Article

What Are the Drivers of Citations?: Application in Tourism and Hospitality Journals

Eunhye Park 1 and Woohyuk Kim 2,*

1 Department of Food and Nutrition, Gachon University, Seongnam 1342, Korea; epark@gachon.ac.kr
2 Department of Consumer Science, Incheon National University, Incheon 22012, Korea
* Correspondence: woohyuk@inu.ac.kr

Abstract: In line with the qualitative and quantitative growth of academic papers, it is critical to understand the factors driving citations in scholarly articles. This study discovered the up-to-date academic structure in the tourism and hospitality literature and tested the comprehensive sets of factors driving citation counts using articles published in first-tier hospitality and tourism journals found on the Web of Science. To further test the effects of research topic structure on citation counts, unsupervised topic modeling was conducted with 9910 tourism and hospitality papers published in 12 journals over 10 years. Articles specific to online media and the sharing economy have received numerous citations and that recently published papers with particular research topics (e.g., rural tourism and eco-tourism) were frequently cited. This study makes a major contribution to hospitality and tourism literature by testing the effects of topic structure and topic originality discovered by text mining on citation counts.

Keywords: citation; journals; tourism; scientific impact; text mining; big data

1. Introduction

Many tourism and hospitality researchers have long sought answers to questions, such as “how are knowledge and theories produced, and how have they evolved to influence academic discipline?” and “to what extent has hospitality and tourism research reached maturity?” [1,2]. As a result of the constant efforts of numerous researchers, the number of related scholarly and academic papers has swiftly increased, and the knowledge of hospitality and tourism has become substantially diversified, evolving rapidly [3–5]. Owing to the interdisciplinary nature of tourism and hospitality, knowledge progression has been attained through the wide application of concepts, ideas, and theories in various disciplines [6]. This leads to a complex academic structure of tourism and hospitality research, creating sub-categories of diverse research topics [7]. Hence, various efforts have been made to uncover the knowledge structure by discovering prevalent research topics in tourism and hospitality research [8–10].

While knowledge development progresses through close relationships between research topics, the growth of academic literature has been facilitated through collaborations among researchers [4,11]. Racherla and Hu [12] posited that citations can represent “the cognitive structure of the scientific communities” (p. 1015). Xiao, et al. [13] also proposed that citations can indicate how knowledge is diffused and utilized in knowledge networks. However, the rapid increase in the quality and quantity of tourism and hospitality research has led fellow researchers to choose only selective articles for citations among numerous published manuscripts [14].

Citations are regarded as a proxy for the scientific impact of individual articles in knowledge networks or scholarly achievements [15,16]. Citations can also demonstrate the strength of associations among researchers within the academic network [17]. However, many studies report that various external factors that may not be relevant to the quality of research can also enhance the visibility of research papers to attract more citations [18].
In addition, the citation patterns may vary according to the sub-categories of research topics owing to the different degrees of theory development and methodological advancement of each research topic [13]. However, little is known about citation patterns in specific research topics and the influence of external factors on citation counts [14,19]. As previous studies primarily adopted manual content analysis to discover research topics and discern academic structure from journal articles [14,19], it is difficult to identify the interrelationships among such research topics. Moreover, only a small sample size (i.e., 300 or less) was utilized for data analysis, and issues with researchers’ subjectivity in coding were raised.

To overcome the aforementioned limitations, this study adopted a machine learning approach to discover research topics from numerous articles and to implement a topic network analysis to discover the up-to-date academic structure. In addition, citation patterns are investigated considering the research topics and other external factors that drive citation counts. Based on the discussions above, this study aimed to test how research topic attributes can influence citation counts, in addition to other citation determinants, such as journal attributes, article structural attributes, author attributes, and reference attributes.

2. Literature Review

2.1. Journal Quality and Special Issue

The scientific impacts of articles or citation counts on articles are strongly associated with journals [20]. In the management discipline, the top seven journals belonging to the first and the second quartiles accounted for more than 60% and 80% of the citations, respectively [21]. Papers published in the top journals are considered good quality because of the journal’s high standards and rigorous review process to screen highly impactful articles. In other words, a journal’s reputation can cue article quality. Subsequently, articles published in reputable journals tend to attract the attention of fellow researchers and have great scientific impact.

Many academic journals publish special issues dedicated to a particular topic [22,23]. Articles focusing on a unique theme may further their influence and significance if they are published in a specialty journal [24]. However, specialty journals tend to have a lower journal impact in comparison with general journals, keeping researchers from choosing specialty journals [25]. Inviting papers for a special issue in reputable journals attracts high-quality manuscripts related to the particular topic because researchers do not need to compromise the journal ranking and gain the attention of fellow researchers with similar research interests [25,26].

2.2. Article Structural Attributes

Previous studies examined the effects of article structures, such as length of article, title, and keywords on citation counts [27–29]. Specifically, long articles may receive more citations than short ones, especially right after they are published in a peer-reviewed journal [27,30]. Long articles are perceived to contain more information than short ones [29]. Thus, the length of an article may be associated with quality [28]. The competition for journal spaces is quite fierce, especially among top-tier journals. Hence, to be accepted by journals, lengthy papers must have good quality that is proportional to their length [28].

The length of the title, which is often measured by the number of words in the title, is another frequently adopted article structure-related factor to predict citation counts. The title length and citation counts are either negatively [30] or not significantly related [31,32]. The association between the number of keywords and the citation counts indicated that papers with many keywords were likely to receive many citations [33].

2.3. Author Attributes

Author attributes are important determinants of predicting citations, and the number of authors and citation counts are positively related [34]. The number of authors may contribute to citations because the quality of articles can be enhanced through knowledge exchange among researchers with diverse expertise or through multiple proofreadings [29].
Furthermore, an article with many authors may receive additional attention from fellow researchers through personal and formal connections with coauthors [35,36].

Previous studies evaluated if authors’ gender influenced citation counts [37,38]. The mixed results indicated no significant gender effect [39–41], high citations with male-authored papers [37,42,43], or high citations with female-authored papers [44,45]. The author’s gender effect on citation counts was mostly dependent on the unique researcher demographics and their research interests. Hence, such an association should be tested in each discipline. The recent study in the tourism and hospitality field demonstrated a clear difference in the number of citations between articles written by male and female researchers, in which male-authored papers received more citations [32].

2.4. Reference Attributes

The number of references and citation counts are positively related [46,47]. Demonstrating the significance of the study and a sound literature review is a key to be accepted in hospitality and tourism journals [48]. Academic papers rely on existing knowledge shared in previous studies to develop new theories or hypotheses, including key articles that are critical to improving the quality of papers [49]. Therefore, the number of references may serve as a quality indicator for scholarly articles because the high number of references may represent extensive literature review, which subsequently attracts fellow researchers to cite the articles [50].

2.5. Maturity of Research Topics

While articles on high-demand or general topics are likely to attract more attention from fellow researchers than less popular or specific topics, topic attributes have been understated, especially in the tourism and hospitality literature. Most academic papers contain multiple research topics that have attracted researchers in various areas [9]. For example, an article testing the use of technology in the foodservice industry was mostly cited by fellow researchers with expertise in information technology or foodservice operations; thus, Antons, Joshi and Salge [46] tested the association between the concentration of research topics and citation counts. Articles with salient core topics were more visible than those with a widespread topic distribution. The study speculated that a concentrated key topic strongly connects with the existing research, such that fellow researchers can easily accept new research topics.

Li, et al. [51] proposed “the evolution of a new subject needs to build on the knowledge accumulation of relevant subjects” (p. 80). Likewise, the maturity of the topics is specific to research subtopics, resulting in different citation behaviors. Therefore, it is also necessary to consider the topic’s originality, which indicates the chronological order in which a particular research topic or method is introduced to the tourism and hospitality literature [52]. Research questions have often been developed based on accumulated knowledge in the field; hence, other researchers may pay attention to recent bibliographic coupling with the newly developed framework and sophisticated methodology [53]. While Antons, Joshi and Salge [46] found that a research paper containing the original ideas received higher citations in the management literature, the opposite was true in other disciplines [54]. That is, novel ideas that differ from existing concepts or methods are challenged and not well accepted by peers and are thus rarely cited [30]. It is also possible that fundamental theory and basic research, rather than a newly developed framework, may be valued in a particular discipline because of their scientific impact sustained over the years [14].

Many systematic review studies have been conducted on hospitality and tourism to understand the progress of research topics [9,17,55–58]. Previous studies have examined the evolution of citations within a particular research topic [58,59] or identified individual research papers that received high citations [56]. More specifically, some studies have focused on bibliometrics-based studies in the field of hospitality and tourism [59–61]. However, empirical evidence regarding the role of research topic attributes on citation counts, including topic originality and concentration, is still lacking [52].
Previous studies examining the effects of research subtopics on citation counts adopted manual content analysis to discover research topics, which required a long time for data analysis and can be challenged in reproducing data due to subjectivity [14,19]. To overcome these limitations, this study tested generally adopted citation factors and research topic attributes discovered by text mining using hospitality and tourism articles.

3. Methods

3.1. Sample

We collected data and article-related metadata from the Web of Science (WoS) database. Journal citation reports—a tool for journal assessment supported by WoS—were used to select hospitality and tourism journals. As for the data collection date, we retrieved the most recent information of 52 SSCI journals in the category of “hospitality, leisure, sport & tourism” in 2019. Among the top 25% of journals within this category (classified as Q1), we excluded two sports-related journals and selected 12 hospitality and tourism-focused journals for the sample.

In March 2021, all full records and cited references of regular journal articles from the 12 journals were downloaded after excluding other types of documents (e.g., editorial materials, book reviews, and biographical items). Data analysis employed a 10-year time window, sampling articles published between January 2011 and the end of 2020. Regardless of the factors that this study aims to explore, articles published in the last months were excluded because they may not get enough citations. Finally, the data analysis included 9,910 papers that consisted of 735,182 cited references and 156,624 pages.

3.2. Measures

3.2.1. Citation Counts

Citation counts were used as a proxy for scientific impact [46]. Articles published in earlier years may have received higher citation counts than more recent papers, which is not necessarily due to their higher scientific impact but to the extended periods in which to receive citations. Therefore, the average citation counts per year were used to handle bias at the time of publication.

We took several steps to calculate adjusted citation counts. First, we recorded the publication year and month for each article. The year of publication was calculated as of the data collection date, March 2021. We calculated adjusted citation counts by dividing the total citation counts by the year of publication.

\[
\text{Adjusted citation counts} = \frac{\text{Total citation counts}}{\text{The number of years since publication}}
\]

Due to many uncited articles, citation scores were highly skewed and did not conform to the standard normal distribution. The square root of the citation scores was used to normalize the data [20,46].

3.2.2. Journal/Article Structural/Author/Reference Attributes

The names of the 12 journals were included as a dummy variable to compare different scientific impacts among them. Another dummy variable was created with a special issue (i.e., coded 0 for regular issues and 1 for special issues). To measure the effects of the article structure, this study accounted for the length of the article, title, and keywords. We measured article length using the number of pages and calculated the number of words in the titles and the keywords. To understand author attributes, the number of authors and that of female authors were calculated. We adopted an automated approach to match the gender of the authors in order to save time for the manual identification of author names. A gender classifier based on the global name dataset was used to perform this step [62]. Finally, the total number of references for each paper was used to examine reference attributes.
3.2.3. Topic Attributes

This study adopted a machine learning technique for automated text mining and a structural topic model (STM) to discover the latent research subtopics (i.e., topic structure) from a vast amount of text data [63]. To build the topic model, the optimal number of topics was set to 40 based on the quantitative index (e.g., residuals and held-out likelihood). In addition, we performed a qualitative review to compare the research topics generated from the topic modeling algorithm with those derived from previous studies that investigated academic structure [56, 64]. Similar to the basic topic model, like the Latent Dirichlet Allocation algorithm, STM also generates two major outputs: (1) the list of top words with the highest probabilities of topics, \( \beta \) and (2) the probability of documents with each topic, \( \theta \) (See Figure 1). The word-topic distribution (\( \beta \)) reveals the most salient 40 research topics from the dataset, and the document-topic proportions (\( \theta \)) demonstrates how closely each document is related to 40 topics. STM was implemented with text data, which combined the title, keywords, and abstract of each article. Before implementing STM, text preprocessing was conducted, such as converting to lowercase, removing non-alphabetic characters and stop words, and lemmatizing. Customized stop words (e.g., study, goal, and limitation) were built for data cleaning, thus eliminating irrelevant words. Bigrams and trigrams were built with phrases that appeared more than 10 times in the corpus.

Then, we conducted topic network analysis and clustering analysis to discover the hierarchical structure of hospitality and tourism research topics [9]. After comparing multiple community detection algorithms, we used a fast-greedy algorithm to determine the membership of each topic. Topic proportions generated from STM indicate the association between journal articles and topics, which are used as a proxy for topic structure. To identify articles that are highly relevant to a particular topic, topic proportions with less than a cutoff value (<0.1) were replaced by zero [65].

Hospitality and tourism research tends to take a multidisciplinary approach [66]. By considering topic centration, this study attempts to identify whether an academic article focusing on a single topic or multiple topics can receive more citations. Topic concentration was calculated using a standard Herfindahl index (HHI) with topic proportion scores [46]. The sum of topic proportions was 1; hence, we multiplied each topic proportion score by 100 before calculating the HHI.

\[
\text{HHI} = p_1^2 + p_2^2 + p_3^2 + \ldots + p_n^2
\]

\( n \) = Number of topics, \( p = \theta \times 100 \)

HHI ranges from zero to 10,000. If HHI is close to 10,000, then the article has a strong focus on a single topic with an exceptionally high topic proportion with the topic. If HHI is close to zero, then the article has a diffused topic distribution, indicating that it focuses on multiple topics.

![Figure 1. Illustration of STM algorithm (Adjusted from Roberts, Stewart and Airoldi [63]).](image-url)
To compare the citation counts of articles on a particular topic at an early stage, we calculated its originality score. Topic originality refers to the relative order of topic discovery among articles on the same topic. For each topic, articles with topic proportions higher than the cutoff values were identified and arranged chronologically. For these articles, higher scores were assigned to articles published in the earlier year, while lower scores were assigned to the recent articles. The rest of the articles with less than 0.1% topic proportions were assigned 0. Because of the many zero data points, the topic originality score was highly skewed. As such, log transformations were performed using the original topic originality score [46].

3.3. Statistical Analyses

We employed STATA 14.2 for further statistical analyses, using descriptive statistics and pairwise correlations. All variables included in the regression analysis were standardized. The skewness of the variables was then corrected. Hence, ordinary least square regressions were applied to predict scientific impacts with structural, author, reference, and topic attributes [46].

4. Results

4.1. Descriptive Statistics

Table 1 presents the descriptive statistics for the major variables. The adjusted WoS citation count for the sampled articles was 4.41 times per year, and the average total citation count was 21.40. The number of papers published in Tourism Management was the highest, followed by International Journal of Hospitality Management and International Journal of Contemporary Hospitality Management.

Table 1. List of journals and number of journal articles in the sample.

| Rank | Journal Name                             | Year Available in WoS from | Average Annual Citation Counts | Average Total Citation Counts | # of Articles Included in the Sample |
|------|------------------------------------------|-----------------------------|-------------------------------|------------------------------|-------------------------------------|
| 1    | Tourism Management                       | 1994                        | 7.33                          | 39.82                        | 1776                                |
| 2    | Journal of Travel Research               | 1982                        | 6.45                          | 33.22                        | 654                                 |
| 3    | International Journal of Hospitality Management | 2008  | 4.93                          | 24.58                        | 1398                                |
| 4    | Annals of Tourism Research               | 2008                        | 5.26                          | 31.89                        | 789                                 |
| 5    | International Journal of Contemporary Hospitality Management | 2009  | 3.96                          | 17.67                        | 1112                                |
| 6    | Journal of Hospitality Marketing & Management | 2012  | 3.23                          | 15.14                        | 401                                 |
| 7    | Journal of Destination Marketing & Management | 2008  | 3.79                          | 15.75                        | 428                                 |
| 8    | Current Issues in Tourism                | 2008                        | 3.77                          | 11.11                        | 765                                 |
| 9    | Journal of Travel & Tourism Marketing    | 2015                        | 3.58                          | 17.97                        | 650                                 |
| 10   | Journal of Sustainable Tourism           | 2008                        | 4.41                          | 21.10                        | 820                                 |
| 11   | Journal of Hospitality & Tourism Research | 2008  | 3.36                          | 16.57                        | 352                                 |
| 12   | Tourism Management Perspectives          | 2015                        | 2.85                          | 11.94                        | 625                                 |

Average 4.41 21.40 814.17

4.2. Research Topics

Table 2 and Figure 1 illustrate the characteristics of the research topics discovered in this study. The dendrogram in Figure 2 depicts the hierarchical structure of the research topics and summarizes the relationships among them. The horizontal axis represents the dissimilarities between topics. The short height of the horizontal axis connecting two topics indicates that these topics have a high correlation and share high similarities, while the long axis indicates a low degree of correlation between them. Eight clusters were discovered with 40 topics, including the standalone cluster (Cluster 2; Cruise). Of the eight clusters, Cluster 4 (destination marketing) was the most popular as the subtopics accounted for 26.5% of the topic proportion. Cluster 3 (tourism planning and development), which
consists of seven topics, had the second-highest topic proportion. In particular, Topic 21 (economic growth) had the highest topic proportion, meaning that much research related to this topic has been conducted.

| Topic Label                        | Top Words                                                                 | Average Topic Proportion |
|-----------------------------------|---------------------------------------------------------------------------|---------------------------|
| T11 training                      | manager, employee, training, management, education                       | 0.023                     |
| T14 career                        | career, work, job satisfaction, efficacy, job                             | 0.018                     |
| T18 job conflict                  | conflict, employee, stress, work, emotional                                | 0.016                     |
| T33 organizational behavior       | employee, organizational, organization, work, leader                      | 0.022                     |
| T4 cruise                         | cruise, goal, passenger, role, traveler                                  | 0.009                     |
| T1 sustainable tourism            | sustainability, sustainable tourism, sustainable, policy, governance      | 0.031                     |
| T17 environment practice          | environmental, green, pro, management, eco                               | 0.022                     |
| T20 rural tourism                 | country, economic, growth, economy, international                         | 0.031                     |
| T21 economic growth               | ecotourism, trail, monitoring, over tourism, carrying capacity            | 0.026                     |
| T36 spatial tourism planning      | spatial, urban, gis, tourist attraction, length stay                      | 0.031                     |
| T37 climate change                | climate, carbon, emission, adaptation, weather                           | 0.017                     |
| T2 destination competitiveness    | destination, competitiveness, management, destination marketing          | 0.029                     |
| T5 social movement                | social, family, power, volunteer tourism, volunteer                      | 0.027                     |
| T6 culture                        | cultural, culture, Chinese, tourist, identity                            | 0.031                     |
| T10 brand                         | brand, image, destination image, destination, brand                       | 0.029                     |
| T12 social issues                 | gender, Taiwan, woman, female, male                                      | 0.011                     |
| T13 tourist perception            | tourist, satisfaction, motivation, tour, visit                           | 0.039                     |
| T25 authenticity & heritage       | experience, authenticity, heritage, museum, memory, criterion, cooperation| 0.040                     |
| T32 entrepreneurship              | entrepreneurial, process, relational                                     | 0.014                     |
| T34 network                       | network, innovation, academic, collaboration, structure                   | 0.044                     |
| T7 community                      | trust, resident, community, quality life, social exchange                 | 0.036                     |
| T19 festival/exhibition/convention | festival, crowd, place attachment, exhibition, theme park               |                           |
| T24 event                         | event, sport, mega, host, game                                          |                           |
| T3 online media                   | online, social media, website, information, user                          | 0.036                     |
| T27 technology                    | robot, app, mobile, technology, virtual                                  | 0.021                     |
| T29 mobility                      | mobility, vacation, destination choice, lifestyle, home                   | 0.021                     |
| T31 travel marketing              | travel, blog, travel agencies, marketing, travel                         | 0.027                     |
| T38 market segmentation           | segment, motivation, preference, cross, market                           | 0.023                     |
| T40 technology acceptance         | technology acceptance, acceptance, situational, service technology, kiosk| 0.005                     |
| T8 finance                        | financial, firm, restaurant, capital, investment                          | 0.025                     |
| T9 hotel management               | hotel, management, strategy, chain, guest                                | 0.041                     |
| T16 revenue management            | price, casino, room, revenue, accommodation                             | 0.020                     |
| T22 CSR                           | responsibility, csr, corporate social, corporate, ethical                 | 0.017                     |
| T30 Customer choice               | time, information, choice, temporal, cost                                | 0.021                     |
| T35 sharing economy               | host, Airbnb, peer, share, sharing economy                              | 0.017                     |
| T26 customer behavior             | customer, loyalty, customer satisfaction, satisfaction, failure           | 0.036                     |
| T28 foodservice operation         | wine, food, eat, food safety, food waste                                 | 0.029                     |
| T39 dining experience             | tip, server, restaurant, dining, services cap                           | 0.019                     |

Table 2. Summary of topic structure ($k = 40$).
Since the outbreak of the novel coronavirus disease (COVID-19), many researchers have shown interest in COVID-19-related issues, and many journals have launched special issues on this topic, leading to the publication of several COVID-19 studies. As COVID-19 was considered to be a major risk with detrimental effects on the industry, related papers share high similarities with previous studies examining the impact of various hazards (e.g., natural disasters or economic risk) on hospitality and tourism settings. As a result, COVID-19 research creates a topic in conjunction with previous risk studies, and the topic is labeled “risk” (see topic 15). The dendrogram shows that topic 15 (risk) is closely related to topics related to technology. This implies that much COVID-19-related research explored technology acceptance and the use of new technology during the pandemic.
4.3. Scientific Impact Prediction

We conducted multiple regression analyses to understand the association between variables and scientific impact (Table 3). The regression results of the article effects indicated that papers with longer pages \((b = 0.05, p < 0.001)\) had a higher scientific impact, which is consistent with previous studies [30,35].

Table 3. Regression analysis.

| Variables                        | Coef. | S.E. |
|----------------------------------|-------|------|
| **Journal attributes**           |       |      |
| Journals                         | Yes   | ***  |
| Special issue                    | 0.18  | ***  | 0.04|
| **Article attributes**           |       |      |
| Page counts                      | 0.05  | ***  | 0.02|
| Title counts                     | 0.00  |      | 0.01|
| Keyword counts                   | 0.01  |      | 0.01|
| **Author attributes**            |       |      |
| Author counts                    | 0.04  | ***  | 0.01|
| Female author counts             | -0.02 | *    | 0.01|
| **Reference attributes**         |       |      |
| Reference counts                 | 0.17  | ***  | 0.01|
| **7. Topic attributes**          |       |      |
| Topic concentration              | 0.05  | ***  | 0.01|
| Topic structure 1                | Yes   | ***  |     |
| Topic originality 2              | Yes   | ***  |     |
| **Cons**                         | -0.21 | **  | 0.08|
| **F**                            | 35.88 | ***  |
| Adjusted R²                      | 0.26  |      |

Note: *\(p < 0.05\), **\(p < 0.01\), ***\(p < 0.001\). 1 & 2 The specific results of topic structure and topic originality are illustrated in Table 4.

When testing the author effects, having more authors and fewer female authors were found to increase citation counts, similar to the findings of Nunkoo, Hall, Rughoobur-Seetah and Teeroovengadum [32], who found that female authors tend to receive fewer citations than male authors. Articles with a more comprehensive list of references tended to have a higher scientific impact \((b = 0.17, p < 0.001)\). Regarding the effects of topic attributes, articles with a strong focus on key research topics tended to receive more citations \((b = 0.05, p < 0.001)\). Several topics positively contributed to the scientific impact, and we found associations between their topic originality and scientific impact. The specific regression coefficients of topic structure and topic originality for the 39 topics are listed in Table 4.

4.4. Effects of Topic Structure and Topic Originality on Scientific Impact

Table 4 indicates the effects of topic structure and topic originality on scientific impact. The positively significant association between topic structure and citation counts implies that research papers focusing on a particular topic may attain many citations, but the negative association indicates that these research topics have not gained much attention from others.
Table 4. Effects of topic structure and originality on scientific impact.

| Cluster | Topic Label | Structure | Originality |
|---------|-------------|-----------|-------------|
|         |             | Coef.     | SE          | Coef. | SE    |
| Cluster 1 | T11 training | 0.00 | 0.02 | −0.03 | * | 0.01 |
|          | T14 career   | 0.03 | 0.02 | −0.02 | 0.02 |
|          | T18 job conflict | 0.05 | * | 0.00 | 0.02 |
|          | T33 organizational behavior | 0.05 | * | −0.01 | 0.01 |
| Cluster 2 | T4 cruise | 0.05 ** | 0.02 | 0.02 | 0.01 |
| Cluster 3 | T1 sustainable tourism | 0.01 | 0.02 | 0.00 | 0.01 |
|          | T17 environment practice | 0.05 * | 0.02 | 0.03 | 0.01 |
|          | T20 rural tourism | 0.02 | 0.02 | −0.02 | * | 0.01 |
|          | T21 economic growth | 0.07 ** | 0.02 | −0.02 | ** | 0.01 |
|          | T23 eco-tourism (over tourism) | 0.02 | 0.02 | −0.03 | * | 0.01 |
|          | T36 spatial tourism planning | 0.04 | * | −0.03 | * | 0.01 |
|          | T37 climate change | 0.06 ** | 0.02 | −0.03 | 0.02 |
| Cluster 4 | T2 destination competitiveness | 0.06 ** | 0.02 | −0.02 | 0.01 |
|          | T5 social movement | −0.02 | 0.02 | −0.02 | 0.01 |
|          | T6 culture | −0.06 ** | 0.02 | −0.01 | 0.01 |
|          | T10 brand | 0.02 | 0.02 | 0.01 | 0.01 |
|          | T12 social issues | 0.03 | 0.02 | −0.05 | * | 0.02 |
|          | T13 tourist perception | 0.09 *** | 0.02 | −0.02 | * | 0.01 |
|          | T25 authenticity & heritage | 0.05 ** | 0.02 | 0.00 | 0.01 |
|          | T32 entrepreneurship | 0.01 | 0.02 | −0.02 | 0.02 |
|          | T34 network | 0.03 | 0.02 | 0.00 | 0.01 |
| Cluster 5 | T7 community | 0.13 *** | 0.02 | −0.03 | 0.01 |
|          | T19 festival/exhibition/convention | 0.04 | 0.02 | −0.01 | 0.02 |
|          | T24 event | 0.01 | 0.02 | −0.03 | 0.02 |
| Cluster 6 | T3 online media | 0.14 *** | 0.02 | 0.00 | 0.01 |
|          | T15 risk | 0.06 ** | 0.02 | −0.04 | * | 0.02 |
|          | T27 technology | 0.06 ** | 0.02 | 0.03 | * | 0.01 |
|          | T29 mobility | 0.01 | 0.02 | −0.02 | 0.01 |
|          | T31 travel marketing | 0.05 | * | −0.01 | 0.01 |
|          | T38 market segmentation | 0.02 | 0.02 | −0.02 | 0.01 |
| Cluster 7 | T8 finance | −0.08 ** | 0.02 | −0.02 | 0.01 |
|          | T9 hotel management | −0.04 * | 0.02 | 0.02 | 0.01 |
|          | T16 revenue management | −0.03 | 0.02 | 0.00 | 0.01 |
|          | T22 CSR | 0.06 | 0.02 | 0.00 | 0.02 |
|          | T30 Customer choice | 0.01 | 0.02 | −0.03 | 0.01 |
|          | T35 sharing economy | 0.11 *** | 0.02 | 0.04 | * | 0.02 |
| Cluster 8 | T26 customer behavior | 0.02 | 0.02 | 0.02 | 0.01 |
|          | T28 foodservice operation | −0.03 | 0.02 | 0.01 | 0.01 |
|          | T39 dining experience | 0.02 | 0.02 | −0.01 | 0.01 |

Note: * p < 0.05, ** p < 0.01, *** p < 0.001.
In Cluster 1 (human resource management), research topics related to job conflict (T8; \( p < 0.01 \)) and organizational behavior (T33; \( p < 0.05 \)) were positively associated with citation counts. The single topic in Cluster 2 (cruise) was also positively associated with citation counts (T4; \( p < 0.01 \)). In Cluster 3 (tourism planning and development), environmental practice (T17; \( p < 0.05 \)), economic growth (T21; \( p < 0.01 \)), spatial tourism planning (T36; \( p < 0.05 \)), and climate change (T37; \( p < 0.01 \)) topics were positively related to citation counts. In Cluster 4 (destination marketing), destination competitiveness (T2; \( p < 0.01 \)), tourist perception (T13; \( p < 0.001 \)), and authenticity and heritage (T25; \( p < 0.01 \)) topics were positively associated with citation counts, while culture (T6; \( p < 0.01 \)) had a negative association. In Cluster 5 (event and convention), community topic (T7; \( p < 0.001 \)) was positively related to citation counts. In Cluster 6 (technology), online media (T3; \( p < 0.001 \)), risk (T15; \( p < 0.01 \)), technology (T35; \( p < 0.01 \)), and travel marketing (T35; \( p < 0.05 \)) topics were positively associated with citation counts. Finally, in Cluster 7 (hotel management), CSR (T22; \( p < 0.01 \)) and sharing economy (T35; \( p < 0.001 \)) were positively associated with citation counts, while finance (T8; \( p < 0.01 \)) and hotel management (T9; \( p < 0.05 \)) had negative associations.

The topic originality coefficients indicated whether early or recently published papers among those on a particular topic were more likely to be cited by others. If the originality coefficient is positively significant, the papers published in the early stage and contained the originality tended to be cited more. If the originality coefficient is negatively significant, recently published papers may regain popularity as these articles have been cited frequently by others. For instance, the training topic had an insignificant association with citation counts, implying that these research papers collectively have not been so popular; however, the significantly negative originality coefficient implies that recent training papers tend to gain more popularity.

In Cluster 3 (tourism planning and development), rural tourism (T20; \( p < 0.05 \)), economic growth (T21; \( p < 0.01 \)), eco-tourism (T23; \( p < 0.05 \)), and spatial tourism planning (T36; \( p < 0.01 \)) had significantly negative coefficients of topic originality, indicating that articles published in recent years tended to receive more citations. The following two topics in Cluster 4 (destination marketing) had significantly negative topic originality coefficients: social issues (T12; \( p < 0.05 \)) and tourist perception (T13; \( p < 0.05 \)). Social issues and tourist perceptions can evolve over time and are, therefore, time-sensitive topics. As a result, more recent papers may have been favored over old papers and cited more. In the case of Cluster 6 (technology), the risk topic (T15; \( p < 0.05 \)) had a significantly negative topic originality coefficient. Among research papers focusing on risk topics, COVID-19 studies account for the majority of recent studies. The originality test results, therefore, reveal that COVID-19 papers have received a lot of attention and have been frequently cited by other researchers.

The following topics had significantly positive topic originality coefficients: technology (T27; \( p < 0.05 \)) and the sharing economy (T35; \( p < 0.05 \)). The results indicate that more pioneering papers on these topics and published in an early stage were cited more often by other researchers compared to recent papers. According to Park, Chae and Kwon [9], information technology research has rapidly advanced in recent years with the development of artificial intelligence and the advent of various social media sites in the hospitality and tourism setting. In addition, sharing economy research is relatively nascent compared to historical research topics, such as human resource management or cultural studies. Thus, researchers interested in these topics seem to examine papers containing original ideas to define the concept and further develop it.

5. Conclusions

This study aims to discover the academic structure of the tourism and hospitality literature by identifying salient research topics and the interrelationships of these topics by analyzing research papers published in top-tier tourism and hospitality journals over the past decade with multiple automated algorithms. Our findings on the predominant research topics in tourism and hospitality research demonstrate the areas of research that
many researchers are interested in. By using the machine learning approach, this study was able to capture emerging and up-to-date research topics, such as COVID-19. In addition, it performed topic network analysis to discover how these research topics are correlated to progress in the research sub-categories.

In addition, different citation patterns were examined depending on the sub-categories of research topics to understand how ideas and knowledge have been exchanged in the academic network. While the academic structure can serve as a snapshot of research maturity in the tourism and hospitality literature, citation patterns can show the path through which research has evolved. For instance, papers related to human resource management have actively cited papers specific to job conflict and organizational behavior. Although training papers published over the past decade have not been very popular, recent training papers have been popularly cited by fellow researchers, implying that this topic has been growing in popularity in recent years. Similar patterns were found in the papers on rural tourism, eco-tourism (over-tourism), social issues, and risk topics since recently published papers have been cited frequently. As for the risk topic, the topic has become very popular due to the growing interest of fellow researchers in the recent major health-related risk, COVID-19. Among the papers concerning technology and sharing economy topics, pioneering papers that introduced the concept to the tourism and hospitality discipline were more actively cited. These findings imply that researchers’ demands and preferences in the literature may vary depending on the research topics—whether they look for the latest articles reflecting the current social and industry issues or prefer fundamental theory in the original concept papers.

This study also investigated the effects of the journal, article, author, reference, and topic attributes on citation counts. Although these attributes are less relevant to the quality of research papers, they are relevant to citation counts; specifically, papers with longer pages, more references, and more authors have received more citations. A gender effect was also significant, showing that papers with more female authors were less cited. This was consistent with the findings of Nunkoo, et al. [67], who proposed that gender can be a latent factor of authorship and collaboration.

According to our findings, only a few research topics (e.g., technology and the sharing economy) represented a benefit due to their originality. This may reflect the characteristics of the hospitality and tourism industries, which are sensitive to the external environment and constantly require innovative ideas [68]. However, this can be a warning that the inflow of new ideas and methodology can be challenged and even discouraged. According to Alvesson and Sandberg [69], the incremental pressure on publishing in top-tier journals with high impact factors has forced researchers to conduct “gap-spotting research,” which can be a double-edged sword. Although strong gap-spotting research can modify existing theories and fill the research gap, it may not bring fundamentally novel ideas or methodologies. Therefore, journal editors should monitor emerging and innovative topics to be accepted in the literature. The findings of this study can be used to alert academia stakeholders by simply reminding them that citation counts can be affected by mere formatting issues or the underlying power dynamics of the tourism and hospitality field. Based on the findings of this study, we propose implications for stakeholders.

5.1. Implications for Academic Scholars

The findings of topic attributes can represent predominant research topics in the tourism and hospitality literature, serving as a reference for young scholars and graduate students. Out of all tested factors, journal selection was most crucial to improving citation counts. Therefore, academic scholars need to choose the journal carefully. Journal reputation itself may enhance the scientific impact of articles. However, producing high-quality papers that meet top journals’ standards is crucial. Another piece of advice is to aim at a special issue to improve the visibility of the manuscript. In addition to journal or topic attributes with the most significant impact on citation counts, factors related to article presentation (i.e., page numbers, word counts in the title, and keywords) contributed to en-
hancing citation counts. Although these factors may not be directly relevant to the quality of academic papers, they tended to play an important role in drawing attention such that papers can be cited more by fellow scholars [33]. Our findings highlighted the importance of having comprehensive references. The significant association between the reference counts and scientific impact may come not necessarily because fellow researchers care about reference counts but because they can be the proxy of in-depth literature review [47]. In the abstract, emphasizing a research topic can be beneficial to enhance citation counts.

As the demand and popularity of research topics evolve over time [9], researchers should consider the association of citation counts with topic structures and originality. This study found that popularity and citation patterns may vary across specific hospitality and tourism research topics, demonstrating research topics with high demand. In addition, journal editors may consider having a special issue related to understudied or emerging topics to encourage the submission of articles corresponding to these topics depending on the vision of the journals [33].

5.2. Implications for University Administrators

Citation counts can be a useful measure by demonstrating how knowledge is diffused in the academic network and may address the shortcomings of simply counting the number of publications [70]. However, university administrators should be careful not to be blinded by the citation counts for the sustainable growth of the hospitality and tourism literature [30]. We found various factors that may be less relevant to research quality (e.g., author and article attribute) that can drive high citations, consistent with a previous study [71]. These findings imply that fewer citations may not be due to the poor quality of the paper but to external factors, such as the format of the paper or the low popularity of the research subject. Moreover, many researchers acknowledge that the quality of papers and journals cannot be solely evaluated by citation counts because some citations can be spurious, superficial, and incorrect [72]. Hence, the aforementioned factors that influence citation counts should be considered and adjusted to evaluate the quality of the paper and its scientific impact. As this study suggests, comparing citations of papers within a particular research subject or methods can be an alternative.

6. Limitations and Future Studies

This study utilized WoS citation counts because of the accuracy and inclusion of influential academic journals; however, these only capture the evaluation of professional fellow researchers in the academic field. Thus, the scientific impacts measured by adjusted citations may not capture opinions from more general groups of people, such as governments, experts, and the public [27]. In addition, as this study analyzed articles from 12 top-tier hospitality and tourism journals over a decade, the inclusion of only top journals makes generalizing the current findings in the hospitality and tourism literature difficult. Moreover, it utilized automated approaches to analyze large datasets (such as the topic modeling approach, the global name dataset) to identify the structures of the articles (i.e., topical structures, the gender of the authors). Although these automated methods enable researchers to save time and generate consistent results, some inaccuracies in the results are inevitable. This study attempted to indirectly demonstrate the changes in citation counts by including topic originality but did not investigate the changes in citation count patterns over time; therefore, future studies should include more journals over a broader period.

Author Contributions: Conceptualization, W.K. and E.P.; methodology, E.P.; software, E.P.; validation, W.K.; formal analysis, E.P.; investigation, W.K.; resources, W.K.; data curation, E.P.; writing—original draft preparation, W.K.; writing—review and editing, E.P.; visualization, E.P.; supervision, W.K.; project administration, W.K.; funding acquisition, W.K. Both authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Incheon National University Research Grant in 2019.
Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.
Data Availability Statement: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.

References
1. Benckendorff, P.; Zehrer, A. A network analysis of tourism research. Ann. Tour. Res. 2013, 43, 121–149. [CrossRef]
2. Echtner, C.M.; Jamal, T.B. The disciplinary dilemma of tourism studies. Ann. Tour. Res. 1997, 24, 868–883. [CrossRef]
3. Gursoy, D.; Sandstrom, J.K. An updated ranking of hospitality and tourism journals. J. Hosp. Tour. Res. 2016, 40, 3–18. [CrossRef]
4. Baggio, R. The science of complexity in the tourism domain: A perspective article. Tour. Rev. 2019, 75, 16–19. [CrossRef]
5. McKercher, B. The future of tourism journals: A perspective article. Tour. Rev. 2020, 75, 12–15. [CrossRef]
6. Belhassen, Y.; Caton, K. Advancing understandings: A linguistic approach to tourism epistemology. Ann. Tour. Res. 2009, 36, 335–352. [CrossRef]
7. Darbellay, F.; Stock, M. Tourism as complex interdisciplinary research object. Ann. Tour. Res. 2012, 39, 441–458. [CrossRef]
8. Cehan, A.; Eva, M.; Iatu, C. A mutliayer network approach to tourism collaboration. J. Hosp. Tour. Manag. 2021, 46, 316–326. [CrossRef]
9. Park, E.; Chae, B.; Kwon, J. Toward understanding the topical structure of hospitality literature: Applying machine learning and traditional statistics. Int. J. Contemp. Hosp. Manag. 2018, 30, 3386–3411. [CrossRef]
10. Tribe, J.; Xiao, H.; Chambers, D. The reflexive journal: Inside the black box. Ann. Tour. Res. 2012, 39, 7–35. [CrossRef]
11. Ali, F.; Park, E.; Kwon, J.; Chae, B. 30 years of contemporary hospitality management Uncovering the bibliometrics and topical trends. Int. J. Contemp. Hosp. Manag. 2019, 31, 2641–2665. [CrossRef]
12. Racherla, P.; Hu, C. A social network perspective of tourism research collaborations. Ann. Tour. Res. 2010, 37, 1012–1034. [CrossRef]
13. Xiao, H.; Li, M.; Lin, E.C.K. Diffusion patterns and knowledge networks: An inductive analysis of intellectual connections in multidisciplinary tourism studies. J. Travel Tour. Mark. 2011, 28, 405–422. [CrossRef]
14. Law, R.; Lee, H.A.; Au, N. Which journal articles are uncited? The case of the Asia Pacific journal of tourism research and the journal of travel and tourism marketing. Asia Pac. J. Tour. Res. 2013, 18, 661–684. [CrossRef]
15. Baldi, S. Normative versus social constructivist processes in the allocation of citations: A network-analytic model. Am. Sociol. Rev. 1998, 63, 829–846. [CrossRef]
16. Small, H.; Boyack, K.W.; Klavans, R. Citations and certainty: A new interpretation of citation counts. Scientometrics 2019, 118, 1079–1092. [CrossRef]
17. Köseoglu, M.A.; Okumus, F.; Rahimi, R. Proposing researcher brand equity index in hospitality and tourism. Tour. Rev. 2019, 74, 990–1002. [CrossRef]
18. Gingras, Y.; Kshetraphati, M. Assessing the effect of the United States’ “citation advantage” on other countries’ scientific impact as measured in the Web of Science (WoS) database. Scientometrics 2018, 114, 517–532. [CrossRef]
19. Kim, C.S.; Bai, B.H.; Kim, P.B.; Chon, K. Review of reviews: A systematic analysis of review papers in the hospitality and tourism literature. Int. J. Hosp. Manag. 2018, 70, 49–58. [CrossRef]
20. Judge, T.A.; Cable, D.M.; Colbert, A.E.; Rynes, S.L. What causes a management article to be cited—Article, author, or journal? Acad. Manag. J. 2007, 50, 491–506. [CrossRef]
21. Podsakoff, P.M.; MacKenzie, S.B.; Bachrach, D.G.; Podsakoff, N.P. The influence of management journals in the 1980s and 1990s. Strateg. Manag. J. 2005, 26, 473–488. [CrossRef]
22. Carnegie, G.D. The special issue: AAAJ and research innovation. Account. Audit. Account. J. 2012, 25, 216–227. [CrossRef]
23. Olk, P.; Griffith, T.L. Creating and disseminating knowledge among organizational scholars: The role of special issues. Organ. Sci. 2004, 15, 120–129. [CrossRef]
24. Adler, N.J.; Harzing, A.-W. When knowledge wins: Transcending the sense and nonsense of academic rankings. Acad. Manag. Learn. Educ. 2009, 8, 72–95. [CrossRef]
25. Conlon, D.E.; Morgeson, F.P.; McNamara, G.; Wiseman, R.M.; Skilton, P.F. From the editors: Examining the impact and role of special issue and regular journal articles in the field of management. Acad. Manag. J. 2006, 49, 857–872. [CrossRef]
26. Perenyi, A.; Losonczy, M. A systematic review of international entrepreneurship special issue articles. Sustainability 2018, 10, 3476. [CrossRef]
27. Bornmann, L. Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. J. Informetr. 2014, 8, 895–903. [CrossRef]
28. Leimu, R.; Koricheva, J. What determines the citation frequency of ecological papers? Trends Ecol. Evol. 2005, 20, 28–32. [CrossRef]
29. Tahamtan, I.; Afshar, A.S.; Ahammadzadkh, K. Factors affecting number of citations: A comprehensive review of the literature. Scientometrics 2016, 107, 1195–1225. [CrossRef]
30. Stremersch, S.; Camacho, N.; Vanneste, S.; Verniers, I. Unraveling scientific impact: Citation types in marketing journals. Int. J. Res. Mark. 2015, 32, 64–77. [CrossRef]
63. Roberts, M.E.; Stewart, B.M.; Airoldi, E.M. A model of text for experimentation in the social sciences. *J. Am. Stat. Assoc.* 2016, 111, 988–1003. [CrossRef]  
64. Park, J.; Wu, B.; Morrison, A.M.; Shen, Y.; Cong, L.; Li, M. The tourism system research categorization framework. *Asia Pac. J. Tour. Res.* 2016, 21, 968–1000. [CrossRef]  
65. Antons, D.; Kleer, R.; Salge, T.O. Mapping the Topic landscape of JPIIM, 1984–2013: In search of hidden structures and development trajectories. *J. Prod. Innov. Manag.* 2016, 33, 726–749. [CrossRef]  
66. Garcia-Lillo, F.; Claver-Cortes, E.; Ubeda-Garcia, M.; Marco-Lajara, B.; Zaragoza-Sáez, P.C. Mapping the “intellectual structure” of research on human resources in the “tourism and hospitality management scientific domain”. *Int. J. Contemp. Hosp. Manag.* 2018, 32, 1041–1066. [CrossRef]  
67. Nunkoo, R.; Thelwall, M.; Ladsawut, J.; Goolaup, S. Three decades of tourism scholarship: Gender, collaboration and research methods. *Tour. Manag.* 2020, 78, 104056. [CrossRef]  
68. Ooi, C.-A.; Hooy, C.-W.; Som, A.P.M. Diversity in human and social capital: Empirical evidence from Asian tourism firms in corporate board composition. *Tour. Manag.* 2015, 48, 139–153. [CrossRef]  
69. Alvesson, M.; Sandberg, J. Has management studies lost its way? Ideas for more imaginative and innovative research. *J. Manag. Stud.* 2013, 50, 128–152. [CrossRef]  
70. Jang, S.S.; Park, K. Hospitality finance research during recent two decades: Subjects, methodologies, and citations. *Int. J. Contemp. Hosp. Manag.* 2011, 23, 479–497. [CrossRef]  
71. Jiang, J.; Woosnam, K.M.; Draper, J. Structural elements of articles and diachronous citation analysis among top-ranking tourism journals (1990–2010). *Tour. Manag. Perspect.* 2014, 12, 48–56. [CrossRef]  
72. Zarezadeh, Z.; Benckendorff, P.; Gretzel, U. Lack of progress in tourist information search research: A critique of citation behaviour and knowledge development. *Curr. Issues Tour.* 2019, 22, 2415–2429. [CrossRef]