

Satisfaction and sustainability concerns in whale-watching tourism: A user-generated content model

Carmelo J. León^{a,b}, Chaitanya Suárez-Rojas^{a,b,c,*}, José Manuel Cazorla-Artiles^a, Matías M. González Hernández^{a,b}

^a Department of Applied Economic Analysis, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain

^b University Institute for Tourism and Sustainable Economic Development (TIDES), University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain

^c Centre for Environmental and Resource Economics (CERE), Umeå, Sweden

ARTICLE INFO

Keywords:

Whale-watching
Satisfaction
Sustainability concerns
UGC
TripAdvisor
Lexicon

ABSTRACT

This study examines the satisfaction and sustainability concerns of whale-watching tourists by analysing user-generated content (UGC) on social media. A satisfaction model was developed and estimated utilising an ordered probit analysis with UGC data from TripAdvisor over the last 13 years that includes a specific whale-watching lexicon. The model addresses most of the physical, human, environmental, experiential, and operational aspects of the activity, including consumers' feelings and sustainability concerns. The significance of the variables in the model was proven with the available empirical data. The findings provide a comprehensive description of the underpinnings of whale watchers' preferences and concerns. The evidence reveals compelling social trends towards higher sustainability concerns influencing satisfaction, providing valuable information for the industry and its decision-makers for understanding preferences for sustainability in whale-watching tourism.

1. Introduction

Whale-watching, involving over 15 million tourists annually, is one of the fastest-growing tourism sectors (Hoyt, 2017). However, this activity also stresses the marine resources it depends upon, threatening the conservation of around 20% of whale and dolphin species (Bedjer et al., 2022). Because of this threat, some scholars have argued that whale-watching needs to radically change if it is to avoid collapse (Suárez-Rojas et al., 2023b). Hence, understanding consumer behaviour and consumer concerns is vitally important to address sound management for sustainability. Consumers' feelings about their interactions with animals not only influence their satisfaction and experiences, but also determine how the industry responds to their preferences (Reynolds & Braithwaite, 2001). More critical is the finding that tourist satisfaction can have a positive impact on sustainable practices across the industry (Rehman et al., 2023).

This point has become pertinent because individual and collective behavioural patterns are highly influenced by social media. The recommendations and feedback that consumers post on social platforms (e.g., TripAdvisor, Booking, Facebook, Instagram, etc.) often have a direct impact on their expectations and decision-making. For instance, over

80% of consumers surf the Internet for information before booking their holidays, and more than 10\$US billion worth of tourist purchases are swayed by online reviews (Reyes-Menendez et al., 2019). These new forms of communication also connect consumers and providers more closely, making it possible for the latter to improve their service provision and marketing strategies (Kozak, 2003; Mariani & Borghi, 2022; Taecharungroj & Mathayomchan, 2019; Wang, 2016).

However, social media may also constitute a 'double-edged sword' for tourism management in terms of sustainability (Lenzi et al., 2020). That is, while it has the potential to increase consumers' awareness about the negative impacts of travelling (Zeng & Gerritsen, 2014), it might also lead to other less ethical and/or environmentally responsible behaviour, such as the proliferation of 'selfies' with wild animals (Lenzi et al., 2020). Hence, it is crucial to monitor and analyse the specific content of 'electronic word-of-mouth' (e-WOM) communication, as well as identify tendencies within social networks, especially when involving highly vulnerable species such as whales.

In recent years, academics have increasingly drawn on user-generated content (UGC) from social media platforms to evaluate consumer experiences (Prakash et al., 2019), with opinion-mining - or 'sentiment analysis' - being one of the most popular quantitative

* Corresponding author. Department of Applied Economic Analysis, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain.

E-mail address: chaitanya.suarez@ulpgc.es (C. Suárez-Rojas).

methods employed (Prakash et al., 2019). However, use of these digital technologies is still incipient in the study of wildlife-based tourism, and particularly in whale-watching (Kredens & Vogt, 2023). Additionally, no previous study has built, to our knowledge, any specific lexicon to parse tourists' sentiments (and opinions) in this field.

Thus, this study analyses tourist behaviour and satisfaction based on UGC, to respond to the following questions: 1) What is the impact of social media on whale-watching tourism? 2) Could consumers' online reviews or opinions help to understand the determinants of whale-watching satisfaction? 3) Are there any sustainability concerns that can be involved in understanding tourist satisfaction with whale-watching tourism?

To address these research questions, we developed a satisfaction model and generated a specific lexicon (based on a literature review, corpus approach, and expert consultation) to assess - by running a set of ordered probit regressions - online consumer satisfaction and revealed current social trends and concerns regarding sustainability in whale-watching.

This study contributes to current knowledge by first providing a tourist satisfaction model built on consumers' online evaluations of whale-watching tours. Second, the model generates a lexicon for analysis, which efficiently identifies the factors that influence consumer satisfaction. Finally, the evidence highlights the role played by sustainability factors in whale-watching tourist satisfaction. Following a two-part literature review on tourist satisfaction and UGC, this study details the methodology and data collection process. Last, we provide the results, discussion, and conclusions.

2. Literature review

2.1. Tourist satisfaction, whale-watching, and sustainability concerns

Tourist satisfaction, understood as the difference between *pre-travel expectations* and *post-travel experience* (Jiang et al., 2017; Pizam et al., 1978), has critical implications for sound tourism management. It provides a key indicator of how the industry can better perform to ensure the quality of tourist experiences and hospitality services. Additionally, it can help inform destination planning, competitive measures, and harmonise the industry's development with conservation of the natural environment (Bentz et al., 2016; Chen et al., 2013; Curtin, 2010; Dolnicar, 2008; Orams, 2000; Simpson, Patroni, Teo, Chan, & Newsome, 2020; Wang, 2016).

Regarding whale-watching tourism, scholars have, for example, asked consumers directly about their opinions and levels of satisfaction with the activity's various elements and services (see, e.g., Cárdenas et al., 2021; La Manna et al., 2020; Tessier-Moreau, 2022; Tkaczynski, 2021). This information can be useful for both confirming that tourist experiences are satisfactory and for implementing sustainable operational practices in accordance with tourists' concerns. Specifically, tourists have been asked about: i) the sighting itself (e.g., the number of whales observed, observation duration, their proximity, cetacean behaviour, etc.); ii) trip features (e.g., boat type, comfort, safety, crowdedness, cost, etc.); iii) the information and educational content provided on board; iv) the provider's engagement with responsible practices (e.g., their commitment, adherence to guidelines, appropriate encounter management, etc.); and v) other external factors (such as the possibility of observing other wildlife and maritime weather conditions). For instance, studies have found that boat type, cost, safety, and information provided on board might be even more important concerns to tourists than other aspects more related to sustainability management such as the proximity of the cetaceans (Bentz et al., 2016; Cárdenas et al., 2021; D'Lima et al., 2018; García-Cegarra & Pacheco, 2017; Orams, 2000).

Nonetheless, the source of tourist satisfaction in whale-watching is still strongly grounded on the provision of close encounters with whales and dolphins, with other valuable elements of the experience often

neglected (Meyer et al., 2022; Orams, 2000). This situation makes the activity potentially less responsive to responsible and/or sustainable practices (Orams, 2000; Ponnampalam, 2011). In this regard, electronic word-of-mouth might represent an opportunity to identify key trends in tourists' concerns and encourage the industry to adopt best practices due to: i) the impact it has on providers' reputation and image (Setiawan et al., 2014) and ii) its influence on tourists' vacation expectations and decision-making processes (Chen et al., 2015; Hernández-Méndez et al., 2015).

Additionally, researchers have found that through online platforms, individuals are able to express their feelings and thoughts freely and spontaneously, thus providing a more direct and straightforward approach than traditional satisfaction surveys to assess tourists' concerns (Cassar et al., 2023; Horney, 2013). In this regard, social media also has the potential to provide researchers access to a broader pool of target respondents, thereby overcoming some common limitations of satisfaction surveys, such as: data accuracy, representativeness, information reliability, and response biases such as social desirability (Araña & León, 2013; Daugherty et al., 2008; Prakash et al., 2019; Shang & Luo, 2022; Song et al., 2019). Thus, by analysing UGC, researchers might identify trends and patterns in tourists' preferences and concerns that may be difficult to detect with traditional survey methods (Daugherty et al., 2008; Lu & Stepchenkova, 2015).

2.2. UGC as a valuable source of information

Due to these advantages over traditional sources of information, social media platforms have been gaining importance in many areas of tourism research (Alegre & Garau, 2010; Prakash et al., 2019; Song et al., 2019). The breadth of information naturally flowing through digital communication platforms is particularly suitable for evaluating the performance of tourist activities and hospitality services and understanding consumer behaviour (Gössling, 2017). Additionally, improvements in computation techniques are assisting researchers in managing large amounts of data with high granularity in terms of temporal and geographical scope (Hernández et al., 2021; Mariani & Borghi, 2022; Van der Zee et al., 2020). Consequently, scholars have found that social media data is extremely valuable for tourism management (Liu et al., 2017; Tokarchuk et al., 2022).

Academics are therefore employing UGC to design management and planning policies at tourist destinations. Applications include: i) measuring destinations' carrying capacity (Tokarchuk et al., 2022); ii) revealing the distribution of tourism hot and cold spots (Van der Zee et al., 2020); iii) identifying the dimensions of tourist attractions (Taecharungroj & Mathayomchan, 2019); iv) understanding the factors that influence revisits (Hernández et al., 2021); and v) analysing hotel management strategies and financial performance (Lui et al., 2018; Xie et al., 2017). Moreover, UGC is also being employed to further understand tourist preferences, behaviour, and satisfaction. In this regard, researchers have examined the role of travel distance in consumer satisfaction (Park et al., 2019) and how the trustworthiness of online reviews affects travel planning decisions (Gurjar et al., 2022). Further insights have helped better understand guests' satisfaction with the 'hotel experience' in terms of the type of hotel, its geographical location, service performance, guest profile (Banerjee & Chua, 2016; Bi et al., 2020; Liu et al., 2017; Padma & Ahn, 2020) and, more recently, whether their environmental concerns are being addressed (Mariani & Borghi, 2022).

In the case of wildlife tourism research, Cong et al. (2014) conducted a content analysis on UGC data from TripAdvisor which revealed that proximity and interaction with animals influenced tourist satisfaction when observing giant pandas at a research base in China. Meanwhile, Prakash et al. (2019) found that online dissatisfaction expressed by tourists regarding their visit to Sri Lanka's National Parks was related to traffic congestion, overcrowding, and the level of operator professionalism (e.g., ethical conduct, norm compliance, visitor safety, etc.).

Meanwhile, [Shang and Luo's \(2022\)](#) machine learning analysis identified that, among other things, the sighting of certain popular species and operators' rule management were crucial to providing a high-quality tourist experience at Indonesia's Sacred Monkey Forest Sanctuary. [Zolfaghari and Choi \(2023\)](#) ran some topic modelling estimations to assess online satisfaction, concluding that *experience quality* attributes (e.g., wildlife sightings/interactions, and photo opportunities) were stronger 'satisfiers' than those related to *service quality* (e.g., crowdedness, tickets, and visitor information) for visitors to Canadian national parks. Despite this accumulating evidence, the application of social media data remains scarce in the field of wildlife-based tourism. Moreover, it has mainly focused on case studies, limiting the generalisability of UGC analysis for the global tourism industry. Accordingly, we aim to test whether whale-watching tourists around the world freely express any concerns about how the activity is carried out, in addition to other issues such as the impact of their emotional state on their satisfaction with the experience. In other words, the objective of this study is to assess the hypothesis that sustainability is an issue of concern among tourists and if managing the activity responsibly has an impact on customer satisfaction beyond other aspects of the trip, such as the boat features or watching more than one species during the experience.

3. Material and methods

3.1. Data

The data source employed for this paper is TripAdvisor (www.tripadvisor.com), one of the leading community-based review platforms for the travel industry ([About TripAdvisor, 2022](#)). The data are based on the extraction of 16,212 online reviews from 5,034 worldwide whale-watching activities advertised on TripAdvisor, spanning 13 years.

Data collection, pre-processing, and analysis drew on the R programming language; specifically, R Selenium (version 1.7.7) and Rvest (version 1.0.2) libraries were applied to web scraping and data download. The former provides a set of functions to interact with Selenium WebDriver that allow the user to automate web navigation. Meanwhile, the Rvest library facilitates data manipulation of HTML files.

The process consisted of searching for the key terms 'whale-watching' and 'dolphin-watching' in the TripAdvisor search engine and selecting the 'Tours & Tickets' activities, ensuring the activity was promoted for commercial purposes. We downloaded data from all the -boat, land and/or air-based - whale-watching activities available on

TripAdvisor from January 2010 to February 2023. From its founding in 2000 in the US (and two years later in Europe), TripAdvisor was launched in the Chinese market in 2009 ([Mariani & Borghi, 2022](#)). Consumer reviews started in 2006 with five million per year. This number doubled by 2007, reaching 10 million reviews ([Alaimo et al., 2020](#)). Meanwhile, in 2015, TripAdvisor began to sell and book experiences ([TripAdvisor, 2017](#)). Hence, to embrace the highest number of whale-watching destinations and obtain a richer data sample, we considered 2010 a suitable starting point for this study. [Fig. 1](#) shows the tendency of the number of reviews on TripAdvisor about whale-watching tours.

TripAdvisor limits searches to 34 web pages of activities - each with 10 items (i.e., a maximum of 340 activities). As we aimed to take most of the worldwide whale-watching activities into consideration, the destination was included as another input in the search engine to ensure that all the available activities for a destination were sampled. To achieve this, we selected the leading and most popular whale-watching spots identified by [Hoyt \(2001\)](#) and [O'Connor et al. \(2009\)](#). In those cases where we observed that the destination search did not include all the existing whale-watching tours - due to the constraint mentioned - we scaled down the destination into more concrete locations, looking for greater granularity. For instance, for the Canary Islands, we did a specific search for Tenerife; whereas, for the Azores' archipelago, we did not focus on any particular island. At the end of this process, we identified 5,034 activities. [Fig. 2](#) shows the geographical scope, coloured according to the number of TripAdvisor reviews for each country of the study and Supplementary Material I presents the list of whale-watching sites by region, country and province.

We analysed the profile of each activity, and most of the relevant characteristics - including the entire list of reviews in every language, the tour price, and its location - were parsed (see [Fig. 3](#) as an example). The list included non-whale-watching tours, as some users had mentioned one of the key terms previously given. Hence, to limit the sample to whale- and dolphin-watching activities exclusively, we removed those observations whose title or description did not contain the key patterns "whale-watching" or "dolphin-watching". Thus, from the initial 5,034 activities, we retained 2,313; totalling 96,806 reviews.

Occasionally, reviews 'shamelessly' praise - or diminish - the service or products of firms. These reviews are usually posted by fictitious profiles or one-time users who disappear after posting very few comments ([Kirilenko et al., 2019](#)). To avoid these fake reviews, we kept only those made by 'trustworthy users.' Following earlier studies, we

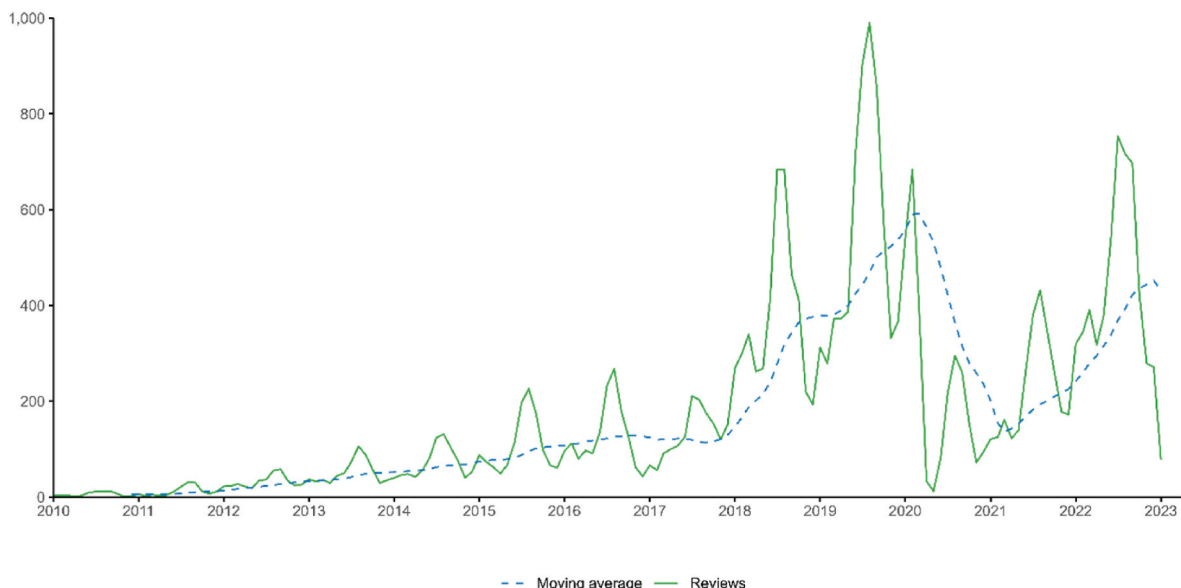


Fig. 1. Evolution (per month) of the number of reviews posted in TripAdvisor about whale-watching tours.

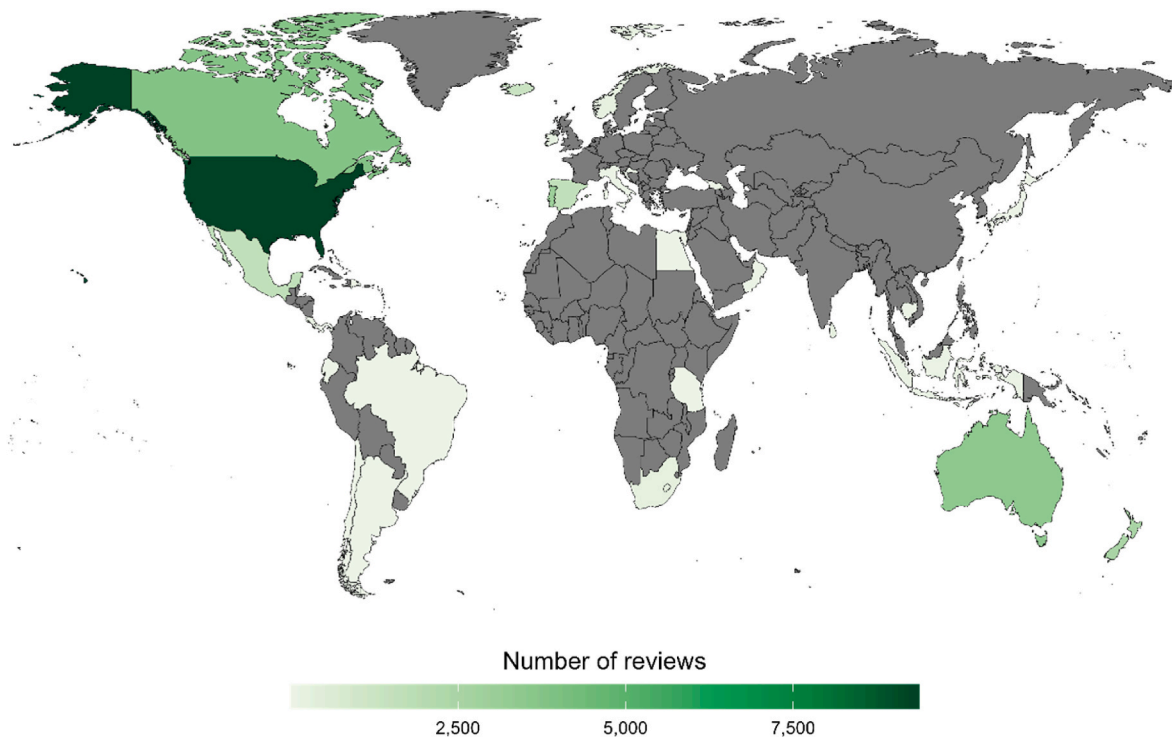


Fig. 2. Map of the whale-watching countries comprising the study. *Note:* The colours represent the density of TripAdvisor reviews for each whale-watching country of study, from those with less than 2500 reviews, such as Argentina or South Africa, to the USA, which has more than 7500 reviews. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

considered ‘trustworthy users’ individuals who have posted at least ten times on TripAdvisor (Feng et al., 2012; Kirilenko et al., 2019; Tokarchuk et al., 2022). This process reduced the sample to 33,642 reviews. Finally, we limited the sample to those written in English (Kirilenko et al., 2019; Mariani & Borghi, 2022; Xiang et al., 2017; Zhao et al., 2019) by employing the language detector R library *clld2* (version 1.2.1). We chose only English reviews for practical reasons. The existing language processing libraries are exclusively available in this language. Additionally, we observed that most reviews in our dataset were written in English, with other languages being underrepresented. Consequently, 28,870 reviews remained, constituting our final dataset. Nevertheless, for some activities the booking process was not available for web scraping. This means that the price could not be downloaded, so when the price variable is considered for analysis, the dataset is limited to 16,211 reviews.

Following data pre-processing, in which we converted the words to lowercase, removed numbers, punctuation marks and English stopwords, we obtained a word list of all the terms contained in the final dataset. This process was undertaken using *Tidytext* (version 0.3.3). This word list provided the basis for the whale-watching lexicon developed in this paper.

3.2. The lexicon

One of the hurdles when measuring online consumer preferences and satisfaction relates to how to model data from social media platform reviews. The researcher might tackle this by reading the reviews one by one, but this procedure is time-consuming. To address UGC accounting, the natural language processing (NLP) method is a valuable alternative, thanks to its ability to analyse and glean the underlying meaning of textual data automatically. In this process, *opinion-mining* is one of the most widely-used techniques (Park et al., 2020; Solangi et al., 2018) for: i) finding product or service features that have been highlighted in the reviews, and ii) deciding whether the comments are positive or negative (Ding et al., 2008).

For this reason, the lexicon-based approach is gaining momentum as a useful tool for assisting text-mining (Bagherzadeh et al., 2021). A lexicon is a textual data quantification tool that uses a dictionary to compare a word that appears in a text and obtains a sentiment (polarity) score for that word to identify groups of words linked to a specific topic (Isabelle et al., 2019).

Lexicons can be divided into binary (e.g., positive, or negative) or multiclass (e.g., emotions) (Mohammad & Turney, 2013), and can be categorised into a corpus- (to find context-specific mappings of attributes and adjectives) and dictionary-based approach (to define adjectives using their attributes) (Fei et al., 2012). To build a lexicon, there are three alternatives: i) a group of experts decides the list of terms and their associated sentiments; ii) a list of non-classified comments serves as a source for extracting a vector of terms, which is later classified by experts into different sentiments; and iii) a list of comments is both classified by experts and according to a statistical measure to associate the words with the different sentiments (Bagherzadeh et al., 2021).

The following process was conducted to build the specific lexicon of the present study. First, an in-depth literature review was conducted to visualise the elements and services of the activity under study and assess consumer satisfaction. This review enabled us to get the first categorisation of the word list retrieved from TripAdvisor. Next, following the corpus approach (Vania et al., 2014), we consulted all reviews to check the meaning of the words. Third, and based on the consensus approach (Pencle & Mălăescu, 2016), the authors organised several internal meetings to define the dimensions of the lexicon accurately.

During this process, terms with different semantic applications were exhaustively checked and discussed to avoid any possible incorrect or ambiguous categorisations and reduce and consolidate the categories (and subcategories) initially defined. For this reason, we introduced the following two rules: 1) include words in their original form, even if they share the same root as a word already included (e.g., *adventur[e]*, *-s*, *-ous*), since users sometimes employ them to talk about different topics or attributes (Bagherzadeh et al., 2021); 2) in particular cases, combine two or more words into one single item (using the ‘bag-of-words’

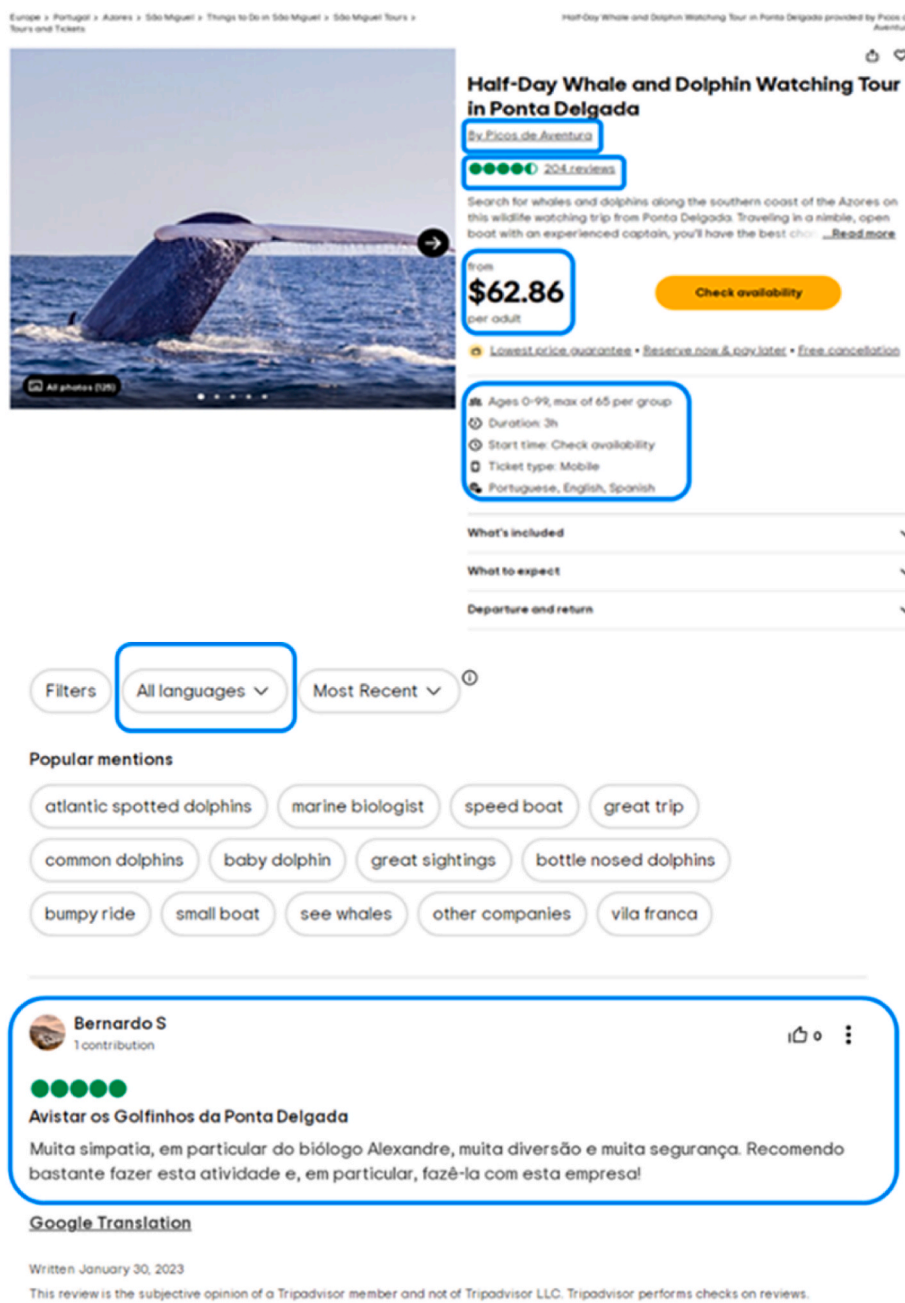


Fig. 3. Example of data from TripAdvisor.

method; e.g., marine-environment, native-environment, natural-environment, etc.), since compound words change their meaning depending on what words they are combined with (Loughran & McDonald, 2011).

3.3. The satisfaction model

3.3.1. Variables

The covariates employed in this study are divided into two groups: 1) control variables and 2) whale-watching variables obtained from the lexicon. The former included: i) *observed average rating*, i.e., the observed rating of the whale-watching activity when the user posts the review; ii) *reviewer experience*, i.e., the number of reviews made by the user - this variable is transformed using logarithms; iii) *length of the review* - this is the number of words in the review, but it enters into the regression using logarithms; iv) *review polarity* or ‘sentiment score’, which is represented

by a continuous variable ranging from -1 (extremely negative) to $+1$ (extremely positive) to show how negative or positive the emotion and content of the review is (Mariani & Borghi, 2022). To obtain this value, we employed the VADER (Valence Aware Dictionary and sEntiment Reasoner) lexicon (Hutto & Gilbert, 2014) through the R library vader (version 0.2.1). This lexicon is a rule-based sentiment analysis explicitly developed in the context of social media. Additionally, the covariate trend was included as a control variable. It reflects the temporal tendency of the online ratings, i.e., it takes value 1 for the first month of the dataset and has a monthly rate of growth. Nonetheless, the key variables in this study are text variables obtained from the lexicon. The role of each category was analysed through the percentage of words belonging to the category over the total length of the review.

Regarding the endogenous variable, we chose the *online user satisfaction rating*. This variable follows a Likert scale with discrete values from 1 to 5, where 1 means ‘terrible experience’ and 5 means ‘excellent

experience'. For this reason, the estimation procedure applied in this paper is an ordered probit model. This model deals with endogenous ordinal variables, i.e., when the variable is categorical and ordered, as is the case with online satisfaction ratings on TripAdvisor.

3.3.2. Model specification

This study carried out a sentiment analysis, as it is a well-known method for revealing relevant topics among the reviews, or even users' attitudes in terms of positive or negative emotions (Mäntylä et al., 2018). Thus, we ran two regressions for estimating tourist satisfaction: one with the four categories and the other with a higher granularity level by utilising the 24 subcategories defined in the lexicon. In other words, the general model specification depends on the attributes level (categories and/or subcategories defined in the whale-watching lexicon). To this end, we considered the whole sample, which was reduced to 22,151 observations due to the missing values.

We also identified any possible divergences in whale-watching attribute valuations and their effects on satisfaction. To achieve this, we ran various segmented models, one in which the price was assumed as a proxy of quality and another in which we distinguished between types of destinations according to their 'specialist' or 'generalist' characteristics. Both variables (price and location) were retrieved from TripAdvisor whale-watching announcements. The destination segmentation was developed in a three-step process: i) conceptualisation (Bentz et al., 2016; Bryan, 1977), ii) literature review (International Whaling Commission, online; O'Connor et al., 2009), and iii) expert validation.

Regarding the former segmentation, activities with high/low prices are perceived as being of high/low quality (Campo & Yagüe, 2009). The review's dataset was disentangled into three different quantiles (i.e., terciles) according to the variable 'price'. However, since not all activities are sold on TripAdvisor, the dataset was reduced to 16,212 reviews. Concerning the destination segmentation, the criterion used to categorise destinations as 'generalist' or 'specialist' relies on the literature signalling out tourists' primary motivation for visiting the destination (if whale-watching is the main motivation, then it is classified as a 'specialist' destination; otherwise, it is considered a 'generalist' destination). In this case, the data set was reduced to 14,038 due to the difficulty of disentangling whether the 'generalist' or a 'specialist' distinction for some destinations.

The general model specification at the categories level is specified as follows:

$$Valence_i^* = \beta_0 + \sum_{j=1}^J \beta_j Z_j + \sum_{h=1}^H \beta_h C_h + \varepsilon_i$$

and at the subcategories level it is defined as:

$$Valence_i^* = \beta_0 + \sum_{j=1}^J \beta_j Z_j + \sum_{k=1}^K \beta_k S_k + \varepsilon_i$$

where Z_j denotes the control variables, C_h denotes the categories of the whale-watching lexicon and S_k is referred to subcategories. The error term for each observation is denoted as ε_i . $Valence_i^*$ is a latent variable in the ordered probit model, which allows us to obtain the predicted online rating through the following rule:

$$Online\ rating_i = r \text{ if } I_{r-1} < Valence_i^* \leq I_r, \text{ for } r = 1, \dots, R$$

where R denotes the number of online rating alternatives in TripAdvisor, i.e., 5 levels of alternatives to rate the experience - from excellent to terrible. The intercepts are denoted as I_r and we assume the lowest is $I_{r-1} = -\infty$ and the highest is $I_R = \infty$. The other R-1 intercepts of the order probit model are obtained in the estimation procedure together with the model regression parameters (Cameron & Trivedi, 2005).

4. Results

4.1. Whale-watching lexicon

The lexicon of the present study was built to identify the salient aspects of tourist satisfaction, which is critical for managing the activity responsibly and contributing to wildlife conservation. One of the main reasons we developed a specific lexicon for whale-watching tourism was to appropriately identify the semantic relationships between the expressions of sentiment and our specific research field. The existing lexicons have been designed for other contexts, such as analysing guests' feedback in the hotel sector (Bagherzadeh et al., 2021), assessing tourists' sentiments towards visiting specific destinations (Liu et al., 2019; Zhang et al., 2022) and their last-chance motivations (Abrahams et al., 2022), or measuring corporate social responsibility disclosures of firms (Pencle & Mălăescu, 2016). Hence, we did not consider them suitable enough for obtaining accurate evidence on the determinants of whale-watching tourist satisfaction.

Following a meticulous process, the new lexicon now composed a total of 1,554 words distributed into four categories and 24

Table 1 Description of the categories included in the whale-watching lexicon.

Category	Description	Subcategories
The trip experience	This category includes most of the physical elements, human resources and organisational characteristics that define the whale-watching activity and trip management.	1) booking/planning 2) ticket price & cancellation 3) boat features 4) onboard crew, staff 5) onboard equipment 6) onboard health & safety 7) onboard food & beverages 8) complementary services, activities & facilities
The watching experience	This comprises most of the aspects involved in the whale observation experience, i.e., the cetacean encounter from both the animal side (e.g., species observed, their behaviour, etc.) and the operator side (e.g., the navigation and approach to the sighting points) and the way the tour is carried out - passive or active. This category also includes other issues related to the observation that are outside operators' control (e.g., weather conditions, seasickness, and/or the opportunity to take photos, etc.).	1) cetaceans 2) cetacean behaviour 3) other wildlife & natural resources 4) photo opportunity 5) absence of navigation & manoeuvring 6) climate & sea conditions 7) seasickness 8) passive observation 9) active observation
The customer	This category includes the entries describing individual 'internal' features (i.e., socio-demographic and psychological characteristics) and outcomes (i.e., their lived experience, revisit intentions or willingness to recommend) to obtain an overall picture of who they are and how they perceive and value the activity.	1) customer profile 2) satisfaction & fidelity 3) experience valuation & emotions 4) crowding perception
Sustainability concerns	How operators (responsibly) manage the tour, provide (attractive, educational) information and utilise innovative tools for interpretation are all critical for tourists' learning experience of marine protection and the sustainable development of whale-watching. With this in mind, the present subcategory was built.	1) responsible behaviour 2) protection & conservation 3) educational component 4) absence of technology-based interpretation & tracking

subcategories, describing most of the elements and services of a whale-watching tour, as summarised in Table 1: i) the trip experience (including the human and physical capital); ii) the watching experience; iii) the customer (including tourists' perceived value of the experience, and their emotional reactions); and iv) sustainability concerns. A broader description of each category and subcategory is presented in Supplementary Material II and the definitive word list (lexicon) is in Supplementary Material III.

Notably, thanks to implementation of the corpus-based approach, we could, for example, categorise 'swimming' in the *Animal behaviour* subcategory instead of *Active observation* (swimming-with activity), since around 85% of the time, users mentioned it to refer to the animal action. For instance:

"... there was a pod of 10–20 swimming around us ..."

"... dolphins were seen swimming ..."

"... three whales swimming together ..."

Likewise, we also found that up to 89% of the time, users are concerned about "crowdedness" (*crowd-ed, -ing, -s; overcrowded*) in terms of how they perceive vessel congestion. Some comments retrieved from the TripAdvisor reviews that support our findings include:

"The boat wasn't too crowded ..."

"Our boat was so crowded ..."

"Boat was pretty crowded, poor viewing availability ..."

4.2. Online consumer satisfaction

4.2.1. General analysis

The results of the ordered probit regressions are shown in Tables 2 and 3. These estimations were run with the aim of understanding consumer satisfaction, with special emphasis on the impact of their concerns on the enjoyment of a responsible, sustainable activity, in addition to the elements defining the tour experience.

As depicted in Table 2, the main categories of the lexicon included in the general satisfaction model - *Sustainability concerns*, *The customer*, *Watching experience*, *The trip* - are significant, and have a positive effect on the user rating. Notably, *Sustainability concerns* has the highest impact on the consumer experience valuation. This means that the whale-watching sustainability discourse has 'on average' a strong positive relationship with customer satisfaction.

Table 3 presents the ordered probit analysis with the subcategories of

Table 2

Ordered probit model using whale-watching categories.

Explicative variables	Coefficient		Std. Err.	Z-score
Categories of the lexicon				
Sustainability concerns	7.5610	***	0.6267	12.07
The customer	6.9877	***	0.3911	17.86
Watching experience	2.9880	***	0.2174	13.74
The trip	1.8226	***	0.2786	6.54
Control variables				
Observed average rating	0.7177	***	0.0281	25.52
Log reviewer experience	-0.0546	***	0.0178	-6.47
Log review length	-0.1150	***	0.0178	-6.47
Review polarity	1.4159	***	0.0260	54.50
Trend	0.0013	***	0.0003	4.27
Intercept -1	1.9523		0.1690	
Intercept-2	2.4442		0.1681	
Intercept-3	3.0211		0.1680	
Intercept-4	3.8234		0.1687	

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.001$.

Observations = 28,400; Pseudo $R^2 = 0.1839$; LR $\chi^2 = 6,687.57$; Log Likelihood = -14,837.75.

Table 3

Ordered probit model using whale-watching subcategories.

Explicative variables	Coefficient		Std. Err.	Z-score
Sustainability concerns				
Responsible behaviour	4.8964	***	1.8358	2.67
Protection & conservation	19.5751	***	7.7186	2.73
Educational component	6.9952	***	0.7256	9.64
Absence of technology-based interpretation & tracking	-10.9853	***	3.1440	-3.49
The customer				
Customer profile	2.4778	***	0.9150	2.71
Satisfaction & fidelity	16.4355	***	1.4410	11.41
Experience valuation & emotions	7.0607	***	0.4748	14.87
Crowding perception	-14.4908	***	3.3510	-4.31
Watching experience				
Cetaceans	5.0591	***	0.4410	11.25
Cetacean behaviour	11.9116	***	1.1533	10.33
Other wildlife & natural resources	4.0280	***	0.5109	7.88
Photo opportunity	5.7968	***	1.7056	3.40
Absence of navigation & manoeuvring	-1.8710	***	0.6402	-2.92
Seasickness	-3.0191	*	1.5432	-1.96
Passive observation	-3.0600	***	0.6954	-4.40
The trip				
Booking planning	3.3633	***	0.5876	5.72
Ticket price & cancellation	-14.2982	***	1.0258	-13.94
Boat features	1.7904	***	0.6403	2.80
Onboard crew-staff	3.8169	***	0.5475	6.97
Onboard food & beverages	2.0647	***	0.7746	2.67
Complementary services-activities-facilities	4.3420	***	1.4314	3.03
Control variables				
Observed average rating	0.6591	***	0.0285	23.12
Log reviewer experience	-0.0474	***	0.0077	-6.20
Log review length	-0.1079	***	0.0188	-5.75
Review polarity	1.3275	***	0.0267	49.67
Trend	0.0010	***	0.0003	3.28
Intercept-1	1.4710		0.1728	
Intercept-2	1.9766		0.1718	
Intercept-3	2.5731		0.1716	
Intercept-4	3.3994		0.1721	

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.001$.

Observations = 28,400; Pseudo $R^2 = 0.2012$; LR $\chi^2 = 7,317.26$; LogLikelihood = -14,522.90.

the lexicon aimed at providing a richer understanding of the underpinnings of consumer satisfaction. In this case, only the variables with a significant effect were included (i.e., *Climate and sea conditions*, *Onboard health & safety*, *Onboard equipment* and *Active observation* are omitted).

All the significant subcategories have a positive effect on whale-watching tourist satisfaction, except for, for example, *Seasickness* and *Crowding perception*, which significantly increases user complaints about the tour, thereby leading to an unsatisfactory evaluation of the experience. Satisfaction with the activity is also negatively affected by those aspects involving management of the tour ticket, i.e., the price, rescheduling and/or cancellation (*Ticket price & cancellation*).

Similarly, *Absence of navigation & manoeuvring* and *Passive observation* also negatively affected overall satisfaction with the tour. *Absence of navigation & manoeuvring* refers to the cruising across and manoeuvring to approach cetaceans, the speed, time and distance invested for the encounter, and the bounciness of the sailing so that a 'bad' experience during the sailing could explain the negative impact this subcategory has on the satisfaction. The same occurs with *Passive observation*, defined as whale watching from a vessel - or land-based - and the closeness of the observation (an in-depth description of each subcategory is found in Supplementary Material II).

Likewise, results show a negative relationship between *Absence of technology-based interpretation & tracking* and satisfaction. In other

words, the lack of good audio and video systems and/or the absence of other disposals, such as hydrophones or other technology to track the animals, causes customers to complain about the whale-watching experience. For instance:

“They need to spend some money and get individual audio sets for people so they can hear what’s being said”.

“Trip would have been enhanced by a simple audio system with an explanation of what we were seeing and why the captain was moving the boat”.

“Since they don’t use technology to track the whales it does take time to spot them”.

“The boat is not equipped with technology to locate whales, so a bit hit or miss”.

Finally, in both estimations, the control variables show the expected sign. The *observed average rating* of the activity (i.e., the rating of the tour that users visualise when they post their review) is significant and positive, revealing a positive relationship between the average rating of the activity by the user and its own valuation on TripAdvisor. The

reviewer experience is significant and negative, i.e., those users with more experience on TripAdvisor seem to be stricter in their valuations. The *length of the review* has a significant and negative relationship with the rating. This means that users tend to include more details in their complaints than in their positive reviews. The *polarity* shows a significant and positive effect on the rating; users give a higher rating when they are in a positive emotional state (Isen, 1987). The effect of the variable ‘trend’ is also significant and positive, i.e., whale-watching activities receive high valuations.

4.2.2. Segmented analysis

Segmentation analysis helps researchers to identify potential differences between groups of tourists that share different patterns of preferences for whale-watching activities based on their specific interests (Malcolm & Duffus, 2008; Suárez-Rojas et al., 2023a; Tkaczynski & Rundle-Thiele, 2019). Here, we conduct two alternative segmentations that illustrate the potential of the approach and shed light on the specific preferences of different types of consumers. The first segmentation is based on the price paid for the whale-watching experience, and the second segmentation is based on the kind of tourist destination in which

Table 4
Ordered Probit for the segmented models.

Explicative variables	Price Segmentation					Destination Segmentation			
	Low price		Aver price		High price	Generalist destination		Specialist destination	
	(T1)		(T2)		(T3)				
Sustainability concerns									
Responsible behaviour	9.0663	**	9.1335	**	4.8818	10.0015	**	8.0007	***
Protection & conservation	9.9257		27.2693	**	21.0863	60.3907	**	15.7317	
Educational component	6.9415	***	6.9164	***	7.2514	9.3611	***	6.1869	***
Absence of technology-based interpretation & tracking	-11.9949	*	-2.2659		-25.2738	-6.0790	**	-8.3858	
The customer									
Customer profile	0.0429		-0.3212		3.2856	6.3463	***	0.6964	
Satisfaction & fidelity	15.3390	***	17.0226	***	17.3061	17.5447	***	15.0991	***
Experience valuation & emotions	6.5002	***	6.5874	***	6.0014	9.5998	***	5.1534	***
Crowding perception	-12.3550	**	-19.5777	**	-10.6064	-12.5187		-8.1794	
The watching experience									
Cetaceans	3.8155	***	5.6691	***	5.5287	6.1805	***	4.7396	***
Cetacean behaviour	11.7337	***	12.4215	***	18.6290	11.8978	***	7.5474	***
Other wildlife & natural resources	4.5995	***	4.2153	***	0.9893	5.6892	***	1.9684	**
Photo opportunity	6.2315		1.8946		9.0426	13.2152	**	4.3643	
Absence of navigation & manoeuvring	-1.1356		-0.7690		-2.0984	1.8835		-3.2866	***
Climate & sea conditions	-5.5852	***	4.8847	**	-0.8241	1.5419		0.3850	
Seasickness	-0.3715		-4.7413	**	5.2465	-2.9213		-5.7574	**
Passive observation	-4.8145	***	-2.7854	**	-4.9928	-2.9379	**	-3.2877	***
Active observation	10.8426		-7.7912	*	10.3917	-2.7968		-3.1468	
The trip									
Booking planning	2.6665	**	3.7036	***	8.8619	3.9247	**	3.6636	***
Ticket price & cancellation	-8.0465	***	-12.8724	***	-26.4019	-5.9738	**	-15.4333	***
Boat features	2.9709	**	3.8769	***	1.3101	2.8947	*	2.4707	**
Onboard crew-staff	4.0254	***	4.8074	***	4.6017	5.0489	***	5.4687	***
Onboard equipment	4.1218		0.5232		-2.7596	5.5641		1.4150	
Onboard health & safety	3.5051		7.7595	*	0.6395	11.5964		3.4410	
Onboard food & beverages	0.5929		-0.0199		3.4791	1.1805		1.7164	
Complementary services-activities-facilities	4.7885		0.4624		9.3987	6.9369	*	5.9373	**
Control variables									
Observed average rating	0.5397	***	0.7428	***	0.8591	0.7083	***	0.7323	***
Log reviewer experience	-0.0718	***	-0.0214		-0.0709	-0.0582	***	-0.0357	***
Log review length	-0.1706	***	-0.1088	***	-0.1086	-0.0700	*	-0.1288	***
Review polarity	1.3408	***	1.1694	***	1.3166	1.5225	***	1.1900	***
Trend	0.0011		0.0017	***	-0.0001	-0.0003		0.0007	
Intercept-1	0.4556		1.9689		2.2145	2.1463		1.5345	
Intercept-2	0.9325		2.4708		2.6763	2.6306		2.0409	
Intercept-3	1.6068		3.1181		3.2741	3.3115		2.6033	
Intercept-4	2.5313		3.9732		4.0086	4.1610		3.4440	
Observations	5,477		7,147		3,588	4,284		9,754	

Note: T1: >0 to <66 USD; T2: ≥66 to <110 USD; T3: ≥110 USD.

*p < 0.1; **p < 0.05; ***p < 0.001.

online reviews were collected, i.e. 'generalist' if the destination receives tourists with various interests beyond whale-watching, and 'specialist' if the destination focuses specifically on whale-watching tourists.

Considering the price, the various price bands for tourist services reflect different consumer preferences and supply conditions that may translate to different experiences and products providing satisfaction to tourists. Thus, there can be different relationships between sustainability preferences and satisfaction with whale-watching tourism that may be uncovered by analysing different price segments.

For the price-based segmented satisfaction model, researchers ran three different regressions in which all control variables and sub-categories of the whale-watching lexicon were included. The price that tourists paid for the whale-watching tour was utilised to segment the sample into terciles, i.e., low (>0 to < 66 \$US), average (≥ 66 to < 110 \$US) and high price (≥ 110 \$US). This segmentation was undertaken to address the challenge of understanding the relationship between *price*, *sustainability concerns* and *satisfaction*, and identify any possible differences in consumer preferences. As Pomeroy et al. (2011) pointed out, premium prices are related to consumers' clear understanding of the value of the tourism product, with those consumers possibly being linked to sustainable tourism niches.

Table 4 shows that all control variables maintain the expected sign, coinciding with results from the general model of Tables 2 and 3. Note that, in this case, the number of observations for each group is different (the number of tours varies between price terciles and destinations, and the number of reviews varies between tours).

The results of the price segmentation show that the customers who paid most for the whale-watching activity are most sensitive to the subcategory *Ticket price & cancellation*, i.e., the aspect most negatively affecting their experience evaluation. Similarly, having a poor communication system or a lack of technology for activity interpretation (*Absence of technology-based interpretation & tracking*) mainly affects those paying a higher price. On the contrary, these customers' evaluations of the experience are not influenced by onboard crowdedness (*Crowding perception*), nor the *Responsible behaviour* of the staff. This result might be due to users assuming that the higher the price they pay for tickets, the more the firms should guarantee them a personalised, unique and responsibly managed tour, i.e., a genuinely sustainable experience.

Additionally, *Boat features* have a greater positive effect on the satisfaction levels of consumers who pay the average tour prices. Notably, amongst segments, the lower the price customers pay for the whale-watching tour, the lesser the impact of the *Educational component* and *Cetacean behaviour* on their satisfaction with the tour.

Table 4 also shows the results of the destination segmentation, indicating that in 'generalist' destinations, tourists' concerns about sustainable practices, i.e. *Responsible behaviour* and *Protection & conservation*, have a stronger impact on their satisfaction with the experience. Likewise, the emotions that the experience aroused in tourists (*Experience valuation & emotions*) and the opportunity to take photos (*Photo opportunity*) also have a higher effect on their satisfaction with 'generalist' destinations compared to 'specialist'. Further, the willingness to repeat the activity (*Satisfaction & fidelity*) in the former destinations also has a higher significant positive relationship with the satisfaction level. The higher impact on satisfaction for the 'generalist' destinations is also noteworthy for those attributes related to the watching experience, e.g., *Cetaceans* and *Cetacean behaviour*.

5. Discussion and conclusions

This research quantitatively evaluates UGC to identify how *the customer* (interests and sentiments), *Sustainability concerns*, and *The trip and Watching experience* influence individuals' satisfaction with the tour experience. The present study is original in assessing the opinions of whale-watching consumers and delivering the first specific UGC lexicon in this field. Evidence from this study will assist researchers and

practitioners in easily monitoring social media platforms and identifying in real-time the tendencies that influence tourist demands and satisfaction. Further, this work differs from previous research in whale-watching tourism (see e.g., Bentz et al., 2016; Cárdenas et al., 2021; Lück & Porter, 2019; Suárez-Rojas et al., 2023a; Tessier-Moreau, 2022; Tkaczynski, 2021) by contributing a novel approach to investigating the role of sustainability concerns in tourist satisfaction, while focusing on the worldwide tourist market over an extended period of time.

Survey-based studies have managed to focus only on partial aspects of the determinants of customer satisfaction. Their evidence has also been limited in terms of geographical scope and time dimension. For instance, consumers' opinions regarding the *Watching experience*, the *Educational component*, and the *Weather and maritime conditions* have been analysed in Bahía de Banderas, Mexico (Cornejo-Ortega et al., 2018), Contadora Island, Panama (Cárdenas et al., 2021), and in the Macaronesia Region (Suárez-Rojas et al., 2023a). However, these studies have neglected other valuable elements of *The trip*, such as the *Booking/planning* or *Boat features* (Cornejo-Ortega et al., 2018) or related to *Sustainability concerns* - e.g., (*operators'*) *Responsible behaviour* - (Cárdenas et al., 2021; Suárez-Rojas et al., 2023a). The latter aspects have been evaluated by other researchers but in other whale-watching sites and at different moments, such as in the Azores (Bentz et al., 2016), New Zealand (Lück & Porter, 2019), Canada (Tessier-Moreau, 2022) and in some Mediterranean destinations (La Manna et al., 2020; Tepsich et al., 2020). Thus, survey-based studies produce results that are case-specific and limit the generalisation of the evidence about the determinants of tourists' satisfaction in whale-watching tourism.

The evidence from this research shows that on social media, consumers spontaneously and voluntarily voice their feelings and impressions about cetacean encounters, the navigation experience and the tour services, as well as their concerns about the sector's sustainability. These expressions are robustly controlled by the price paid for the tour. The results demonstrate that research utilising UGC is valuable for tourist satisfaction assessment and understanding how customers demand authentic responsible and sustainable practices. Therefore, they may help the industry to demonstrate sounder management that focuses on a reconciliation of current customer demands and sustainable practices.

5.1. Theoretical contributions

One of the main theoretical contributions of this study relates to the development of a model for estimating online satisfaction, supported by a specific lexicon for the field of whale-watching tourism that accounts for - in addition to the tour's attributes - consumers' *sustainability concerns*.

In particular, the new lexicon goes a step further in applying qualitative and quantitative content analysis techniques such as *word frequency counting* and/or *machine learning*. For instance, in contrast to the latter, the lexicon-based approach is presented as a more straightforward tool that does not require a large body of data or long-term training (Liu et al., 2019; Mukhtar et al., 2018). It also overcomes the limitations of using generic lexicons that are often responsible for miscoded data and inconsistent results (Bagherzadeh et al., 2021). Likewise, it allows the identification of areas of concern, trends and behavioural patterns that have not been identified through traditional surveys (Lee & Park, 2023; Lukyanenko et al., 2016; Wang et al., 2011; Ziegler et al., 2018).

In a major advance, we found a gap between researchers' beliefs about the elements influencing tourist satisfaction (traditionally asked about in satisfaction surveys) and the grounding factors that are important to consumers (found in this study by analysing UGC). While academics have commonly asked individuals about overcrowding at whale-watching sites, consumers 'talk from the heart' about vessel crowdedness as a factor of (dis)satisfaction. Further, this is the first time that a comprehensive depiction of the attributes impacting customer satisfaction has been obtained on a worldwide scale and over such a long time frame. This evidence has enabled us to address questions that

constrained survey-based studies have been unable to measure.

The general and segmented satisfaction models run in this research have led to a greater understanding of consumer behaviour. Concerning the general estimations, the control variables show that: i) there exists a positive relationship between the average satisfaction rate and giving on-line evaluations, ii) users write longer reviews when they want to express dissatisfaction, iii) they give higher ratings when they are in a positive emotional state, and iv) the more active they are on TripAdvisor, the stricter they are with their evaluations. Regarding the segmented model estimations, in which the dependent variable referred to the *price* as a proxy of *activity quality*, the coefficients of ordered probit regressions also resulted in the expected sign, confirming the predictive power of the variables (i.e., lexicon subcategories). Thus, this new lexicon has been robustly proven to be well-developed and has the potential to be replicable. That is, it could be easily adapted to other related sectors and topics of concern in the field of wildlife tourism activities, thereby increasing the usefulness of the lexicon.

5.2. Practical implications

Regarding implications, the results are consistent with earlier evidence about the activity's attributes influencing consumer evaluations of the whale-watching experience. For instance, we found that *Seasickness* (Chuang et al., 2020; Kessler et al., 2014; Orams, 2000) and *Ticket price & cancellation* (Chen et al., 2013; Fuchs & Weiermair, 2003; Tkaczynski, 2021) negatively impact the enjoyment of the activity and the satisfaction rating. According to some researchers, rescheduling or cancelling the trip due to adverse environmental circumstances constrain the perceived (monetary) value of the activity (Chen et al., 2013; Fuchs & Weiermair, 2003; Tkaczynski, 2021). Similarly, *Absence of navigation & manoeuvring* and *Passive observation* also negatively affect customer satisfaction with the experience. In this regard, Bentz et al. (2016) suggested that high overall satisfaction could be due to favourable navigation and sighting conditions. Hence, the contrary scenario, i.e., the absence of favourable *whale-watching encounter management* and *observation conditions*, makes users more likely to complain.

Crowding perception also negatively impacts the 'activity enjoyment' and satisfaction rating. In other words, vessel congestion decreases tourist satisfaction with the experience. To our knowledge, only Torres-Matovelle and Molina-Molina (2019) also analysed the effect of on-board congestion on *experience satisfaction* – but by employing a survey-based method. Traditionally, scholars have preferably focused on crowding perception in the sighting area (Bentz et al., 2015; Torres-Matovelle & Molina-Molina, 2019; Avila-Foucat et al., 2013).

Interestingly, customers often complain about the lack of investment in technology (*Absence of technology-based interpretation & tracking*). Separately, previous literature has highlighted that technological innovation adds value, with customers willing to pay for that additional element – e.g., hydrophones or submarine cameras accompanied by screens on board – in order to feel closer to the animals and experience an (emotional) connection with wildlife (Finkler et al., 2019; Jacobs & Harms, 2014; Lopez & Pearson, 2017; Shapiro, 2006). Analysis of UGC in this research shows that the speed with which these technological advances have been rolled out have led consumers to believe that this should be an essential service on the tour, and therefore it should be expected as a guarantee of quality.

Our price-based segmented satisfaction analysis shows that whale-watching consumers are price sensitive. This finding is consistent with the consumer satisfaction theory, signalling a negative relationship between price and satisfaction, in contrast to the relation between price and quality (positive) (Athanasopoulos, 2000; Campo & Yagüe, 2008; Dodds et al., 1991). Further, we identified different groups of consumers for different price ranges, with some customers who might be characterized as 'generalists', and others as 'specialists' (Malcolm & Duffus, 2008; Suárez-Rojas et al., 2023a). For instance, while customers who pay higher prices tend to be more concerned about the use of

technological innovation in interpreting a particular activity (significant, negative effect of *Absence of technology-based interpretation & tracking*) and have a genuine interest in marine wildlife (e.g., *Cetacean behaviour* subcategory), those paying less are more interested in a comfortable and pleasant onboard experience (i.e., *Boat features*).

Concerning destination-based segmentation, we found that the satisfaction of those tourists who carry out the activity in 'generalist' destinations is more positively influenced by the protection and conservation efforts, the emotions aroused by the experience, and the opportunity to take photos (i.e., *Protection & conservation; Experience valuation & emotions; Photo opportunity*). This evidence can be explained because whale-watching at 'generalist' destinations is preferably carried out by novel tourists who see for the first time the 'spectacle' of whales in nature (Suárez-Rojas et al., 2023a).

This study constitutes a salient source of information for whale-watching providers by contributing a tool for the analysis of the determinants of tourist satisfaction based on UGC data. The differences identified between the various groups of consumers confirm the importance of the expected relationship between price and sustainability concerns and how it influences satisfaction. In this regard, there is scope for addressing sustainability concerns to reduce tourists' complaints, thereby improving the projected image, reputation, and profitability of the activities (Salah et al., 2023; Setiawan et al., 2014). At the same time, operators may invest in innovative tools to both meet expectations and quality demands for those paying more, and help develop greater emotional connections with cetaceans for those paying less. This kind of investment may favour a further harmonisation of the human-animal relationship (Kredens & Vogt, 2023; Suárez-Rojas et al., 2023a).

5.3. Limitations and future research

The present work is not exempt from some limitations. From a methodological point of view, the UGC nature of the data makes it difficult to consider some important control variables in the satisfaction model, such as users' age, gender, or place of residence, that would provide more accurate insights. In this study, as our database contained insufficient information –there was too much missing data concerning these covariates—we decided not to include them in the model estimations. Second, as we were looking for the effect of tour price on evaluations (segmented model), our sample only comprised activities for which tickets are sold on TripAdvisor. This means that our study is limited to those particular commercial activities, and so other firms that are promoted on the social media platform – but not for economic purposes (i.e., they do not sell tours) – are thereby excluded. Similarly, our study only considers reviews in English; including entries written in other languages might bring more valuable insights from larger datasets. In addition to these limitations, although the specific lexicon for whale-watching provided more accurate information compared to the use of pre-existing generic lexicons, it does not enable us to analyse consumers' previous knowledge about whale-watching, limiting the findings to the sense of words from the text itself (Cassar et al., 2023).

From a practical point of view, although we consider the use of TripAdvisor reviews to be a good proxy for understanding the underpinnings of whale-watching tourist satisfaction, caution needs to be considered when it comes to the managerial implications. TripAdvisor reviews have a positive impact on perceived credibility and usefulness for other users (Filieri et al., 2021). However, it has also been found that customers' trustworthiness on this platform has fallen due to fake and promotional reviews posted on the platform (Filieri et al., 2021). In this regard, future research should consider reviewer trust or helpfulness in categorising the reviews since not all have the same impact on potential visitors (Toral, Martínez-Torres, & Gonzalez-Rodríguez, 2018).

Information about tourists' behaviour and decision making from social media is considered a reliable and representative alternative to other sources, overcoming some of the hurdles in the statistical

measurement of the field based on conventional surveys (Ma & Kirilenko, 2020). However, it still presents a significant limitation related to obtaining a complete profile of the ‘authors’ of the reviews and whether they may represent the tourist segment under study (Van der Zee & Bertocchi, 2018). In addition, it is still unclear if reviews reflect tourist behaviour or limit themselves to a biased snapshot of the tourist experience (Cassar et al., 2023). The same authors (Cassar et al., 2023) concluded that further research could be devoted to triangulating UGC from TripAdvisor with other big data sources and empirical studies about consumer behaviour.

The approach utilised in this paper also limits the ability to find coherent or conflicting preferences for whale-watching activities since it is fundamentally based on word counts for specific attributes. An alternative approach that addresses how words and phrases are combined could lead to a more thorough characterisation of the underlying or disaggregated preferences of tourists. In addition, further research might seek to understand, for instance, whether users’ sociodemographic profile has an impact on whale-watching ratings and the evaluation of the various elements (categories and subcategories) defining the activity.

Accounting for sociodemographic variables could also be helpful for consumer clustering and identifying latent patterns influencing their satisfaction with the experience (Garner & Kim, 2022). This line of research would enable researchers and practitioners to respond to tailored and personalised consumer demands while improving service quality and boosting the position of whale-watching firms in the competitive tourism market (Ali et al., 2021). Further, clustering customers by the whale-watching destination will also be valuable for managing the destination image (Teles et al., 2024). Finally, future research might also focus on tracking other social media platforms, such as Google Reviews, to obtain more comprehensive and updated market information. According to Filieri et al. (2021), Google Reviews is increasingly growing, gaining momentum over TripAdvisor and even positioning itself as the favourite platform for mobile searches.

Funding details

This work has been carried out with funding by MARCET II Project (MAC/4.6c/392), approved in the second Call of the 2014-2020 Interreg MAC Program with ERDF-EU Funds and the Project TED2021-131848B-I00 from Ministry of Science and Innovation, State Research Agency, with European Union Next Generation funding. It has been also supported by the postgraduate research scholarship APCR2020010004 from the Canarian Agency for Research, Innovation and the Information Society - ACISI (Government of the Canary Islands), co-financed by the European Social Fund.

Declaration of interest statement

No potential conflict of interest was reported by the authors.

Impact statement

It is widely recognised that whale-watching tourism hounds and stresses the marine life it depends upon. These impacts have been of increasing concern because new forms of electronic communication can influence certain unethical and irresponsible behaviours. The result is that whale-watching providers may be led to carry out the provision of some services that are demanded by tourists but damage ecosystems. This study focuses on understanding consumers’ satisfaction with whale-watching tours as influenced by sustainability concerns. To this end, we analyse review comments by whale-watching tourists posted on TripAdvisor between 2010 and 2023. These online reviews are analysed in order to assess the various aspects of the tours (physical, human, environmental, experiential, and operational) that influence consumers’ online satisfaction rating and sustainability concerns. The results show

that sustainability concerns, as expressed on social media, have a positive effect on tourists’ whale-watching satisfaction. This could help operators to plan for more sustainable practices focusing on tourists’ preferences. Further, the identification of emerging social trends about sustainability concerns is important for managing the activity by all stakeholders involved in the industry.

CRedit authorship contribution statement

Carmelo J. León: Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization. **Chaitanya Suárez-Rojas:** Writing – review & editing, Writing – original draft, Project administration, Funding acquisition, Data curation, Conceptualization. **José Manuel Cazorla-Artiles:** Writing – original draft, Software, Methodology, Formal analysis, Data curation. **Matías M. González Hernández:** Visualization, Supervision, Project administration, Funding acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.tourman.2024.105019>.

References

- AboutTripAdvisor. (2022). *Media center, about TripAdvisor*. Visited on 2023-02-21.
- Abrahams, Z., Hoogendoorn, G., & Fitchett, J. M. (2022). Glacier tourism and tourist reviews: An experiential engagement with the concept of “Last Chance Tourism”. *Scandinavian Journal of Hospitality and Tourism*, 22(1), 1–14.
- Alaimo, C., Kallinikos, J., & Valderrama, E. (2020). Platforms as service ecosystems: Lessons from social media. *Journal of Information Technology*, 35(1), 25–48.
- Alegre, J., & Garau, J. (2010). Tourist satisfaction and dissatisfaction. *Annals of Tourism Research*, 37(1), 52–73.
- Ali, T., Marc, B., Omar, B., Soulimane, K., & Larbi, S. (2021). Exploring destination’s negative e-reputation using aspect based sentiment analysis approach: Case of Marrakech destination on TripAdvisor. *Tourism Management Perspectives*, 40, Article 100892.
- Araña, J. E., & León, C. J. (2013). Correcting for scale perception bias in tourist satisfaction surveys. *Journal of Travel Research*, 52(6), 772–788.
- Athanassopoulos, A. D. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of Business Research*, 47(3), 191–207.
- Ávila-Foucat, V. S., Vargas, A. S., Jordan, A. F., & Flores, O. R. (2013). The impact of vessel crowding on the probability of tourists returning to whale watching in Banderas Bay, Mexico. *Ocean & Coastal Management*, 78, 12–17.
- Bagherzadeh, S., Shokouhyar, S., Jahani, H., & Sigala, M. (2021). A generalizable sentiment analysis method for creating a hotel dictionary: Using big data on TripAdvisor hotel reviews. *Journal of Hospitality and Tourism Technology*, 12(2), 210–238.
- Banerjee, S., & Chua, A. Y. K. (2016). In search of patterns among traveller’s hotel ratings in TripAdvisor. *Tourism Management*, 53, 125–131.
- Bejder, L., Higham, J. E., & Lusseau, D. (2022). *Tourism and research impacts on marine mammals: A bold future informed by research and technology. Ethology and behavioral ecology of marine mammals: The evolving human factor* (pp. 255–275). Cham, Switzerland: Springer Nature.
- Bentz, J., Lopes, F., Calado, H., & Dearden, P. (2016). Enhancing satisfaction and sustainable management: Whale watching in the Azores. *Tourism Management*, 54, 465–476.
- Bentz, J., Rodrigues, A., Dearden, P., Calado, H., & Lopes, F. (2015). Crowding in marine environments: Divers and whale watchers in the Azores. *Ocean & Coastal Management*, 109, 77–85.
- Bi, J. W., Liu, Y., Fan, Z. P., & Zhang, J. (2020). Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry. *Tourism Management*, 77, Article 104006.
- Bryan, H. (1977). Leisure value systems and recreational specialization: The case of trout fishermen. *Journal of Leisure Research*, 9(3), 174–187.
- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.
- Campo, S., & Yagüe, M. J. (2008). Research note: Effects of price on tourist satisfaction. *Tourism Economics*, 14(3), 657–661.
- Campo, S., & Yagüe, M. (2009). Exploring non-linear effects of determinants on tourists’ satisfaction. *International Journal of Culture, Tourism and Hospitality Research*, 3(2), 127–138.
- Cárdenas, S., Gabela-Flores, M. V., Amrein, A., Surrey, K., Gerber, L. R., & Guzmán, H. M. (2021). Tourist knowledge, pro-conservation intentions, and tourist concern for the impacts of whale-watching in Las Perlas Archipelago, Panama. *Frontiers in Marine Science*, 8, Article 627348.
- Cassar, M., Konietzny, J., & Caruana, A. (2023). Customer encounter satisfaction and narrative force: An investigation of user-generated content on TripAdvisor. *Scandinavian Journal of Hospitality and Tourism*, 1–22.

- Chen, C. H., Nguyen, B., Klaus, P. P., & Wu, M. S. (2015). Exploring electronic word-of-mouth (eWOM) in the consumer purchase decision-making process: The case of online holidays—evidence from United Kingdom (UK) consumers. *Journal of Travel & Tourism Marketing*, 32(8), 953–970.
- Chen, Y., Zhang, H., & Qiu, L. (2013). A review on tourist satisfaction of tourism destinations. In *LISS 2012: Proceedings of 2nd international conference on logistics, informatics and service science* (pp. 593–604). Berlin Heidelberg: Springer.
- Chuang, L. Z., Chen, C. L., Kung, C. W., & Shih, Y. C. (2020). A nuisance at sea: Decoding tourists' comfort on whale watching vessels. *Ocean & Coastal Management*, 184, Article 104915.
- Cong, L., Wu, B., Morrison, A. M., Shu, H., & Wang, M. (2014). Analysis of wildlife tourism experiences with endangered species: An exploratory study of encounters with giant pandas in Chengdu, China. *Tourism Management*, 40, 300–310.
- Cornejo-Ortega, J. L., Chavez-Dagostino, R. M., & Malcolm, C. D. (2018). Whale watcher characteristics, expectation-satisfaction, and opinions about whale watching for private vs community-based companies in Bahía de Banderas, Mexico. *International Journal of Sustainable Development and Planning*, 13(5), 790–804.
- Curtin, S. (2010). Managing the wildlife tourism experience: The importance of tour leaders. *International Journal of Tourism Research*, 12(3), 219–236.
- Daugherty, T., Eastin, M. S., & Bright, L. (2008). Exploring consumer motivations for creating user-generated content. *Journal of Interactive Advertising*, 8(2), 16–25.
- Ding, X., Liu, B., & Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining. In *Proceedings of the 2008 international conference on web search and data mining* (pp. 231–240).
- D'Lima, C., Everingham, Y., Diedrich, A., Mustika, P. L., Hamann, M., & Marsh, H. (2018). Using multiple indicators to evaluate the sustainability of dolphin-based wildlife tourism in rural India. *Journal of Sustainable Tourism*, 26(10), 1687–1707.
- Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of price, brand, and store information on buyers' product evaluations. *Journal of Marketing Research*, 28(3), 307–319.
- Dolnicar, S. (2008). Market segmentation in tourism. In A. G. Woodside, & D. Martin (Eds.), *Tourism management, analysis, behaviour and strategy* (pp. 129–150). Oxfordshire: CABI International.
- Fei, G., Liu, B., Hsu, M., Castellanos, M., & Ghosh, R. (2012). A dictionary-based approach to identifying aspects implied by adjectives for opinion mining. In *Proceedings of COLING 2012* (pp. 309–318).
- Feng, S., Xing, L., Gogar, A., & Choi, Y. (2012). Distributional footprints of deceptive product reviews. *Proceedings of the International AAAI Conference on Web and Social Media*, 6(1).
- 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.
- Finkler, W., Higham, J. E., León, B., & Aitken, R. (2019). Bridging the void: Science communication videos for sustainable whale watching. *International Journal of Science Education, Part B*, 9(4), 312–326.
- Fuchs, M., & Weiermair, K. (2003). New perspectives of satisfaction research in tourism destinations. *Tourism Review*, 58(3), 6–14.
- García-Cegarra, A. M., & Pacheco, A. S. (2017). Whale-watching trips in Peru lead to increases in tourist knowledge, pro-conservation intentions and tourist concern for the impacts of whale-watching on humpback whales. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 27(5), 1011–1020.
- Garner, B., & Kim, D. (2022). Analyzing user-generated content to improve customer satisfaction at local wine tourism destinations: An analysis of yelp and TripAdvisor reviews. *Consumer Behavior in Tourism and Hospitality*, 17(4), 413–435.
- Gössling, S. (2017). Tourism, information technologies and sustainability: An exploratory review. *Journal of Sustainable Tourism*, 25(7), 1024–1041.
- Gurjar, P., Kaurav, R. P. S., & Thakur, K. S. (2022). Trustworthiness of user-generated content in travel planning decisions of tourists. *Journal of Positive School Psychology*, 3645–3660.
- 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*, 23, Article 104236.
- Hernández-Méndez, J., Muñoz-Leiva, F., & Sánchez-Fernández, J. (2015). The influence of e-word-of-mouth on travel decision-making: Consumer profiles. *Current Issues in Tourism*, 18(11), 1001–1021.
- Horney, K. (2013). *Our inner conflicts: A constructive theory of neurosis* (Vol. 17). Routledge.
- Hoyt, E. (2001). Whale watching 2001: Worldwide tourism numbers, expenditures, and expanding socioeconomic benefits. *International Fund for Animal Welfare, Yarmouth Port, MA, USA*, 1–vi, 1–158.
- Hoyt, E. (2017). The global status and true value of whale watching. *Whales in a changing ocean. Whale and dolphin conservation conference*. Tonga.
- Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conferences on Web and Social Media*, 8(1), 212–225. <https://doi.org/10.1609/icwsm.v8i1.14550>
- International Whaling Commission (online). *Whale Watching Handbook*. A comprehensive, online tool for regulators, industry and the general public. Available at: <https://wwhandbook.iwc.int/en/about>.
- Isabelle, G., Maharani, W., & Asror, I. (2019). Analysis on opinion mining using combining lexicon-based method and multinomial Naïve Bayes. In *2018 international conference on industrial enterprise and system engineering (ICoESE 2018)* (pp. 214–219). Atlantis Press.
- Izen, A. M. (1987). Positive effect, cognitive processes, and social behaviour. Berkowitz (Ed.). *Advances in Experimental Social Psychology*, 20, 203–253.
- Jacobs, M. H., & Harms, M. (2014). Influence of interpretation on conservation intentions of whale tourists. *Tourism Management*, 42, 123–131.
- Jiang, J., Zhang, J., Zhang, H., & Yan, B. (2017). Natural soundscapes and tourist loyalty to nature-based tourism destinations: The mediating effect of tourist satisfaction. *Journal of Travel & Tourism Marketing*, 1–13.
- Kessler, E., Harcourt, R., & Bradford, W. (2014). Will whale watchers sacrifice personal experience to minimize harm to whales? *Tourism in Marine Environments*, 10(1–2), 21–30.
- Kirilenko, A. P., Stepchenkova, S. O., & Hernandez, J. M. (2019). Comparative clustering of destination attractions for different origin markets with network and spatial analyses of online reviews. *Tourism Management*, 72, 400–410.
- Kozak, M. (2003). Measuring tourist satisfaction with multiple destination attributes. *Tourism Analysis*, 7(3–4), 229–240.
- Kredens, C., & Vogt, C. A. (2023). A user-generated content analysis of tourists at wildlife tourism attractions. *Frontiers in Sustainable Tourism*, 2, Article 1090749.
- La Manna, G., Melis, G., Rako-Gospic, N., Basta, J., Mackelworth, P., Holcer, D., ... Leeb, K. (2020). Sustainable dolphin watching tours as a tool to increase public awareness of marine conservation—a comparative analysis between two Mediterranean destinations and implications for management. *Journal of Ecotourism*, 19(4), 345–361.
- Lee, J. S., & Park, S. (2023). A cross-cultural anatomy of destination image: An application of mixed-methods of UGC and survey. *Tourism Management*, 98, Article 104746.
- Lenzi, C., Speiran, S., & Grasso, C. (2020). "Let me take a selfie": Implications of social media for public perceptions of wild animals. *Society and Animals*, 31, 64–83. <https://doi.org/10.1163/15685306-bja10023>
- Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. *Tourism Management*, 71, 337–347.
- Liu, Y., Teichert, T., Rossi, M., Li, H., & Hu, F. (2017). Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 417,784 user-generated reviews. *Tourism Management*, 59, 554–563.
- Lopez, G., & Pearson, H. C. (2017). Can whale watching be a conduit for spreading educational and conservation messages? A case study in Juneau, Alaska. *Tourism in Marine Environments*, 12(2), 95–104.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65.
- Lu, W., & Stepchenkova, S. (2015). User-generated content as a research mode in tourism and hospitality applications: Topics, methods, and software. *Journal of Hospitality Marketing & Management*, 24(2), 119–154.
- Lück, M., & Porter, B. A. (2019). Experiences on swim-with-dolphins tours: An importance–performance analysis of dolphin tour participants in Kaikoura, New Zealand. *Journal of Ecotourism*, 18(1), 25–41.
- Lui, T. W., Bartosiak, M., Piccoli, G., & Sathya, V. (2018). Online review response strategy and its effects on competitive performance. *Tourism Management*, 67, 180–190.
- Lukyzenko, R., Parsons, J., Wiersma, Y., Sieber, R., & Maddah, M. (2016). Participatory design for user-generated content: Understanding the challenges and moving forward. *Scandinavian Journal of Information Systems*, 28(1), 2.
- Ma, S., & Kirilenko, A. P. (2020). Climate change and tourism in English-language newspaper publications. *Journal of Travel Research*, 59(2), 352–366.
- Malcolm, C. D., & Duffus, D. (2008). Specialization of whale watchers in British Columbia waters. In J. E. S. Higham, & M. Lück (Eds.), *Marine wildlife and tourism management, insights from the natural and social sciences* (pp. 109–129). Oxfordshire: CABI International.
- Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis – a review of research topics, venues, and top cited papers. *Computer Science Review*, 27, 16–32.
- Mariani, M., & Borghi, M. (2022). Exploring environmental concerns on digital platforms through big data: The effect of online consumers' environmental discourse on online review ratings. *Journal of Sustainable Tourism*, 1–20.
- Meyer, L., Barry, C., Araujo, G., Barnett, A., Brunnschweiler, J. M., Chin, A., ... Huveneers, C. (2022). Redefining provisioning in marine wildlife tourism. *Journal of Ecotourism*, 21(3), 210–229.
- Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), 436–465.
- Mukhtar, N., Khan, M. A., & Chiragh, N. (2018). Lexicon-based approach outperforms supervised machine learning approach for Urdu sentiment analysis in multiple domains. *Telematics and Informatics*, 35(8), 2173–2183.
- O'Connor, S., Campbell, R., Cortez, H., & Knowles, T. (2009). *Whale watching worldwide: Tourism numbers, expenditures and expanding economic benefits, a special report from the international fund for animal welfare*.
- Orams, M. B. (2000). Tourists getting close to whales, is it what whale-watching is all about? *Tourism Management*, 21(6), 561–569.
- Padma, P., & Ahn, J. (2020). Guest satisfaction & dissatisfaction in luxury hotels: An application of big data. *International Journal of Hospitality Management*, 84, Article 102318.
- Park, E., Kang, J., Choi, D., & Han, J. (2020). Understanding customers' hotel revisiting behaviour: A sentiment analysis of online feedback reviews. *Current Issues in Tourism*, 23(5), 605–611.
- Park, S., Yang, Y., & Wang, M. (2019). Travel distance and hotel service satisfaction: An inverted U-shaped relationship. *International Journal of Hospitality Management*, 76, 261–270.
- Pencle, N., & Mălăeșcu, I. (2016). What's in the words? Development and validation of a multidimensional dictionary for CSR and application using prospectuses. *Journal of Emerging Technologies in Accounting*, 13(2), 109–127.

- Pizam, A., Neumann, Y., & Reichel, A. (1978). Dimensions of tourist satisfaction with a destination area. *Annals of Tourism Research*, 5(3), 314–322.
- Pomeroy, A., Noble, G., & Johnson, L. W. (2011). Conceptualising a contemporary marketing mix for sustainable tourism. *Journal of Sustainable Tourism*, 19(8), 953–969.
- Ponnampalam, L. S. (2011). Dolphin watching in Muscat, Sultanate of Oman: Tourist perceptions and actual current practice. *Tourism in Marine Environments*, 7(2), 81–93.
- 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.
- Rehman, A. U., Abbas, M., Abbasi, F. A., & Khan, S. (2023). How tourist experience quality, perceived price reasonableness and regenerative tourism involvement influence tourist satisfaction: A study of Ha'il region, Saudi Arabia. *Sustainability*, 15(2), 1340.
- Reyes-Menendez, A., Saura, J. R., & Martinez-Navalon, J. G. (2019). The impact of e-WOM on hotels management reputation: Exploring tripadvisor review credibility with the ELM model. *IEEE Access*, 7, 68868–68877.
- Reynolds, P. C., & Braithwaite, D. (2001). Towards a conceptual framework for wildlife tourism. *Tourism Management*, 22(1), 31–42.
- Salah, M. H. A., Abdou, A. H., Hassan, T. H., El-Amin, M. A. M. M., Kegour, A. B. A., Alboray, H. M. M., ... Mohammed, E. F. A. (2023). Power of eWOM and its antecedents in driving customers' intention to revisit: An empirical investigation on five-star eco-friendly hotels in Saudi Arabia. *Sustainability*, 15(12), 9270.
- Setiawan, P. Y., Troena, E. A., & Armanu, N. (2014). The effect of e-WOM on destination image, satisfaction and loyalty. *International Journal of Business and Management Invention*, 3(1), 22–29.
- Shang, Z., & Luo, J. M. (2022). Topic modelling for wildlife tourism online reviews: Analysis of quality factors. *Current Issues in Tourism*, 1–15.
- Shapiro, K. R. (2006). *Whale watch passengers' preferences for tour attributes and marine management in Maui*. Simon Fraser University.
- Simpson, G. D., Patroni, J., Teo, A. C., Chan, J. K., & Newsome, D. (2020). Importance-performance analysis to inform visitor management at marine wildlife tourism destinations. *Journal of Tourism Futures*, 6(2), 165–180.
- Solangi, Y. A., Solangi, Z. A., Aarain, S., Abro, A., Mallah, G. A., & Shah, A. (2018). Review on natural language processing (NLP) and its toolkits for opinion mining and sentiment analysis. In *2018 IEEE 5th international conference on engineering technologies and applied sciences (ICETAS)* (pp. 1–4).
- Song, S., Kawamura, H., Uchida, J., & Saito, H. (2019). Determining tourist satisfaction from travel reviews. *Information Technology & Tourism*, 21, 337–367.
- Suárez-Rojas, C., González Hernández, M., & León, C. J. (2023). Segmented importance-performance analysis in whale-watching: Reconciling ocean coastal tourism with whale preservation. *Ocean & Coastal Management*, 233, Article 106453.
- Suárez-Rojas, C., González Hernández, M., & León, C. J. (2023). Sustainability in whale-watching: A literature review and future research directions based on regenerative tourism. *Tourism Management Perspectives*, 47, Article 101120.
- Taecharungroj, V., & Mathayomchan, B. (2019). Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. *Tourism Management*, 75, 550–568.
- Teles, R., Martins, H., Pinheiro, A., & Gonçalves, E. (2024). Tourists' perception of tourist destinations: The case study of Nazaré (Portugal). *Sustainability*, 16(4), 1387.
- Tepsich, P., Borroni, A., Zorgno, M., Rosso, M., & Moulins, A. (2020). Whale watching in the Pelagos sanctuary: Status and quality assessment. *Frontiers in Marine Science*, 7, Article 596848.
- Tessier-Moreau, A. (2022). Évolution des profils, attentes et de la satisfaction des participants aux croisières d'observation des mammifères marins dans le parc marin du Saguenay-Saint-Laurent. *étude comparative de 1999 et 2021*. Doctoral dissertation, Université du Québec à Rimouski.
- Tkaczynski, A. (2021). I can't get no satisfaction: Or can I? Satisfying Australian whale-watching tourists. *Tourism in Marine Environments*, 16(3), 153–165.
- Tkaczynski, A., & Rundle-Thiele, S. (2019). Identifying whale-watching tourist differences to maximize return on investment. *Journal of Vacation Marketing*, 25(3), 390–402.
- Tokarchuk, O., Barr, J. C., & Cozzio, C. (2022). How much is too much? Estimating tourism carrying capacity in urban context using sentiment analysis. *Tourism Management*, 91, Article 104522.
- Toral, S. L., Martínez-Torres, M. R., & Gonzalez-Rodriguez, M. R. (2018). Identification of the unique attributes of tourist destinations from online reviews. *Journal of Travel Research*, 57(7), 908–919.
- Torres-Matovelle, P., & Molina-Molina, G. (2019). Evaluation of crowding and tourist satisfaction in the practice of humpback whale-watching, the case of Puerto López-Ecuador. *Cuadernos de Gestión*, 19(2), 185–208.
- TripAdvisor. (2017). *Q1 2017 results*. TripAdvisor presentation.
- Van der Zee, E., & Bertocchi, D. (2018). Finding patterns in urban tourist behaviour: A social network analysis approach based on TripAdvisor reviews. *Information Technology & Tourism*, 20(1), 153–180.
- Van der Zee, E., Bertocchi, D., & Vanneste, D. (2020). Distribution of tourists within urban heritage destinations: A hot spot/cold spot analysis of TripAdvisor data as support for destination management. *Current Issues in Tourism*, 23(2), 175–196.
- Vania, C., Ibrahim, M., & Adriani, M. (2014). Sentiment lexicon generation for an under-resourced language. *International Journal of Computational Linguistics and Applications*, 5(1), 59–72.
- Wang, Y. (2016). *More important than ever: Measuring tourist satisfaction*. Queensland, Australia: Griffith University. Griffith Institute for Tourism.
- Wang, L., Youn, B. D., Azarm, S., & Kannan, P. K. (2011). Customer-driven product design selection using web based user-generated content. *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 54822, 405–419.

- Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51–65.
- Xie, K. L., So, K. K. F., & Wang, W. (2017). Joint effects of management responses and online reviews on hotel financial performance: A data-analytics approach. *International Journal of Hospitality Management*, 62, 101–110.
- Zeng, B., & Gerritsen, R. (2014). What do we know about social media in tourism? A review. *Tourism Management Perspectives*, 10, 27–36.
- Zhang, Y., Song, J., Sciacca, A., Chan, J., & Qi, X. (2022). Novel sentiment lexica derived from user generating content by Chinese tourists in Pacific Islands. *Sustainability*, 14(23), Article 15833.
- Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, 76, 111–121.
- Ziegler, J. A., Silberg, J. N., Araujo, G., Labaja, J., Ponzo, A., Rollins, R., & Dearden, P. (2018). A guilty pleasure: Tourist perspectives on the ethics of feeding whale sharks in Oslob, Philippines. *Tourism Management*, 68, 264–274.
- Zolfaghari, A., & Choi, H. C. (2023). Elevating the park experience: Exploring asymmetric relationships in visitor satisfaction at Canadian national parks. *Journal of Outdoor Recreation and Tourism*, 43, Article 100666.



Dr. Carmelo J. León González is full professor at the University of Las Palmas de Gran Canaria. He is also Director of the Institute of Tourism and Sustainable Economic Development (TiDES Institute) and the UNESCO Chair in Tourism and Sustainable Economic Development ULPGC.



Dr. Chaitanya Suárez Rojas holds a postdoctoral researcher position at the University of Las Palmas de Gran Canaria (Spain). She is a researcher at the Institute of Tourism and Sustainable Economic Development (TiDES) and associated with the Centre for Environmental and Resource Economics (Sweden).



Dr. José Manuel Cazorla Artilles is lecturer at the University of Las Palmas de Gran Canaria. He also held a postdoctoral researcher position at MARCET II Project (MAC/4.6c/392).



Dr. Matías M. González Hernández is professor at the University of Las Palmas de Gran Canaria. He is also researcher in charge of the Training and Development Cooperation Division of the Institute of Tourism and Sustainable Economic Development (TiDES Institute).