International Journal of Hospitality Management 104 (2022) 103245

Contents lists available at ScienceDirect

International Journal of Hospitality Management

journal homepage: www.elsevier.com/locate/ijhm

Impact of the rating system on sentiment and tone of voice: A Booking.com and TripAdvisor comparison study

Paulo Rita [a,*], Ricardo Ramos [b,e], Maria Teresa Borges-Tiago [c], David Rodrigues [d]

[a] NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Portugal
[b] Instituto Politécnico de Coimbra, ESTGOH, Oliveira do Hospital, Portugal Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal
[c] Department of Business, University of the Azores, Portugal
[d] NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Portugal
[e] CICEE – Centro de Investigação em Ciências Económicas e Empresariais, Universidade Autónoma de Lisboa, Lisboa, Portugal

ARTICLE INFO

Keywords:
Online reviews
Sentiment analysis
Hotel category
Nationality
TripAdvisor
Booking.com

ABSTRACT

Online reviews have bridged the gap between traditional word-of-mouth and viral communication, influencing peer's decision processes. Analyzing tourists' online reviews helps hotels address overall customer (dis)satisfaction. Using sentiment analysis to understand reviewers' satisfaction and analyze the voice tone and expressed feelings, this research attempts to enlarge hotel, platform, and tourist trilogy's knowledge. A total of 38,292 reviews posted on Booking.com and TripAdvisor from 191 hotels were analyzed. Results indicated that the rating system influenced reviewer's sentiment, even though sentiment did not increase alongside the hotel category, leading to using a sterner tone of voice. Differences were acknowledged according to reviewers' nationality. The most positive feelings were expressed on TripAdvisor linked to staff-tourist encounters whereas Booking.com presented more negative feelings, especially linked to overcharging and billing issues. These outcomes can guide managers in establishing priorities to improve service and meet customers' expectations.

1. Introduction

The relevance of tourism for today's economy is undeniable, and the forecasts only seem to anticipate lasting importance for the coming decades (Kontogianni and Alepis, 2020). Hotels are essential assets in the tourism sector. The relationship between the performance of the tourism sector and hotels is a close one, being that the first influences the latter both directly and indirectly (Mucharreira et al., 2019).

The growth in Information and Communication Technologies impacted the sector due to the emergence of social media and online platforms where consumers share their experiences (Bizirgianni and Dionysopoulou, 2013). This new reality changed the way hotels managed their communication in an online environment (Casado-Díaz et al., 2020) and revolutionized the way consumers plan and book their trips. 90% of the consumers claim that they consider electronic word-of-mouth (eWOM) before performing their final purchase decision (Akhtar et al., 2019).

The characteristics of an online platform is considered one of the main determinants for customer behavior (Parboteeah et al., 2009). Various attributes, such as the rating system can directly influence a reviewer behavior (Mariani and Borghi, 2018; Mellinas et al., 2016). Consequently, influences their narrative around the hotel's experience and their perception of the service quality. In turn, since this information is perceived as trustworthy by other consumers, it can influence expectations of future customers before booking their reservation. Therefore, it seems necessary to consider the characteristics on an online tourism plaform as an essential tool to measure service quality (Brochado et al., 2019; Oliveira et al., 2019). Previous studies aimed to understand the impact of an online travel platform (OTP) rating system on the consumer behavior. For instance, Cena et al. (2017) aimed to understand the impact of rating scales on user ratings. Another study developed by Chen (2017) had the objective to uncover the influence of rating systems on users perception of information quality, cognitive effort, enjoyment and continuance intention. Casaló et al. (2015) aimed to investigate how hotel ratings influence customers regarding an hotel experience and booking intentions. However, so far, no study aimed to understand the impact of OTPs rating systems on the guests' sentiment and tone of voice of their reviews.

Therefore, this investigation aims to understand if the Booking.com and TripAdvisor's rating system influence the reviewers' sentiment and

* Corresponding author.
E-mail address: prita@novaims.unl.pt (P. Rita).

https://doi.org/10.1016/j.ijhm.2022.103245
Received 7 May 2021; Received in revised form 11 April 2022; Accepted 6 May 2022
Available online 25 May 2022
0278-4319/

tone of voice according to the hotel category and customers' nationality. Simultaneously, it aims to comprehend if the hotel category and nationality influence the reviewers' sentiment.

To achieve this objective, a sentiment analysis (SA) of 38,292 online reviews of 191 Lisbon hotels was conducted. With sentiment analysis, it was possible to retrieve the subjective information in the review, like opinions, appraisals, emotions, or attitudes regarding specific service features, which could not be achieved when considering only quantitative metrics (e.g., rating). The content can be classified as positive, negative, or neutral, and within each of these categories, it can specify the tone of voice adopted. The tone of voice is defined as "*the way a person is speaking to someone*" (Merriam-Webster, 2021) when expressing its feelings towards the subject, experience, or person. In the digital context, the user's tone of voice is expressed through its writing, as Xu (2020) noted when analyzing the customer emotion embedded in hotel online reviews. Data will be analyzed through text mining technique, specifically sentiment and tone of voice analysis, and three ANOVA and one Student's t-test will be conducted to confirm the raised hypotheses. This study makes a conceptual and practical increment to academia and industry by uncovering the influence of OTPs rating system have on the guests' sentiment and tone of voice, which not only can improve firms adequancy to customers preferences, but also can catered emotion-based knowledge to AI readers.

2. Literature review and research hypotheses

An OTP establishes the connection between the tourism service providers and consumers, providing consumers the convenience to purchase tourism services (Zhang et al., 2021) or access to information considered relevant for the customer (Furtado et al., 2022). In each OTP there is a rating system that helps the customer in its decision making (Israeli, 2002). However, OTPs adopt different scales and rating systems. For instance, Boooking employes a 2.5–10 scale, considering the mean of six individual dimensions (Martin-Fuentes et al., 2018a). TripAdvisor adopts a 1–5 scale and other dimensions can be evaluated afterward. The latter are not considered in the former evaluation. These rating systems influence the perceptions of information quality (Chen, 2017) that, in turn, will influence customer behavior (Mariani and Borghi, 2018; Mellinas et al., 2016).

Customer satisfaction results from a comparison between what was experienced and the initial expectations, a pivotal dimension to promote an increase in hotel demand (Xu and Li, 2016). Satisfaction is a function of consumers' expectations before the purchase and the perceived disconfirmation of those expectations after consuming the product (Oliver, 1980); while, expectations do not influence satisfaction, as previously believed (Anderson and Sullivan, 1993). Disconfirmation plays its role in the equation, but these authors proposed service quality as a new variable and argued that service quality influences satisfaction and repurchase intentions. Moreover, failing to meet the consumers' expectations has a higher impact on customer satisfaction than exceeding them, which indicates the importance of continuously providing a high-quality service (Anderson and Sullivan, 1993).

Depending on the hotel category, different attributes are valued by clients. In low-end hotels, the accommodation infrastructure and employee expertise improve customer satisfaction, whereas, in mid-range hotels, clients appreciate the security and room quality (Nunkoo et al., 2020). For high-end hotels, clients highlight the customer-to-customer interaction and reduce the waiting time (Nunkoo et al., 2020) while also paying attention to cleanliness, comfort, pleasant view, and amenities (Padma and Ahn, 2020). Sleep quality is the most relevant attribute to improve customer satisfaction in low-end hotels (Bi et al., 2020). The consumers' most valued attributes are also influenced by the type of the trip - business vs. leisure, and by the guest's nationality (Bi et al., 2020; Galati and Galati, 2019).

The competition in the hotel industry is fierce (Lee, 2015), and firms can reach competitive advantage by choosing a strategy to compete through price or differentiation. In the first, firms compete to offer the service at a competitive price, while in the latter, firms create a service with unique features and higher quality (Xia et al., 2020). Service quality is vital to promote customer satisfaction and word-of-mouth, increasing the repurchase intentions of the consumers (Roy et al., 2018) while developing customer loyalty, which allows for standing apart from the competitors (Roy et al., 2018; Saleem and Sarfraz Raja, 2014; Xia et al., 2020). Whereas competing through low prices might hinder the hotels' ability to ensure long-term sustainability (Kandampully and Suhartanto, 2000). It can be easily copied by other players in the market (Lawton, 1999). Providing high-quality service means standing apart from the competitors and developing customer loyalty (Xia et al., 2020). Thus, service quality and customer satisfaction are the main goals of any service provider (Sureshchandar et al., 2002) and the foundations of customer loyalty (Saleem and Sarfraz Raja, 2014). In turn, a more loyal customer to a particular brand is more likely to repurchase from that brand (Chinomona and Maziriri, 2017). The hotel industry is a clear example where the customer acquisition price is higher than the cost of retaining the current customers and where several initiatives must be developed to promote customer satisfaction and loyalty (Dominici and Guzzo, 2010).

Considering the benefits mentioned above from ensuring customer satisfaction and promoting a high-quality service and the challenges that services pose when compared with products, it is vital to consider the customers' opinions when measuring service quality. Hotel star-rating system establishes a classification for the same type of accommodations according to their common physical and service characteristics (Huang et al., 2018). For instance, if a hotel provides high-quality service and performance, it justifies a higher star rating. The star rating is considered a critical metric to estimate the quality of the hotel and as the first guidance for consumers (Huang et al., 2018; Mohsin et al., 2019). It can influence expectations, which in turn influence satisfaction. The expectancy-disconfirmation theory (Oliver, 1980) explains that satisfaction is higher if the experience is higher than expected. In the hotel industry, Rhee and Yang (2015) suggest clients' different expectations considering the different ratings, while Bulchand-Gidumal et al. (2011) suggest a positive relationship between star-rating and the customers' satisfaction. In a study developed by Qu et al. (2000), it was found that the level of satisfaction increased with the number of stars. These results were confirmed by Martin-Fuentes (2016) that found that the hotel's service quality can be derived from the hotel category, while satisfaction around the service provided is related to each hotel's number of stars. Since the rating systems may influence customer behavior in the online environment (Mariani and Borghi, 2018; Mellinas et al., 2016), this study hypothesized the following:

2.1. H1a: Depending on the rating system, the sentiment varies according to the hotels' category

Most earlier literature on tourism focused on expectation formation and linked to tangible facilities and intangible services offered by different hotel categories (Ariffin and Maghzi, 2012; Costa et al., 2004). At the most fundamental level, the customer expectations can be considered guidelines for service evaluation (Yang and Aggarwal, 2019). Afterward, the hotel experience can be shared online on multiple platforms, assuming a positive or negative valence (Yen and Tang, 2015). As pointed out by Kim et al. (2019), the higher the level of expectations consumers have regarding the hotel facilities and services, the higher the probability of feeling less satisfied if the service provided does not meet their initial expectations. In this case, consumers will tend to adopt a less positive tone of voice in their reviews. Nonetheless, suppose hotel managers are able to provide a level of performance that equals or surplus customers' expectations. In that case, the outcome will be positive, and the tone of voice adopted in the review. Rhee and Yang (2015) found that reviews can be classified in a two by two matrix representing these different trade-offs and types of comments made. Xu

(2020), analyzing Expedia.com and Booking.com, found evidence that consumers expressed emotions in online reviews when pleased with the experience, denoting different tones of voice in the writing content. Moreover, since the customer behavior is influenced by the rating system (Mariani and Borghi, 2018; Mellinas et al., 2016), combined with customer experience expectations, these arguments led to the hypothesis H1b:

2.2. H1b: Depending on the rating system, the tone of voice varies according to the hotels' category

Culture significantly affects customer behavior (Wang et al., 2019). Customers with different cultural background may have different preferences and perceptions. For instance, according to the Hofstede's cultural theory, customers from countries with high uncertainty avoidance, such as Greece or Portugal, rely on word of mouth from trusted scores. In turn, customers from countries with low uncertainty avoidance, such as UK or USA, believe in commercial marketing sources (Litvin, 2019). Cultural background influence not only uncertainty avoidance, but also power distance, individualism, masculinity, long term orientation, and indulgence (Jia, 2020). These cultural differences influence perception of a hotel experience that, in turn, is reflected in the evaluation shared in an OTP.

Accordingly, different nationalities lead to different evaluations. Several studies attempted to examine the role that nationality plays in satisfaction. Ngai et al. (2007) found differences in the complaint behavior between Asian and non-Asian guests. Asian guests were more likely to engage in private complaining than non-Asian guests, therefore less vocal and persistent about their complaints. As the authors postulated, the critical insight is the influence that culture has on complaint behavior. Kim et al. (2018) also found differences in the reviews posted by customers from different cultural backgrounds, claiming that Westerners' reviews tend to be more positive and analytical.

Moreover, guests from the United States tend to value more opinions from their compatriots (Kim et al., 2018). Based on these studies, we can assume that the sentiment depends on the reviewers' nationality. Considering that the rating system of the OTP influence customer behavior (Akram et al., 2018; Kwon et al., 2002), the following research hypothesis was created:

2.3. H2: Depending on the rating system, the sentiment varies according to the nationality of the reviewer

Mellinas et al. (2015) found that Booking.com employed a 2.5–10 scale rather than a 1–10 scale, inflating the final score. Moreover, this inflation is more significant in hotels with lower scores (Mellinas et al., 2016). Mariani and Borghi (2018) supported this fact, confirming the left-skewness of the review scores on Booking.com, and declared that the distribution is more left-skewed as the hotel category increases.

Booking.com and TripAdvisor's classification system is different (Martin-Fuentes et al., 2018a). In Booking.com, the rating is calculated by considering the mean of six individual dimensions. In contrast, in TripAdvisor, the score is attributed directly, and the particular dimensions can be evaluated independently afterward, not being taken into account for the final score. Therefore, hypothesis three was developed:

H3a. The sentiment in Booking.com is expected to be higher.

Borges-Tiago et al. (2021) found some similarities between Booking.com and TripAdvisor when looking at the content of the positive tone of voice reviews. Hence, the following hypothesis is proposed:

H3b. The tone of voice in Booking.com and TripAdvisor is expected to be the same.

3. Methodology

3.1. Sample

The sample considered in this study comprised the guests who stayed in at least one of the 204 Lisbon hotels registered in the Registo Nacional de Turismo (RNT: https://registos.turismodeportugal.pt/) that posted their opinion in both Booking.com and TripAdvisor. RNT is a platform that includes information about tourism entities in Portugal, such as hotels and travel agencies (Turismo de Portugal, 2020). Since 2017, Portugal won the World Travel Award as Europe's leading destination, and Lisbon won Europe's Leading City Break Destination 2019 (WTA, 2020). The Lisbon metropolitan area of its capital city had over 8 million guests in 2019, 73% of whom were foreigners (INE, 2019), and has become an ideal destination for worldwide tourists (Cró and Martins, 2018).

Data were collected from both Booking.com and TripAdvisor. Booking.com is an OTA considered the world's leading service in hotel reservations (Mariani and Borghi, 2018; Mellinas et al., 2016). This online platform is part of Booking Holdings Inc. It has a market capitalization of over 85 billion US Dollars (Finance, 2021a) and is available in over 43 different idioms with an offer above the 29 million accommodations in more than 2.5 million properties (Booking.com, 2021). Booking.com is present in over 225 countries, with beyond 1.5 million reservations per day (Booking.com, 2020). TripAdvisor is one of the top players on a web-based consumer opinion platform (Martin-Fuentes, 2016). It is also a public-traded company and has a market capitalization of nearly 5 billion US Dollars (Finance, 2021b). On average, it has 463 million visits per month and over 860 million reviews on more than 8.7 million accommodations, restaurants, airlines, cruises, and other experiences to make the best of their trip. This platform is accessible in 49 markets in 28 different idioms to spread eWOM and help customers plan a trip (TripAdvisor, 2021).

The hotels reviewed by customers on both platforms were compared. The ones with no correspondence in one of the platforms were removed. Therefore, 12 hotels were eliminated and led to a final dataset of 38,292 reviews from 191 hotels.

3.2. Data collection

This study's data collection process comprised the extraction and compilation of online reviews from Booking.com and TripAdvisor for the selected sample. A Python script was developed using Anaconda ® - an open-source software with over 20 million users worldwide (Anaconda, 2020) – to *web scrape* every URL.

The first step was to identify the URL of the hotel on both platforms. Second, the Web pages were *web scraped*, and the results were exported into an individual *CSV* file per platform. Finally, the individual files were compiled into a unique *CSV* file that comprised the final dataset. The script used *BeautifulSoup*, a Python library, to retrieve data from HTML. Each review collected for this research included a rating from 1 to 5, in the case of TripAdvisor, and 1–10 in Booking.com. Additionally, the reviewer's opinion, name, and nationality were collected, and hotels' name, location, rating, and the number of reviews. The variables extracted from both platforms are presented in Table 1.

The first two variables – hotel name and hotel location – cross-checked the information between *RNT*, TripAdvisor, and Booking.com.

Table 1
Variables extracted from Booking.com and TripAdvisor.

Variables extracted from Booking.com and TripAdvisor		
Hotel name	Language of the review	Rating of the review
Hotel location	Date of the review	Title of the review
Hotel rating	Author of the review	Body of the review
Number of reviews	Nationality of the reviewer	

After assuring the correspondence, the hotel category and city obtained in *RNT* were considered in the analysis. Given that the rating scales used in TripAdvisor and Booking.com differed, the ratings were normalized (Martin-Fuentes et al., 2018a) to ensure comparability between both scales. The normalization was performed using a Python script and the *Scikit learn* library. The sample comprised 38,292 reviews in 191 hotels located in Lisbon, Portugal, summarized in Table 2.

The number of reviews posted on Booking.com was 2.48 times higher than the reviews on TripAdvisor. This number difference is expected since the Booking.com approach to stimulating reviews is more intrusive (Martin-Fuentes et al., 2018a). Moreover, in Booking.com, the minimum number of reviews per hotel was 8, whereas, in TripAdvisor, that number is 1. The reviewers were from 131 different countries, 87 of which were shared on both platforms. In TripAdvisor, there were 109 different nationalities among the reviewers. In 2672 reviews, the nationality of the reviewer was not available.

3.3. Data analysis

For data analysis, a text mining approach was conducted, specifically a SA. A *Student's t-test* and two one-way *ANOVA* were performed using IBM® SPSS® Statistics software to test the hypothesis. Text mining is used to extract knowledge from a substantial amount of non-structured textual data, uncovering patterns and relevant information (Ramos et al., 2019). SA analyzes people's opinions, evaluations, attitudes, and emotions towards products, services, or organizations, revealing a positive, neutral or negative reaction (Rita et al., 2020). SA can be applied to consumer's reviews (Calheiros et al., 2017) to obtain feedback from what can be improved by the seller to ensure a better consumer buying experience, to measure the positioning and awareness of a particular brand, and to identify opportunities for new products or services in the market, among others.

To conduct the SA, the comments were analyzed using Semantria, a SA software from Lexalytics. The software includes an Excel plugin that uses the reviews to perform the analysis. Each review is a document, and the sentiment conveyed through the review is classified as negative, neutral, or positive (Rita et al., 2020). Through the combination of natural language processing and machine learning, the reviews' classification varies according to the information expressed by words, concepts, or phrases within context (Pampulevski et al., 2020). Semantria software uses a proprietary sentiment library with an extensive collection of adjectives (e.g., good, awful) and phrases (e.g., nice place, terrible bed). To each is attributed a score (positive, neutral, or negative), according to their sentiment value (Öztürk and Ayvaz, 2018; Yu et al., 2021). The existing literature includes several studies that used Semantria (Rodrigues et al., 2016; Yu et al., 2021). Semantria also includes industry packs as features. Industry packs are industry-specific dictionaries developed to improve the overall sentiment in content related to each specific industry (Lexalytics, 2020). The "Hospitality – Hotels industry pack" was used in Semantria's plugin for Excel. The SA outcome through Semantria returns a polarity of sentiments accordingly to the scale represented in Table 3.

Table 2
Sample structure.

	Booking.com	TripAdvisor	Total
Number of reviews	27,289	11,003	38,292
Min. reviews per hotel	8	1	–
Max. reviews per hotel	866	499	–
Avg. reviews per hotel	142.87	57.61	–
Number of nationalities of the reviewers	112	109	
Number of 1-star hotels			6
Number of 2-star hotels			20
Number of 3-star hotels			53
Number of 4-star hotels			80
Number of 5-star hotels			32

Additionally, the WordStat, a SA software from Provalis, was used to unveil the dimensions that can emerge under these broad categories. The linguistic variety of the online reviews can lead to misclassifications, when applying a general sentiment word list, as shown by Loughran and McDonald (2011). These authors developed a scale that allows to assess the tone of voice used on online reviews, beyond simply classifying into negative, neutral or positive discourse. This approach has been used in several studies, mainly in the financial field, allowing a fine tuning of discourse when approaching machine and AI readership (Loughran and McDonald, 2016).

The process conducted for data collection and data analysis can be observed in Fig. 1.

According to the central limit theorem, the data distribution tends to be normal in the case of large samples (Ghasemi and Zahediasl, 2012). Therefore, this research assumes the normality of the data in each of the statistical analyses performed.

To test H1, two one-way ANOVA (Férnandez and Bedia, 2004; Martin-Fuentes et al., 2018b) were performed to measure whether there were statistically significant differences between the means of the hotel categories in terms of the sentiment score and the review rating after normalization. A one-way ANOVA was used to test H2, assessing whether statistically significant differences between each region's mean scores. Finally, to test H3, following the approach used by Martin-Fuentes et al. (2018a), a *Student's t*-test was performed considering the scale after normalization in both platforms. SPSS Statistics software (v.25, IBM SPSS) was used to conduct the experiments and Wordstat software to conduct content analysis.

4. Results and discussion

This research comprised two stages: in the first stage, a SA was performed on the sentiment conveyed through the reviews posted in Booking.com and TripAdvisor by Lisbon hotels' guests through Wordstat®Provalis Research®. Statistical analysis was then conducted through IBM® SPSS® Statistics to test the second stage's proposed hypotheses.

4.1. Sentiment Analysis

Sentiment analysis can comprehend two types of research approaches - lexicon-based and machine learning. The first was adopted, considering the positions and correlations of the words and unveiling the polarity of the content. In the first analysis, the recurrence of terms from all guests was analyzed to scrutinize the text's data. A total of 16,395 different terms were detected. The ten most frequent terms can be observed in Table 4.

The list includes terms related to the service's attributes, such as staff, breakfast, and clean. Some of the most frequent terms are consistent with previous research (Geetha et al., 2017) like "hotel" and "staff". Kandampully and Suhartanto (2000) also highlighted the importance of food, beverage, and house cleaning, which are connected with the terms "breakfast" and "clean". Guests frequently mentioned the word "staff", a factor found to be essential to influence the guests' overall satisfaction towards the service provided (Choi and Chu, 2001).

To assess the emotions and feelings reflected in tourists' reviews, a SA was conducted using a sentiment dictionary on Wordstat software

Table 3
Default sentiment scale used in Semantria Adapted from: Lexalytics (2020).

Sentiment	Range
Negative	< -0.05
Neutral	$[-0.05$ to $0.22]$
Positive	> 0.22

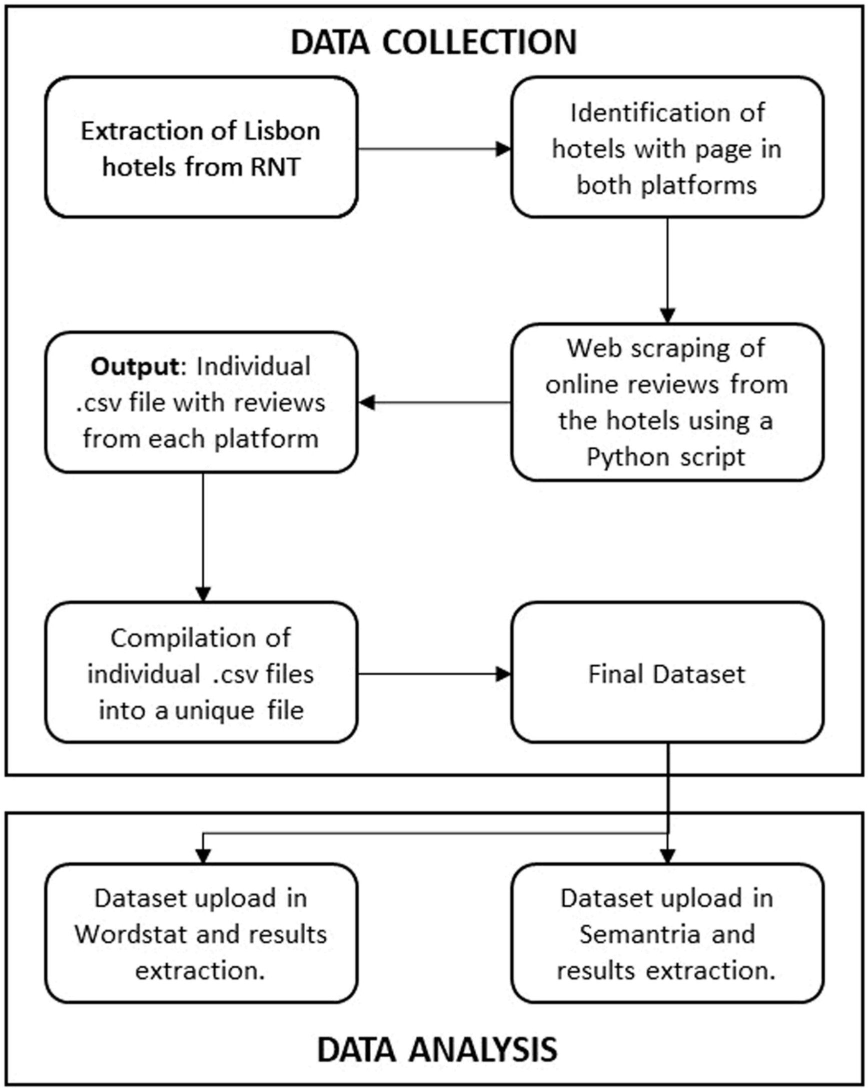


Fig. 1. Data collection and data analysis process.

Table 4
Top 10 most frequent words.

Word	Frequency
room	31,738
hotel	28,560
staff	20,487
location	18,336
breakfast	18,210
Good	15,024
Great	12,598
Stay	10,725
Clean	9,907
Nice	9,522

created for this study's purpose and composed based on the scale developed by Ekman (1992) and updated by Sabini and Silver (2005). This dictionary comprises seven emotional dimensions, reflecting 400 words linked to feelings associating words with emotions like happiness, love, surprise, sadness, anger, fear, and disgust. Then an aspect-based sentiment analysis was composed to understanding the aspects or features underlying or linked to the sentiments expressed. According to Do et al. (2019), aspect-based sentiment analysis comprehends three major tasks. The first consisted of opinion target extraction, that in the present case are the reviews extracted from both online travel platforms. For these authors, the study of sentiment analysis can be conducted as a document, sentence, and aspect or entity. In the present case, each review is considered as sentences. Thus the second step of the aspect-based sentiment analysis consists of breaking the comment into small and independent sentences, removing the stop words since they don't add value to the research, and conducting a spelling. Subsequently, the content is cataloged into two main categories, considered as aspect category detection: aspects or features categories and sentiment expressed are identified; and classified in terms of adoption frequency in the comments per source: Booking.com or TripAdvisor. For instance, the statement "The rooftop was a disappointment" contained under the category of aspects rooftop and in the type of sentiment disappointment/sadness. This technique is quite commonly used to analyze customer reviews shared on e-commerce sites, such as Amazon, or rating platforms such as Yelp (Do et al., 2019). In order to overcome the non-local dependency problems, that considers the possibility of two interconnect words being apart in a sentence, and the word sense disambiguation problems, that considers the possibility that in a sentence, two subsegments are presenting different sentiments, in the present work the sentiment analysis was conducted with full sentences (Nandwani and Verma, 2021).

Fig. 2 summarizes the outcome of these analyses for the top 4 emotional dimensions that covered 83.7% of all emotional discourse.

Love and happiness sentiments, accounted together, are more

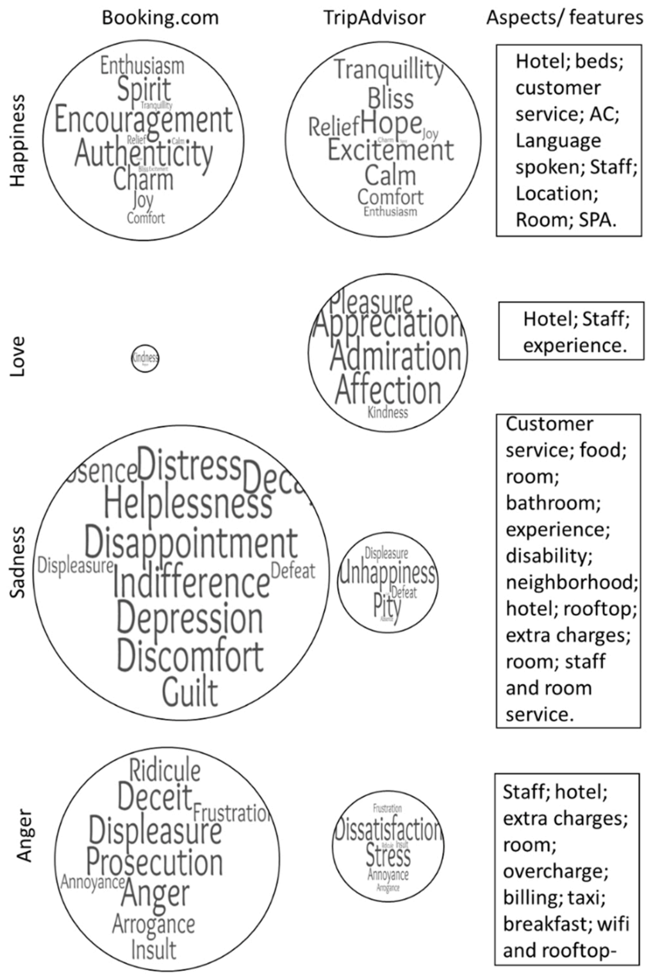


Fig. 2. Sentiment expressed linked to hotel service aspects or features.

present in TripAdvisor since the difference between Booking.com and TripAdvisor is slight in terms of happiness expression (Booking.com – 659; TripAdvisor – 641), and the love expressions in Booking.com are entirely residual. At the same time, anger and sadness are more present in Booking.com travelers' reviews. The outcome shows that tourist-employee interactions drive the most positive and more negative emotions. In this analysis of the emotional dimensions, some common feelings such as empathy appeared linked to hotel features but in the negative form. The anger is associated with the financial component since it comprehends billing issues, overcharged, and extra charges. Fig. 2 allows inferring that the tone of voice is not identical on both platforms.

The sentiment score acknowledged by the sentiment analysis ranged from −1.875 to 2.595. The number of positive reviews is overwhelming, in line with previous work (Zervas et al., 2021). However, there is a considerable difference of satisfied, neutral, and dissatisfied customers between both platforms (Table 5).

Table 5
Sentiment score characterization in TripAdvisor and Booking.com.

	Booking.com	TripAdvisor
Lowest sentiment score	-1.560	-1.875
Highest sentiment score	2.595	1.830
Relative number of reviews with a positive sentiment	54.84%	81.24%
Relative number of reviews with a neutral sentiment	29.82%	11.20%
Relative number of reviews with a negative sentiment	15.34%	7.56%

According to the achieved results, the sentiment of the reviewers had different behaviors in the two platforms. The five-star hotels had higher average sentiment in both platforms in the reviews: 0.296 and 0.4526 in Booking.com and TripAdvisor, respectively. Notwithstanding, in Booking.com, the three-star hotels were ranked second, followed by the four- and one-star hotels. On both platforms, the two-star hotels scored the worst in terms of average sentiment. In TripAdvisor, the one-star hotels were the second in terms of average sentiment, followed by the four- and three-star hotels, respectively (Table 6). This study's results are consistent with Geetha et al. (2017), who found that premium hotels' reviews tend to be more optimistic than budget hotels. Table 6 shows that, except for two-star hotels, the average sentiment increased as the hotel category also increased.

Americas (AME) scored the highest average sentiment in terms of the scores per region, whereas the Middle East and Africa (MEA) scored the lowest. Moreover, Asia and South Pacific also scored lower than the Americas, which is consistent with the research from Kim et al. (2018), who asserted that Westerners tend to be more positive in their online reviews than Eastern societies.

Nevertheless, based on this study, one could expect the standard deviation in the average sentiment to be smaller in the AME than in Asia and South Pacific (ASP). However, the standard deviation is similar in all regions.

Table 6 also shows that the average sentiment in TripAdvisor was consistently higher when compared with Booking.com. This finding aligns with the performance of TripAdvisor in research from Xiang et al. (2017), who reached the distribution of the sentiment score in TripAdvisor, Yelp, and Expedia for hotels in Manhattan. TripAdvisor also scored higher in terms of average sentiment score than the remaining platforms, and the distribution was also skewed towards the positive side.

This research also discovered that the review rating after normalization was higher in Booking.com than in TripAdvisor. Moreover, TripAdvisor's sentiment was always higher than in Booking.com – approximately 1.64 times more, on average. In the case of one-star hotels, albeit TripAdvisor's sentiment was 101% higher than in Booking.com, the average review rating after normalization was only 0.74% higher. Although the average sentiment was 67% higher in TripAdvisor in two-star hotels, the review rating was 9.54% smaller than Booking.com.

Different dimensions can emerge under the broad categories of positive, neutral, and negative sentiment on a review, reflecting the user' tone of voice (Loughran and McDonald, 2011). These authors added to the subdimensions of uncertainty, litigiousness, and strong modal words.

Applying a sentiment dictionary in WordStat with these dimensions, a graph correlation structure was found, with a $R^2 = 0.7806$, supporting the idea that there is space for different sentiment nuances within the positive and negative discourse (Fig. 3). The outcome validates that litigious and constraining comments are far more negative than those only linked with not-so-pleasant experiences.

Considering that the two platforms were under comparison, this analysis was conducted comparing both platforms (Fig. 4). The outcome produced found significant differences in what concerns constraining and litigious discourse.

Looking closer to the data, TripAdvisor presents a more relevant set of comments anchored in INTERESTING or SUPERFLUOUS opinions, while Booking.com has more NEGATIVE and UNCERTAINTY-related comments.

We conducted a multi-level binary SA to find if the hotel category and review rating influenced the tone of voice used in an individual review and to assess the existence of tone of voice differences between regions and rating systems adopted (Table 7). Regarding hotel star rating, an association between hotel category and tone of voice was observed. Once more, differences were confirmed related to constraining and litigious discourse, also complemented by superfluous

Table 6
Average sentiment per category, platform, and region.

	#	Avg. Sentiment (Std. Deviation)	AME		ASP		EUR		MEA		OTHERS[a]	
			Booking.com	TripAdvisor	Booking.com	TripAdvisor	Booking.com	TripAdvisor	Booking.com	TripAdvisor	Booking.com	TripAdvisor
1-Star	519	0.2255 (0.3333)	0.2542 (0.3749)	0.5604 (0.3088)	0.1357 (0.2696)	0.3936 (0.1504)	0.2061 (0.3277)	0.3437 (0.2542)	0.1709 (0.2908)			0.5328 (0.2515)
2-Star	2,865	0.1992 (0.3549)	0.2006 (0.3588)	0.2547 (0.3463)	0.1967 (0.3410)	0,2978 (0.2762)	0.1813 (0.3590)	0.3226 (0.3041)	0.1383 (0.3431)	0.3140 (0.2699)	0.4237 (0.1401)	0.3321 (0.353)
3-Star	9,140	0.2746 (0.3327)	0.2840 (0.3385)	0.3590 (0.2962)	0.2671 (0.3426)	0,2819 (0.2812)	0,2494 (0.33)	0.3992 (0.2861)	0.2111 (0.3567)	0.2940 (0.3634)	0.1103 (0.2069)	0.3873 (0.3086)
4-Star	17,920	0.2988 (0.3462)	0.2872 (0.3626)	0.4068 (0.2854)	0.2408 (0.3433)	0,3502 (0.2905)	0,2451 (0.3576)	0.4044 (0.2804)	0.2181 (0.3453)	0.4253 (0.2583)	0.1813 (0.4374)	0.4224 (0.3211)
5-Star	7,848	0.3717 (0.3343)	0.3457 (0.3574)	0.4786 (0.2623)	0.2944 (0.3337)	0,4187 (0.2749)	0,2841 (0.37)	0.4257 (0.2709)	0.2612 (0.3529)	0.4645 (0.2477)	0.2964 (0.6016)	0.4603 (0.3083)
TOTAL	38,292	0.2995 (0.3441)	0.2869 (0.3564)	0.4274 (0.2832)	0.2506 (0.3415)	0.3608 (0.2865)	0.2447 (0.3522)	0.4069 (0.2797)	0.22 (0.3498)	0.4179 (0.2745)	0.2116 (0.4051)	0.4289 (0.3171)

[a] In some reviews, the nationality of the reviewer is not disclosed, or it is impossible to identify the country. The category "Others" is created to allocate those reviews

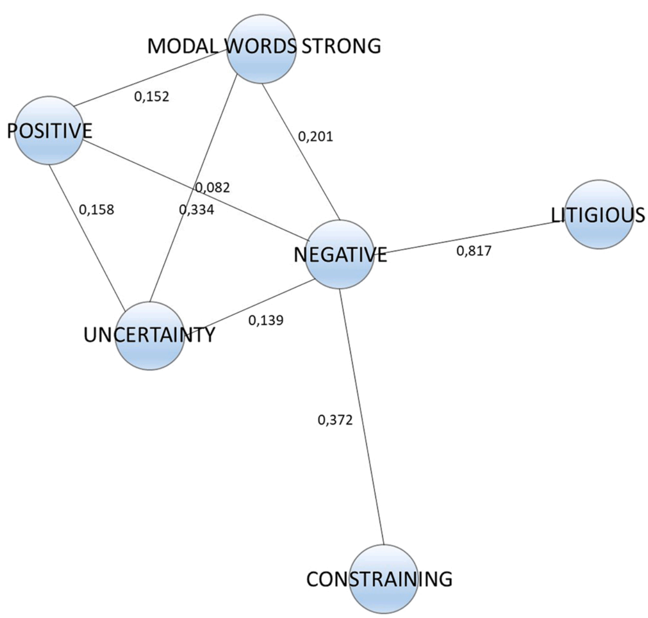


Fig. 3. Correlation graph between sentiment subdimensions in the sample.

comments.

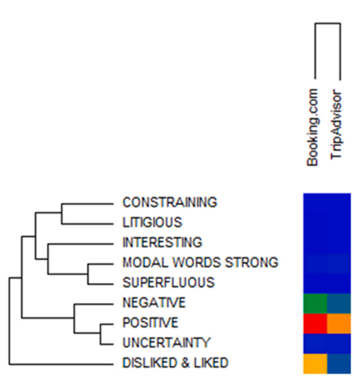
Table 6 also shows the results of the average sentiment of the various regions of tourists by the rating system. Regardless of the region, TripAdvisor's sentiment was always higher than in Booking.com. To understand the significance of these differences, Table 7 presents the decomposing of the tone of voice differential by rating system and region. Once more, the significant differences were related to constraining, litigious and superfluous discourse. Americas score the highest in terms of constraining discourse and the lowest at litigious discourse. Although the Middle East and Africa (MEA) scored the lowest in average sentiment it presents the second-highest value in terms of constraining discourse.

The observation of the subdimensions litigious shows that the number is higher in the five-star hotels, even when the rating posted online is not the lower (Fig. 5). The high number of litigious reviews performed by five-star hotel guests may reveal the disconformity between consumer expectations and consumer experience. The following comment demonstrates it: "At the very least a 5-star hotel should offer a clean, comfortable and QUIET room.as a guest, I am obligated to pay my invoice.the hotel is obligated to provide the services for which it was contracted.I PAID.the hotel DID NOT provide the service.". A similar outcome was found in what concerns the hotel review rating and the tone of voice used.

In some reviews, it is possible to find a good rating in the platform's main dimensions and a litigious tone of voice in the following comment, implying that the platforms' standard rating dimensions may not fully address all the clients' concerns.

Considering that Booking.com is an OTA specialized in hotel booking and TripAdvisor, a travel review platform, it is expected to find differences concerning value for money and payment processes (see, Fig. 2). All comments related to money, payments processes, and value for money were retrieved from the database to analyze differences between platforms. Moreover, the tone of voice used led to conclude that there are differences between the two platforms, related to positive and negative comments (Table 8).

The differences seem to be stronger in negative comments (generally negative, negations, negative when negated) presented in the litigious



	Booking.com	TripAdvisor	Chi2	P (2-tails)
CONSTRAINING	0.08%	0.14%	0.523	0.469
DISLIKED & LIKED	33.56%	15.16%	1390715	0.000
INTERESTING	0.05%	0.17%	45.117	0.000
LITIGIOUS	0.05%	0.10%	3.61	0.057
MODAL WORDS STRONG	1.01%	2.49%	226.661	0.000
NEGATIVE	15.97%	18.51%	1929.205	0.000
POSITIVE	47.41%	60.93%	3047.942	0.000
SUPERFLUOUS	0.03%	0.09%	11.407	0.001
UNCERTAINTY	1.84%	2.42%	99.867	0.000

Fig. 4. Results from SA by platform.

Table 7
Tone of voice by region versus platfarm and by hotel number of stars.

Booking.com versus TripAdvisor							
	AME	**ASP**	**EUR**	**MEA**	**OTHERS**	**Chi2**	**P (2-tails)**
CONSTRAINING	2,94	1,69	1,46	1,89	1,4	12,01	0,213
DISLIKED & LIKED	789,38	888,91	830,91	803,89	716,99	13884,00	0,000
INTERESTING	0	0,95	0,8	1,27	1,84	55,38	0,000
LITIGIOUS	0,49	1,27	1,22	1,03	1,33	6,55	0,683
MODAL WORDS STRONG	15,7	21,79	21,05	26,15	27,71	278,93	0,000
NEGATIVE	380,71	415,79	356,29	373,5	357,32	2029,31	0,000
POSITIVE	905,66	901,29	1035,31	1030,57	1047,36	3715,82	0,000
SUPERFLUOUS	0,1	0,42	0,63	0,92	1,08	17,56	0,041
UNCERTAINTY	53,48	44,96	46	42,78	42,65	120,89	0,000
	*	**	***	****	*****	**Chi2**	**P (2-tails)**
CONSTRAINING	2.47	1.63	1.85	2.01	1.78	1.871	0.760
DISLIKED & LIKED	701.21	756.41	668.77	521.41	357.72	1082807	0.000
INTERESTING	0.41	1.39	1.18	1.82	2.4	22.222	0.000
LITIGIOUS	0.41	1.39	1.4	1.27	1.51	3.086	0.544
MODAL WORDS STRONG	21.39	23.64	24.79	30.37	35.64	111.021	0.000
NEGATIVE	374.67	404.08	346.55	332.17	290.14	486.579	0.000
POSITIVE	965.25	935.47	1060.98	1022.26	1007.51	167.625	0.000
SUPERFLUOUS	0	0.65	0.84	1.12	1.18	7.418	0.115
UNCERTAINTY	48.53	43.85	43.64	39.8	35.97	42.092	0.000

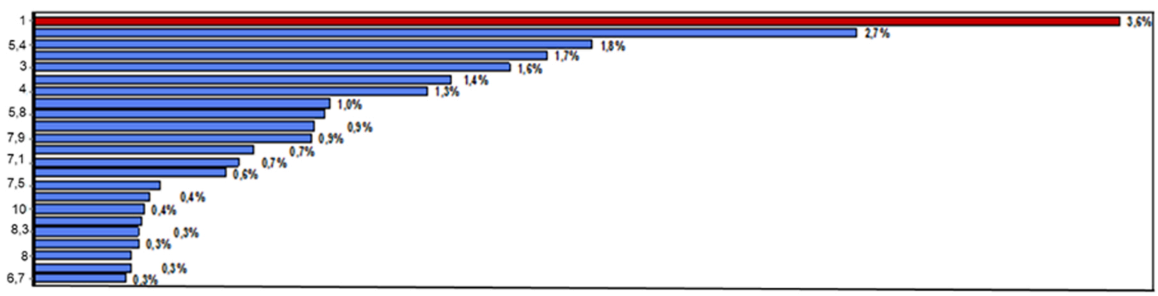


Fig. 5. Percentage of cases with litigious language by hotel rating.

Table 8
The tone of voice adopted in issues linked to payment issues in both platforms.

Discourse link	**VFM Booking.com**	**VFM TripAdvisor**	**Chi2**	**P (2-tails)**
DOUBLE NEGATION	3	7	6,016	0,014
EXCEPTIONS	119	25	16,681	0,000
GENERALLY NEGATIVE	58	31	0,080	0,777
NEGATIONS	8	4	0,000	0,995
NEGATIVE WHEN NEGATED	40	22	0,119	0,730
NEGATIVE WORDS	318	64	47,662	0,000
POSITIVE WORDS	814	360	3,994	0,046
REAL_BAD	318	63	48,793	0,000
REAL_GOOD	814	360	3,994	0,046
TO_ASSESS	49	31	1,023	0,312

discourse, so an additional analysis was conducted to unveil the differences concerning the subjects that originate the litigious comments (Fig. 6).

The topics that originate more litigious discourse from customers in Booking.com are related to payment issues (overcharge, value for money, and credit card issues). At the same time, on TripAdvisor, the topics are connected to the overcharge issues and room conditions.

4.2. Hypotheses Testing

To test the developed hypotheses, statistical tests were conducted in IBM® SPSS® Statistics. Two one-way ANOVA (Férnandez and Bedia, 2004; Martin-Fuentes et al., 2018b) were performed to measure statistically significant differences in the hotel categories sentiment score and the review rating after normalization. The results showed that, with a confidence level of 99% ($p < 0.01$), there were statistically significant differences between the hotel categories in terms of the sentiment score and the review rating, allowing us to reject the null hypothesis that the mean score of the five different hotel categories is equal – $F(38287,4) = 168.046$, $p = 0.000$ and $F(38287,4) = 175.719$, $p = 0.000$, respectively. The output from both analyses is included in Table 9. Therefore, H1a was partially accepted. We can assert that there were statistically significant differences between the hotel categories regarding sentiment and review rating.

The outcome of the tone of voice analysis conducted (Table 7), shows that there is no direct connection between the hotel category and the tone of voice used. For instance, in the ligitious discourse, the five-star hotels present a higher value, implying that customers are somewhat more disappointed with higher hotel' performance. Additionally, as noted in Fig. 4, the difference between the tone of voice used in Booking.com and TripAdvisor reviews are only significant when it concerns constraining and litigious discourse. In this sense, H1b was rejected since there is insufficient support that the (higher) the hotel category is, the (better) the tone of voice used in the review. Our results are more aligned with the assessment made by Kim et al. (2019) regarding consumers dissatisfaction. As Xu (2020) noted there is a relationship between the tone of voice used in the reviews and consumers' emotions and satisfaction, which varies with initial customer expectations.

Moreover, the hotels from the highest category were the ones with the highest average sentiment. However, the average sentiment did not increase consistently with the hotel category. This study's findings align with Férnandez and Bedia's (2004) research, where the quality of service was not measured by the hotel category but by meeting the

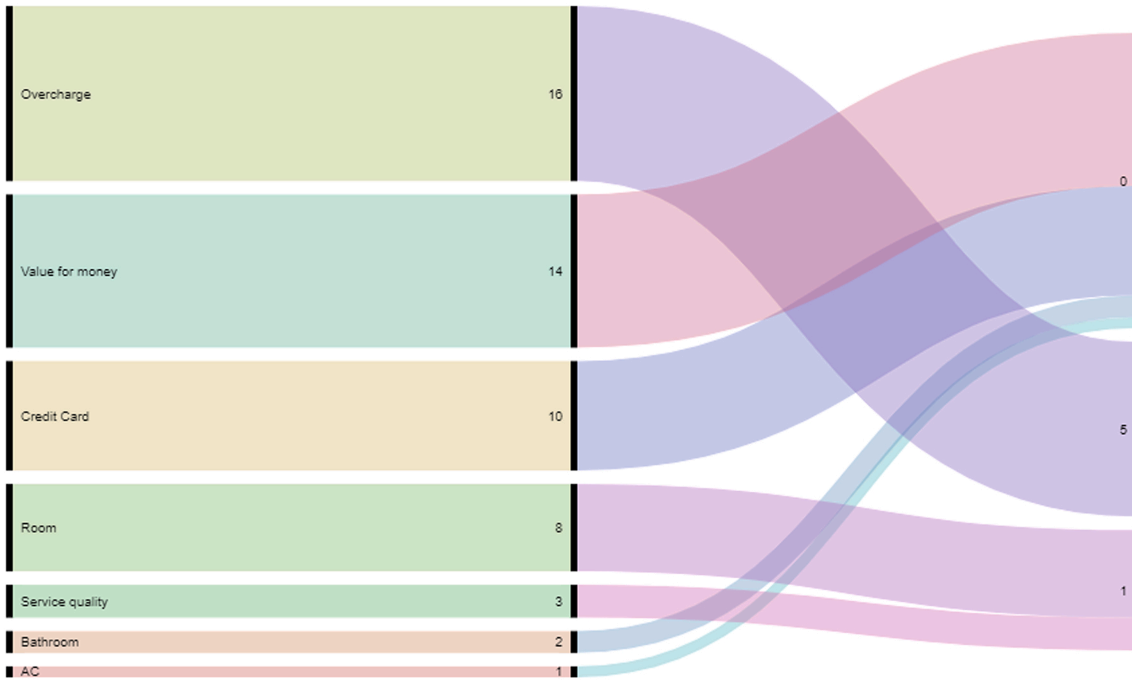


Fig. 6. Dispersion of subjects within the litigious comments.

Table 9
Results from ANOVA - sentiment per hotel category and review rating per hotel category.

ANOVA*sentiment*					
	Sum of Squares	***df***	***Mean Square***	***F***	***Sig***
Between Groups	78.207	4	19.552	168.046	0.000
Within Groups	4,454.594	38,287	0.116		
Total	**4,532.800**	**38,291**			
ANOVA*review_rating*					
	Sum of Squares	***df***	***Mean Square***	***F***	***Sig***
Between Groups	25.277	4	6.319	175.719	0.000
Within Groups	1376.889	38287	0.036		
Total	**1,402.166**	**38,291**			

customers' expectations.

To test H2, this research mapped each country with a region –AME, ASP, Europe (EUR), and the MEA – following the approach proposed by Banerjee and Chua (2016). Moreover, a single factor ANOVA was used to test statistically significant differences between the mean scores. The results of the analysis are summarized in Table 9, showing that, with a confidence level of 99% ($p < 0.01$), there were statistically significant differences between the mean scores of each region in terms of sentiment score, allowing to reject the null hypothesis that the mean score for each region (AME, ASP, EUR, MEA) was equal – $F(35616,3) = 81.391$, $p = 0.000$. There were statistically significant differences between the reviewer's nationality regarding the sentiment conveyed through the review (Table 10), even when the tone of voice adopted was not truly different between rating systems (Table 7). Therefore, H2 was partially accepted. The cross-cultural literature has evolved over the years, providing a plethora of country-specific and cross-cultural comparative studies in several domains, using a variety of cultural theories: Hofstede's cultural dimensions theory, Schwartz's cultural values, analytical/holistic thinking style, among others (Kusawat and Teerakapibal, 2022). Some authors have tightened these outcomes with Hofstede's cultural theory. Jia (2020) reported that Asian customers tend to evaluate an experience higher and less indulgent, reflecting collectivism and in favor of social norms. In turn, US customers are less willing to evaluate an experience higher, confirming individualism and freedom to express themselves. Stamolampros et al. (2019) also found that customers tend to be more satisfied with services that are more closely associated with their own cultural values. In the present study, since users were aggregated by region and not by country, it's not possible to establish a direct connection between the findings and Hofstede's cultural theory, but it's possible to see that American and European are more likely to express positive sentiments towards an offer located on Europe.

Table 10
Results from ANOVA to test H2.

ANOVA*region*					
	Sum of Squares	***df***	***Mean Square***	***F***	***Sig***
Between Groups	28.696	3	9.565	81.391	0.000
Within Groups	4,185.756	35,616	0.118		
Total	**4,214.452**	**35,619**			

Finally, H3 proposed that the sentiment conveyed by the reviewers on both platforms would be the same. Moreover, in line with previous research, the review ratings in Booking.com were expected to be inflated compared to TripAdvisor. In this regard, the first difference was that contrary to Bjørkelund et al. (2012) – wherein over 600,000 reviews none scored less than 2.5 in Booking.com - the dataset used in this research included reviews ranging from 1 to 10, which might indicate that Booking.com was trying to correct the inflation of the scores within the platform (Table 11). Following the approach used by Martin-Fuentes et al. (2018a), a *Student's t-test* was performed considering the scale after normalization in both platforms. Contrary to Martin-Fuentes et al. (2018a), the obtained results did not allow this research to reject the null hypothesis that both means were equal. Therefore, this research found no statistically significant evidence that the ratings from Booking.com were inflated.

Moreover, in line with the previous finding regarding reviews with a score lower than 2.5 in Booking.com, the results might indicate that Booking.com was trying to correct their scores inflation. H3a was then rejected. Nevertheless, it is worth noticing that the reviews' sentiment was higher in TripAdvisor by approximately 65%, indicating that the sentiment and the review rating were not aligned.

Regarding the feelings expressed in both platforms, results show that love-related words are more commonly used in TripAdvisor, while on Booking.com, more anger traits can be found (See, Fig. 2). It was also found that the tone of voice adopted had differences between the two platforms, present in both positive and negative comments. These differences were more significant in comments classified as litigious discourse and linked to money issues (overcharging or billing). H3b was then also rejected.

5. Conclusions

Multiple studies and evidence have shown reviewers' power in influencing peers' journey in the tourism field (Akhtar et al., 2019). Additionally, tourist online reviews can be insightful to implement an agile and improved hotel customer service (Brochado et al., 2019; Oliveira et al., 2019). Although previous studies have examined the importance and consequences of customers online reviews (Cena et al., 2017; Casaló et al., 2015), there has been only limited attention given to the rating system's influence on customers' reviews (Borges-Tiago et al., 2021; Xu, 2020).

Given the heightened user-generated content available on different online rating platforms, tourism and hospitality firms become interested in identifying customer preferences and emotions from distinctive experiences shared on the different rating systems. Kusawat and Teerakapibal (2022) recalled that an analysis that takes into account not only the ratings but also the emotions expressed is quite a value, as it can help marketers to overcome misunderstandings, communication errors, and even hurt feelings that can occur as a consequence of not conducting a suitable SA of the content shared. When adding to SA the tone of voice, it also can improve the sentiment analysis capabilities associated with the machine and AI readership-base interactions, that are increasingly being used by tourism and hospitality firms (Go et al., 2020; Pillai and Sivathanu, 2020).

Since online customer reviews are here to stay, a more comprehensive understanding of the platforms adopted by tourists and the tone of voice and sentiment expressed in the reviews is needed. Thus, this research studied the influence of hotel category, review rating, and reviewer demographics on the reviewer comments shared on both Booking.com and TripAdvisor. Moreover, this study responds to calls for further research related to the customer preferences for hotels expressed in comments on aspect-level sentiment and provides a more in-depth understanding of the relevance of tone of voice present on those comments.

The hotel star-ratings system was composed to reflect physical and service characteristics offered by the hotel (Huang et al., 2018) and it guides customers' expectation formation (Kim et al., 2019; Martin-Fuentes, 2016; Rhee and Yang, 2015). Contrary to hypothesized, the results show statistically significant differences between the hotel category and the expressed sentiment, showing that sentiment did not increase as the hotel category increased. The findings seemed to align with Fernández and Bedia (2004), who claimed that the service's quality depended more on meeting customers' expectations than delivering a luxury experience. Moreover, the results are aligned with the outcome of the study conducted by Kim et al. (2019), which claims that the higher the hotel rating and expectations formed, the higher the probability of not overcoming customer expectations and generating unsatisfaction. Additionally, the most positive sentiment found, "Love", was related to customer-staff encounters and small details of the experience, while the most negative, "Anger", had a wider range of reasons.

Travelers are using different travel rating platforms to share their experiences online (Xu, 2020), which by the nature of their business model are perceived as different (Borges-Tiago et al., 2021). The customers' motivations to share their experience online also may vary by these rating systems (Kim et al., 2018) and produce distinctive discourses.

The most positive feelings were expressed on TripAdvisor, while the most negative ones had predominance in Booking.com. The outcome remains the same when looking at the scores instead of comments since reviews lower than 2.5 are on Booking.com. This outcome is not aligned with previous evidence (Bjørkelund et al., 2012; Martin-Fuentes et al., 2018a). Moreover, it differs from the results obtained in previous studies (Mariani and Borghi, 2018; Martin-Fuentes et al., 2018a; Mellinas et al., 2016, 2015) that found inflated review ratings in Booking.com when compared with those from TripAdvisor. Besides not finding pieces of evidence of inflated review ratings, this work found significant differences between these two platforms regarding the tone of voice adopted in the review, in what concerns constraining and litigious discourse. The subjects that originate more litigious discourse from customers in Booking.com are related to payment issues, whereas TripAdvisor is also related to the rooms' conditions.

For hotels located in Lisbon, this research also discovered that more than 80% of the reviews written in English and posted on TripAdvisor conveyed a positive sentiment, in line with Zervas et al. (2021) study. In contrast, only 54% reported a positive sentiment in Booking. Furthermore, the average sentiment was consistently higher in TripAdvisor than in Booking.com. However, in the case of two-, three-, and four-star hotels, the review rating after normalization (which allows for a comparison between both platforms) was higher in Booking.com. Moreover, it was found that the tone of voice used varies with hotel category, with the five-star hotels presenting a higher number of litigious comments. These results may reveal the disconformity between consumer expectations and consumer experience and need to be carefully considered by hoteliers. This investigation also found that the sentiment expressed in the reviews depended on the nationality of the reviewer, supporting previous work (Alvarez and Hatipoɫlu, 2014; Kim et al., 2018). They posited that the differences in the consumer's reviews depended on the reviewers' nationality and cultural background.

This study contributes to the hospitality literature by explaining the influence that online tourism platforms have on the sentiment and tone of voice expressed by the reviewers. By testing the first hypothesis of this study, which posits that the tone of voice varies according to the hotel category and rating system, the current study extends the theoretical understanding of customer expectation formation related to the hotel star-rating systems. The results were aligned with previous works.

The fine-grained sentiment analysis approach adopted also advances knowledge regarding customers' preferences. Considering that concept

Table 11
Results of the Student's t-test assuming unequal variances.

*S*tudent's t-test assuming unequal variances				
		t	df	Sig. (2-tailed)
Review_Rating	Equal variances not assumed	1.368	15,507.839	0.171
Group Statistics				
	Platform	N	Mean	Std. Deviation
Review_Rating	Booking.com	27,289	0.853809227	0.167674522
	TripAdvisor	1,1003	0.85381713	0.240220161

of the tone of voice is still underexplored in the tourism literature, this work also contributes by highlighting its value to the field. The results of the current study advance the previous literature by suggesting that the hotel category and reviewers' nationality can influence on sentiment and tone of voice adopted; and that these vary depending on the online tourism platform used.

For practitioners, this research suggests differences in sentiment and tone of voice in different online tourism platforms. Simultaneously, hotel managers and marketers should be aware that different nationalities have different reactions. Therefore, the experience provided should be adapted to satisfy their target audience's needs and adjusted to the guests' different nationalities and cultural backgrounds. This knowledge can provide insights for the creation of distinct marketing strategies considering the information above. As tourism and hospitality firms embrace technology, the use of deep learning tools to frame sentiment and tone perception will increase. Accordingly, practitioners and managers can utilize the aspect-level sentiment and tone of voice findings to improve their capabilities to identify the key hotel features at customers' eyes and the sentiment values activated by them. Hoteliers and managers should, however, keep in mind that tone of voice may influence other hotel features not explored in this research. The last contribution is related to AI adoption (eg. Chatbots); the data catered to AI readers can enable more friendly and accurate responses when customers are less satisfied.

In terms of limitations, this research only considered reviews written in English. Although many reviews were considered when analyzing the English language, other languages could be analyzed to depict more comprehensive results (e.g., Portuguese, Spanish, or French). Therefore, a limitation and recommendation for future endeavors are to consider other languages as well, applying a multilingual sentiment analysis. Another limitation of this study was that only Booking.com and TripAdvisor were considered. The analysis could be extended to other online travel platforms (e.g., Expedia, Yelp or Ctrip). Moreover, the sample only comprised guest reviews from hotels located in Lisbon. Thus, another recommendation for future studies is to broaden the sample to other cities to obtain extensive coverage and to understand whether the results hold in different cities or regions of the world.

Also, additional research could be conducted to understand how tourists' sentiments and opinions were affected by the COVID-19 pandemic. For example, one could measure other dimensions (e.g., cleanliness, safety) that most influenced the consumers' pre-and post-COVID-19 pandemic. This topic requires attention since customer reviews are beneficial to both peers and managers. Nevertheless, because competitors are paying attention to it, these reviews can also be responsible for securing the firm's online visibility in search rankings.

Acknowledgements

Paulo Rita was supported by FCT (Fundação para a Ciência e a Tecnologia) under UIDB/04152/2020—Centro de Investigação em Gestão de Informação (MagIC).

References

Akhtar, N., Sun, J., Akhtar, M.N., Chen, J., 2019. How attitude ambivalence from conflicting online hotel reviews affects consumers' behavioural responses: The moderating role of dialecticism☆. J. Hosp. Tour. Manag. 41, 28–40. https://doi.org/10.1016/j.jhtm.2019.09.003.

Akram, U., Hui, P., Kaleem Khan, M., Tanveer, Y., Mehmood, K., Ahmad, W., 2018. How website quality affects online impulse buying. Asia Pac. J. Mark. Logist. 30, 235–256. https://doi.org/10.1108/APJML-04-2017-0073.

Alvarez, M.D., Hatipoĺlu, B., 2014. Nationality differences in user-generated reviews in the hospitality industry. Adv. Cult. Tour. Hosp. Res 8, 23–29. https://doi.org/10.1108/S1871-317320140000008000.

Anaconda, 2020. Anaconda | Individual Edition [WWW Document].

Anderson, E.W., Sullivan, M.W., 1993. The antecedents and consequences of customer satisfaction for firms. Mark. Sci. 12, 125–143. https://doi.org/10.1287/mksc.12.2.125.

Ariffin, A.A.M., Maghzi, A., 2012. A preliminary study on customer expectations of hotel hospitality: influences of personal and hotel factors. Int. J. Hosp. Manag. 31, 191–198. https://doi.org/10.1016/j.ijhm.2011.04.012.

Banerjee, S., Chua, A.Y.K., 2016. In search of patterns among travellers' hotel ratings in TripAdvisor. Tour. Manag 53, 125–131. https://doi.org/10.1016/j.tourman.2015.09.020.

Bi, J.W., Liu, Y., Fan, Z.P., Zhang, J., 2020. Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry. Tour. Manag. 77, 104006 https://doi.org/10.1016/j.tourman.2019.104006.

Bizirgianni, I., Dionysopoulou, P., 2013. The influence of tourist trends of youth tourism through social media (SM) & information and communication technologies (ICTs). Proc. Soc. Behav. Sci. 73, 652–660. https://doi.org/10.1016/j.sbspro.2013.02.102.

Bjørkelund, E., Burnett, T.H., Nørvåg, K., 2012. A study of opinion mining and visualization of hotel reviews, in: ACM International Conference Proceeding Series. ACM Press, New York, New York, USA, pp. 229–238. https://doi.org/10.1145/2428736.2428773.

Booking.com, 2020. Hotels in 226 countries worldwide [WWW Document].

Booking.com, 2021. Hotel Affiliate Program by Booking.com - Earn money on your website [WWW Document]. URL https://www.booking.com/affiliate-program/v2/index.html (accessed 2.18.21).

Borges-Tiago, M., Arruda, C., Tiago, F., Rita, P., 2021. Differences between TripAdvisor and Booking.com in branding co-creation. J. Bus. Res. 123, 380–388. https://doi.org/10.1016/j.jbusres.2020.09.050.

Brochado, A., Rita, P., Oliveira, C., Oliveira, F., 2019. Airline passengers' perceptions of service quality: themes in online reviews. Int. J. Contemp. Hosp. Manag 31, 855–873. https://doi.org/10.1108/IJCHM-09-2017-0572.

Bulchand-Gidumal, J., Melián-González, S., González López-Valcárcel, B., 2011. Improving hotel ratings by offering free Wi-Fi. J. Hosp. Tour. Technol. 2, 235–245. https://doi.org/10.1108/17579881111173776.

Calheiros, A.C., Moro, S., Rita, P., 2017. Sentiment classification of consumer-generated online reviews using topic modeling. J. Hosp. Mark. Manag. 26, 675–693.

Casado-Díaz, A.B., Andreu, L., Beckmann, S.C., Miller, C., 2020. Negative online reviews and webcare strategies in social media: effects on hotel attitude and booking intentions. Curr. Issue Tour. https://doi.org/10.1080/13683500.2018.1546675.

Casaló, L.V., Flavián, C., Guinalíu, M., Ekinci, Y., 2015. Do online hotel rating schemes influence booking behaviors? Int. J. Hosp. Manag. 49, 28–36. https://doi.org/10.1016/j.ijhm.2015.05.005.

Cena, F., Gena, C., Grillo, P., Kuflik, T., Vernero, F., Wecker, A.J., 2017. How scales influence user rating behaviour in recommender systems. Behav. Inf. Technol. 36, 985–1004. https://doi.org/10.1080/0144929X.2017.1322145.

Chen, C.-W., 2017. Five-star or thumbs-up? The influence of rating system types on users' perceptions of information quality, cognitive effort, enjoyment and continuance intention. Internet Res 27, 478–494. https://doi.org/10.1108/IntR-08-2016-0243.

Chinomona, R., Maziriri, E.T., 2017. The influence of brand awareness, brand association and product quality on brand loyalty and repurchase intention: a case of male consumers for cosmetic brands in South Africa, www.jbrmr.com A Journal of the Academy of Business and Retail Management. ABRM.

Choi, T.Y., Chu, R., 2001. Determinants of hotel guests' satisfaction and repeat patronage in the Hong Kong hotel industry. Int. J. Hosp. Manag. 20, 277–297. https://doi.org/10.1016/S0278-4319(01)00006-8.

Costa, G., Glinia, E., Goudas, M., Antoniou, P., 2004. Recreational services in resort hotels: Customer satisfaction aspects. J. Sport Tour. 9, 117–126. https://doi.org/10.1080/14775080410001732541.

Cró, S., Martins, A.M., 2018. Hotel and hostel location in Lisbon: looking for their determinants. Tour. Geogr. 20, 504–523. https://doi.org/10.1080/14616688.2017.1360386.

Dominici, G., Guzzo, R., 2010. Customer satisfaction in the hotel industry: a case study from sicily. Int. J. Mark. Stud. 2. https://doi.org/10.5539/ijms.v2n2p3.

Do, H.H., Prasad, P.W.C., Maag, A., Alsadoon, A., 2019. Deep learning for aspect-based sentiment analysis: a comparative review. Expert Syst. Appl. 118, 272–299.

Ekman, P., 1992. Are there basic emotions? Psychol. Rev. 99, 550–553. https://doi.org/10.1037/0033-295X.99.3.550.

Férnandez, M., Bedia, A., 2004. Is the hotel classification system a good indicator of hotel quality? Tour. Manag 25, 771–775. https://doi.org/10.1016/j.tourman.2004.06.007.

Finance, Y., 2021b. TripAdvisor, Inc. [WWW Document]. Yahoo! Financ. URL https://finance.yahoo.com/quote/TRIP?p=TRIP&.tsrc=fin-srch (accessed 2.10.21).

Finance, Y., 2021a. Booking Holdings Inc. [WWW Document]. Yahoo! Financ. URL https://finance.yahoo.com/quote/BKNG?p=BKNG&.tsrc=fin-srch (accessed 2.10.21).

Furtado, A., Ramos, R.F., Maia, B., Costa, J.M., 2022. Predictors of hotel clients' satisfaction in the cape verde islands. Sustainability 14, 2677. https://doi.org/10.3390/su14052677.

Galati, F., Galati, R., 2019. Cross-country analysis of perception and emphasis of hotel attributes. Tour. Manag 74, 24–42. https://doi.org/10.1016/j.tourman.2019.02.011.

Geetha, M., Singha, P., Sinha, S., 2017. Relationship between customer sentiment and online customer ratings for hotels - an empirical analysis. Tour. Manag. 61, 43–54. https://doi.org/10.1016/j.tourman.2016.12.022.

Ghasemi, A., Zahediasl, S., 2012. Normality tests for statistical analysis: a guide for non-statisticians. Int. J. Endocrinol. Metab. 10, 486–489. https://doi.org/10.5812/ijem.3505.

Go, H., Kang, M., Suh, S.C., 2020. Machine learning of robots in tourism and hospitality: interactive technology acceptance model (iTAM)–cutting edge. Tour. Rev. 75 (4), 625–636. https://doi.org/10.1108/TR-02-2019-0062.

Huang, W.-J., Chen, C.-C., Lai, Y.M., 2018. Five-star quality at three-star prices? opaque booking and hotel service expectations. J. Hosp. Mark. Manag. 27, 833–854. https://doi.org/10.1080/19368623.2018.1448315.
INE, 2019. Tourism Statistics 2019 [WWW Document]. INE. URL https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=133574&PUBLICACOESmodo=2 (accessed 2.20.21).
Israeli, A.A., 2002. Star rating and corporate affiliation: their influence on room price and performance of hotels in Israel. Int. J. Hosp. Manag. 21, 405–424. https://doi.org/10.1016/S0278-4319(02)00037-3.
Jia, S. (Sixue), 2020. Motivation and satisfaction of Chinese and U.S. tourists in restaurants: a cross-cultural text mining of online reviews. Tour. Manag 78, 104071. https://doi.org/10.1016/j.tourman.2019.104071.
Kandampully, J., Suhartanto, D., 2000. Customer loyalty in the hotel industry: the role of customer satisfaction and image. Int. J. Contemp. Hosp. Manag. 12, 346–351. https://doi.org/10.1108/09596110010342559.
Kim, B., Kim, S., (Sam), Heo, C.Y., 2019. Consequences of customer dissatisfaction in upscale and budget hotels: focusing on dissatisfied customers' attitude toward a hotel. Int. J. Hosp. Tour. Adm. 20, 15–46. https://doi.org/10.1080/15256480.2017.1359728.
Kim, J.M., Jun, M., Kim, C.K., 2018. The effects of culture on consumers' consumption and generation of online reviews. J. Interact. Mark. 43, 134–150. https://doi.org/10.1016/j.intmar.2018.05.002.
Kontogianni, A., Alepis, E., 2020. Smart tourism: State of the art and literature review for the last six years. Array 6, 100020. https://doi.org/10.1016/j.array.2020.100020.
Kusawat, P., Teerakapibal, S., 2022. Cross-cultural electronic word-of-mouth: a systematic literature review. J. Mark. ESIC. https://doi.org/10.1108/SJME-06-2021-0116. Ahead of Print.
Kwon, O.B., Kim, C.-R., Lee, E.J., 2002. Impact of website information design factors on consumer ratings of web-based auction sites. Behav. Inf. Technol. 21, 387–402. https://doi.org/10.1080/0144929021000050256.
Lawton, T.C., 1999. The limits of price leadership: Needs-based positioning strategy and the long-term competitiveness of Europe's low fare airlines. Long. Range Plann. 32, 573–586. https://doi.org/10.1016/s0024-6301(99)00074-6.
Lee, S.K., 2015. Quality differentiation and conditional spatial price competition among hotels. Tour. Manag 46, 114–122. https://doi.org/10.1016/j.tourman.2014.06.019.
Lexalytics, 2020. Industry Packs [WWW Document]. Lexalytics. URL https://semantria-docs.lexalytics.com/docs/industry-packs.
Litvin, S.W., 2019. Hofstede, cultural differences, and TripAdvisor hotel reviews. Int. J. Tour. Res. 21, 712–717. https://doi.org/10.1002/jtr.2298.
Loughran, T., McDonald, B., 2011. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. J. Financ. 66, 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x.
Loughran, T., McDonald, B., 2016. Textual analysis in accounting and finance: a survey. J. Account. Res. 54 (4), 1187–1230. https://doi.org/10.1111/1475-679X.12123.
Mariani, M.M., Borghi, M., 2018. Effects of the Booking.com rating system: bringing hotel class into the picture. Tour. Manag. 66, 47–52. https://doi.org/10.1016/j.tourman.2017.11.006.
Martin-Fuentes, E., 2016. Are guests of the same opinion as the hotel star-rate classification system? J. Hosp. Tour. Manag. 29, 126–134. https://doi.org/10.1016/j.jhtm.2016.06.006.
Martin-Fuentes, E., Mateu, C., Fernandez, C., 2018a. Does verifying uses influence rankings? analyzing booking.com and tripadvisor. Tour. Anal. 23, 1–15. https://doi.org/10.3727/108354218×15143857349459.
Martin-Fuentes, E., Mateu, C., Fernandez, C., 2018b. Are users' ratings on tripadvisor similar to hotel categories in Europe? Cuad. Tur. https://doi.org/10.6018/turismo.42.13.
Mellinas, P., María-Dolores, S., García, J., 2015. Booking.com: the unexpected scoring system. Tour. Manag. 49, 72–74. https://doi.org/10.1016/j.tourman.2014.08.019.
Mellinas, P., María-Dolores, S., García, J., 2016. Effects of the Booking.com scoring system. Tour. Manag. 57, 80–83. https://doi.org/10.1016/j.tourman.2016.05.015.
Merriam-Webster, 2021. Tone of voice [WWW Document]. Merriam-Webster. URL https://www.merriam-webster.com/dictionary/tone of voice (accessed 9.8.21).
Mohsin, A., Rodrigues, H., Brochado, A., 2019. Shine bright like a star: hotel performance and guests' expectations based on star ratings. Int. J. Hosp. Manag. 83, 103–114. https://doi.org/10.1016/j.ijhm.2019.04.012.
Mucharreira, P.R., Antunes, M.G., Abranja, N., Justino, M.R.T., Quirós, J.T., 2019. The relevance of tourism in financial sustainability of hotels. Eur. Res. Manag. Bus. Econ. 25, 165–174. https://doi.org/10.1016/j.iedeen.2019.07.002.
Nandwani, P., Verma, R., 2021. A review on sentiment analysis and emotion detection from text. Soc. Net. Anal. Min. 11 (1), 1–19.
Ngai, E.W.T., Heung, V.C.S., Wong, Y.H., Chan, F.K.Y., 2007. Consumer complaint behaviour of Asians and non-Asians about hotel services: an empirical analysis. Eur. J. Mark. 41, 1375–1391. https://doi.org/10.1108/03090560710821224.
Nunkoo, R., Teeroovengadum, V., Ringle, C.M., Sunnassee, V., 2020. Service quality and customer satisfaction: the moderating effects of hotel star rating. Int. J. Hosp. Manag. 91. https://doi.org/10.1016/j.ijhm.2019.102414.
Oliveira, C., Brochado, A., Moro, S., Rita, P., 2019. Consumer perception of tourist experience through online reviews: the islands of the senses of Cape Verde. Worldw. Hosp. Tour. Themes. https://doi.org/10.1108/WHATT-09-2019-0052.
Oliver, R.L., 1980. A cognitive model of the antecedents and consequences of satisfaction decisions. J. Mark. Res. 17, 460–469. https://doi.org/10.1177/002224378001700405.
Öztürk, N., Ayvaz, S., 2018. Sentiment analysis on Twitter: a text mining approach to the Syrian refugee crisis. Telemat. Inform. 35, 136–147. https://doi.org/10.1016/j.tele.2017.10.006.
Padma, P., Ahn, J., 2020. Guest satisfaction & dissatisfaction in luxury hotels: an application of big data. Int. J. Hosp. Manag. 84. https://doi.org/10.1016/j.ijhm.2019.102318.
Pampulevski, V., Giaquinto, J.R., Rametta, M., Toscani, M., Barone, J., Nadal, J.C., 2020. Sentiment of media coverage and reputation of the pharmaceutical industry. Ther. Innov. Regul. Sci. 54, 220–225. https://doi.org/10.1007/s43441-019-00048-8.
Parboteeah, D.V., Valacich, J.S., Wells, J.D., 2009. The influence of website characteristics on a consumer's urge to buy impulsively. Inf. Syst. Res. 20, 60–78.
Pillai, R., Sivathanu, B., 2020. Adoption of AI-based chatbots for hospitality and tourism. Int. J. Contemp. Hosp. Manag. 32 (10), 3199–3226. https://doi.org/10.1108/IJCHM-04-2020-0259.
Qu, H., Ryan, B., Chu, R., 2000. The importance of hotel attributes in contributing to travelers' satisfaction in the Hong Kong hotel industry. J. Qual. Assur. Hosp. Tour. 1, 65–83. https://doi.org/10.1300/J162v01n03_04.
Ramos, R.F., Rita, P., Moro, S., 2019. From institutional websites to social media and mobile applications: A usability perspective. Eur. Res. Manag. Bus. Econ. 25, 138–143. https://doi.org/10.1016/j.iedeen.2019.07.001.
Rhee, H.T., Yang, S.B., 2015. Does hotel attribute importance differ by hotel? focusing on hotel star-classifications and customers' overall ratings. Comput. Hum. Behav. 50, 576–587. https://doi.org/10.1016/j.chb.2015.02.069.
Rita, P., Ramos, R.F., Moro, S., Mealha, M., Radu, L., 2020. Online dating apps as a marketing channel: a generational approach. Eur. J. Manag. Bus. Econ. 30, 1–17. https://doi.org/10.1108/EJMBE-10-2019-0192.
Rodrigues, R.G., das Dores, R.M., Camilo-Junior, C.G., Rosa, T.C., 2016. Sentihealth-cancer: a sentiment analysis tool to help detecting mood of patients in online social networks. Int. J. Med. Inf. 85, 80–95. https://doi.org/10.1016/j.ijmedinf.2015.09.007.
Roy, S.K., Shekhar, V., Lassar, W.M., Chen, T., 2018. Customer engagement behaviors: the role of service convenience, fairness and quality. J. Retail. Consum. Serv. 44, 293–304. https://doi.org/10.1016/j.jretconser.2018.07.018.
Sabini, J., Silver, M., 2005. Ekman's basic emotions: why not love and jealousy? Cogn. Emot. 19, 693–712. https://doi.org/10.1080/02699930441000481.
Saleem, H., Sarfraz Raja, N., 2014. The impact of service quality on customer satisfaction, customer loyalty and brand image: evidence from hotel industry of Pakistan. Middle East J. Sci. Res. 19, 706–711. https://doi.org/10.5829/idosi.mejsr.2014.19.5.21018.
Stamolampros, P., Korfiatis, N., Kourouthanassis, P., Symitsi, E., 2019. Flying to quality: cultural influences on online reviews. J. Travel Res. 58 (3), 496–511.
Sureshchandar, G.S., Rajendran, C., Anantharaman, R.N., 2002. The relationship between service quality and customer satisfaction – a factor specific approach. J. Serv. Mark. 16, 363–379. https://doi.org/10.1108/08876040210433248.
TripAdvisor, 2021. US Press Center | About Tripadvisor [WWW Document]. TripAdvisor. URL https://tripadvisor.mediaroom.com/US-about-us (accessed 2.21.21).
Turismo de Portugal, 2020. RNT - Registo Nacional de Turismo [WWW Document].
Wang, Y., Wang, Z., Zhang, D., Zhang, R., 2019. Discovering cultural differences in online consumer product reviews. J. Electron. Commer. Res. 20, 169–183.
WTA, 2020. Europe's Leading Destination 2020 [WWW Document]. World Travel Award. URL https://www.worldtravelawards.com/award-europes-leading-destination-2020 (accessed 2.10.21).
Xia, H., Vu, H.Q., Law, R., Li, G., 2020. Evaluation of hotel brand competitiveness based on hotel features ratings. Int. J. Hosp. Manag. 86. https://doi.org/10.1016/j.ijhm.2019.102366.
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. Tour. Manag. 58, 51–65. https://doi.org/10.1016/j.tourman.2016.10.001.
Xu, X., 2020. Examining consumer emotion and behavior in online reviews of hotels when expecting managerial response. Int. J. Hosp. Manag. 89, 102559 https://doi.org/10.1016/j.ijhm.2020.102559.
Xu, X., Li, Y., 2016. The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: a text mining approach. Int. J. Hosp. Manag. 55, 57–69. https://doi.org/10.1016/j.ijhm.2016.03.003.
Yang, L.W., Aggarwal, P., 2019. No small matter: how company size affects consumer expectations and evaluations. J. Consum. Res 45, 1369–1384. https://doi.org/10.1093/jcr/ucy042.
Yen, C.-L., Tang, C.-H., 2015. Hotel attribute performance, eWOM motivations, and media choice. Int. J. Hosp. Manag. 46, 79–88. https://doi.org/10.1016/j.ijhm.2015.01.003.
Yu, T., Rita, P., Moro, S., Oliveira, C., 2021. Insights from sentiment analysis to leverage local tourism business in restaurants. Int. J. Cult. Tour. Hosp. Res. https://doi.org/10.1108/IJCTHR-02-2021-0037.
Zervas, G., Proserpio, D., Byers, J.W., 2021. A first look at online reputation on Airbnb, where every stay is above average. Mark. Lett. 32, 1–16. https://doi.org/10.1007/s11002-020-09546-4.
Zhang, X., Liu, Y., Dan, B., 2021. Cooperation strategy for an online travel platform with value-added service provision under demand uncertainty. Int. Trans. Oper. Res. 28, 3416–3436. https://doi.org/10.1111/itor.12977.