

Exploring tourists' discourse on sustainability and its impact on satisfaction: Insights from online reviews of rural and urban tourist attractions in Spain

Maria Fernanda Bernal Salazar ^{1*}, Elisa Baraibar-Diez ² and Jesús Collado Agudo ³

¹ Santander Financial Institute (SANFI), University of Cantabria, 39005 Santander, Cantabria, Spain. E-mail: maria-fernanda.bernal@alumnos.unican.es

² Business Administration Department, University of Cantabria, Avda. Los Castros s/n, 39005 Santander, Cantabria, Spain; E-mail: elisa.baraibar@unican.es

³ Santander Financial Institute (SANFI), University of Cantabria, 39005 Santander, Cantabria, Spain. E-mail: jesus.collado@unican.es

*Corresponding author

Abstract

This research explores how tourists' discourse on sustainability (economic, environmental, social, and cultural) is associated with their satisfaction with, and evaluation of, tourist attractions when writing online reviews. It examines the relationship between the importance tourists attach to sustainability and their level of satisfaction, as reflected in the ratings given to tourist attractions. Using a mixed-methods approach that combines data science techniques with statistical testing, a total of 207706 reviews of major tourist attractions in Spain are processed. The Spanish BERT-based deep learning language model BETO is employed to classify the reviews, and the hypotheses are tested through ordered logistic regression models. Additionally, the study compares how tourists' sustainability discourse is associated with their level of satisfaction with attractions located in rural and urban areas. The results show that the cultural and environmental dimensions are associated with greater satisfaction, whereas the social and economic dimensions tend to receive lower ratings. Trips with friends or solo travel receive better evaluations than those taken as a couple, and domestic tourists are more critical than international ones. Environmental sustainability has a greater impact in rural areas, while the cultural dimension stands out in urban contexts.

Keywords: sustainability, tourism, tourist satisfaction, online reviews, data science

Citation: Bernal Salazar, M. F., Baraibar-Diez, E., & Collado Agudo, J. (2026). Exploring tourists' discourse on sustainability and its impact on satisfaction: Insights from online reviews of rural and urban tourist attractions in Spain. *European Journal of Tourism Research*, 42, 4214. <https://doi.org/10.54055/ejtr.v42i.4012>.

Publication history:

Received: 19/11/2024; Revised: 07/04/2025; 05/05/2025; 29/06/2025; 19/07/2025; 06/08/2025; Accepted: 13/08/2025;

Published online: 01/02/2026; Volume date: 01/03/2026

Coordinating editor: Faruk Seyitoglu



© 2026 The Author(s)

This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

1. Introduction

Online reviews are a key tool for both marketing and tourism research (D'Acunto *et al.*, 2020). They represent one of the most relevant forms of user-generated content (UGC), as they directly influence consumer decisions and the economic performance of destinations and tourism businesses (Bigné *et al.*, 2019; Jalilvand & Samiei, 2012; Liu & Park, 2015; Xiang & Gretzel, 2010). Furthermore, online ratings reflect levels of satisfaction with tourism services and products (Bulchand-Gidumal *et al.*, 2013; Schuckert *et al.*, 2015), indicating the gap between tourists' prior expectations and their actual experience (Jiang *et al.*, 2018; Pizam *et al.*, 1978).

Given their importance, online reviews have become a valuable alternative for analysing consumer awareness of sustainable practices in the tourism sector (D'Acunto *et al.*, 2020). The discourse on sustainability in tourist reviews is often implicit and linked to specific experiences. Tourists tend to value aspects such as cleanliness, the surrounding environment, the treatment received, and respect for local culture, but they only mention sustainability when it directly affects their satisfaction (Väisänen *et al.*, 2023; Wu & Yang, 2023). Recently, some studies have begun applying data science and natural language processing (NLP) techniques to this type of content to extract useful insights into specific topics (Cai *et al.*, 2024; D'Acunto *et al.*, 2023, 2024); however, the potential of these techniques has not yet been fully leveraged in tourism research (Egger, 2022; Mariani *et al.*, 2018; Mariani & Baggio, 2022), particularly for analysing tourists' sustainability discourse and its relationship with satisfaction (Brazyté *et al.*, 2016; Gerdt *et al.*, 2019; Wu & Yang, 2023; Yu *et al.*, 2017).

This research analyses how tourists' discourse on sustainability is associated with their satisfaction and their evaluation of Spain's main tourist attractions when writing online reviews. The study explores the relationship between the importance tourists assign to sustainability and their level of satisfaction, as reflected in the rating given to tourist attractions, using data science techniques. Sustainability discourse is understood to encompass aspects of economic, social, cultural, and environmental dimensions, as these contribute to promoting the development of sustainable tourism (Agyeiwaah *et al.*, 2017). This includes any information that a tourist expresses – either explicitly or implicitly – in their review. In line with previous studies on satisfaction and online review ratings (Bulchand-Gidumal *et al.*, 2013; Schuckert *et al.*, 2015), the rating assigned by a tourist to an attraction is considered an indicator of satisfaction. The analyses are conducted both at national level (Spain) and by area type (rural or urban), as significant differences may exist between rural and urban tourism areas regarding sustainability dimensions (Rasoolimanesh *et al.*, 2017).

The contribution of this study is significant in several respects. First, it expands the body of research applying data science techniques to online reviews in the analysis of sustainability, following the emerging research trend in this field (Cai *et al.*, 2024). Specifically, it enriches the existing literature by exploring the relationship between sustainability and tourist satisfaction using this type of data, an approach rarely adopted in previous studies (Bigne *et al.*, 2024; Brazyté *et al.*, 2016; Gerdt *et al.*, 2019; Marchi *et al.*, 2023; Yu *et al.*, 2017). Second, to the best of our knowledge, this is the first study to compare sustainability discourse in tourists' online reviews between rural and urban areas using data science techniques applied to UGC. Third, this study is pioneering in examining the sustainability discourse in tourists' reviews of destinations in Spain, as no comparable research combining this approach has yet been identified in this geographical context. This added value provides stakeholders with useful insights into which sustainability-related aspects are mentioned in tourists' reviews and how these relate to attraction ratings, thus supporting informed decision-making and continuous improvement.

The article is structured as follows. It begins with a literature review, outlining the theoretical foundations on sustainability discourse, tourist satisfaction, and prior research using data science techniques. The methodology section then details the data science methods and statistical tests used.

Next, the results are presented, including a comparison between rural and urban contexts. Finally, the findings are discussed and the conclusions are presented, along with the study's theoretical and practical implications and its limitations.

2. Literature review

2.1. *Social identity theories and expectation disconfirmation*

To understand how tourists' discourse on sustainability is associated with their satisfaction, this study draws on two theoretical frameworks: social identity theory and expectation disconfirmation theory. According to social identity theory (Tajfel & Turner, 1979), individuals construct part of their self-concept based on their membership in social groups with which they share values and to which they attribute emotional significance (Tajfel, 1978). In the tourism context, this identity may be expressed through online reviews, where visitors not only describe their experiences but also project their beliefs and symbolic affiliations. Thus, when tourists emphasize aspects related to sustainability or social responsibility, they convey their interest in these values (Bigne *et al.*, 2024). This behaviour is also driven by various motivations, such as an altruistic concern for other travellers, the desire to engage in a virtual community, the pursuit of recognition as experts, or the intention to influence the destination or service provider by expressing satisfaction or dissatisfaction (Chen *et al.*, 2011; Tian, 2013).

Expectation disconfirmation theory (Oliver, 1980) posits that customer satisfaction and post-consumption behaviours depend on the discrepancy between prior expectations and the actual experience with a service or product. When perceived performance exceeds customer expectations, a positive disconfirmation occurs, leading to satisfaction; conversely, if performance falls short of expectations, negative disconfirmation arises, resulting in dissatisfaction. This theory also provides a lens through which to interpret how tourists express their evaluation of a service in the discourse in their reviews. According to Jha and Shah (2021), disconfirmation significantly influences the way an online review is written and the credibility it conveys to future readers. Moreover, satisfied customers often develop a favourable attitude toward a product or service, which increases the likelihood of repeat purchases and the generation of positive reviews (Li *et al.*, 2020; Namkung & Jang, 2007). In this regard, the analysis of sustainability-related discourse in tourists' online reviews enables the exploration not only of their satisfaction levels but also of the social identities constructed around sustainable values.

2.2. *Sustainability dimensions in tourists' discourse*

Sustainability is structured around four core dimensions – economic, environmental, social, and cultural – which together promote the development of sustainable tourism (Agyeiwaah *et al.*, 2017). Although some studies have focused on these dimensions, many address only one of them (D'Acunto *et al.*, 2023; Mishra & Gupta, 2019; Sánchez-Fernández *et al.*, 2016), highlighting the importance of analysing them in an integrated manner from the perspective of the tourist experience. In this context, the economic dimension is linked to financial well-being and the sustained growth of tourism businesses and destinations, as reflected in operational profitability, cost management, and customer satisfaction (Väisänen *et al.*, 2023). Tourists assess the value for money, which directly influences their satisfaction levels and, in the long term, the economic viability of the destination (Ye *et al.*, 2014). In online reviews, this dimension is often expressed through evaluations of “value for money” (Väisänen *et al.*, 2023). Moreover, tourism spending drives local development and contributes to improving residents' quality of life (Jasrotia *et al.*, 2019; Suhartanto, 2018). The environmental dimension, in turn, focuses on protecting the natural environment for the benefit of both people and other species, through sustainable practices related to resource consumption, cleanliness, transportation, and the use of eco-friendly materials (Jones *et al.*, 2016; Väisänen *et al.*, 2023).

Social sustainability refers to human and labour rights, well-being, inclusion, accessibility, and safety (Swarbrooke, 1999; UNEP, 2005). In service contexts, this dimension can be divided into physical factors – such as safe infrastructure, accessibility, local purchasing, and healthy food options (Dewhurst &

Thomas, 2003; Garay & Font, 2012; Jones *et al.*, 2014) – and social aspects, including local employment, sustainability education, equal treatment, and the management of tourist overcrowding (Väisänen *et al.*, 2023). Finally, cultural sustainability focuses on the preservation of heritage and local traditions. It is also expressed through physical factors, such as traditional architecture or gastronomy, and social elements, such as the active participation of the local community in the tourism experience (Madanaguli *et al.*, 2023; Väisänen *et al.*, 2023).

2.3. Online reviews and data science: Analysing sustainability and tourist satisfaction

The analysis of UGC provides valuable insights for researchers in understanding tourists' preferences, behaviours, and satisfaction (León *et al.*, 2025). Online reviews stand out because their rating serves as an indicator of overall satisfaction, and their content offers detailed information that may influence future consumer demand (Xie *et al.*, 2014; Zhou *et al.*, 2014). Gerdt *et al.* (2019) point out that, unlike traditional surveys, online reviews represent a more original and less biased data source. Surveys, which are widely used in studies on customer satisfaction, tend to overestimate the relevance of sustainability in tourism due to social desirability bias, as participants aim to project a favourable self-image (Roxas & Lindsay, 2012).

The application of data science techniques has strengthened tourism research by generating new knowledge, enabling the analysis of large volumes of data, and overcoming the sample limitations of survey-based studies (Li *et al.*, 2017; Xiang *et al.*, 2015; Yang *et al.*, 2015). Recent studies have begun to apply advanced methods and NLP to analyse UGC (Chai *et al.*, 2021; Rey-Moreno *et al.*, 2023; Viñán-Ludeña & de Campos, 2022), such as online reviews, in order to extract relevant information (Cai *et al.*, 2024; D'Acunto *et al.*, 2023, 2024). However, studies specifically addressing the relationship between sustainability and tourist satisfaction using this type of data remain limited (Bigne *et al.*, 2024; Brazyté *et al.*, 2016; Gerdt *et al.*, 2019; Marchi *et al.*, 2023; Yu *et al.*, 2017).

On the other hand, Conti *et al.* (2023) highlight the importance of tourists' perceptions of sustainability in the design of sustainable strategies, particularly from a multidimensional perspective. However, there are still a limited number of studies that have employed data science or big data techniques to analyse the various dimensions of sustainability or corporate social responsibility as reflected in online reviews (Bigne *et al.*, 2024; D'Acunto *et al.*, 2020; Marchi *et al.*, 2023). Previous studies demonstrate that the different dimensions of sustainability are associated in distinct ways with tourist satisfaction. The environmental and cultural dimensions are generally associated with high levels of satisfaction, as tourists value elements such as the natural environment, the authenticity of a destination, historical heritage, and cultural interaction (Chen & Rahman, 2018; Li *et al.*, 2020; Väisänen *et al.*, 2023; Wu & Yang, 2023). These experiences foster emotional connections with a destination, increase the intention to return, and lead to favourable reviews (Kim *et al.*, 2017; Mariani & Borghi, 2023; Nguyen Huu *et al.*, 2024). Furthermore, eco-friendly practices such as recycling and waste reduction have also been shown to positively affect satisfaction (Merli *et al.*, 2019; Prud'homme & Raymond, 2013).

In contrast, the social and economic dimensions tend to be associated with negative experiences that lead to dissatisfaction, adverse emotions, and low ratings (D'Acunto *et al.*, 2020; Rita *et al.*, 2022). Factors such as staff rudeness, long waiting times, a lack of accessibility, overcrowding, high prices, and poor service negatively affect perceived quality and the value-for-money ratio (Kim *et al.*, 2020; Meneses *et al.*, 2023; Prakash *et al.*, 2019). As a result, negative comments often focus on these dimensions, reflecting critical aspects that deteriorate the visitor experience.

2.4. Sustainability differences in rural and urban tourist areas

Tourism activities in rural and urban areas offer different experiences derived from their resources and contexts. While rural tourism is based on contact with nature, local culture, and traditional lifestyles

(World Tourism Organization, 2024), urban tourism is characterized by its cultural, social, and technological diversity (World Tourism Organization, 2019). Various studies have shown that the dimensions of sustainability are manifested differently depending on the type of destination. Rasoolimanesh *et al.* (2017) and Väisänen *et al.* (2023) emphasize that economic, social, cultural, and environmental impacts vary between rural and urban contexts, requiring a more in-depth comparative analysis. In rural destinations, for example, environmental and cultural sustainability take on greater importance, as tourists value the connection with nature, heritage, and personal well-being (Wang *et al.*, 2022; Wu & Yang, 2023; Zheng *et al.*, 2022). Moreover, this type of tourism contributes to local development, helps counteract depopulation, and improves quality of life – provided that its value is effectively communicated (Cucari *et al.*, 2019; Guizzardi *et al.*, 2022). It also offers a restorative alternative to urban stress (Kastenholz *et al.*, 2012; Kumar & Shekhar, 2020), and rural tourists, particularly ecotourists, display a willingness to pay more for sustainable services that align with their values (Kazeminia *et al.*, 2016).

In contrast, cultural and economic dimensions tend to predominate in urban destinations. Sustainability in this context is associated with the revitalization of urban spaces, the improvement of infrastructure, and the management of services that enhance satisfaction (Lee *et al.*, 2014; Qiu *et al.*, 2024). Elements such as cultural centres, urban lighting, signage, public art, and performances are also valued as part of a sustainable and satisfying experience. Additionally, Kusumah (2023) points out that environmental, cultural, and socio-economic imagery positively influences tourist satisfaction and loyalty in urban destinations.

3. Hypotheses and research question

Based on previous literature, four research hypotheses are proposed to examine how the thematic presence of sustainability-related content (economic, environmental, social, and cultural) in tourists' online reviews is associated with their evaluation of tourist attractions, as reflected in the satisfaction ratings they provide.

H1: *Reviews with a higher thematic presence of environmental sustainability aspects are associated with higher satisfaction ratings.*

H2: *Reviews with a higher thematic presence of cultural sustainability aspects are associated with higher satisfaction ratings.*

H3: *Reviews with a higher thematic presence of economic sustainability aspects are associated with lower satisfaction ratings.*

H4: *Reviews with a higher thematic presence of social sustainability aspects are associated with lower satisfaction ratings.*

Additionally, with the aim of exploring potential differences in sustainability discourse according to the territorial context in which tourist attractions are located (rural or urban), the following research question is proposed:

RQ1: *Are there differences in the contribution of sustainability dimensions to tourist satisfaction according to the territorial context (rural vs urban municipality) of the attraction?*

It is important to clarify that, in this study, the expression “higher thematic presence of sustainability discourse” refers to the frequency with which content related to a specific dimension of sustainability (economic, environmental, cultural, or social) appears in a review, as identified by the BETO language model. This measure does not account for the polarity of the content (i.e. whether the mention is

positive or negative). Satisfaction is measured solely through the review rating and is not interpreted as an evaluative perception of sustainability.

4. Research methodology

This study employs content analysis as its primary technique – a well-established method in the social sciences for identifying patterns and extracting meaning from communicative outputs, using both qualitative and quantitative approaches (Piñuel, 2002). Specifically, it analyses the text of tourists' online reviews, which is one of the most common formats in this type of research (Stemler, 2000). Following recent approaches (Anagnostopoulou *et al.*, 2020; Garay, 2019; Hansen, 2020; Huertas *et al.*, 2021; Yu & Sun, 2019; W. Zhang & Fukami, 2024), a computerized quantitative content analysis is adopted, which allows for the transformation of unstructured textual data into measurable variables for subsequent statistical analysis. Figure 1 exhibits the research methodology of the study, consisting of three phases: 1) data collection; 2) classification; and 3) modelling. The programming language used for coding is Python and the statistical program is Stata.

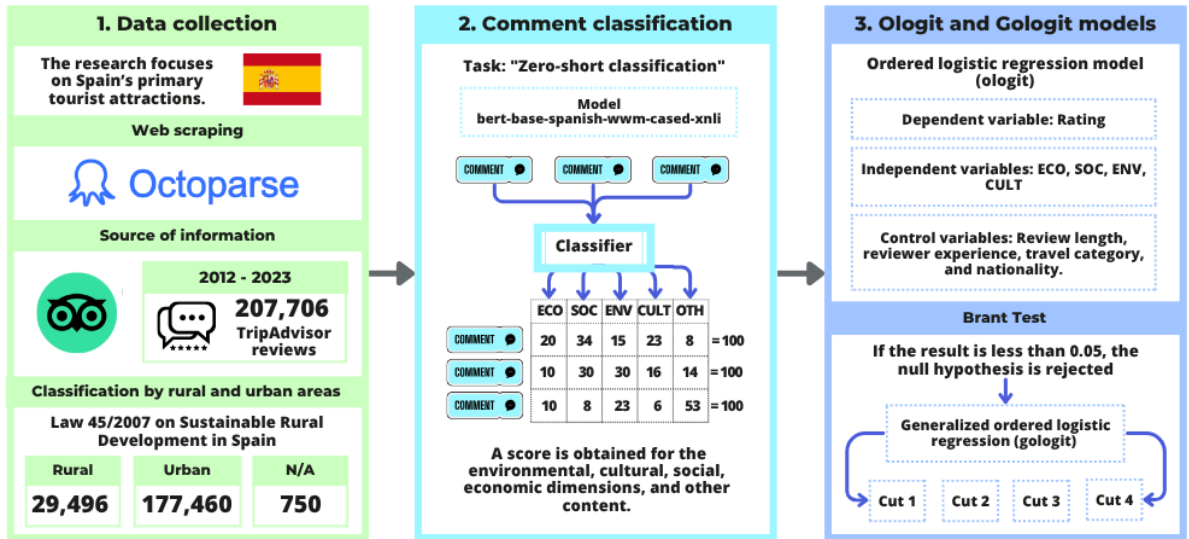


Figure 1. Research methodology.

4.1. Data collection

The source of information is the TripAdvisor website, as it continues to be one of the leading UGC platforms due to its perceived ease of use, credibility, and usefulness (Filieri *et al.*, 2021). To ensure a balanced territorial representation, the selection criterion was to include the 15 top-ranked tourist attractions in each of Spain's 17 autonomous communities. This number was chosen after a preliminary review of the available content, which confirmed that all communities had at least that number of attractions with a sufficient volume of reviews for analysis. The selection was based on TripAdvisor's popularity index, which ranks attractions according to the quantity, quality, and frequency of reviews. Moreover, previous studies have shown that the popularity and rating of a tourist attraction influence tourists when deciding whether or not to visit it (Hernández *et al.*, 2021; Kemperman *et al.*, 2019).

Data were collected between December 2023 and February 2024, and up to the first 2000 comments were retrieved (when available) using the web scraping tool Octoparse. A total of 313683 comments were obtained from 255 attractions, covering the period from 2007 to 2024. For each review, data were collected on location, date, travel category, nationality of the visitor, rating (from one to five), and number of contributions on TripAdvisor. Some fields, such as the nationality and travel category, were not always completed by users. The category "business" was removed because its motivations differ

from those of leisure tourism. In addition, the dates of the comments were filtered by including only the years 2012 to 2023, resulting in a final sample of 207706 comments.

Since one of the objectives of the study is to analyse potential differences in the sustainability discourse in tourist reviews according to the territorial context of the attraction visited, tourist attractions were classified as rural or urban based on their administrative-geographic location. This typology was based on the characteristics of the municipality in which each attraction is located, following the criteria established by Spain's Law 45/2007 on Sustainable Rural Development. Municipalities were considered rural if they had fewer than 30000 inhabitants and a population density below 100 inhabitants per km².

This approach enables the exploration of sustainability discourse in relation to the territorial setting of the attraction, rather than based on individual tourist characteristics such as origin or motivations. It is important to note that this classification is not aimed at capturing the experiential nature of rural or urban tourism, but rather follows an objective, reproducible, and comparative framework across regions. As a result, 44 municipalities were classified as rural, associated with 29496 reviews, and 78 municipalities were categorized as urban, with a total of 177460 reviews. In addition, three municipalities, corresponding to 750 reviews, could not be classified due to a lack of specific location information. Table 1 presents the number of selected attractions, the distribution of rural and urban municipalities, and the volume of reviews collected for each autonomous community.

Table 1. *Distribution of municipalities, area type, and reviews by autonomous community*

	Autonomous Community	No. of Attractions	No. of Municipalities / Area Type						Reviews				
			Total	Rural		Urban		Un-classified		Total	Rural	Urban	Un-classified
0	Andalusia	15	7	-	0%	7	100%	-	0%	15827	-	15827	87
1	Aragon	15	8	6	75%	2	25%	-	0%	13213	5986	7227	
2	Principality of Asturias	15	10	6	60%	3	30%	1	10%	10804	5407	5310	
3	Balearic Islands	15	12	1	8%	11	92%	-	0%	11046	497	10549	
4	Basque Country	15	7	2	29%	5	71%	-	0%	12534	652	11882	-
5	Canary Islands	15	14	6	43%	8	57%	-	0%	11053	3280	7773	
6	Cantabria	15	9	2	22%	7	78%	-	0%	11344	1711	9633	
7	Castile and León	15	8	1	13%	7	88%	-	0%	16073	192	15881	
8	Castile-La Mancha	15	4	2	50%	2	50%	-	0%	11696	727	10969	-
9	Catalonia	15	2	-	0%	2	100%	-	0%	16322		16322	
10	Community of Madrid	15	1	-	0%	1	100%	-	0%	17453	-	17453	
11	Extremadura	15	8	5	63%	3	38%	-	0%	10346	3238	7108	
12	Galicia	15	7	1	14%	5	71%	1	14%	14891	1352	13431	108
13	La Rioja	15	9	4	44%	5	56%	-	0%	5668	1625	4043	555
14	Chartered Community of Navarre	15	9	6	67%	2	22%	1	11%	7251	3416	3280	
15	Region of Murcia	15	6	1	17%	5	83%	-	0%	5856	134	5722	
16	Valencian Community	15	4	1	25%	3	75%	-	0%	16329	1279	15050	
Total		255	125	44		78		3		207706	29496	177460	750

4.2. Classification of comments using artificial intelligence

4.2.1. Classification model and approach

To classify sustainability-related content in tourists' reviews, an emerging machine learning approach known as "zero-shot learning" is applied. This method enables an artificial intelligence model to categorize classes or concepts without having been explicitly trained in them (Alhoshan *et al.*, 2023). This study employs the "BETO cased" model (Cañete *et al.*, 2020), a version of BERT adapted to the Spanish language, which was selected for four main reasons: first, it has been successfully applied in previous research (Chai *et al.*, 2021; Rey-Moreno *et al.*, 2023; Viñán-Ludeña & de Campos, 2022; Zhang *et al.*, 2023); second, its bidirectional architecture facilitates a comprehensive understanding of sentence

context, which is particularly useful for tasks such as text classification, automatic summarization, and language generation; third, it was trained on a large Spanish-language corpus using the Whole Word Masking technique, thereby enhancing the model's contextual accuracy (Cañete *et al.*, 2020); and fourth, it has performed better than other comparable models, such as BETO uncased and mBERT (Cañete *et al.*, 2020), in various NLP tasks (Wu & Dredze, 2019).

Specifically, the fine-tuned version “bert-base-spanish-wwm-cased-xnli”, developed by Recognai (2021) and available on the open-source Hugging Face platform, was used. This model has been fine-tuned with data from the XNLI (Cross-lingual Natural Language Inference) data set, designed to train systems capable of inferring semantic relationships between sentence pairs across different languages. The inclusion of Spanish in this corpus reinforces its suitability for semantic inference tasks applied to tourism-related texts.

4.2.2. Sustainability scoring assignment

Each review was classified according to five thematic labels: economic, social, environmental, cultural, and an additional category that groups general content. These labels were directly input into the BETO cased model using the zero-shot classification approach, which enabled the model to autonomously determine the degree of semantic affinity between each comment and the respective categories, without requiring additional training or detailed label descriptions.

The model assigned a probability of association to each category, which was then normalized to a scale ranging from 0 to 100. This produced a percentage distribution for each review, enabling the identification of the proportion of discourse related to each sustainability dimension. Since a single review can address multiple topics, a given comment may receive simultaneous scores across multiple dimensions.

Table 2 illustrates examples of the outcome of this process and reports the scores assigned to each sustainability dimension: economic (ECO), social (SOC), environmental (ENV), cultural (CULT) and other aspects (OTHERS). It shows, for instance, how an economic mention can result in a high score for that dimension, even if the comment also addresses other aspects. This behaviour reflects the semantic inference logic applied by the model.

4.2.3. Methodological considerations and limitations

After applying the model, a manual review of 120 reviews was conducted, selected through stratified sampling based on the confidence level and sustainability dimension (high > 90, ambiguous 70–85, low 50–70), in order to assess the accuracy of the classifications and quantify potential false positives. This validation revealed an overall accuracy of 75.8%, with variations across dimensions: cultural (100%), economic (70%), social (66.7%), and environmental (66.7%). This exploratory analysis confirmed the overall validity of the approach and helped identify patterns in the model's behaviour and possible deviations in specific cases.

Despite its usefulness, this approach presents some limitations. First, the model tends to assign high scores to dimensions that contain explicit terms – such as “price” or “euros” – even when these represent only a minor portion of the review. For example, in the review “Although 7 euros may seem expensive, the visit is worth it; they provide an audio guide to better understand the Cathedral”, the mention of price results in a high economic score (93.82%), despite the content also including cultural elements. Second, the validation confirmed that the model establishes semantic associations that may lead to false positives (12.5% of the cases analysed). Expressions such as “superbly maintained” or “very clean” were interpreted as being linked to environmental sustainability due to their conceptual proximity to terms like “conservation” and “environmental care”, although the authors were referring to aesthetic or

general maintenance aspects. This pattern was particularly evident in the environmental and social dimensions, where general terms were misinterpreted as sustainable practices.

Table 2. *Examples of comments with their classification and punctuation*

Comment	ECO	SOC	ENV	CULT	OTHERS
There is a tour with which you can visit 5 sites, it is highly recommended, it is quite cheap, they sell it at the tourist office.	95.97	0.96	0.70	0.85	1.52
Although 7 euros may seem expensive, the visit is worth it; at the entrance you will be provided with an audio guide to better understand this Cathedral.	93.82	1.32	0.73	1.98	2.15
Improve safety standards, social distance in shows. It produces discussions among spectators. The seats not to be occupied should be marked.	1.62	89.19	0.87	1.38	6.94
Everything in the Gothic Quarter has its charm: churches hidden among alleys, street musicians who make you smile, impressive cathedrals, people of all kinds. Overall, one of the best neighbourhoods in Barcelona to get lost.	4.54	68.95	3.39	14.07	9.05
Without a doubt, if you want to visit a natural environment representative of the Mediterranean forest, you must explore this wonderful national park.	0.64	1.10	94.55	1.60	2.11
Great environment, you can hear the river Segura, there is space for barbecue, and you can bring your dog. And also, the Sanctuary is beautiful.	1.47	1.55	93.22	1.90	1.86
The Cathedral has an important cultural value, which is recognised and highlighted by its restoration and conservation. Inside, they explain the meaning of each thing, making known its historical and cultural importance.	0.51	1.45	0.81	95.53	1.69
A classic example of our cultural mix. The Sephardi Museum, with excellent supporting texts available in the gift shop, helps visitors to better understand the visit and to deepen their history.	0.61	2.13	0.74	86.68	9.85
Other readers and visitors have described it better than I. I, frankly, loved it. Two highlights: the arm of Saint Vincent and the Chapel of the Holy Chalice. Otherwise, it is a temple must visit.	0.75	2.32	0.80	2.19	93.94
It takes you to another time when things were done with more ingenuity than technology. Cannons are something from another world.	1.54	2.51	1.16	2.54	92.24
Culture everywhere, testimony of Roman Hispania. It is a visit that cannot be missed if your goal is to understand the history of Spain and its roots.	0.39	5.30	0.36	89.79	4.16
Beautiful museum, superbly maintained, great atmosphere!!! The entrance is very reasonable for everything it offers! Recommended.	13.62	0.83	47.57	37.21	0.76

Third, it is important to clarify that the classification performed does not necessarily imply that the tourist expresses a conscious or reflective stance on sustainability. The model detects content thematically linked to sustainable dimensions – environmental, social, cultural, or economic – based on semantic proximity, without distinguishing intentionality or polarity. Manual validation identified 14 ambiguous cases (11.7%) in which this distinction was particularly difficult to determine. This approach

enables the capture of relevant semantic signals in 75.8% of cases, but it also entails the risk of interpreting general expressions or personal evaluations as sustainability-related content, which is acknowledged as a methodological limitation.

Moreover, although the model was not specifically adapted to the context of sustainable tourism in Spain, its direct application enabled a robust analysis, with particularly high accuracy in the cultural dimension (100%). Nevertheless, future research could benefit from more refined adjustments that reduce the rate of false positives in the environmental and social dimensions, thus enhancing the sensitivity of the model to the cultural and thematic nuances of the tourism sector.

4.3. *Ologit and Gologit models*

To analyse how the thematic presence of sustainability-related content in the reviews is associated with the rating given by tourists – considered an indicator of their satisfaction – ordered logistic regression models (ologit) are applied, since the dependent variable is ordinal (on a scale ranging from 1 to 5). This model, proposed by McCullagh (1980), enables estimation of the probability that an observation falls into a particular ordinal category based on a set of explanatory variables. The independent variables include the scores assigned to each sustainability dimension (economic, social, cultural, and environmental), previously generated using the BETO model. Following D'Acunto *et al.* (2023), several control variables are also included: type of trip (friends, solo, family, couples), tourist nationality (domestic or international), review length (number of words), and user experience (number of contributions).

Correlations among the independent variables (social, economic, environmental, and cultural) were verified, and the categorical variables (type of trip and nationality) were converted into dummy variables, using “couples” as the reference category for type of trip. Continuous variables with high dispersion were transformed using the natural logarithm to reduce the influence of extreme values (contributions and word count). Before estimating the models, the proportional odds assumption was tested using the Brant test, which evaluates whether the relationship between the independent variables and the dependent variable remains constant across its categories. Since the assumption was violated in all cases, generalized ordered logistic regression (gologit) was applied, allowing the proportionality assumption to be relaxed only for the variables that require it. Estimation was carried out using the *gologit2* module in Stata (Williams, 2006), with the *autofit* option.

To ensure the robustness of the results and to analyse potential differences based on the grouping of categories and control variables, four different models were estimated:

- Models 1 and 2 (M1 and M2) use the original five categories of the dependent variable (1 to 5).
- Models 3 and 4 (M3 and M4) group the categories into three levels: low (1–2), medium (3–4), and high (5).
- M1 and M3 include only two control variables: type of trip and nationality.
- M2 and M4 incorporate all four control variables: type of trip, nationality, review length, and reviewer experience.

The estimates of the logistic regression models (logit) are obtained through maximum likelihood estimation. Unlike linear regression, the traditional coefficient of determination (R^2) is not applicable; instead, Pseudo R^2 is used, which should not be interpreted as the percentage of explained variance, but rather as a relative measure of model fit: higher values indicate better predictive power within the same data set (Hemmert *et al.*, 2016; UCLA, 2024).

Two key indicators for interpreting the results are the coefficients and the odds ratios. A positive coefficient indicates that higher values in the independent variable increase the likelihood of receiving

a higher rating, and vice versa. The odds ratio indicates how many times more likely an event is to occur (e.g. giving a higher rating) when the independent variable increases by one unit, compared to a reference group.

Table 3 presents the distribution of reviews by rating category. Since most reviews are concentrated in the highest values (4 and 5), the use of alternative models with grouped categories allows the robustness of the results to be tested under different configurations. Table 4 summarizes all the variables and configurations used.

Table 3. *Number of reviews per category of the dependent variable*

Category	Nº of reviews	M1 and M2	M3 and M4
1	2503	2503	5209
2	2706	2706	67587
3	12693	12693	134910
4	54894	54894	
5	134910	134910	
Total	207706	207706	207706

Table 4. *Variables included in the analysis*

Construct	Variable name	Type of variable	Operationalization	M1	M2	M3	M4
Rating Score	RATING ₁	Dependent variable	Categorical variables from 1 to 5	X	X		
Rating Score	RATING ₂	Dependent variable	Categorical variables from 1 to 3			X	X
Economic content of the review	ECO	Independent variable	The natural logarithm of the calculated score of the economic content	X	X	X	X
Social content of the review	SOC	Independent variable	The natural logarithm of the calculated social content score	X	X	X	X
Environmental content of the review	ENV	Independent variable	The natural logarithm of the calculated environmental content score	X	X	X	X
Cultural content of the review	CULT	Independent variable	The natural logarithm of the calculated score of cultural content	X	X	X	X
Origin	ORI	Control variable	Variable dummy: Domestic tourist (o foreigner)	X	X	X	X
Travel category	FRI, SOLO, FAM	Control variable	Variable dummy: Friends, Solo and Family (o couples)	X	X	X	X
Length of the review	WC	Control variable	The natural logarithm of the word count of each review		X		X
Contributions from reviewers	CONT	Control variable	The natural logarithm of the total number of reviews published by the reviewer		X		X

5. Results

5.1. Association between sustainability discourse and the ratings given by tourists

Table 5 presents the results of the logit models applied to the full sample (207706 reviews). Overall, the results are consistent across the four estimated models, supporting the robustness of the findings. However, Models M2 and M4, which include a greater number of control variables, perform better in

terms of model fit, as indicated by their higher Pseudo R² values. It is important to note that, unlike the R² coefficient in linear regression models, Pseudo R² does not represent the percentage of variance explained. Nevertheless, it enables the comparison of models estimated on the same data set, with higher values indicating better relative predictive power. In this case, the models that include additional controls – such as review length and user experience explain better the variability in the ratings given by tourists.

Table 5. *Results of gologit models*

Model	M1		M2		M3		M4		
Pseudo R2	0.0219		0.0316		0.0223		0.0335		
Variable	Cut	Coefficient	Odds Ratio	Coefficient	Odds Ratio	Coefficient	Odds Ratio	Coefficient	Odds Ratio
Constant	Cut 1	2.3823***	10.8299	4.8488***	127.5879	1.4490***	4.2591	4.1258***	61.9174
	Cut 2	1.3149***	3.7245	3.9547***	52.1787	-0.5270***	0.59037	0.3433***	1.4095
	Cut 3	0.2103***	1.2341	1.8208***	6.1769				
	Cut 4	-0.5290***	1.0312	0.3364***	1.3999				
ECO	Cut 1	-0.1303***	0.8778	-0.1830***	0.8327	-0.1325***	0.8759	-0.1854***	0.8308
	Cut 2	-0.1303***	0.8778	-0.1900***	0.8270	-0.1325***	0.8759	-0.1349***	0.8738
	Cut 3	-0.1303***	0.8778	-0.1401***	0.8693				
	Cut 4	-0.1303***	0.8778	-0.1344***	0.8743				
SOC	Cut 1	-0.6039***	0.5467	-0.5192***	0.5950	-0.4198***	0.6572	-0.3894***	0.6775
	Cut 2	-0.4105***	0.6633	-0.3517***	0.7035	-0.0087	0.9913	-0.0114	0.9887
	Cut 3	-0.1465***	0.8637	-0.1422***	0.8675				
	Cut 4	-0.0098	0.9902	-0.0117	0.9883				
ENV	Cut 1	0.5642***	1.7581	0.5014***	1.6511	0.4772***	1.6115	0.4274***	1.5333
	Cut 2	0.4995***	1.6479	0.4314***	1.5394	0.1600***	1.1735	0.1442***	1.1551
	Cut 3	0.3562***	1.4279	0.3183***	1.3748				
	Cut 4	0.1599***	1.1734	0.1438***	1.1547				
CULT	Cut 1	1.0300***	2.8010	0.8383***	2.3125	0.9533***	2.5944	0.8024***	2.2310
	Cut 2	0.9587***	2.6082	0.7973***	2.2196	0.3290***	1.3896	0.3254***	1.3846
	Cut 3	0.6831***	1.9800	0.6295***	1.8767				
	Cut 4	0.3295***	1.3902	0.3247***	1.3836				
FRI	Cut 1	0.0952***	1.0998	0.0831***	1.0867	0.0158	1.0159	0.0826***	1.0861
	Cut 2	0.0952***	1.0998	0.0831***	1.0867	0.0983***	1.1032	0.0826***	1.0861
	Cut 3	0.0952***	1.0998	0.0831***	1.0867				
	Cut 4	0.0952***	1.0998	0.0831***	1.0867				
SOLO	Cut 1	-0.0599	0.9419	-0.1125	0.8936	-0.0133	0.9868	0.0051	1.0052
	Cut 2	0.0208	1.0210	-0.0107	0.9894	0.1979***	1.2188	0.1965***	1.2171
	Cut 3	0.1034**	1.1090	0.1146***	1.1214				
	Cut 4	0.1965***	1.2171	0.1947***	1.2150				
FAM	Cut 1	-0.1943***	0.8234	-0.0199	0.9803	-0.1609***	0.8514	-0.0181	0.9820
	Cut 2	-0.1302***	0.8779	-0.0199	0.9803	0.0073	1.0073	-0.0181	0.9820
	Cut 3	-0.0630***	0.9390	-0.0199	0.9803				
	Cut 4	0.0061	1.0061	-0.0199	0.9803				
ORI	Cut 1	-0.2650***	0.7672	-0.1705**	0.8432	-0.1994***	0.8192	-0.1098**	0.8960
	Cut 2	-0.1885***	0.8282	-0.1091**	0.8966	0.0309*	1.0314	0.0201	1.0203
	Cut 3	-0.0657***	0.9364	-0.0356	0.9650				
	Cut 4	0.0307**	1.0312	0.0198	1.0200				
WC	Cut 1			-0.7353***	0.4794			-0.7320***	0.4809
	Cut 2			-0.6984***	0.4974			-0.1165***	0.8901
	Cut 3			-0.3941***	0.6742				
	Cut 4			-0.1133***	0.8929				
CONT	Cut 1			0.2238***	1.2508			0.1563***	1.1692
	Cut 2			0.1427***	1.1534			-0.0858***	0.9178
	Cut 3			0.0228***	1.0231				
	Cut 4			-0.0861***	0.9175				

Significance levels are denoted as *** for a p-value < 0.001, ** for a p-value < 0.01, and * for a p-value < 0.05.

The probability of receiving higher ratings decreases when the discourse in the reviews includes economic and social aspects (ECO and SOC), whereas reviews with a greater presence of environmental (ENV) and cultural (CULT) themes tend to receive higher scores, indicating that mentioning these aspects is positively associated with the given rating. In addition, the nature of the trip plays a key role. People travelling with friends (FRI) or solo travellers (SOLO), as opposed to travelling as a couple, tend to give more favourable ratings to tourist attractions, whereas domestic users (ORI) show a tendency

to be more critical, often giving lower ratings than international travellers (Table 5, Models 1 and 3, respectively).

In addition, reviews characterized by a higher word count (WC) tend to be associated with lower ratings, suggesting a tendency to elaborate on negative experiences, which often require more detailed explanations. This confirms the results of Filieri *et al.* (2019), as they show that negative ratings are associated with long reviews. Similarly, expert users (CONT) tend to assign higher ratings to tourist attractions. This behaviour could reflect the fact that, with more experience, users generally develop a more critical but positive appreciation of certain aspects of attractions. However, these results differ from those found in previous research, which tends to associate greater experience with more negative reviews, as the effect of perceived destination quality is weaker among expert travellers (Wang *et al.*, 2017; Zhang *et al.*, 2016).

Among the four models analysed, M2 emerges as the most robust option for reflecting the results of this research. It also presents one of the highest Pseudo R² values, indicating better predictive performance than the other models. When examining the odds ratios of M2 (see Table 5), relevant effects can be observed in certain sustainability dimensions: for example, a one-unit increase in the social dimension (SOC) reduces the probability of a review falling into a higher rating category at the first threshold (Cut 1) – that is, moving from a rating of 1 to 2 – by 41%. In contrast, the cultural dimension (CULT) shows a notably positive influence: a one-unit increase in this variable raises the probability of receiving a higher rating at the same threshold by 131%. Figure 2 complements this information by visualizing the estimated coefficients for Model M2. Statistically significant values at each threshold are highlighted with vertical lines, making it easier to identify the magnitude and direction of the effect of each variable.

For validating the hypotheses, and given that the proportional odds assumption is not met, it was decided to accept a hypothesis when the coefficients of all the cut-off points show statistical significance in accordance with the proposed relationship, whether positive or negative. Therefore, H1, H2, and H3 are accepted, since the coefficients of ENV and CULT are positive and statistically significant, while those of ECO are negative and statistically significant. This demonstrates that a greater presence of discourse on environmental and cultural aspects is positively associated with the ratings given to tourist attractions; however, reviews containing more discourse on the economic dimension tend to obtain lower ratings. These findings are consistent with previous research such as that of Mariani and Borghi (2023), Rita *et al.* (2022), and Väisänen *et al.* (2023).

Conversely, H4, corresponding to the independent variable SOC, is rejected, since although all its coefficients are negative, the coefficient of the last intercept (cut-off 4) is not statistically significant. Therefore, it is not possible to confirm a negative association between the presence of discourse on the social dimension and the rating given by tourists. These results differ from those in the study by D'Acunto *et al.* (2020); however, the proposed relationship with respect to the rating was fulfilled, being negative in all coefficients. In other words, comments with greater social content exhibit a tendency to be associated with lower ratings; however, this relationship was not statistically validated across all thresholds of the model. This is because a gologit model shows how the independent variables affect the probability of transitioning between the different categories of the dependent variable at each cut-off point. The differences between the intersections do not necessarily reflect overall changes in the relationships between the (dependent and independent) variables, as it could be seen that these relationships are not parallel. Consequently, the social dimension is not significantly associated with the probability of a tourist attraction being given a rating of four instead of five.

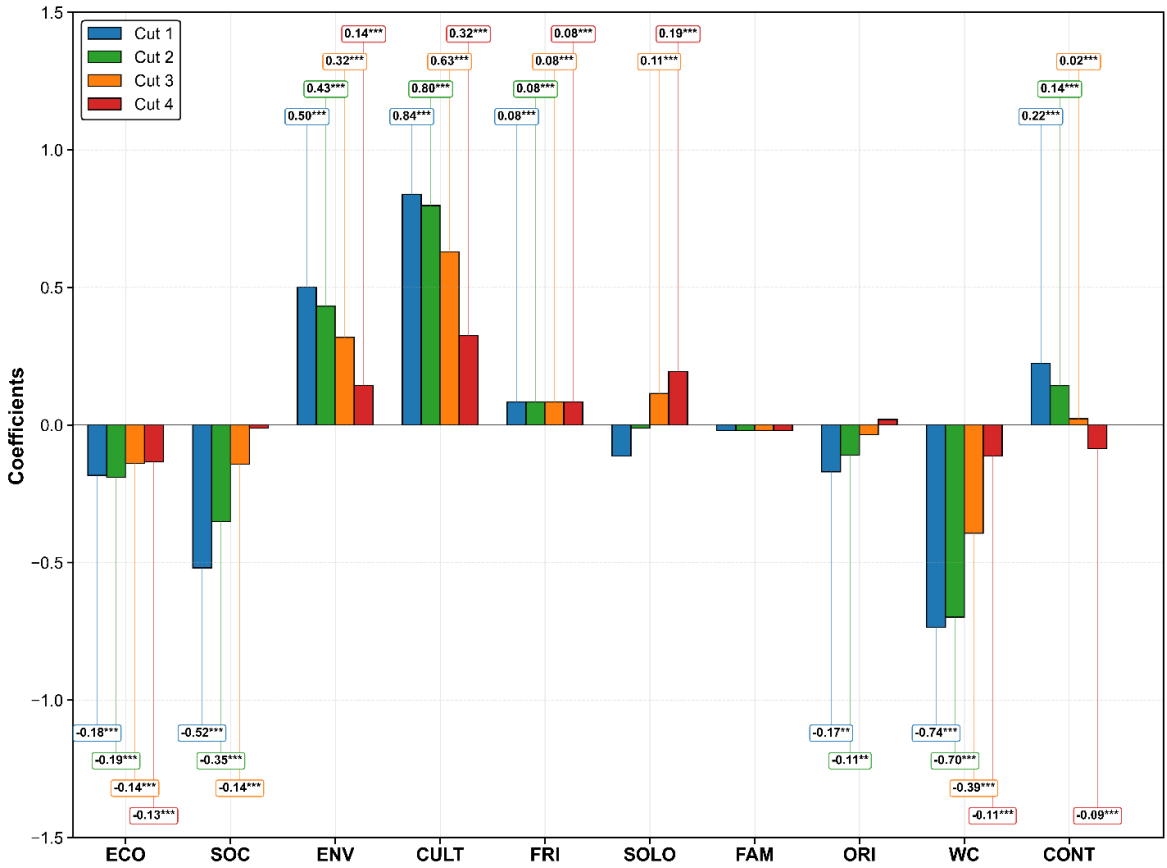


Figure 2. Coefficients of M2.

5.2. Differences in sustainability discourse between rural and urban areas

Considering the models based on the territorial context (rural or urban) of the tourist attractions yields interesting results (see Tables 6 and 7). As before, Models 2 and 4 achieve the best results, as they have the highest Pseudo R-squares. While the probability of giving a higher rating is also decreased by the discourse associated with the economic (ECO) and social (SOC) dimensions in urban attractions, the probability is higher when the reviews contain social references but not when the review contains economic ones. The environmental (ENV) and cultural (CULT) dimensions in the reviews have a similar effect on rural and urban tourist attractions, although the environmental dimension shows a higher impact on rural attractions and the cultural dimension on urban ones.

Trips made with friends (FRI) present higher scores for urban tourist attractions, in contrast to trips made by couples. However, the results for rural attractions are less consistent and not significant, suggesting that the company of friends does not have an impact on rural attraction ratings. Solo travellers (SOLO) tend to be more critical and demanding than couples in rural areas, giving higher ratings in urban areas. Family trips (FAM) tend to be associated with lower ratings in rural areas, whereas they do not have a significant effect in urban contexts, indicating that this type of trip does not consistently influence the evaluation of these attractions. On the other hand, domestic travellers (ORI) give lower ratings to urban environments than foreigners do, which could suggest that nationals have higher expectations of these tourist attractions. In contrast, this variable has no influence on the ratings of rural attractions, as its effect is not significant.

Table 6. Rural and urban M1 and M2 gologit models

Model		M1				M2			
Pseudo R2		0.0169		0.0233		0.0255		0.0341	
Variable	Cut	Coefficient		Odds Ratio		Coefficient		Odds Ratio	
		Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Constant	Cut 1	2.3961***	2.3415***	10.9798	10.3973	6.2360***	4.7690	510.8377	117.8068
	Cut 2	1.4242***	1.2807***	4.1547	3.5991	4.6923***	3.9417	109.108	51.5074
	Cut 3	0.3068	0.1612***	1.3590	1.1750	2.1347***	1.8596	8.4548	6.4209
	Cut 4	-0.5186***	-0.5729**	0.5952	0.5639	0.5332***	0.3118	1.7044	1.3658
ECO	Cut 1	-0.0847***	-0.1333***	0.9188	0.8752	-0.1578*	-0.1946***	0.8540	0.8231
	Cut 2	-0.0847***	-0.1333***	0.9188	0.8752	-0.1569***	-0.1969***	0.8547	0.8213
	Cut 3	-0.0847***	-0.1333***	0.9188	0.8752	-0.0643*	-0.1469***	0.9377	0.8633
	Cut 4	-0.0847***	-0.1333***	0.9188	0.8752	-0.0993***	-0.1359***	0.9055	0.8730
SOC	Cut 1	-0.7878***	-0.5771***	0.4549	0.5615	-0.6436***	-0.4999***	0.5254	0.6066
	Cut 2	-0.5715***	-0.3855***	0.5647	0.6801	-0.4338***	-0.3432***	0.6480	0.7095
	Cut 3	-0.1328***	-0.1455***	0.8756	0.8646	-0.0833*	-0.1482***	0.9201	0.8622
	Cut 4	0.0132	-0.0088	1.0133	0.99123	0.0218	-0.0092	1.0220	0.9908
ENV	Cut 1	0.7008***	0.5359***	2.0154	1.7090	0.5770***	0.4703***	1.7806	1.6004
	Cut 2	0.6130***	0.4738***	1.8459	1.6061	0.4932***	0.4036***	1.6376	1.4972
	Cut 3	0.4289***	0.3381***	1.5355	1.4022	0.3640***	0.2966***	1.4390	1.3453
	Cut 4	0.2346***	0.1446***	1.2644	1.1555	0.2055***	0.1272***	1.2282	1.1357
CULT	Cut 1	1.0222***	1.040***	2.7793	2.8302	0.7920***	0.8313***	2.2079	2.2963
	Cut 2	0.9397***	0.9683***	2.5591	2.6336	0.7478***	0.8056***	2.1123	2.2380
	Cut 3	0.5735***	0.7078***	1.7744	2.0295	0.4828***	0.6557***	1.6206	1.9265
	Cut 4	0.2537***	0.3506***	1.2888	1.4200	0.2101***	0.3526***	1.2338	1.4228
FRI	Cut 1	0.0081	0.1138***	1.0081	1.1205	0.0022	0.1012***	1.0022	1.1065
	Cut 2	0.0081	0.1138***	1.0081	1.1205	0.0022	0.1012***	1.0022	1.1065
	Cut 3	0.0081	0.1138***	1.0081	1.1205	0.0022	0.1012***	1.0022	1.1065
	Cut 4	0.0081	0.1138***	1.0081	1.1205	0.0022	0.1012***	1.0022	1.1065
SOLO	Cut 1	0.0359	-0.0091	1.0365	0.9909	-0.7116*	-0.0128	0.4909	0.9872
	Cut 2	0.0359	0.0449	1.0365	1.0460	-0.1874	0.0583	0.8291	1.0600
	Cut 3	0.0359	0.1322***	1.0365	1.1413	-0.0226	0.1643***	0.9776	1.1786
	Cut 4	0.0359	0.2236***	1.0365	1.2506	0.0491	0.2167***	1.0503	1.2420
FAM	Cut 1	-0.0795**	-0.1831***	0.9236	0.8327	-0.0708**	-0.0089	0.9317	0.9911
	Cut 2	-0.0795**	-0.1218***	0.9236	0.8853	-0.0708**	-0.0089	0.9317	0.9911
	Cut 3	-0.0795**	-0.0553**	0.9236	0.9461	-0.0708**	-0.0089	0.9317	0.9911
	Cut 4	-0.0795**	0.0215	0.9236	1.0218	-0.0708**	-0.0089	0.9317	0.9911
ORI	Cut 1	0.0149	-0.3046***	1.0150	0.7374	0.0173	-0.1818**	1.0175	0.8337
	Cut 2	0.0149	-0.2305***	1.0150	0.7941	0.0173	-0.1328**	1.0175	0.8756
	Cut 3	0.0149	-0.0814***	1.0150	0.9218	0.0173	-0.0531*	1.0175	0.9483
	Cut 4	0.0149	0.0208	1.0150	1.0211	0.0173	0.0059	1.0175	1.0059
WC	Cut 1					-0.8686***	-0.7493***	0.4195	0.4727
	Cut 2					-0.7668***	-0.7152***	0.4645	0.4891
	Cut 3					-0.4567***	-0.4136***	0.6334	0.6613
	Cut 4					-0.1921***	-0.1105***	0.8252	0.8953
CONT	Cut 1					0.1310***	0.2577***	1.1400	1.2939
	Cut 2					0.1354***	0.1575***	1.1450	1.1706
	Cut 3					0.0542***	0.0241***	1.0557	1.0244
	Cut 4					-0.0284***	-0.0971***	0.9720	0.9074

Significance levels are denoted as *** for a p-value < 0.001, ** for a p-value < 0.01, and * for a p-value < 0.05.

Table 7. Rural and urban M₃ and M₄ gologit models

Model		M ₃				M ₄			
Pseudo R ₂		0.0187		0.0236		0.0288		0.0359	
Variable	Cut	Coefficient		Odds Ratio		Coefficient		Odds Ratio	
		Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Constant	Cut 1	1.4289***	1.4008***	4.1742	4.0585	4.7178***	4.1624***	111.926	64.2286
	Cut 2	-0.5135***	-0.5690***	0.5983	0.5661	0.5284***	0.3201***	1.6963	1.3772
ECO	Cut 1	-0.0911***	-0.1352***	0.9129	0.8736	-0.1014***	-0.1929***	0.9036	0.8246
	Cut 2	-0.0911***	-0.1352***	0.9129	0.8736	-0.1014***	-0.1365***	0.9036	0.8723
SOC	Cut 1	-0.6221***	-0.4008***	0.5368	0.6698	-0.5027***	-0.3754***	0.6049	0.6870
	Cut 2	0.0164	-0.0074	1.0165	0.9926	0.0243	-0.0092	1.0246	0.9908
ENV	Cut 1	0.6107***	0.4531***	1.8417	1.5732	0.5088***	0.3957***	1.6632	1.4855
	Cut 2	0.2339***	0.1443***	1.2636	1.1552	0.2065***	0.1276***	1.2294	1.1360
CULT	Cut 1	0.9966***	0.9644***	2.7090	2.6233	0.8058***	0.8001***	2.2385	2.2259
	Cut 2	0.2514***	0.3496***	1.2858	1.4186	0.2098***	0.3535***	1.2334	1.4241
FRI	Cut 1	0.0031	0.1146***	1.0031	1.1214	-0.0041	0.1017***	0.9959	1.1071
	Cut 2	0.0031	0.1146***	1.0031	1.1214	-0.0041	0.1017***	0.9959	1.1071
SOLO	Cut 1	0.0351	0.0435	1.0357	1.0445	0.0392	0.0831	1.0399	1.0867
	Cut 2	0.0351	0.2240***	1.0357	1.2511	0.0392	0.2197***	1.0399	1.2457
FAM	Cut 1	-0.0795**	-0.1244***	0.9236	0.8831	0.1512	-0.0061	1.1632	0.9939
	Cut 2	-0.0795**	0.0219	0.9236	1.0221	-0.0812**	-0.0061	0.9220	0.9939
ORI	Cut 1	0.0265	-0.2474***	1.0268	0.7809	0.0295	-0.1472***	1.0299	0.863
	Cut 2	0.0265	0.0213	1.0268	1.0215	0.0295	0.0061	1.0299	1.0061
WC	Cut 1					-0.8517***	-0.7556***	0.4267	0.4697
	Cut 2					-0.1934***	-0.1141***	0.8242	0.8922
CONT	Cut 1					0.1631***	0.1736***	1.1772	1.1895
	Cut 2					-0.0285***	-0.0968***	0.9719	0.9077

The need to explain negative experiences does not appear to be related to the type of setting, although the length of the review (WC) is more strongly associated with lower ratings in rural environments than in urban tourist attractions. Furthermore, experienced travellers (CONT) generally rate attractions positively in both rural and urban settings, although they tend to give even higher ratings in urban ones, which suggests that as users' experience increases, their perception of value in urban attractions is enhanced, possibly due to a greater diversity of services or a tourism offer that better meets their expectations. To facilitate the interpretation of the results, Figure 3 displays the estimated coefficients of Model M₂ broken down by cut-off point (Cut 1–4), for the full sample as well as for rural and urban areas. In addition, lines and labels indicating levels of statistical significance have been added, enabling visual identification of which effects are relevant in each territorial context.

6. Discussion

The results reinforce the existing literature that examines the relationship between sustainability and tourist satisfaction based on online reviews (Bigne *et al.*, 2024; Brazyté *et al.*, 2016; Gerdt *et al.*, 2019; Marchi *et al.*, 2023; Yu *et al.*, 2017). Moreover, they confirm that the dimensions of sustainability are associated in different ways with tourist satisfaction. In particular, our findings indicate that when reviews include references to environmental and cultural sustainability, the probability of assigning a higher rating to that tourist attraction also increases.

These findings are consistent with studies that emphasize the transformative value of sustainability as part of the tourism experience (Moliner-Tena *et al.*, 2023), as well as with research showing that a sustainable atmosphere significantly influences customer satisfaction (Chang & Lin, 2022; Zhao *et al.*, 2025). As regards the cultural dimension, its importance as a pull factor is confirmed (Correia *et al.*, 2013), particularly in urban contexts, through elements such as museums, heritage sites, public art, and cultural programming (Lee *et al.*, 2014; Qiu *et al.*, 2024). The environmental dimension, in turn, is associated with benefits related to well-being and mental health (Akhoundogli & Buckley, 2023;

Buckley, 2019, 2020, 2023; Buckley *et al.*, 2021; Cooper & Buckley, 2022), reinforcing its value as a restorative experience in contrast to the urban environment (Kastenholz *et al.*, 2012; Kumar & Shekhar, 2020).

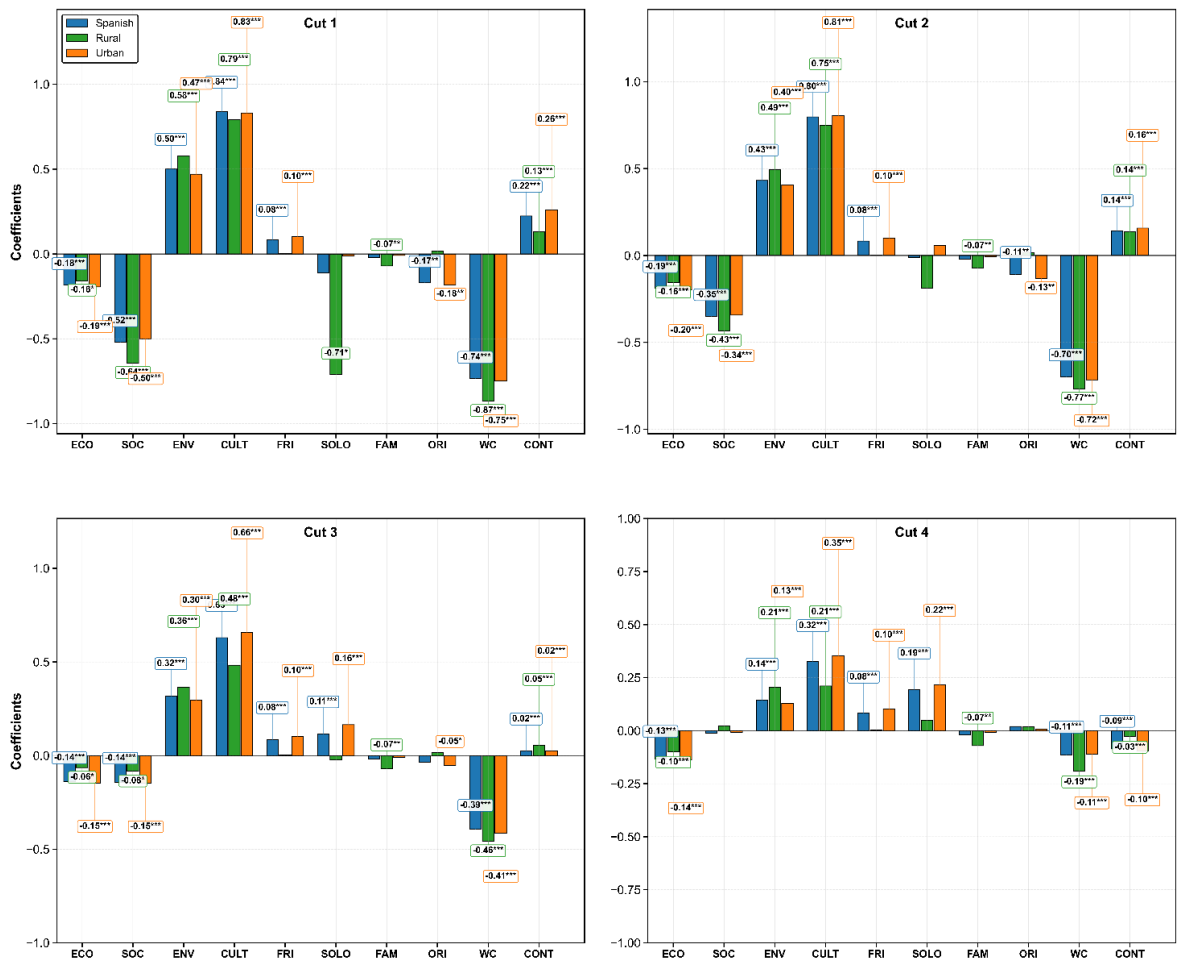


Figure 3. M2 coefficients for Spain, rural and urban areas.

With respect to the economic dimension, the results show a negative association with the rating given by tourists. While studies such as those by Ye *et al.* (2014) and Väisänen *et al.* (2023) highlight the importance of value for money in tourist evaluations, this study indicates that mentions of prices or costs are not always associated with greater satisfaction. In fact, the economic dimension tends to reduce the likelihood of higher ratings, particularly in urban areas. This may be explained by greater price sensitivity or a perceived imbalance between cost and experience. In contrast, this dimension carries less weight in rural areas, possibly due to a higher willingness to pay among tourists who are more committed to sustainable values, such as ecotourists (Kazeminia *et al.*, 2016).

As for the social dimension, the results do not confirm the existence of a significant and consistent relationship with tourist satisfaction, in contrast to what was observed for the environmental and cultural dimensions. This finding aligns with previous studies such as the ones by D'Acunto *et al.* (2020) and Rita *et al.* (2022), which associate the social dimension with lower ratings on online review

platforms. This dimension encompasses aspects such as customer and infrastructure safety, support for local employability, and healthy food and activity options, as well as customer education on sustainability (Dewhurst & Thomas, 2003; Garay & Font, 2012; Jones *et al.*, 2014; Väisänen *et al.*, 2023).

The findings support the need for a more balanced and comprehensive sustainability communication, as proposed by Čuić Tanković and Mušanović (2022) and Marchi *et al.* (2023). Integrating sustainability into communication and marketing strategies can enhance satisfaction and foster visitor loyalty (Garrod *et al.*, 2023); moreover, the importance of adapting communication strategies based on the type of destination is highlighted, as sustainability discourse and its association with tourist ratings vary between urban and rural contexts. Finally, the findings are in line with those of authors such as León *et al.* (2025) and Moreno-Brito *et al.* (2024), who argue that cultural and contextual differences – such as the tourist's origin or language, or the depth of discourse – can influence how sustainability is expressed. Although this study does not explore these variations in depth, we acknowledge their importance as a key avenue for future research that is essential for personalizing communication and enriching the tourist experience.

7. Conclusions, implications, limitations, and future lines of research

7.1. Conclusions

This study analyses 207706 reviews of Spain's main tourist attractions published on TripAdvisor between 2012 and 2023. Through automated content analysis and the application of generalized ordered logistic regression models (gologit), the sustainability discourse of tourists is examined to identify how the thematic presence of the sustainability dimensions – economic, social, environmental, and cultural – is associated with satisfaction levels reflected in the ratings given. The results show that a higher proportion of environmental and cultural discourse is associated with higher ratings, whereas a greater presence of economic and social discourse is linked to lower ones.

These associations vary depending on the territorial context. Reviews that mention environmental aspects are more likely to receive higher ratings in rural destinations, while references to cultural elements are associated with higher levels of satisfaction in urban settings. In terms of visitor profiles, trips with friends tend to generate higher ratings for urban attractions than do trips as a couple. On the other hand, solo travellers have a tendency to be more critical in rural environments, whereas in urban areas they give higher ratings than couples. Consequently, this study demonstrates that analysing sustainability discourse through online reviews is a valuable tool for understanding how the different sustainability dimensions are related to tourist satisfaction, taking into account both the territorial context and the type of visitor.

7.2. Theoretical implications

This study contributes to enriching the current tourism literature by expanding research that applies innovative methodologies based on data science techniques to the analysis of user-generated content (UGC) (Cai *et al.*, 2024). Specifically, it explores the relationship between sustainability and tourist satisfaction using online reviews, which is an emerging line of research in the academic field (Bigne *et al.*, 2024; Brazyté *et al.*, 2016; Gerdt *et al.*, 2019; Marchi *et al.*, 2023; Yu *et al.*, 2017). In this regard, the methodological value of the approach employed is noteworthy, as it combines natural language processing models with statistical analysis, offering a replicable and adaptable framework for future studies focused on sustainable tourism and automated review analysis. This contribution is particularly relevant given that, despite the growing number of studies based on big data and data science over the past decade (António & Rita, 2023; Boyd & Crawford, 2012; Cai *et al.*, 2024; Mariani & Baggio, 2022), these methodologies have not yet been fully integrated into tourism research (Egger, 2022; Mariani *et al.*, 2018; Mariani & Baggio, 2022). Thus, the study reinforces the need to incorporate new approaches that leverage the potential of unstructured data, such as online reviews.

Secondly, this study contributes to improving the comparative understanding of sustainability discourse by analysing how its different dimensions are expressed depending on the type of destination, thus complementing previous research (Moliner-Tena *et al.*, 2023; Rasoolimanesh *et al.*, 2017). The findings reveal meaningful patterns that can guide decision-making by various stakeholders, enabling the adaptation of strategies and the reinforcement of the most valued sustainability dimensions according to the territorial context (rural or urban) and visitor profile. Finally, and along the same lines, the study provides added value by focusing on the analysis of online reviews from tourists who visited attractions in Spain – a geographical context that, to the best of our knowledge, has not been previously addressed in similar research. In doing so, it offers contextualized empirical evidence that enhances the understanding of how sustainability dimensions are linked to tourist satisfaction within the Spanish tourism sector.

7.3. *Practical implications*

The findings provide valuable insights for tourism managers, confirming that comments related to cultural and environmental dimensions are associated with higher levels of satisfaction, suggesting that strengthening these attributes can enhance the visitor experience. Urban attractions can highlight their cultural offerings through heritage interpretation, urban design, or artistic expressions, while rural sites may emphasize their natural surroundings, tranquillity, or authenticity – aligning with the motivations of their respective audiences. Even attractions not centred on nature or culture could incorporate sustainable elements through ambience, interpretive storytelling, or the responsible management of resources, thereby promoting a more balanced form of sustainability.

In contrast, the economic and social dimensions are associated with lower ratings, suggesting that there is an opportunity to improve how these aspects are communicated and integrated into the tourist experience. Managers could communicate more clearly how they contribute to the local economy or the well-being of the community. For example, messages highlighting local employment, universal accessibility, safety, or support for local producers could be included in signage, on official websites, or during the visitor experience. It is also recommended to develop communication campaigns focused on the overall value of the experience – beyond price – and to establish participatory mechanisms in which residents can co-create sustainable initiatives alongside visitors.

Likewise, self-assessment of the projected image may be key to identifying potential gaps between the sustainability that is communicated and the way these aspects are recognized or mentioned by tourists in their online reviews. Adjusting narratives and communication channels based on UGC analysis can help convey, more coherently, the actual sustainability efforts undertaken by the attraction. Strategies such as using QR codes in physical spaces to showcase applied sustainable practices, developing digital informational guides, and training staff to communicate these values during the visitor experience can be effective tools.

7.4. *Limitations and future research*

Finally, this study presents several limitations. First, the data were drawn exclusively from the TripAdvisor platform, so future research could incorporate additional sources to validate the findings. Second, tourists' nationality was classified in a binary way (domestic/foreign) without exploring more detailed cultural nuances that could be associated with differences in the sustainability discourse reflected in the reviews. Third, the study does not examine the underlying motivations that lead tourists to associate certain sustainability dimensions with their level of satisfaction. Therefore, it is recommended to conduct complementary qualitative research or to apply more advanced natural language processing techniques that enable deeper discourse patterns to be identified. Fourth, although the classification of tourist attractions as rural or urban is based on objective administrative criteria, we acknowledge that this approach does not necessarily reflect the experiential nature of tourism in each

setting. Visitor perception, the type of activity, or landscape characteristics may not align with territorial boundaries, which represents a methodological limitation to be considered in future research.

Acknowledgments

This work is part of the Research Project "TED2021-131314B-I00 SOSTENIBILIDAD CORPORATIVA Y TURISMO INTELIGENTE EN COMUNIDADES RURALES: INFLUENCIA EN EL DESARROLLO ECONÓMICO Y SOCIAL DEL TERRITORIO", funding by MICIU/AEI /10.13039/501100011033 and European Union (EU) NextGenerationEU/ PRTR.

References

- Agyeiwaah, E., McKercher, B., & Suntikul, W. (2017). Identifying core indicators of sustainable tourism: A path forward? *Tourism Management Perspectives*, 24, 26–33. <https://doi.org/10.1016/J.TMP.2017.07.005>
- Akhoundoglu, M., & Buckley, R. (2023). Outdoor tourism to escape social surveillance: health gains but sustainability costs. *Journal of Ecotourism*, 22(1), 4–24. <https://doi.org/10.1080/14724049.2021.1934688>
- Alhoshan, W., Ferrari, A., & Zhao, L. (2023). Zero-shot learning for requirements classification: An exploratory study. *Information and Software Technology*, 159, 107202. <https://doi.org/10.1016/J.INFSOF.2023.107202>
- Anagnostopoulou, S. C., Buhalis, D., Kountouri, I. L., Manousakis, E. G., & Tsekrekos, A. E. (2020). The impact of online reputation on hotel profitability. *International Journal of Contemporary Hospitality Management*, 32(1), 20–39. <https://doi.org/10.1108/IJCHM-03-2019-0247>
- António, N., & Rita, P. (2023). Twenty-two years of International Journal of Hospitality Management: A bibliometric analysis 2000–2021. *International Journal of Hospitality Management*, 114, 103578. <https://doi.org/10.1016/j.ijhm.2023.103578>
- Bigné, E., Oltra, E., & Andreu, L. (2019). Harnessing stakeholder input on Twitter: A case study of short breaks in Spanish tourist cities. *Tourism Management*, 71, 490–503. <https://doi.org/10.1016/J.TOURMAN.2018.10.013>
- Bigne, E., Zanfardini, M., & Andreu, L. (2024). Destination online reviews: lexicons and thematic arguments on social responsibility. *Tourism Review*, 80(4), 944–965. <https://doi.org/10.1108/TR-10-2023-0725>
- Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Brazyté, K., Weber, F., & Schaffner, D. (2016). Sustainability Management of Hotels: How Do Customers Respond in Online Reviews? *Journal of Quality Assurance in Hospitality & Tourism*, 18(3), 282–307. <https://doi.org/10.1080/1528008X.2016.1230033>
- Buckley, R. (2020). Nature tourism and mental health: parks, happiness, and causation. *Journal of Sustainable Tourism*, 28(9), 1409–1424. <https://doi.org/10.1080/09669582.2020.1742725>
- Buckley, R. (2023). Tourism and Mental Health: Foundations, Frameworks, and Futures. *Journal of Travel Research*, 62(1), 3–20. <https://doi.org/10.1177/00472875221087669>
- Buckley, R. C. (2019). Therapeutic mental health effects perceived by outdoor tourists: A large-scale, multi-decade, qualitative analysis. *Annals of Tourism Research*, 77, 164–167. <https://doi.org/10.1016/J.ANNALS.2018.12.017>
- Buckley, R., Zhong, L., & Martin, S. (2021). Mental health key to tourism infrastructure in China's new megapark. *Tourism Management*, 82, 104169. <https://doi.org/10.1016/J.TOURMAN.2020.104169>
- Bulchand-Gidumal, J., Melián-González, S., & González Lopez-Valcarcel, B. (2013). A social media analysis of the contribution of destinations to client satisfaction with hotels. *International Journal of Hospitality Management*, 35, 44–47. <https://doi.org/10.1016/J.IJHM.2013.05.003>

- Cai, Y., Li, G., Wen, L., & Liu, C. (2024). Intellectual landscape and emerging trends of big data research in hospitality and tourism: A scientometric analysis. *International Journal of Hospitality Management*, 117, 103633. <https://doi.org/10.1016/j.ijhm.2023.103633>
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2020). Spanish Pre-trained BERT Model and Evaluation Data. *ArXiv*. <https://doi.org/10.48550/arXiv.2308.02976>
- Chai, C., Song, Y., & Qin, Z. (2021). A Thousand Words Express a Common Idea? Understanding International Tourists' Reviews of Mt. Huangshan, China, through a Deep Learning Approach. *Land*, 10(6), 549. <https://doi.org/10.3390/LAND10060549>
- Chang, T. Y., & Lin, Y. C. (2022). The Effects of Atmosphere on Perceived Values and Customer Satisfaction toward the Theme Hotel: The Moderating Role of Green Practice Perception. *Sustainability*, 14(15), 9153. <https://doi.org/10.3390/SU14159153>
- Chen, H., & Rahman, I. (2018). Cultural tourism: An analysis of engagement, cultural contact, memorable tourism experience and destination loyalty. *Tourism Management Perspectives*, 26, 153–163. <https://doi.org/10.1016/j.TMP.2017.10.006>
- Chen, Y., Fay, S., & Wang, Q. (2011). The Role of Marketing in Social Media: How Online Consumer Reviews Evolve. *Journal of Interactive Marketing*, 25(2), 85–94. <https://doi.org/10.1016/j.INTMAR.2011.01.003>
- Conti, S., Dias, Á., & Pereira, L. (2023). Perceived City Sustainability and Tourist Behavioural Intentions. *Smart Cities*, 6(2), 692–708. <https://doi.org/10.3390/SMARTCITIES6020033>
- Cooper, M. A., & Buckley, R. (2022). Tourist Mental Health Drives Destination Choice, Marketing, and Matching. *Journal of Travel Research*, 61(4), 786–799. <https://doi.org/10.1177/0047287521101548>
- Correia, A., Kozak, M., & Ferradeira, J. (2013). From tourist motivations to tourist satisfaction. *International Journal of Culture, Tourism, and Hospitality Research*, 7(4), 411–424. <https://doi.org/10.1108/IJCTHR-05-2012-0022>
- Cucari, N., Wankowicz, E., & Esposito De Falco, S. (2019). Rural tourism and Albergo Diffuso: A case study for sustainable land-use planning. *Land Use Policy*, 82, 105–119. <https://doi.org/10.1016/j.LANDUSEPOL.2018.11.050>
- Čuić Tanković, A., & Mušanović, J. (2022). Exploring direct and indirect effects of sustainability communication on destination reputation. *Journal of Destination Marketing & Management*, 25, 100729. <https://doi.org/10.1016/j.JDMM.2022.100729>
- D'Acunto, D., Filieri, R., & Amato, S. (2023). Hotels' environmentally-framed eWOM. The moderating role of environmental culture. *Tourism Management*, 98, 104776. <https://doi.org/10.1016/j.tourman.2023.104776>
- D'Acunto, D., Filieri, R., & Amato, S. (2024). Who is sharing green eWOM? Big data evidence from the travel and tourism industry. *Journal of Sustainable Tourism*, 32(11), 2380–2402. <https://doi.org/10.1080/09669582.2024.2328103>
- D'Acunto, D., Tuan, A., Dalli, D., Viglia, G., & Okumus, F. (2020). Do consumers care about CSR in their online reviews? An empirical analysis. *International Journal of Hospitality Management*, 85, 102342. <https://doi.org/10.1016/j.ijhm.2019.102342>
- Dewhurst, H., & Thomas, R. (2003). Encouraging Sustainable Business Practices in a Non-regulatory Environment: A Case Study of Small Tourism Firms in a UK National Park. *Journal of Sustainable Tourism*, 11(5), 383–403. <https://doi.org/10.1080/09669580308667212>
- Egger, R. (2022). *Tourism on the verge. Applied data science in tourism: Interdisciplinary approaches, methodologies, and applications*. Springer.
- Filieri, R., Acikgoz, F., Ndou, V., & Dwivedi, Y. (2021). Is TripAdvisor still relevant? The influence of review credibility, review usefulness, and ease of use on consumers' continuance intention. *International Journal of Contemporary Hospitality Management*, 33(1), 199–223. <https://doi.org/10.1108/IJCHM-05-2020-0402>

- Filieri, R., Raguseo, E., & Vitari, C. (2019). What moderates the influence of extremely negative ratings? The role of review and reviewer characteristics. *International Journal of Hospitality Management*, 77, 333–341. <https://doi.org/10.1016/J.IJHM.2018.07.013>
- Garay, L. (2019). #VisitSpain. Breaking down affective and cognitive attributes in the social media construction of the tourist destination image. *Tourism Management Perspectives*, 32, 100560. <https://doi.org/10.1016/J.TMP.2019.100560>
- Garay, L., & Font, X. (2012). Doing good to do well? Corporate social responsibility reasons, practices and impacts in small and medium accommodation enterprises. *International Journal of Hospitality Management*, 31(2), 329–337. <https://doi.org/10.1016/J.IJHM.2011.04.013>
- Garrod, B., Zhao, A. L., & Koenig-Lewis, N. (2023). A greener way to stay: The role of perceived sustainability in generating loyalty to Airbnb. *International Journal of Hospitality Management*, 110, 103432. <https://doi.org/10.1016/J.IJHM.2023.103432>
- Gerdts, S. O., Wagner, E., & Schewe, G. (2019). The relationship between sustainability and customer satisfaction in hospitality: An explorative investigation using eWOM as a data source. *Tourism Management*, 74, 155–172. <https://doi.org/10.1016/j.tourman.2019.02.010>
- Guizzardi, A., Stacchini, A., & Costa, M. (2022). Can sustainability drive tourism development in small rural areas? Evidences from the Adriatic. *Journal of Sustainable Tourism*, 30(6), 1280–1300. <https://doi.org/10.1080/09669582.2021.1931256>
- Hansen, T. (2020). Media framing of Copenhagen tourism: A new approach to public opinion about tourists. *Annals of Tourism Research*, 84, 102975. <https://doi.org/10.1016/J.ANNALS.2020.102975>
- Hemmert, G. A. J., Schons, L. M., Wieseke, J., & Schimmelpfennig, H. (2016). Log-likelihood-based Pseudo-R² in Logistic Regression. *Sociological Methods & Research*, 47(3), 507–531. <https://doi.org/10.1177/00491241166638107>
- Hernández, J. M., Santana-Jiménez, Y., & González-Martel, C. (2021). Factors influencing the co-occurrence of visits to attractions: The case of Madrid, Spain. *Tourism Management*, 83, 104236. <https://doi.org/10.1016/J.TOURMAN.2020.104236>
- Huertas, A., Ferrer-Rosell, B., Marine-Roig, E., & Cristobal-Fransi, E. (2021). Treatment of the Airbnb controversy by the press. *International Journal of Hospitality Management*, 95, 102762. <https://doi.org/10.1016/J.IJHM.2020.102762>
- Jalilvand, M. R., & Samiei, N. (2012). The impact of electronic word of mouth on a tourism destination choice: Testing the theory of planned behavior (TPB). *Internet Research*, 22(5), 591–612. <https://doi.org/10.1108/10662241211271563>
- Jasrotia, S. S., Mishra, H. G., & Koul, S. (2019). Brick or Click? Channel Choice Disruptions in Travel Industry. *Asia-Pacific Journal of Management Research and Innovation*, 15(1–2), 16–26. <https://doi.org/10.1177/2319510X19829308>
- Jha, A. K., & Shah, S. (2021). Disconfirmation effect on online review credibility: An experimental analysis. *Decision Support Systems*, 145, 113519. <https://doi.org/10.1016/J.DSS.2021.113519>
- Jiang, J., Zhang, J., Zhang, H., & Yan, B. (2018). Natural soundscapes and tourist loyalty to nature-based tourism destinations: the mediating effect of tourist satisfaction. *Journal of Travel & Tourism Marketing*, 35(2), 218–230. <https://doi.org/10.1080/10548408.2017.1351415>
- Jones, P., Hillier, D., & Comfort, D. (2014). Sustainability in the global hotel industry. *International Journal of Contemporary Hospitality Management*, 26(1), 5–17. <https://doi.org/10.1108/IJCHM-10-2012-0180>
- Jones, P., Hillier, D., & Comfort, D. (2016). Sustainability in the hospitality industry: Some personal reflections on corporate challenges and research agendas. *International Journal of Contemporary Hospitality Management*, 28(1), 36–67. <https://doi.org/10.1108/IJCHM-11-2014-0572>
- Kastenholz, E., Carneiro, M. J., Peixeira Marques, C., & Lima, J. (2012). Understanding and managing the rural tourism experience — The case of a historical village in Portugal. *Tourism Management Perspectives*, 4, 207–214. <https://doi.org/10.1016/J.TMP.2012.08.009>

- Kazeminia, A., Hultman, M., & Mostaghel, R. (2016). Why pay more for sustainable services? The case of ecotourism. *Journal of Business Research*, 69(11), 4992–4997. <https://doi.org/10.1016/J.JBUSRES.2016.04.069>
- Kemperman, A., Arentze, T., & Aksenov, P. (2019). Tourists' City Trip Activity Program Planning: A Personalized Stated Choice Experiment. *Tourism, Hospitality and Event Management*, 53–70. https://doi.org/10.1007/978-3-030-11160-1_4
- Kim, J. H., Wang, Y., & Song, H. (2020). Understanding the causes of negative tourism experiences. *Current Issues in Tourism*, 24(3), 304–320. <https://doi.org/10.1080/13683500.2020.1711711>
- Kim, S. H., Lee, K., & Fairhurst, A. (2017). The review of “green” research in hospitality, 2000–2014: Current trends and future research directions. *International Journal of Contemporary Hospitality Management*, 29(1), 226–247. <https://doi.org/10.1108/IJCHM-11-2014-0562>
- Kumar, S., & Shekhar. (2020). Technology and innovation: Changing concept of rural tourism-A systematic review. *Open Geosciences*, 12(1), 737–752. <https://doi.org/10.1515/geo-2020-0183>
- Kusumah, E. P. (2023). Sustainable tourism concept: tourist satisfaction and destination loyalty. *International Journal of Tourism Cities*, 10(1), 166–184. <https://doi.org/10.1108/IJTC-04-2023-0074>
- Lee, Y. K., Lee, C. K., Choi, J., Yoon, S. M., & Hart, R. J. (2014). Tourism's role in urban regeneration: examining the impact of environmental cues on emotion, satisfaction, loyalty, and support for Seoul's revitalized Cheonggyecheon stream district. *Journal of Sustainable Tourism*, 22(5), 726–749. <https://doi.org/10.1080/09669582.2013.871018>
- León, C. J., Suárez-Rojas, C., Cazorla-Artiles, J. M., & González Hernández, M. M. (2025). Satisfaction and sustainability concerns in whale-watching tourism: A user-generated content model. *Tourism Management*, 106, 105019. <https://doi.org/10.1016/J.TOURMAN.2024.105019>
- Li, H., Liu, Y., Tan, C. W., & Hu, F. (2020). Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews. *International Journal of Contemporary Hospitality Management*, 32(5), 1713–1735. <https://doi.org/10.1108/IJCHM-06-2019-0581>
- Li, X., Pan, B., Law, R., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management*, 59, 57–66. <https://doi.org/10.1016/J.TOURMAN.2016.07.005>
- Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, 47, 140–151. <https://doi.org/10.1016/J.TOURMAN.2014.09.020>
- Madanaguli, A., Dhir, A., Joseph, R. P., Albishri, N. A., & Srivastava, S. (2023). Environmental sustainability practices and strategies in the rural tourism and hospitality sector: a systematic literature review and suggestions for future research. *Scandinavian Journal of Hospitality and Tourism*, 23(1), 1–28. <https://doi.org/10.1080/15022250.2023.2174179>
- Marchi, V., Marasco, A., & Apicerni, V. (2023). Sustainability communication of tourism cities: A text mining approach. *Cities*, 143, 104590. <https://doi.org/10.1016/J.CITIES.2023.104590>
- Mariani, M., & Baggio, R. (2022). Big data and analytics in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, 34(1), 231–278. <https://doi.org/10.1108/IJCHM-03-2021-0301>
- Mariani, M., Baggio, R., Fuchs, M., & Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, 30(12), 3514–3554. <https://doi.org/10.1108/IJCHM-07-2017-0461>
- Mariani, M., & Borghi, M. (2023). Exploring environmental concerns on digital platforms through big data: the effect of online consumers' environmental discourse on online review ratings. *Journal of Sustainable Tourism*, 31(11), 2592–2611. <https://doi.org/10.1080/09669582.2022.2033982>
- McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42(2), 109–127. <https://doi.org/10.1111/J.2517-6161.1980.TB01109.X>
- Meneses, R., Brito, C., Lopes, B., & Correia, R. (2023). Satisfaction and Dissatisfaction in Wine Tourism: A User-Generated Content Analysis. *Tourism and Hospitality Research*. 25(1), 120–134. <https://doi.org/10.1177/14673584231191989>

- Merli, R., Preziosi, M., Acampora, A., Lucchetti, M. C., & Ali, F. (2019). The impact of green practices in coastal tourism: An empirical investigation on an eco-labelled beach club. *International Journal of Hospitality Management*, 77, 471–482. <https://doi.org/10.1016/j.ijhm.2018.08.011>
- Mishra, A., & Gupta, A. (2019). Green hotel servicescape: attributes and unique experiences. *Current Issues in Tourism*, 22(20), 2566–2578. <https://doi.org/10.1080/13683500.2018.1502259>
- Moliner-Tena, M. Á., Monferrer-Tirado, D., Estrada-Guillén, M., & Vidal-Meliá, L. (2023). Sustainable service ecosystems from the transformative value perspective: A study in tourism destinations. *Corporate Social Responsibility and Environmental Management*, 30(1), 342–357. <https://doi.org/10.1002/CSR.2358>
- Moreno-Brito, Y. L., Wang, J., & Kim, H. S. (2024). Unraveling the Nexus Between Environmental Presence and Customer Satisfaction: A Cross-Cultural Analysis of Online Reviews of Eco-Friendly Accommodations. *International Journal of Tourism Research*, 26(5), e2780. <https://doi.org/10.1002/IJTR.2780>
- Namkung, Y., & Jang, S. C. (2007). Does Food Quality Really Matter in Restaurants? Its Impact On Customer Satisfaction and Behavioral Intentions. *Journal of Hospitality & Tourism Research*, 31(3), 387–409. <https://doi.org/10.1177/1096348007299924>
- Nguyen Huu, T., Nguyen Ngoc, H., Nguyen Dai, L., Nguyen Thi Thu, D., Ngo Truc, L., & Nguyen Trong, L. (2024). Effect of tourist satisfaction on revisit intention in Can Tho City, Vietnam. *Cogent Business & Management*, 11(1), 2322779. <https://doi.org/10.1080/23311975.2024.2322779>
- Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, 17(4), 460–469. <https://doi.org/10.1177/002224378001700405>
- Piñuel, J. L. (2002). Epistemology, methodology and content analysis techniques. *Estudios de Sociolingüística*, 3(1), 1–42.
- Pizam, A., Neumann, Y., & Reichel, A. (1978). Dimensions of tourist satisfaction with a destination area. *Annals of Tourism Research*, 5(3), 314–322. [https://doi.org/10.1016/0160-7383\(78\)90115-9](https://doi.org/10.1016/0160-7383(78)90115-9)
- Prakash, S. L., Perera, P., Newsome, D., Kusuminda, T., & Walker, O. (2019). Reasons for visitor dissatisfaction with wildlife tourism experiences at highly visited national parks in Sri Lanka. *Journal of Outdoor Recreation and Tourism*, 25, 102–112. <https://doi.org/10.1016/j.jort.2018.07.004>
- Prud'homme, B., & Raymond, L. (2013). Sustainable development practices in the hospitality industry: An empirical study of their impact on customer satisfaction and intentions. *International Journal of Hospitality Management*, 34(1), 116–126. <https://doi.org/10.1016/j.ijhm.2013.03.003>
- Qiu, N., Li, H., Pan, C., Wu, J., & Guo, J. (2024). The study on the relationship between perceived value, satisfaction, and tourist loyalty at industrial heritage sites. *Heliyon*, 10(17), e37184. <https://doi.org/10.1016/j.heliyon.2024.E37184>
- Rasoolimanesh, S. M., Ringle, C. M., Jaafar, M., & Ramayah, T. (2017). Urban vs. rural destinations: Residents' perceptions, community participation and support for tourism development. *Tourism Management*, 60, 147–158. <https://doi.org/10.1016/j.tourman.2016.11.019>
- Recognai. (2021). Model: bert-base-spanish-wwm-cased-xnli. <https://huggingface.co/Recognai/bert-base-spanish-wwm-cased-xnli>
- Rey-Moreno, M., Sánchez-Franco, M. J., De La, M., & Rey-Tienda, S. (2023). Examining transaction-specific satisfaction and trust in Airbnb and hotels. An application of BERTopic and Zero-shot text classification. *Tourism & Management Studies*, 19(2), 21–37. <https://doi.org/10.18089/tms.2023.190202>
- Rita, P., Ramos, R., Borges-Tiago, M. T., & Rodrigues, D. (2022). Impact of the rating system on sentiment and tone of voice: A Booking.com and TripAdvisor comparison study. *International Journal of Hospitality Management*, 104, 103245. <https://doi.org/10.1016/j.ijhm.2022.103245>
- Roxas, B., & Lindsay, V. (2012). Social Desirability Bias in Survey Research on Sustainable Development in Small Firms: an Exploratory Analysis of Survey Mode Effect. *Business Strategy and the Environment*, 21(4), 223–235. <https://doi.org/10.1002/BSE.730>

- Sanchez-Fernandez, R., Iniesta-Bonillo, M. A., & Cervera-Taulet, A. (2016). Environmental sustainability in the Mediterranean destinations: a latent class segmentation analysis. *Environmental Engineering & Management Journal*, 7. <https://eemj.eu/index.php/EEMJ/article/view/2981>
- Schuckert, M., Liu, X., & Law, R. (2015). A segmentation of online reviews by language groups: How English and non-English speakers rate hotels differently. *International Journal of Hospitality Management*, 48, 143–149. <https://doi.org/10.1016/J.IJHM.2014.12.007>
- Stemler, S. (2000). An overview of content analysis. *Practical Assessment, Research, and Evaluation*, 7(1):17 <https://doi.org/10.7275/Z6FM-2E34>
- Suhartanto, D. (2018). Tourist satisfaction with souvenir shopping: evidence from Indonesian domestic tourists. *Current Issues in Tourism*, 21(6), 663–679. <https://doi.org/10.1080/13683500.2016.1265487>
- Swarbrooke, J. (1999). Sustainable tourism management. In Cabi. CABI Publishing.
- Tajfel, H. (1978). Differentiation between social groups. Academic Press.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In S. Worchel & W. G. Austin (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterrey, CA: BrooksCole.
- Tian, Y. (2013). Engagement in online hotel reviews: A comparative study. *Discourse, Context & Media*, 2(4), 184–191. <https://doi.org/10.1016/J.DCM.2013.10.002>
- UCLA. (2024). FAQ: What are Pseudo R-squareds? <https://stats.oarc.ucla.edu/other/mult-pkg/faq/general/faq-what-are-pseudo-r-squareds/> Accessed 4 March 2024
- UNEP. (2005). Making Tourism More Sustainable - A Guide for Policy Makers. Books; World Tourism Organization (UNWTO). <https://doi.org/10.1811/9789284408214>
- Väisänen, H. M., Uusitalo, O., & Ryyänen, T. (2023). Towards sustainable servicescape – tourists' perspectives of accommodation service attributes. *International Journal of Hospitality Management*, 110, 103449. <https://doi.org/10.1016/J.IJHM.2023.103449>
- Viñán-Ludeña, M. S., & de Campos, L. M. (2022). Discovering a tourism destination with social media data: BERT-based sentiment analysis. *Journal of Hospitality and Tourism Technology*, 13(5), 907–921. <https://doi.org/10.1108/JHTT-09-2021-0259>
- Wang, T. L., Tran, P. T. K., & Tran, V. T. (2017). Destination perceived quality, tourist satisfaction and word-of-mouth. *Tourism Review*, 72(4), 392–410. <https://doi.org/10.1108/TR-06-2017-0103>
- Wang, X., Wang, L., Niu, T., & Song, M. (2022). An Empirical Study on the Satisfaction of Rural Leisure Tourism Tourists Based on IPA Method. *Computational Intelligence and Neuroscience*, 2022(1), 7113456. <https://doi.org/10.1155/2022/7113456>
- Williams, R. (2006). Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables. *The Stata Journal*, 6(1), 58–82. <https://doi.org/10.1177/1536867X0600600104>
- World Tourism Organization. (2019). UNWTO Tourism Definitions. UNWTO, Madrid. <https://doi.org/10.1811/9789284420858>
- World Tourism Organization. (2024). Rural tourism. <https://www.unwto.org/rural-tourism> Accessed 1 Feb 2024
- Wu, J., & Yang, T. (2023). Service attributes for sustainable rural tourism from online comments: Tourist satisfaction perspective. *Journal of Destination Marketing & Management*, 30, 100822. <https://doi.org/10.1016/J.JDMM.2023.100822>
- Wu, S., & Dredze, M. (2019). Beto, Bentz, Becas: The Surprising Cross-Lingual Effectiveness of BERT. ArXiv, 833–844. <https://doi.org/10.18653/v1/d19-1077>
- Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, 31(2), 179–188. <https://doi.org/10.1016/J.TOURMAN.2009.02.016>
- Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, 44, 120–130. <https://doi.org/10.1016/J.IJHM.2014.10.013>
- Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management*, 43, 1–12. <https://doi.org/10.1016/J.IJHM.2014.07.007>

- Yang, X., Pan, B., Evans, J. A., & Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, 46, 386–397. <https://doi.org/10.1016/J.TOURMAN.2014.07.019>
- Ye, Q., Li, H., Wang, Z., & Law, R. (2014). The Influence of Hotel Price on Perceived Service Quality and Value in E-Tourism: An Empirical Investigation Based on Online Traveler Reviews. *Journal of Hospitality and Tourism Research*, 38(1), 23–39. <https://doi.org/10.1177/1096348012442540>
- Yu, C. E., & Sun, R. (2019). The role of Instagram in the UNESCO's creative city of gastronomy: A case study of Macau. *Tourism Management*, 75, 257–268. <https://doi.org/10.1016/J.TOURMAN.2019.05.011>
- Yu, Y., Li, X., & Jai, T.-M. (2017). The impact of green experience on customer satisfaction: evidence from TripAdvisor. *International Journal of Contemporary Hospitality Management*, 29(5), 1340–1361. <https://doi.org/10.1108/IJCHM-07-2015-0371>
- Zhang, H., Liu, R., & Egger, R. (2023). Unlocking Uniqueness: Analyzing Online Reviews of Airbnb Experiences Using BERT-based Models. *Journal of Travel Research*. 63(7), 1688–1708. <https://doi.org/10.1177/00472875231197381>
- Zhang, W., & Fukami, S. (2024). Quantitative content analysis of codes of conduct for ecotourism in Japan. *Journal of Hospitality and Tourism Management*, 58, 209–213. <https://doi.org/10.1016/J.JHTM.2024.01.002>
- Zhang, Z., Zhang, Z., & Yang, Y. (2016). The power of expert identity: How website-recognized expert reviews influence travelers' online rating behavior. *Tourism Management*, 55, 15–24. <https://doi.org/10.1016/J.TOURMAN.2016.01.004>
- Zhao, Z., Cheng, M., Lv, K., & Huang, L. (2025). What matters most? A sustainable or a conventional hotel experience? The interaction effect on guest satisfaction: A mixed-methods approach. *International Journal of Hospitality Management*, 124, 103988. <https://doi.org/10.1016/J.IJHM.2024.103988>
- Zheng, W., Qiu, H., Morrison, A. M., Wei, W., & Zhang, X. (2022). Rural and Urban Land Tourism and Destination Image: A Dual-Case Study Approach Examining Energy-Saving Behavior and Loyalty. *Land*, 11(2), 146. <https://doi.org/10.3390/LAND11020146>
- Zhou, L., Ye, S., Pearce, P. L., & Wu, M. Y. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. *International Journal of Hospitality Management*, 38, 1–10. <https://doi.org/10.1016/J.IJHM.2013.12.004>