Marcelino Sánchez-Rivero / Luis Murillo-González / María Cristina Rodríguez-Rangel

Tourism Management Through Natural Language Processing and Sentiment Analysis. A Case Study of the Main Natural Areas of Extremadura, Spain

Abstract

Nowadays, digital scenarios and global connectivity have changed the way we travel; visitors can comment and share their experiences from different sources of information, and this action can be of essential value in tourism research. Consequently, these factors can help design new marketing, management, or business strategies and make tourist destinations more competitive. Multiple techniques can process big data and machine learning analysis, such as natural language processing (NLP), which can support tourist resources in analyzing this information. This paper proposes an NLP technique to present an emotion detection analysis based on social web reviews about the main natural spaces in Extremadura, Spain. Sentiment and emotions scores have been computed based on the Plutchik model using a lexicon database to understand visitor experience. Empirical evidence suggests some recommendations regarding tourism management to improve the visitor experience and reduce negative sentiments and emotions.

Keywords: natural language processing, machine learning, artificial intelligence, sentiment analysis, tourism, natural spaces, Spain

Funding

This publication is part of the research carried out for the project entitled "Análisis de factores críticos para el desarrollo turístico de Extremadura (IB-18015)". It is funded by the Ministry of Economy, Science and The Digital Agenda of the Junta de Extremadura and by the European Regional Development Fund (ERDF).

1. Introduction

Spain is one of the most important tourist destinations in the world, with more than 80 million people visiting the country in 2019 (Hernandez-Maskivker et al., 2021). Despite the importance of Spanish tourism on the international stage, there are still regions with extraordinary potential to improve the competitiveness of tourism in Spain. One of these regions is the Community of Extremadura, which can offer new alternatives related to the environment, culture, or activities (Sánchez-Martín et al., 2018).

Over the last two decades, the Community of Extremadura has become an immensely popular tourist destination among national and international tourists across Europe. This phenomenon has significantly impacted the source of income that tourism brings to the region and has allowed higher levels of economic and social

Marcelino Sánchez-Rivero, PhD, Senior and Tenured Professor, University of Extremadura, Department of Economics, Badajoz, Spain; ORCID ID: https://orcid.org/0000-0003-3988-6278; e-mail: sanriver@unex.es

Luis Murillo-González, MSc, Corresponding Author, Researcher, University of Extremadura, Department of Economics, Badajoz, Spain; ORCID ID: https://orcid.org/0000-0002-9164-3540; e-mail: luismg@unex.es

María Cristina Rodríguez-Rangel, PhD, Substitute Lecturer, University of Extremadura, Department of Economics, Badajoz, Spain; ORCID ID: https://orcid.org/0000-0002-3526-5039; e-mail: mcrisrod@unex.es

Case Study
Vol. 73/ No. 1/ 2025/ 169 - 185
UDC: 338.48:811.1/8-047.44 (460); https://doi.org/10.37741/t.73.1.12

development derived from increased tourism services (Sánchez-Martín et al., 2019). The Community of Extremadura strives for a sustainable future encompassing social, financial, and environmental dimensions. At the same time, tourists and businesses have a growing demand for digital resources, such as web-based tourism packages, to enhance their outreach. (Sánchez-Oro et al., 2021).

These new digital scenarios suggest new business opportunities where social networks, travel agencies and different economic agents play pivotal roles in influencing consumer behaviour. Nevertheless, the double interaction between consumers and economic actors could be a problem because the actions of users and customers on the network hold a significant value (Huertas et al., 2015).

Comments and experiences shared online can provide important credibility to potential consumers (Torrado Rodríguez & Blanco Gregory, 2018). Across various communication channels, these comments and experiences can focus on tangible elements such as monuments, museums and restaurants or intangible elements such as sentiments or emotions (Marine-Roig & Huertas, 2016).

Understanding visitors' intangible and emotional aspects can assist the tourism sector in improving the services offered for more significant economic and social benefit to the region or country (Prakash & Aruna-Kumari, 2019). Therefore, analysing the emotions and sentiments generated in digital environments becomes fundamental to achieving these objectives.

Based on the above reflections, this paper proposes to analyse, using Natural Language Processing, the feelings (sentiments and emotions) expressed in the comments made by visitors on various digital platforms of the leading natural areas of the Community of Extremadura.

Finally, this analysis aims to investigate whether it is possible to enhance the presence of positive sentiments and emotions in each destination to improve (i) the online reputation of tourist destinations and (ii) user experience. The empirical evidence gathered from visitor experiences suggests some recommendations in terms of tourism management that can contribute positively to increasing the competitiveness of tourism in the Community of Extremadura and the rest of Spain.

2. Literature review

2.1. Creation and dissemination of tourism content online

The information generated online from tourism activity has increased in recent years due to the expansion of social networks and the Internet (Powers, 2021; Yang et al., 2017). Consequently, it has become a vital intangible asset within the tourism industry worldwide, and the primary content generated on these tourism platforms consists of reviews and comments visitors leave about their experiences. This phenomenon has also contributed to improving the performance of tourism services, increasing customer satisfaction levels, and influencing consumers' experiences (Gretzel, 2006; Paredes et al., 2021).

The literature review of online content indicates that one of the most accessible sources of this type of information is eWOM (electronic word of mouth), which is the digital version of word of mouth (WOM) (Chatterjee, 2001). WOM refers to interpersonal communication between consumers concerning their subjective experiences with a company (Richins, 1983).

One of the most significant theoretical advances in eWOM analysis applied to the tourism industry has been the design of new commercial strategies (Álvarez-Carmona et al., 2022) to improve online reputation. Tourism destinations must align their commercial strategy to meet these demands by analysing and enhancing feelings and emotions as they play an essential role in the reputation of online tourism environments (Chhetri et al., 2004; Hernández Mogollón et al., 2020).

The literature review indicates that if tourism destinations improve their online reputation by 1%, the average daily rate will increase by 0.89%, occupancy levels by 0.54% and up to 1.42% in the profit margin per available room (RevPAR) (Anderson, 2012).

2.2. Natural language processing and tourism

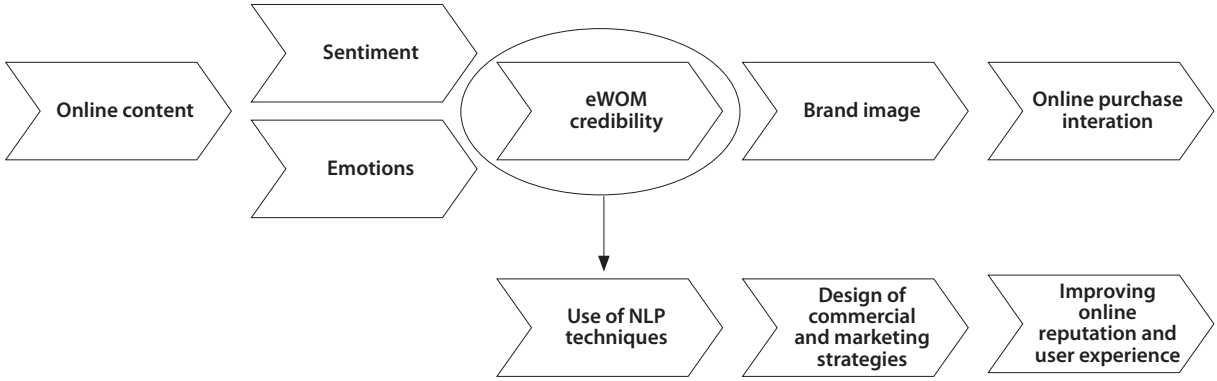
The volume of comments and opinions derived from tourism activity is enormous. Still, the economic and social sphere needs to understand these sentiments, how these attributes relate to tourism destinations and whether there is a positive or negative aspect (Georgescu & Bogoslov, 2019; Liu, 2012).

Recent applications of NLP have been used in the tourism sector to improve marketing and commercial strategies, chatbots, sentiment analysis, and hotel recommendations, among others (Bulchand-Gidumal, 2022; García-Pablos et al., 2016; Prameswari et al., 2017; Putri et al., 2019).

Chatbots have been primarily used by large companies such as Amazon with Siri or Google NLP (Godbole et al., 2021; Khalil & Pipa, 2022), but the tourism sector has used chatbots to implement virtual travel assistants, conversational systems, and robots (Tussyadiah & Miller, 2019).

Similarly, sentiment analysis has been truly relevant in the tourism literature due to theoretical advances that have contributed significantly to various aspects, such as the improving online reputation (Guerrero-Rodríguez et al., 2023; Mariani et al., 2018; Shaikh & Kulkarni, 2020). In addition, sentiment analysis has contributed to developing theories that support strategic decision-making in different business domains (Bordoloi & Biswas, 2023). These advances have proven helpful in product development, experience improvement, marketing strategies, competitive intelligence, content recommendation, regulatory compliance, and semantic search (Álvarez-Carmona et al., 2022).

Figure 1
Theoretical relationship



Source: Own elaboration.

Figure 1 shows the relationship between the sentiments and emotions observed in the online content that the eWOM transmits to other consumers and how these impact the brand or destination. At the same time, this way of communicating between users can be analysed through NLP to improve the user experience and reputation of a brand or destination through commercial strategies.

On the other hand, the analysis of comments and emotions requires algorithmic techniques, scripting, and high-speed computing processes. However, artificial intelligence can be costly regarding processing time or data mining. Techniques such as Natural Language Processing (NLP) can make it possible due to the relationships that can develop between computers and human language (Jafari et al., 2021).

NLP can support the tourism sector in analysing intangible assets by mining and automatically classifying texts. As mentioned above, NLP allows the construction of lines of work and new business strategies by considering the information extracted from the users and the destinations studied. This is one of the strengths that NLP offers tourists and the different economic agents operating in the tourism sector (Álvarez-Carmona et al., 2021).

Similarly, automatic text classification can extend and improve analyses in tourism research. This method can transform qualitative information into quantitative information. Furthermore, combining the results of these computational processes with other methodologies such as time series analysis, econometric process, or cluster analysis is possible.

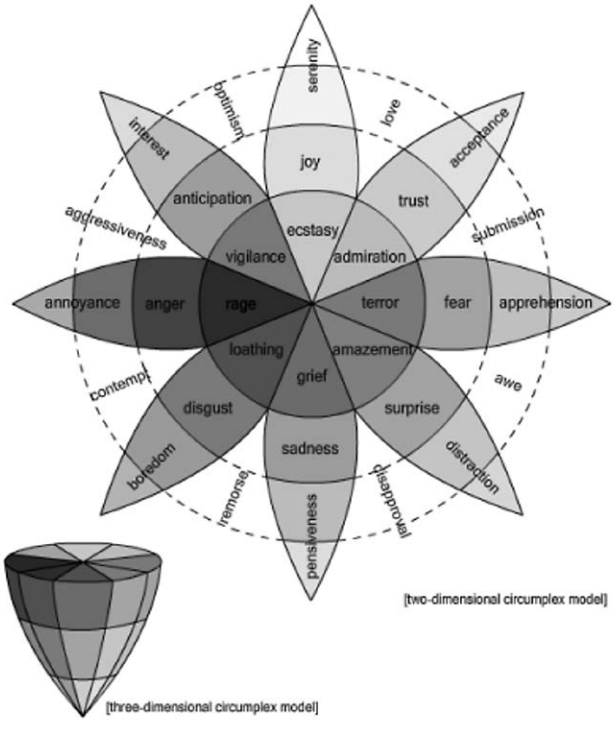
Nowadays, many NLP and sentiment analysis libraries are found within tourism and business management analysis. This paper has proposed using two of them, the Wordcloud package (Fellows, 2018) and the Syuzhet package (Jockers, 2015) of R, whose function is detailed in the methodological section of this paper.

2.3. Sentiment analysis

Sentiments and emotions have always been innately present in human activity, even when cultures have not been related to each other. It has been accepted that a set of basic emotions exists (Ekman & Friesen, 2003). Many professionals have worked in this field (Ekman, 1992; Frijda, 1988; Parrott, 2001; Plutchik, 1980), but this paper will focus on the proposals of Ekman (1992) and Plutchik (1980).

Paul Ekman classifies six basic human emotions: fun, sadness, anger, fear, disgust, and surprise. However, Plutchik identifies a total of eight basic emotions, which include the previous six and adds two more emotions: trust and anticipation. Trust, fun, anticipation, and surprise correspond with a positive sentiment, while anger, disgust, fear, and sadness correspond with a negative sentiment.

Figure 2
Plutchik's Wheel of emotion



Source: Plutchik and Conte (1997).

Plutchik proposes a model based on a wheel of emotions organised around these eight basic emotions (Figure 2). In addition to the basic emotions, Plutchik argues that each primary emotion has an opposite primary emotion, according to the pairwise grouping fun/sadness, fear/anger, anticipation/surprise, and disgust/trust. Emotions can be combined (for example, the sum of fun and trust creates love), and the intensity of these emotions is represented in Figure 1 from the inside out.

There are many ways to extract sentiments and emotions from the comments. This paper seeks to extract and identify that intangible and emotional part through sentiment lexicons. The most popular resources for this purpose are the following: Opinion Lexicon (Hu & Liu, 2004), General Inquirer (Stone & Hunt, 1963), NRC EmoLex (Mohammad & Turney, 2010), and LIWC (Linguistic Inquiry and Word Count) (Pennebaker et al., 2001).

On the other hand, these libraries base their processes on the resources of the NRC Word-Emotion Association Lexicon. The NRC Emotion Lexicon is a list of English words associated with Plutchik's eight basic emotions, allowing comments to be linked with an emotional analysis. The NRC has around 14,000 words labelled according to Plutchik's model. The set of words is in English, but it is possible to find more than 40 languages available. These annotations have been manually annotated through collaboration and using the automatic classification of Mohammad and Turney (2010, 2013).

Table 1
EmoLex: Representation of the word dark

Palabra	Categoria	Asociacion
dark	anger	0
dark	anticipation	0
dark	disgust	0
dark	fear	0
dark	jay	0
dark	negative	0
dark	positive	0
dark	sadness	1
dark	siprise	0
dark	trust	0

Source: Escortell Pérez et al. (2017).

For example, the word "dark" is associated only with the emotion "sadness", where it is related to a value of 1, while the others are associated with 0 (see Table 1).

3. Methodology

The object of study is the leading natural destinations of the Community of Extremadura, where around 30% is protected by international, national, or regional designations (González-Barahona et al., 2021; Sánchez-Martín et al., 2020). These natural spaces are the most popular destinations and the most highly rated by visitors in this region, and these destinations are:

Monfragüe National Park, Garganta de los Infiernos Nature Reserve, Tagus Internacional Nature Reserve, Cornalvo Nature Reserve, Mina la Jayona Natural Monument, Barruecos Natural Monument, Cueva de Fuentes de León Natural Monument and Cueva del Castañar Natural Monument.

These destinations have a solid position in the international context. One of these figures is UNESCO, which has recognised these places as a unique destination with an extraordinary environment and biodiversity (González-Barahona et al., 2021). These destinations protect biodiversity and Mediterranean sustainability and are a reference for improving and recovering other areas that human activity has destroyed or destroyed.

Although these destinations have similar characteristics and offers, the tourist management can differ in each place.

This paper proposes as an initial hypothesis that specific comments, feelings and emotions affect (i) online reputation and (ii) the user experience is different in each natural space. Once the initial hypothesis is verified, it will be possible to design innovative marketing and commercial strategies to improve online reputation and user experience.

Additionally, this paper proposes an NLP process to extract the information and, consequently, sentiment analysis to verify this hypothesis (i) and (ii). This paper divides the methodological process into two parts. Firstly, the comments are processed to optimise the information to be computed, and secondly, a sentiment analysis is performed using the R Syuzhet library (Jockers, 2015).

3.1. Optimisation of information

ReviewPro obtained comments on the main natural areas of Extremadura between 1st January 2017 and 31st September 2020. This software can provide the volume of comments left by users that determine the online reputation for each of these natural destinations. A total of 2,350 comments were collected: Monfragüe National Park (1,017), Cueva del Castañar Natural Monument (37), Cuevas de Fuentes de León (27), Barruecos Natural Monument (48), Mina la Jayona Natural Monument (280), Cornalvo Nature Reserve (132), Tagus Internacional Nature Reserve (80), Garganta de los Infiernos Nature Reserve (729).

Once the information had been obtained, the text was filtered and homogeneously translated into English. During this filtering, duplicate comments were removed, and the process was completed by eliminating unnecessary spaces, punctuation marks and other elements that could complicate the analysis. The methodological analysis seeks to provide the information in the most homogeneous way possible to efficiently process an algorithm that can identify the human sentiments and emotions expressed by tourists in each of the natural spaces analysed.

To carry out the above process, this paper has proposed a tm library (Feinerer et al., 2008). This library helps support the methodological process because all the original information has been processed as a corpus.

3.2. Sentiment analysis

Analysing sentiment is one of the most critical tasks of NLP (Dhuria, 2015). Once the data has been optimised, this paper proposes two methods to verify the hypothesis (i) and (ii). On the one hand, the destinations are compared between two types of word clouds (frequency word cloud and emotions word cloud); on the other hand, the presence of these emotions and sentiments is observed in the total of the comments.

Both methodological processes make it possible to verify if reputation is affected positively or negatively and if the user experience differs. The empirical evidence depends on the frequency and characteristics of comments about each natural destination.

3.3. Wordcloud

Initially, the two types of word clouds are presented; the first type represents the number of most frequent words, while the second type groups the words according to these emotions. These word clouds represent the at least 100 most frequent words. When a specific word has a high frequency in the comment, it appears with a large size and centred.

Both word clouds can be represented through Wordcloud (Fellows, 2018), RColorBrewer (Neuwirth, 2014) and wordcloud2 (Lang, 2018) libraries, which are available in R. However, to represent the second type of word cloud, which links each word to its respective emotion, obtaining the presence of these emotions and sentiments is necessary.

For this reason, this paper has decided to compute this information through the Syuzhet library version 1.0.1 (Jockers, 2015), which has implemented NRC as a free code through the language in the R to obtain these frequencies (Mohammad, 2012).

Once the information has been processed, a set of associations (scores) is obtained, indicating if the words observed are associated with the sentiments or emotions. These scores can represent the quantitative information of the process and create the second type of word cloud, where each word is defined according to the correct emotion.

3.4. Distribution of sentiments and emotions

Using the different arguments of the R Syuzhet package, it is possible to obtain the relative presence of each sentiment and emotion in the total volume of comments.

This work presents two types of tables; one shows the presence of sentiments (positive and negative), and the other shows the presence of the eight basic emotions (fun, sadness, anger, fear, disgust, surprise, trust, and anticipation). For both tables, the intensity of the colour is defined by the maximum and minimum values of these sentiments and emotions in each natural space analysed.

Note that the NRC dictionary indicates the values in both tables, so the number of associations highly depends on this dictionary. Furthermore, this computational process only shows the primary emotion combinations; if another emotion is required, that needs to follow another method.

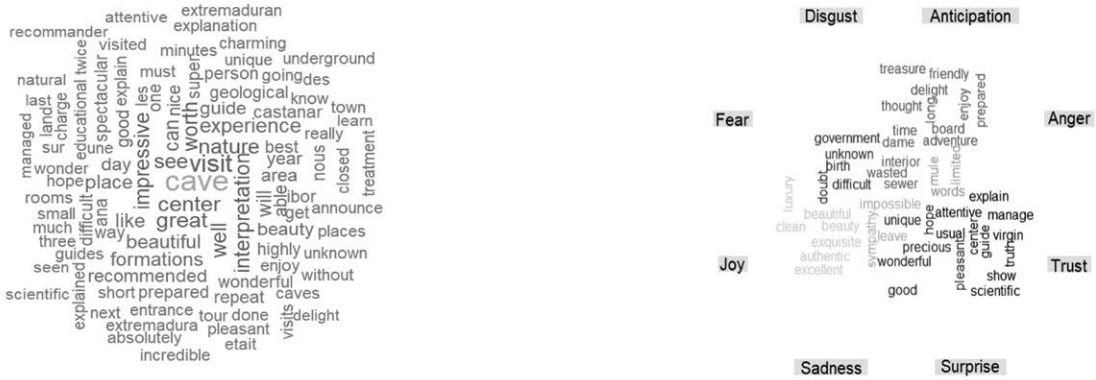
4. Results

4.1. Wordcloud

Figure 3
Frequency and emotions of the comments at Monfragüe National Park

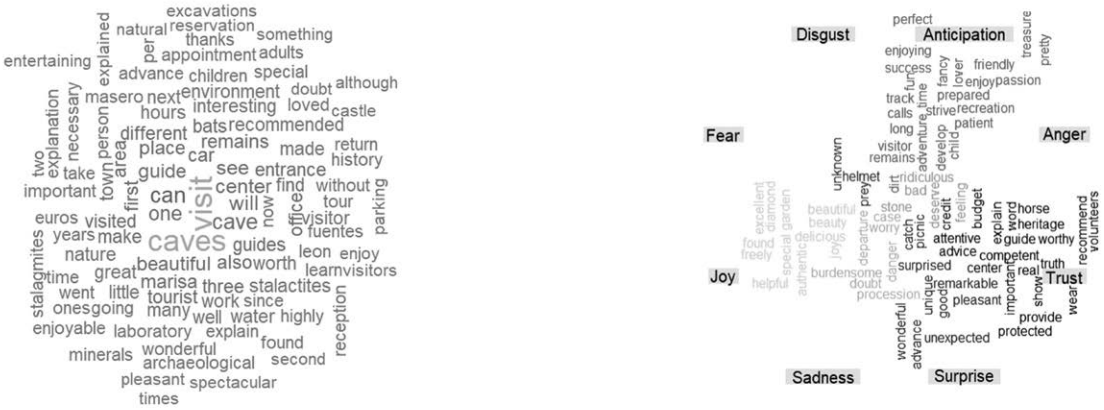
Source: Own elaboration.

Figure 4
Freauency and emotions of the comments at Cueva del Castañar Natural Monument



Source: Own elaboration.

Figure 5
Frequency and emotions of the comments at Cueva de Fuentes de León Natural Monument

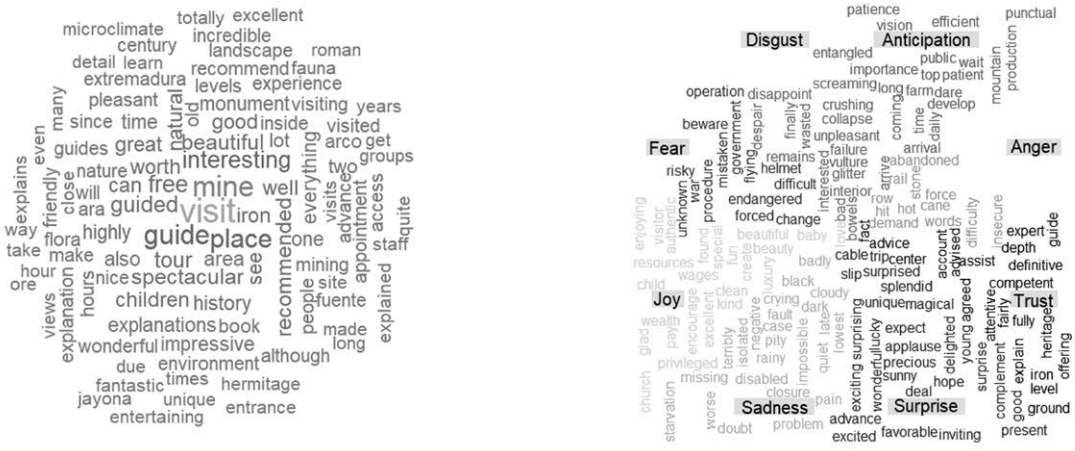


Source: Own elaboration.

Figure 6
Frequency and emotions of the comments at Barruecos Natural Monument

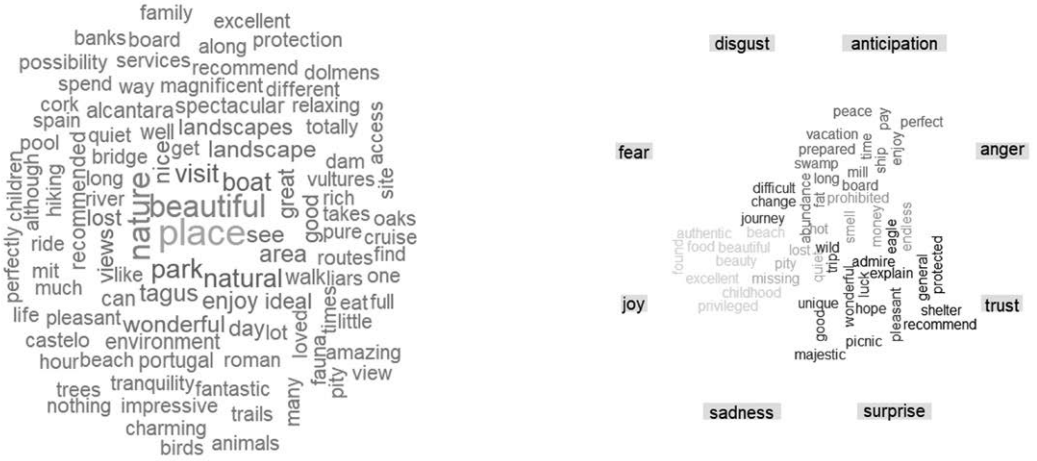
Source: Own elaboration.

Figure 7
Frequency and emotions of the comments at Mina la Jayona Natural Monument



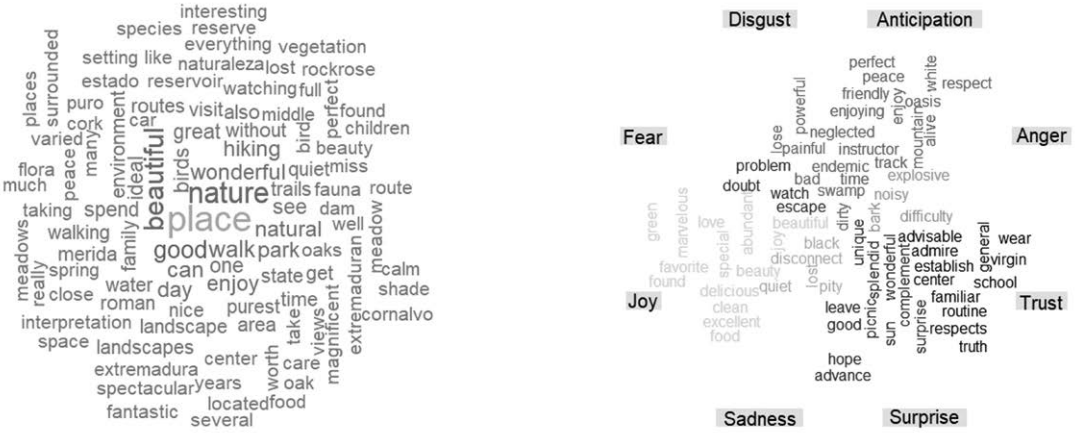
Source: Own elaboration.

Figure 8
Frequency and emotions of the comments at Cornalvo Nature Reserve



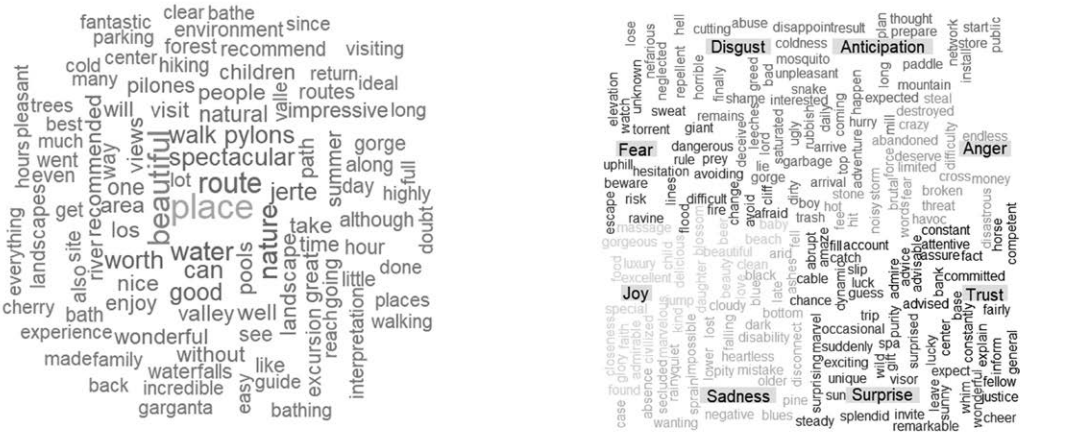
Source: Own elaboration.

Figure 9
Frequency and emotions of the comments at Tagus Internacional Nature Reserve



Source: Own elaboration.

Figure 10
Frequency and emotions of the comments at Garganta de los Infiernos Nature Reserve



Source: Own elaboration.

These figures show the most frequent words and the type of emotion in each of these words. Note that the volume of comments does not increase the presence of sentiments and feelings, but the meaning of these comments does.

The results show distinct characteristics in each of the word clouds. There are destinations where sentiments and emotions are incredibly low compared to others, directly affecting different perceptions of the brand and personal experience. Users create distinct images of each destination based on their activities and experiences.

Monfragüe National Park has registered the highest frequency of words related to hiking routes, landscapes, and botany. Visitors have valued Salto del Gitano, the castle of Monfragüe, outdoor activities or birdwatching.

The words that define the positive sentiments of these comments are related to the quality of services and the views of the park, which make it a unique place. Similarly, negative feelings are defined by words such as fire, rubbish, or shooting.

The comments observed at Cueva del Castañar Natural Monument discuss the visits, the interpretation centre, and nature. Positive sentiments are identified with the cleanliness, the quality of the place, and the enjoyment of the experience where words like "centre", "impressive", and "clean" have appeared in the comments. Negative sentiments are related to mentions of the sewers and the low quality of some tourist services.

The comments observed at Cuevas de Fuentes de León discuss the lessons on minerals and the visit, with words like "guide" or "stalactites" highlighted. In general, visitors have rated the place positively; comments talk about the beauty of the place, gastronomy, and food quality. On the other hand, there are reports about hazards or rubbish that can be found in some parts of the destination.

Barruecos Natural Monument shows comments focusing on the environment, rocks, and the wild animals that live there. There are many references to the museum, the Game of Thrones series, or alternative destinations such as the Monumental City of Cáceres. Positive sentiments show the quietness and quality of services offered by the site and the contribution of the Game of Thrones series. At the same time, negative sentiments talk about areas with rubbish, which is not a good advertisement for the destination.

The comments observed at Mina la Jayona Natural Monument talk about the guided tour and the botany of the destination. Botany is a unique aspect that is only observed in this natural area. On the other hand, negative sentiments show how visitors are disappointed with the time spent on services in this natural area; for example, the word "loss" appears associated with the word "time".

The comments observed at Cornalvo Natural Park speak of the diversity of the birdlife, the routes, and their distinct levels. Note that these comments are also observed in Monfragüe National Park. Still, visitors have described Cornalvo Nature Reserve as a family-friendly destination where you can enjoy gastronomy and biodiversity. However, visitors have highlighted the negative aspects, such as noise and waste. Compared to other natural areas, these negative aspects are more pronounced, with words like "negligence" or "painful" appearing in the comments.

Activities related to hiking, water, bird watching, visits to Roman monuments, and connections between Spain and Portugal are appreciated at the Tagus Internacional Nature Reserve. Positive comments describe an excellent destination where visitors recommend places like the beach or a boat trip. Negative remarks and sentiments mention the smells coming from the river.

Finally, Garganta de los Infiernos Nature Reserve is the destination with the highest number of comments related to water and biodiversity. Routes through natural pools or waterfalls are the best rated and commented on by visitors. The positive sentiments observed in the comments are about fun, quality, and nature. However, negative sentiments talk about forest fires, litter, and mosquitoes.

4.2. Emotional presence

The results show different values of emotional presence in each of the destinations. Overall, the online reputation and visitor experience are significantly strengthened due to the more positive emotions observed in the comments. At the same time, negative sentiments and emotions describe an image that does not resemble the reality of the destination.

Table 2 shows the presence of sentiment (positive and negative) for each of the natural destinations analysed concerning the total volume of comments. The value is expressed in percentages, and the range of values goes from 3% to 17%. The most considerable evidence in Table 2 is the presence of green, which indicates positive sentiments. Positive sentiments are generally more common than negative sentiments in the comments and destinations.

Table 2
Presence of positive and negative sentiments over total comments

Natural space	Positive	Negative
Monfragüe National Park	10%	5%
Cueva del Castañar Natural Monument	17%	3%
Cuevas de Fuentes de León	16%	4%
Barruecos Natural Monument	15%	3%
Mina la Jayona Natural Monument	13%	6%
Cornalvo Nature Reserve	15%	5%
Tagus Internacional Nature Reserve	14%	4%
Garganta de los Infiernos Nature Reserve	13%	7%

Source: Own elaboration.

Cueva del Castañar Natural Monument and Cuevas de Fuentes de León are the destinations with the highest positive comments, with 17% and 16%, respectively. Additionally, the natural area with the lowest proportion of positive comments is Monfragüe National Park, with 10%. This means a difference of 7% with the best reference destination, Cueva del Castañar Natural Monument.

Two of the natural areas with the least negative sentiment in their comments are Cueva del Castañar Natural Monument and Barruecos Natural Monument, with 3% in both cases. On the other hand, Garganta de los Infiernos and Mina la Jayona Natural Monument present 7% and 6%, respectively.

Table 3 and Table 4 present the percentage of each emotion concerning the total number of comments in each of the natural areas analysed. The simultaneous comparison of these tables and the grouped word clouds can help to understand the main differences between the natural areas analysed.

The emotion "Trust" presents the highest average in natural areas analysed. Cueva del Castañar Natural Monument has 9%, while Monfragüe National Park has the lowest value with 5%. Through these values and the word clouds of the destinations analysed with a low presence of trust, this work proposes to improve the didactic environment to increase the presence of this emotion in the comments.

The emotion "Fun" is similar in four of the eight destinations (Cuevas de Fuentes de León, Barruecos Natural Monument, Cornalvo Nature Reserve and Tagus Internacional Nature Reserve), with a presence of 7% in each. Suppose Monfragüe National Park (with a 4% presence of this emotion over the total) wants to increase this value. In that case, this destination should promote its gastronomic culture and the quality and beauty of its landscape.

On the other hand, the highest presence of "Anticipation" has been observed in Cueva del Castañar Natural Monument, Tagus Internacional Nature Reserve, Cuevas de Fuentes de León, and Barruecos Natural Monument, with values between 7% and 6%. However, Monfragüe National Park only has 4% of this emotion

in the comments. Therefore, to increase this emotion, the services offered should improve the quality and timing of services.

Finally, "surprise" is the lowest emotion observed in the main destinations in the Community of Extremadura. Seven register 3% of this emotion, while Monfragüe National Park is at 2%. In these destinations, the words "unique", "surprise", and "luck" appear regularly. Therefore, to increase the presence of this emotion, this paper proposes to improve the comparative advantages of each destination.

Table 3
Presence of positive emotions over total number of comments

Natural space	Trust	Fun	Anticipation	Surprise
Monfragüe National Park	5%	4%	4%	2%
Cueva del Castañar Natural Monument	9%	6%	7%	3%
Cueva de Fuentes de León Natural Monument	8%	7%	6%	3%
Barruecos Natural Monument	7%	7%	6%	3%
Mina la Jayona Natural Monument	7%	5%	5%	3%
Cornalvo Nature Reserve	8%	7%	5%	3%
Tagus Internacional Nature Reserve	7%	7%	6%	3%
Garganta de los Infiernos Nature Reserve	6%	5%	5%	3%

Source: Own elaboration.

Table 4
Presence of negative emotions over total number of comments

Natural space	Anger	Disgust	Fear	Sadness
Monfragüe National Park	2%	1%	2%	2%
Cueva del Castañar Natural Monument	2%	1%	1%	2%
Cueva de Fuentes de León Natural Monument	1%	1%	2%	2%
Barruecos Natural Monument	0%	1%	1%	2%
Mina la Jayona Natural Monument	2%	1%	3%	3%
Cornalvo Nature Reserve	2%	2%	3%	3%
Tagus Internacional Nature Reserve	1%	1%	2%	2%
Garganta de los Infiernos Nature Reserve	2%	2%	4%	4%

Source: Own elaboration.

On the other hand, the analysis of negative sentiments shows that the emotion "Anger" is around 1% in each destination. The most significant value is observed at Barruecos Natural Monument, which is around 0%, indicating an absence of this sentiment in this destination. Word clouds have detected a close relationship between noise, shots, and fear, particularly in destinations such as Monfragüe National Park. Barruecos Natural Monument does not detect the presence of these words, making it a reference for reducing this emotion.

Similarly, "Repulsion" presents the same values as the previous emotion, between 1% and 2%. Garganta de los Infiernos Nature Reserve and Cornalvo Nature Reserve are the destinations with the highest presence of this emotion. The high presence of this emotion is caused by litter, smells, and dirt, which eventually appear in the comments of these natural spaces. To reduce this emotion, the tasks and courses should focus on improving cleanliness and hazards in these destinations.

Another negative emotion to reduce is "fear". The value is higher than the two previous emotions. Garganta de los Infiernos Nature Reserve and Mina la Jayona Natural Monument are the destinations where this emotion is most frequent, with 4% and 3%, respectively. On the other hand, Cueva del Castañar Natural Monument and Barruecos Natural Monument have the lowest value (1%). The main difference between the destinations is that the comments include words like "risk", "caution" and "fire". To reduce the presence of this emotion, these destinations should increase the number of risk signals and eliminate all hazards.

Finally, "Sadness" is between 3% and 4% in these destinations. Garganta de los Infiernos Nature Reserve has the highest value with 4%. In contrast, Monfragüe National Park, Cueva del Castañar Natural Monument, Cuevas de Fuentes de León, Barruecos Natural Monument and Tagus Internacional Nature Reserve have 2%.

5. Conclusions

The results have verified the initial hypotheses. Firstly, the comments have shown that sentiments and emotions have affected online reputation, depending on whether these sentiments have been positive or negative. Secondly, each user's experience is different due to the objective facts experienced in each of the natural destinations of the Community of Extremadura.

The destinations with the best presence of positive sentiments also have a low presence of negative comments. Regarding feelings, Cueva del Castañar Natural Monument and the Cuevas de Fuentes de León are the destinations that best manage sentiments and emotions. At the same time, the Monfragüe National Park shows the worst level.

These natural destinations can increase and improve the presence of emotions such as "trust", "fun", "anticipation", or "surprise" through the selection of a destination as a reference where a positive visitor experience has a high presence.

Based on the above observations, this paper proposes some improvements to equalise, increase or decrease the presence of sentiments and emotions to improve (i) the online reputation of tourist destinations and (ii) user experience, which can be summarised as follows:

Firstly, the leading natural destinations should equalise the presence of "Surprise" to the other emotions. A high presence of this emotion can increase the total number of positive comments in the short and medium term. In the same way, if destinations manage to reduce the presence of emotions such as "fear" or "sadness", negative comments will be significantly reduced.

Similarly, the analysis shows a possibility of increasing positive sentiments. Monfragüe National Park and Cueva del Castañar Natural Monument can improve some emotions like "Fun" and "Anticipation", but the best way to obtain a balance with the rest of them is to assign the emotion "Confidence" as the best reference (see Figure 3 to receive more information).

Moreover, this paper recommends that Cuevas de Fuentes de León, Barruecos Natural Monument and Cornalvo Nature Reserve should increase the presence of "Anticipation". At the same time, Mina Jayona and Garganta de los Infiernos Nature Reserve should improve the presence of "Fun" and "Surprise" to significantly increase the total number of positive comments.

Words like "rubbish" and "dirt" have been mentioned multiple times. Natural spaces should reduce or eliminate this problem, associated with the emotion "Repulsion" and negative sentiment. This problem should be solved with a high priority because some destinations have especially important tourist or governmental certifications in the international community.

Cuevas de Fuentes de León and Barruecos Natural Monument have a low presence of the emotion "Anger". These destinations can be a reference for Monfragüe National Park, Cueva del Castañar Natural Monument, Cornalvo Nature Reserve and Tagus Internacional Nature Reserve to reduce the presence of this emotion and the total negative feelings.

The results show that online reputation is affected, and user experiences differ in each tourism destination. These differences have several theoretical and practical implications.

One theoretical implication is given through the theories of consumer and destination perception, which, based on eWOM, allows a small number of highly positive or negative comments to have a disproportionate impact on the reputation of a destination. According to the literature review, any variation in the presence of sentiment and emotions that increases online reputation by 1% will positively improve each tourism destination's average daily rate, occupancy levels and profit margins.

Similarly, the practical implications derived from the analysis of objective opinions (not influenced by personal factors such as climate or company but influenced by the quality of the destination) in the comments allow for the design of strategies to manage online reputation. These strategies can maximise the positive impact and minimise the effect of negative comments, understand trends and patterns, and identify areas of improvement and strengths of the destination.

Finally, these results highlight the crucial importance of proactively managing online reputation and user experiences in tourism destinations to understand the influence of comments on visitor perceptions and the economic results of the destination. The ability to implement strategies that encourage positive reviews and mitigate the impact of negative comments can significantly affect occupancy, fares, and profitability. In a digital environment where the voice of the consumer carries significant weight, online reputation is an invaluable asset for the sustained success of any tourism destination.

References

Álvarez-Carmona, M.A., Aranda, R., Arce-Cárdenas, S., Fajardo-Delgado, D., Guerrero-Rodríguez, R., López-Monroy, A.P., Martínez-Miranda, J., Pérez-Espinoza, H., & Rodríguez-González, A. (2021). Overview of Rest-Mex at IberLEF 2021: Recommendation System for Text Mexican Tourism. *Procesamiento del Lenguaje Natural, 67*, 163-172. http://dx.doi.org/10.26342/2021-67-14

Álvarez-Carmona, M.A., Aranda, R., Rodríguez-González, A.Y., Fajardo-Delgado, D., Sánchez, M.G., Pérez-Espinosa, H., Martínez-Miranda, J., Guerrero-Rodríguez, R., Bustio-Martínez, L., & Díaz-Pacheco, A. (2022). Natural language processing applied to tourism research: A systematic review and future research directions. *Journal of King Saud University - Computer and Information Sciences, 34*(10), 10125-10144. https://doi.org/10.1016/j.jksuci.2022.10.010

Anderson, C. (2012). The impact of social media on lodging performance. *Cornell Hospitality Report, 12*(15), 1-12.

Bordoloi, M., & Biswas, S.K. (2023). Sentiment analysis: A survey on design framework, applications, and future scopes. *Artificial Intelligence Review, 56*, 12505-12560. https://doi.org/10.1007/s10462-023-10442-2

Bulchand-Gidumal, J. (2022). Impact of artificial intelligence in travel, tourism, and hospitality. In Z. Xiang, M. Fuchs, U. Gretzel, & W. Höpken (Eds.), *Handbook of e-Tourism*. Springer. https://doi.org/10.1007/978-3-030-48652-5_110

Chatterjee, P. (2001). Online reviews – Do consumers use them? *Advances in Consumer Research, 28*, 129-134.

Chhetri, P., Arrowsmith, C., & Jackson, M. (2004). Determining hiking experiences in nature-based tourist destinations. *Tourism Management, 25*(1), 31-43. https://doi.org/10.1016/S0261-5177(03)00057-8

Dhuria, S. (2015). Sentiment analysis: An approach in natural language processing for data extraction. *International Journal of Innovations in Engineering and Technology, 2*(4), 27-31.

Ekman, P. (1992). An argument for basic emotions. *Cognition Emotion, 6*(3-4), 169–200. https://doi.org/10.1080/02699939208411068

Ekman, P., & Friesen, W.V. (2003). *Unmasking the face. A guide to recognizing emotions from facial clues.* Malor Books.

Escortell Pérez, M.A., Giménez Fayos, M., & Rosso, P. (2017). El impacto de las emociones en el análisis de la polaridad en textos con lenguaje figurado en Twitter [The impact of emotions on polarity analysis in texts with figurative language on Twitter]. *Procesamiento del Lenguaje Natural, 58*, 85-92.

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software, 25*(5), 1–54. https://doi.org/10.18637/jss.v025.i05

Fellows, I. (2018). *Wordcloud: Word clouds*. https://cran.r-project.org/web/packages/wordcloud/wordcloud.pdf

Frijda, N.H. (1988). The laws of emotion. *American Psychologist, 43*(5), 349–358. https://doi.org/10.1037/0003-066X.43.5.349

García-Pablos, A., Cuadros, M., & Linaza, M.T. (2016). Automatic analysis of textual hotel reviews *Information Technology & Tourism, 16*(1), 45-69. https://doi.org/10.1007/s40558-015-0047-7

Georgescu, M.R., & Bogoslov, I.A. (2019). Importance and opportunities of sentiment analysis in developing e-learning systems through social media. *DIEM: Dubrovnik International Economic Meeting, 4*(1), 83-93.

Godbole, M., Agarwal, A., & Sahay, B.S. (2021). Application of AI/ML/NLP technology into the business process modelling. *International Journal of Advanced Research in Engineering and Technology (IJARET), 12*(5), 37-50. https://doi.org/10.34218/IJARET.12.5.2021.004

González-Barahona, P., Barrena-González, J., Lagar-Timón, D., Lozano-Parra, J., & Pulido-Fernández, M. (2021). Identification of tourist resources for design thematic routes in the natural corridor of Armorican Arch of Cáceres (Extremadura, SW Spain). *International Journal of Geoheritage and Parks, 9*(1), 69-81. https://doi.org/10.1016/j.ijgeop.2020.12.001

Gretzel, U. (2006). Consumer generated content: Trends and implications for branding. *e-Review of Tourism Research (eRTR), 4*(3), 9-11.

Guerrero-Rodríguez, R., Álvarez-Carmona, M.Á., Aranda, R., & López-Monroy, A.P. (2023). Studying online travel reviews related to tourist attractions using NLP methods: The case of Guanajuato, Mexico. *Current Issues in Tourism, 26*(2), 289-304. https://doi.org/10.1080/13683500.2021.2007227

Hernández Mogollón, J.M., Pasaco González, B.S., & Campón Cerro, A.M. (2020). Aplicación del enfoque experiencial a la innovación de marketing de destinos turísticos: Recomendaciones para su implementación en Ecuador [Application of the experiential approach to marketing innovation in tourist destinations: Recommendations for its implementation in Ecuador]. *Innovar, 30*(77), 63-76. https://doi.org/10.15446/innovar. v30n77.87429

Hernandez-Maskivker, G., Fornells, A., Teixido-Navarro, F., & Pulido, J.I. (2021). Exploring mass tourism impacts on locals: A comparative analysis between Barcelona and Sevilla. *European Journal of Tourism Research, 29*, Article 2908. http://dx.doi.org/10.54055/ejtr.v29i.2427

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 168-177). Association for Computing Machinery. https://doi.org/10.1145/1014052.1014073

Huertas, A., Setó-Pàmies, D., & Míguez-González, M.-I. (2015). Comunicación de destinos turísticos a través de los medios sociales [Communication of tourist destinations through social media]. *Profesional De La información, 24*(1), 15-21. https://doi.org/10.3145/epi.2015.ene.02

Jafari, P., Al Hattab, M., Mohamed, E., & AbouRizk, S. (2021). Automated extraction and time-cost prediction of construction using natural language processing and simulation. *Applied Sciences, 11*(13), Article 6188. https://doi.org/10.3390/app11136188

Jockers, M. (2015). *Syuzhet: Extract sentiment and plot arcs from text*. https://cran.r-project.org/web/packages/syuzhet/syuzhet.pdf

Khalil, F., & Pipa, G. (2022). Is deep-learning and natural language processing transcending the financial forecasting? Investigation through lens of news analytic process. *Computacional Economics, 60*, 147-171. https://doi.org/10.1007/s10614-021-10145-2

Lang, D. (2018). *Wordcloud2: Create word cloud by 'htmlwidget'*. https://cran.r-project.org/web/packages/wordcloud2/wordcloud2.pdf

Liu, B. (2012). *Sentiment analysis and opinion mining. Synthesis lectures on human language technologies*. Springer. https://doi.org/10.1007/978-3-031-02145-9

Mariani, M., Baggio, R., Fuchs, M., & Höpken, 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

Marine-Roig, E., & Huertas, A. (2016). La comunicación de los destinos turísticos y sus marcas a través de los medios sociales [Communication of tourist destinations and their brands through social media]. In A. Huertas (Ed.), *La comunicación de los destinos turísticos y sus marcas a través de los medios sociales* (Capítulo 2). University Rovira i Virgili.

Mohammad, S. (2012). *Emotional tweets*. https://aclanthology.org/S12-1033.pdf

Mohammad, S., & Turney, P.D. (2010). *Emotions evoked by common words and phrases: Using mechanical Turk to create an emotion lexicon.* https://aclanthology.org/W10-0204.pdf

Mohammad, S., & Turney, P.D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence, 29*(3), 436–465. https://doi.org/10.48550/arXiv.1308.6297

Neuwirth, E. (2014). *RColorBrewer: ColorBrewer palettes*. https://cran.r-project.org/web/packages/RColorBrewer/RColorBrewer.pdf

Paredes, O., Melo, D., Guamán, A.R., García, M., & Guamán-Guevara, F. (2021). Innovative solutions of non-technological and technological nature needed to improve tourism services, Tungurahua Province, Ecuador. *Tourism: An International Interdisciplinary Journal, 69*(4), 559-577. http://dx.doi.org/10.37741/t.69.4.6

Parrott, W.G. (2001). *Emotions in social psychology: Essential readings*. Psychology Press Ldt.

Pennebaker, J.W., Francis, M.E., & Booth, R.J. (2001). *Linguistic inquiry and word count: Liwc 2001*. Lawrence Erlbaum Associates.

Plutchik, R. (1980). A general psychoe volutionary theory of emotion. In R. Plutchik (Ed.), *Theories of emotion* (pp. 3-33). Academic Press. https://doi.org/10.1016/B978-0-12-558701-3.50007-7

Plutchik, R., & Conte, H.R. (1997). *Circumplex models of personality and emotions*. American Psychological Association. https://doi.org/10.1037/10261-000

Powers, Z. (2021). The proliferation of the online sales of tourism activities. *Tourism: An International Interdisciplinary Journal, 69*(4). https://doi.org/10.37741/t.69.4.10

Prakash, P., & Aruna-Kumari, D. (2019). Business intelligence analytics using sentiment analysis-a survey. *International Journal of Electrical and Computer Engineering (IJECE), 9*(1), 613-620. http://dx.doi.org/10.11591/ijece.v9i1.pp613-620

Prameswari, P., Surjandari, I., & Laoh, E. (2017). Opinion mining from online reviews in Bali tourist area. In *Proceedings of the 2017 3rd international conference on science in information technology (ICSITech)* (pp. 226-230). IEEE. https://doi.org/10.1109/ICSITech.2017.8257115

Putri, F.P., Meidia, H., & Gunawan, D. (2019). Designing intelligent personalized chatbot for hotel services. In *Proceedings of the 2019 2nd international conference on algorithms, computing and artificial intelligence* (pp. 468–472). Association for Computing Machinery. https://doi.org/10.1145/3377713.3377791

Richins, M.L. (1983). Negative word-of-mouth by dissatisfied consumers: A pilot study. *Journal of Marketing, 47*(1), 68–78. http://dx.doi.org/10.2307/3203428

Sánchez-Martín, J.M., Reginfo-Gallego, J.I., & Martín-Delgado, L.M. (2018). Tourist mobility at the destination toward protected areas: The case-study of Extremadura. *Sustainability, 10*(12), Article 4853. https://doi.org/10.3390/su10124853

Sánchez-Martín, J.M., Rengifo-Gallego, J.I., & Blas-Morato, R. (2019). Hot spot analysis versus cluster and outlier analysis: An enquiry into the grouping of rural accommodation in Extremadura (Spain). *ISPRS International Journal of Geo-Information, 8*(4), 176. https://doi.org/10.3390/ijgi8040176

Sánchez-Martín, J.M., Rengifo-Gallego, J.I., & Sánchez-Rivero, M. (2020). Protected areas as a centre of attraction for visits from World Heritage Cities: Extremadura (Spain). *Land, 9*(2), Article 47. https://doi.org/10.3390/land9020047

Sánchez-Oro, M.S.O., García, Y.G., & Vecino, N.R. (2021). *La digitalización de las empresas turísticas extremeñas* [The digitalization of Extremadura tourism companies]. https://www.turismoextremadura.com/.content/observatorio/2021/EstudiosYMemoriasAnuales/Panel_digitalizacion_empresas_turisticas_2021.pdf

Shaikh, A., & Kulkarni, S.B. (2020). Natural language processing applications for tourism sector. *Journal of Computer Engineering, 22*(1), 27-35. https://doi.org/10.9790/0661-2201032735

Stone, P.J., & Hunt, E.B. (1963). A computer approach to content analysis: Studies using the General Inquirer system. In *Proceedings of the May 21-23, 1963, spring joint computer conference* (pp. 241-256). Association for Computing Machinery.

Torrado Rodríguez, G., & Blanco Gregory, R. (2018). Uso de Internet y redes sociales para la práctica turística. Caso de Extremadura [Use of the Internet and social networks for tourism practice. Case of Extremadura]. *Almenara, 11*, 66-82.

Tussyadiah, I., & Miller, G. (2019). Perceived impacts of artificial intelligence and responses to positive behaviour change intervention. In J. Pesonen & J. Neidhardt (Eds.), *Information and communication technologies in tourism* (pp. 359-370). Springer. https://doi.org/10.1007/978-3-030-05940-8_28

Yang, S.B., Shin, S.H., Joun, Y., & Koo, C. (2017). Exploring the comparative importance of online hotel reviews' heuristic attributes in review helpfulness: A conjoint analysis approach. *Journal of Travel and Tourism Marketing, 34*, 963-985. https://doi.org/10.1080/10548408.2016.1251872

Submitted: October 18, 2022
Revised: July 23, 2024
Accepted: October 17, 2024