Vietnam"]
cluster3 = ["Argentina", "Brazil", "Chile", "Colombia", "Costa_Rica", "Cuba", "Mexico", "Ecuador",
"Netherlands", "Peru", "Spain", "Portugal", "Switzerland"]
cluster4 = ["Denmark", "Canada", "Iceland", "Finland", "Jamaica", "Malta", "Mauritius", "New_
Zealand", "Norway", "Sweden", "Trinidad_and_Tobago", "United_Kingdom"]
cluster5 = ["Botswana", "Ghana", "Ireland", "Kenya", "Namibia", "South_Africa", "Nigeria",
"Tanzania", "United_States", "Zimbabwe"]
cluster6 = ["Egypt", "Algeria", "France", "Jordan", "Morocco", "Lebanon", "Oman", "Qatar", "Saudi_
Arabia", "United_Arab_Emirates"]
cluster7 = ["Azerbaijan", "Iran", "Bangladesh", "Pakistan", "Malaysia", "Sri_Lanka", "TOΦrkiye",
"Uzbekistan"]
cluster8 = ["India", "Ethiopia", "Nepal"]

# Define a function to extract the country from the affiliation
def extract_country(affiliation):
    if pd.isna(affiliation):
        return "Unknown"
    for country in cluster1:
        if country in affiliation:
            return country
    for country in cluster2:
        if country in affiliation:
            return country
    for country in cluster3:
        if country in affiliation:
            return country
    for country in cluster4:
        if country in affiliation:
            return country
    for country in cluster5:
        if country in affiliation:
            return country
    for country in cluster6:
        if country in affiliation:
            return country
    for country in cluster7:
        if country in affiliation:
            return country
    for country in cluster8:
        if country in affiliation:
            return country
    return "Unknown"

```

Table A1. *Cont.*

```

# Apply the function to create the "Country" column
df['Country'] = df['Affiliations'].apply(extract_country)

# Define clusters based on countries
clusters = {
    "Cluster 1": df[df['Country'].apply(lambda x: x in cluster1)],
    "Cluster 2": df[df['Country'].apply(lambda x: x in cluster2)],
    "Cluster 3": df[df['Country'].apply(lambda x: x in cluster3)],
    "Cluster 4": df[df['Country'].apply(lambda x: x in cluster4)],
    "Cluster 5": df[df['Country'].apply(lambda x: x in cluster5)],
    "Cluster 6": df[df['Country'].apply(lambda x: x in cluster6)],
    "Cluster 7": df[df['Country'].apply(lambda x: x in cluster7)],
    "Cluster 8": df[df['Country'].apply(lambda x: x in cluster8)],
}

# Function to extract keywords from a text
def extract_keywords(text, num_keywords=10):
    # Define a CountVectorizer for text analysis
    vectorizer = CountVectorizer()

    # Convert the text to vectors
    X = vectorizer.fit_transform([text])

    # Calculate the sum of word frequencies
    word_freq = list(zip(vectorizer.get_feature_names_out(), X.sum(axis=0).tolist()[0]))

    # Sort keywords by frequency and select the top ones
    sorted_word_freq = sorted(word_freq, key=lambda x: x[1], reverse=True)
    top_keywords = [word[0] for word in sorted_word_freq[:num_keywords]]

    return top_keywords

# Function to check if a word is a noun
def is_noun(word):
    # Use WordNet to determine the word's part of speech
    synsets = wordnet.synsets(word)
    for synset in synsets:
        if synset.pos() == 'n': # 'n' means noun
            return True
    return False

# Function to extract keywords from articles in a cluster
def extract_keywords_from_articles(articles, num_keywords=50):
    all_keywords = []
    corpus = articles['Title'] + ' ' + articles['Abstract'] + ' ' + articles['Author Keywords'] + ' ' + articles['Index Keywords']

    # Replace NaN values with an empty string
    corpus = corpus.fillna("")

    vectorizer = CountVectorizer()
    X = vectorizer.fit_transform(corpus)

    word_freq = list(zip(vectorizer.get_feature_names_out(), X.sum(axis=0).tolist()[0]))
    sorted_word_freq = sorted(word_freq, key=lambda x: x[1], reverse=True)

    # Filter results to include only nouns
    noun_keywords = [word[0] for word in sorted_word_freq if is_noun(word[0])]

```

Table A1. *Cont.*

```

# Select the top noun keywords
top_noun_keywords = noun_keywords[:num_keywords]

return top_noun_keywords

# Call the function to extract the top keywords per cluster
result = {}
for cluster, articles in clusters.items():
    keywords = extract_keywords_from_articles(articles)
    result[cluster] = keywords
# Print the top keywords per cluster
for cluster, keywords in result.items():
    print(f"Cluster: {cluster}")
    print(f"Keywords: {' '.join(keywords)}")
    print()

```

Appendix B

Table A2. Automated article categorization script based on keywords.

```

# Define the lists of keywords for each cluster
cluster1_keywords = ["Community-based tourism", "sustainable tourism", "sustainable tourism development", "COVID-19", "local participation", "stakeholder", "innovation", "strategic approach", "marketing", "perception"]
cluster2_keywords = ["biodiversity", "climate change", "conservation", "environmental impact", "sustainable development", "tourism", "protected area", "decision making", "planning"]
cluster3_keywords = ["economic development", "economic growth", "urban planning", "urbanization", "spatiotemporal analysis", "GIS", "China", "Spain"]

# Create a function for assigning articles to clusters
def assign_cluster(article_text):
    if any(keyword in article_text for keyword in cluster1_keywords):
        return "Cluster 1 (Sustainable Tourism Development and Governance Approaches)"
    elif any(keyword in article_text for keyword in cluster2_keywords):
        return "Cluster 2 (Environmental Conservation and Sustainable Tourism)"
    elif any(keyword in article_text for keyword in cluster3_keywords):
        return "Cluster 3 (Urban Development and Economic Growth)"
    else:
        return "Uncategorized"

# Apply the function to each article and count the clusters
data['Cluster'] = data['Abstract'].apply(assign_cluster)
cluster_counts = data['Cluster'].value_counts()

# Display the results
for cluster, count in cluster_counts.items():
    print(f"{cluster}: {count} articles")

```

References

1. Braun, E.; Kavaratzis, M.; Zenker, S. My city—My brand: The different roles of residents in place branding. *J. Place Manag. Dev.* **2013**, *6*, 18–28. [\[CrossRef\]](#)
2. Krabokoukis, T. Exploring the State of Research on Tourism Sustainability: A Bibliometric Analysis in the Post-COVID Era. *Highlights Sustain.* **2023**, *2*, 50–61. [\[CrossRef\]](#)
3. Santos-Rojo, C.; Llopis-Amorós, M.; García-García, J.M. Overtourism and sustainability: A bibliometric study (2018–2021). *Technol. Forecast. Soc. Chang.* **2023**, *188*, 122285. [\[CrossRef\]](#)
4. Sánchez-Camacho, C.; Carranza, R.; Martín-Consuegra, D.; Díaz, E. Evolution, trends and future research lines in corporate social responsibility and tourism: A bibliometric analysis and science mapping. *Sustain. Dev.* **2022**, *30*, 462–476. [\[CrossRef\]](#)

5. Pahrudin, P.; Liu, L.-W.; Li, S.-Y. What Is the Role of Tourism Management and Marketing toward Sustainable Tourism? A Bibliometric Analysis Approach. *Sustainability* **2022**, *14*, 4226. [\[CrossRef\]](#)
6. Serrano, L.; Sianes, A.; Ariza-Montes, A. Using Bibliometric Methods to Shed Light on the Concept of Sustainable Tourism. *Sustainability* **2019**, *11*, 6964. [\[CrossRef\]](#)
7. Polyzos, S. *Urban Development*, 2nd ed.; Kritiki Publishing: Athens, Greece, 2023.
8. Cracolici, M.F.; Nijkamp, P. The attractiveness and competitiveness of tourist destinations: A study of Southern Italian regions. *Tour. Manag.* **2009**, *30*, 336–344. [\[CrossRef\]](#)
9. Seguí-Amortegui, L.; Clemente-Almendros, J.A.; Medina, R.; Gala, M.G. Sustainability and Competitiveness in the Tourism Industry and Tourist Destinations: A Bibliometric Study. *Sustainability* **2019**, *11*, 6351. [\[CrossRef\]](#)
10. Madzík, P.; Falát, L.; Copuš, L.; Valeri, M. Digital transformation in tourism: Bibliometric literature review based on machine learning approach. *Eur. J. Innov. Manag.* **2023**, *26*, 177–205. [\[CrossRef\]](#)
11. Gössling, S. Technology, ICT and tourism: From big data to the big picture. *J. Sustain. Tour.* **2021**, *29*, 849–858. [\[CrossRef\]](#)
12. Buhalis, D.; O’connor, P.; Leung, R. Smart hospitality: From smart cities and smart tourism towards agile business ecosystems in networked destinations. *Int. J. Contemp. Hosp. Manag.* **2023**, *35*, 369–393. [\[CrossRef\]](#)
13. Sun, Y.; Song, H.; Jara, A.J.; Bie, R. Internet of Things and Big Data Analytics for Smart and Connected Communities. *IEEE Access* **2016**, *4*, 766–773. [\[CrossRef\]](#)
14. Jarneving, B. Bibliographic coupling and its application to research-front and other core documents. *J. Informetr.* **2007**, *1*, 287–307. [\[CrossRef\]](#)
15. Raymond, C.M.; Fazey, I.; Reed, M.S.; Stringer, L.C.; Robinson, G.M.; Evely, A.C. Integrating local and scientific knowledge for environmental management. *J. Environ. Manag.* **2010**, *91*, 1766–1777. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Jendritzky, G.; de Dear, R.; Havenith, G. UTCI—Why another thermal index? *Int. J. Biometeorol.* **2012**, *56*, 421–428. [\[CrossRef\]](#)
17. Lee, T.H. Influence analysis of community resident support for sustainable tourism development. *Tour. Manag.* **2013**, *34*, 37–46. [\[CrossRef\]](#)
18. D’Amato, D.; Droste, N.; Allen, B.; Kettunen, M.; Lähtinen, K.; Korhonen, J.; Leskinen, P.; Matthies, B.D.; Toppinen, A. Green, circular, bio economy: A comparative analysis of sustainability avenues. *J. Clean. Prod.* **2017**, *168*, 716–734. [\[CrossRef\]](#)
19. Koens, K.; Postma, A.; Papp, B. Is Overtourism Overused? Understanding the Impact of Tourism in a City Context. *Sustainability* **2018**, *10*, 4384. [\[CrossRef\]](#)
20. Gretzel, U.; Werthner, H.; Koo, C.; Lamsfus, C. Conceptual foundations for understanding smart tourism ecosystems. *Comput. Hum. Behav.* **2015**, *50*, 558–563. [\[CrossRef\]](#)
21. 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.; et al. A critical analysis of the impacts of COVID-19 on the global economy and ecosystems and opportunities for circular economy strategies. *Resour. Conserv. Recycl.* **2021**, *164*, 105169. [\[CrossRef\]](#)
22. Scheyvens, R.; Banks, G.; Hughes, E. The Private Sector and the SDGs: The Need to Move Beyond ‘Business as Usual’: The Private Sector and the SDGs: Moving Beyond ‘Business-as-Usual’. *Sustain. Dev.* **2016**, *24*, 371–382. [\[CrossRef\]](#)
23. Styliadis, D.; Biran, A.; Sit, J.; Szivas, E.M. Residents’ support for tourism development: The role of residents’ place image and perceived tourism impacts. *Tour. Manag.* **2014**, *45*, 260–274. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.