Hotel customer segmentation and sentiment analysis through online reviews: an analysis of selected European markets

Segmentação de clientes hoteleiros e análise de sentimento através de avaliações online: uma análise de mercados europeus selecionados

Anderson S. Oliveira
School of Management, Hospitality and Tourism (ESGHT), Universidade do Algarve, Portugal, a63247@ualg.pt

Ana I. Renda
School of Management, Hospitality and Tourism (ESGHT), Universidade do Algarve & Centre for Tourism Research, Development and Innovation (CiTUR) & Centre for Tourism, Sustainability and Well-being (CinTurs), Portugal, arenda@ualg.pt

Marisol B. Correia
School of Management, Hospitality and Tourism (ESGHT), Universidade do Algarve & Centre for Tourism Research, Development and Innovation (CiTUR) & Centre for Tourism, Sustainability and Well-being (CinTurs) & CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Portugal, mcorreia@ualg.pt

Nuno Antonio
NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa & Centre for Tourism Research, Development and Innovation (CiTUR), Portugal, nantonio@novaims.unl.pt

Received: 20.06.2021; **Revisions required:** 16.10.2021; **Accepted:** 17.01.2022

Abstract

This study aims to verify how distinct markets evaluate hotels in the Algarve through the analysis of online reviews, in order to identify if satisfaction and dissatisfaction attributes are similar among some of the main markets of overnight stay tourists in the region. Online reviews of hotels in the Algarve, written in English, French as well as Portuguese and posted on Tripadvisor by British, French and Portuguese residents from January 2019 to December 2019 are analysed. After the analysis of 8,596 online textual reviews, the results demonstrated that not only satisfaction and dissatisfaction rates towards hotel attributes differ according to the language, but also that customers from different countries place dissimilar emphasis on hotel attributes. Besides extending the current research on the use of online reviews, the findings of this study also assist hoteliers to identify improvement opportunities. Although many studies on marketing segmentation through data mining have been conducted, this paper analyses the customer satisfaction of relevant tourist markets and suggests up-to-date practical implications for hoteliers.

Keywords: Online reviews, data mining, sentiment analysis, Tripadvisor, hotel management.

Resumo

Este estudo tem como objetivo verificar como mercados distintos avaliam hotéis no Algarve através da análise de comentários online, a fim de identificar se os atributos de satisfação e insatisfação são semelhantes entre alguns dos principais mercados turísticos da região. São analisadas avaliações online de hotéis no Algarve, escritas em inglês, francês e português e publicadas no Tripadvisor por residentes britânicos, franceses e portugueses de janeiro de 2019 a dezembro de 2019. Após a análise de 8.596 avaliações textuais online, os resultados demonstraram que não apenas as taxas de satisfação e insatisfação em relação aos atributos hoteleiros diferem de acordo com a língua, mas também que clientes de diferentes países colocam ênfase diferente nos atributos do hotel. Além de ampliar a pesquisa atual sobre o uso de revisões online, os resultados deste estudo também auxiliam os hoteleiros a identificar oportunidades de melhoria. Embora muitos estudos sobre segmentação de marketing por meio da mineração de dados tenham sido realizados, este artigo analisa a satisfação dos clientes dos mercados turísticos relevantes e sugere implicações práticas atualizadas para os hoteleiros.

Palavras-chave: Avaliações online, mineração de dados, análise de sentimento, Tripadvisor, gestão hoteleira.

1. Introduction

Driven by technological advances and a stronger world economy, the international tourism industry presented sustained growth for the 9th consecutive year in 2018. The overall number of international tourist arrivals and tourism receipts raised 5.4% and 4.4% respectively worldwide. On top of that, international tourist arrivals increased globally (World Tourism Organization, 2019). However, it is crucial to mention that tourist arrivals dropped by 56% between January and May 2020, when compared to the same period of the previous year, due to the unprecedented global economic crisis caused by the COVID-19 pandemic, which led to borders closures and lockdowns in several countries and cities (World Tourism Organization, 2020).

Still, the European continent, which accounts for half of the world's international arrivals, registered 710 million foreign visitors in 2018, representing an increase of 5%. Portugal reached the impressive mark of 22.8 million international arrivals in 2018, a 7.5% growth compared to the previous year, representing 3.2% of all European international arrivals market share (World Tourism Organization, 2019).

Portugal's robust performance in tourism also dynamizes the hotel industry in the country. Portugal has currently 1400 hotels spread across the country and its two autonomous regions, 91 new units compared to the previous year and 719 new units when compared to 2009 (PORDATA, 2019). Moreover, 54.2 million overnight stays were registered among the Portuguese hotels, corresponding to 83.6% of the total overnight stays in 2018. Foreigners stood for 70.6% of the total number of overnight stays and are represented by four major markets: the UK (19.5% of all international overnight stays), Germany (13.5%), Spain (10.2%) and France (9.8%). The domestic market also plays a vital role, since it generated 19.9 million overnight stays in Portugal, which corresponds to 29.4% of the total and a 6.5% growth in comparison to 2017. Regarding Portugal's tourist regions, the Algarve topped the list of overnight stays (30.2%), followed by the Lisbon Metropolitan Area (25.9%) (Instituto Nacional de Estatística, 2019).

Considering the cultural traits, language and other particularities of each of the aforementioned markets, hoteliers must be aware that different origins might lead to different levels of satisfaction. Therefore, it becomes essential to tailor the offer of hotel products and services, besides developing different marketing strategies for each segment in order to stand out in a competitive market (Ahani, Nilashi & Ibrahim, 2019a; Ahani, Nilashi, Yadegaridehkordi, Sanzogni, Tarik et. al 2019c; Francesco & Roberta, 2019; Mariani & Predvoditeleva, 2019; Xu, 2018; Xu, Wang, Li & Haghighi, 2017; Zhou, Ye, Pearce & Wu, 2014).

In order to assist hoteliers in achieving this goal, this study aims to verify how distinct markets evaluate hotels in the Algarve through the analysis of online reviews, in order to identify if satisfaction and dissatisfaction attributes are similar among some of the main markets of overnight stay tourists in the region. This work will take into consideration hotels located in the Algarve, since the region is the main Portuguese destination in total number of overnight stays and registered more than one-third of the non-resident stays in hotel units. According to Registo Nacional de Turismo (2019), a Portuguese organisation responsible for gathering tourism-related data in the country, 170 hotels are currently located in the area. The region also accounts for 36.3% of the accommodation capacity (beds) available in the country. Additionally, an average of 289 beds are offered per hotel unit, the highest volume among all Portuguese regions (Instituto Nacional de Estatística, 2019).

By doing so, this paper wants to uncover if guests from three of the most representative tourist markets of the Algarve, namely, the British, French and Portuguese residents, differ considerably in terms of overall satisfaction and demands in the region, besides providing managerial implications for hoteliers and contributing to the extension of studies in this field.

Therefore, this study is presented as follows. First, a review of the literature is conducted. Second, each stage of the proposed methodology is explained. After, the results of the research are analysed, presented and discussed. Finally, theoretical and practical implications are provided along with research limitations and recommendations for future studies.

2. Literature review

2.1 Online reviews and group characteristics

As formerly observed by Oliveira et al. (2020), previous studies have made use of online reviews to analyse different groups of travellers in the same destination. Antonio, De Almeida, Nunes, Batista and Ribeiro (2018) and Phillips, Antonio, De Almeida and Nunes (2020) analysed reviews from a set of Portuguese hotels in three different dialects and observed that cultural backgrounds as well as geographic distance influence hotel online reviews as travellers of different origins may have distinct expectancies. Ahani et al. (2019a; 2019c) collected reviews from 5-star hotels in Wellington, New Zealand, and four and five-star hotels in the Canary Islands, Spain, respectively. The findings of both researches detected various degrees of satisfaction with dissimilar preferences among customers and recommended hoteliers to segment travellers' preference and satisfaction through data mining as a way to improve the quality of hotel products and services.

Besides collecting reviews from four different world capitals, Francesco and Roberta (2019), investigated if travellers from different countries put different emphasis on hotel attributes and if these attributes were perceived differently. After analysing reviews on Tripadvisor written by Italian, American and Chinese travellers, the results also suggested considerable dissimilarities in the sentiments customers were having regarding the hotels. According to the findings provided by Li, Liu, Tan and Hu (2020), guests' expectations towards hotel performance might differ with respect to their geography. Although some similarities were found, different responses from local and international travellers regarding hotel attributes were registered.

Mariani and Predvoditeleva (2019) examined the role and influence of guests' cultural traits and perceived experience through online review ratings of Russian hotels and their findings show that different online customer groups can be clustered into segments, as they display different online behaviours and give different online evaluations. Similar online review behaviour, was also observed by Xu (2018). Table 1 presents the main studies observed.

Table 1 – Previous studies focused on market segmentation through online reviews

Source	Title	Aims	Method	Database Sample Size/ Geography	Key Findings	Implications
Phillips et al. (2020)	The influence of geographic and psychic distance on online hotel ratings	To examine the relationship between distance measures and online hotel reviews written in Portuguese, Spanish and English.	Text mining	Booking.com Tripadvisor 34,622 Portugal	Travellers of different origins may have significantly different expectations. Travellers with less psychic and geographic distance give a lower rating score than travellers with greater distance.	New insights into how geographic and psychic distance influence online hotel ratings. Practical implications suggest hotels listen to travellers in order to perform service improvements.
Ahani et al. (2019a)	Travellers segmen-tation and choice prediction through online reviews: The case of Wellington's hotels in New Zealand	To develop a method for 5-star hotels segmentation, and travellers' choice forecast through online reviews using machine learning methods.	Multi-criteria decision making (MCDM); Technique for order of preference by similarity to ideal solution (TOPSIS)	Tripadvisor 5,944 New Zealand	Results provide an overview of different segments of travellers'preferences towards 5-star hotels attributes.	Application of online reviews to identify guests'preferences. It also can support hoteliers in their decision-making process about customer satisfaction.
Ahani et al. (2019c)	Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels	To provide an understandings about customers' satisfaction and preferences using travellers' generated content in online hotel reviews.	Self-organizing map (SOM); TOPSIS; MCDM	Tripadvisor 9,128 Spain	Customers may declare dissimilar service desires that are generally complicated to manage. Also, having dissimilar preferences affect customers' satisfaction linked to hotel features.	Help hotels to understand customer's preferences and satisfaction.
Francesco and Roberta (2019)	Cross-country analysis of perception and emphasis of hotel attributes	To investigate whether there are differences in the way Italians, Americans, and Chinese travellers perceive and emphasize several hotel attributes.	Text link analysis	Booking.com 9,000 USA, UK, UAE and China	Each group of travellers emphasized hotel attributes differently. Travellers belonging to different countries place different emphasis on hotel attributes.	Help hotel managers in determining the optimal allocation of financial resources to improve customer satisfaction.
Mariani and Predvoditeleva (2019)	How do online reviewers' cultural traits and perceived experience influence hotel online ratings?	To examine the role and influence of online reviewers 'cultural traits and perceived experience on review ratings of Russian hotels.	Censored regression; semi-structured interviews	Booking.com 75,000 Russia	Cultural traits exert a significant influence on hotel online ratings. Reviewers' perceived experience in online reviewing is negatively related to online ratings.	Theoretical insights for hospitality management and marketing literature. Assisting hotels to improve their understanding of consumers and boost their online ratings.
Antonio et al. (2018)	Hotel online reviews: different languages, different opinions	To understand how guests from different origins assess hotels in online reviews.	Sentiment analysis	Booking.com Tripadvisor 23,353 Portugal	Different behaviours towards online reviews depending on guests' cultural background.	Helps hoteliers to understand what customers with different backgrounds criticize or value.
Xu (2018)	Does traveller satisfaction differ in various travel group compositions?	To investigate the online customer review behaviour and determinants of overall satisfaction with hotels of travellers in various travel group compositions.	Latent semantic analysis (LSA); text regression	Booking.com 4,800 USA	Not all positive and negative textual factors from reviews influenced their overall satisfaction. Determinants of traveller satisfaction differ in different travel group compositions.	Using online reviews to identify customer perception. It helps hoteliers to understand customer perception towards their products and services to better meet their customer's expectations.

Source: Own elaboration.

All aforementioned studies registered dissimilarities among guest's satisfaction and preferences when customers presented different cultural traits and backgrounds, nationalities or group composition. Accordingly, this study intends to verify if the four chosen European markets also display disparities concerning satisfaction and dissatisfaction factors about hotel products and services offered in the same region.

2.2 Hotel attributes and satisfaction

Many previous studies aimed to identify customer's sentiment towards specific hotel attributes. Ban et al. (2019) performs opinion mining techniques over 6,500 reviews to identify key hotel attributes that influence customer satisfaction. According to the findings of that particular study, staff service stands out as a fundamental feature to enhance customer experience and satisfaction, overcoming facilities, showing that intangible service has a great impact on customer experience. Also, Almeida and Pelissari (2019), Annisa and Surjandari (2019), Chaves, Gomes and Pedron (2012), Li et. Al (2020), Xu (2019, 2020), Ying, Chan and Qi (2020), identify a range of intangible services that are considered essential by guests. Although many of the attributes found by these studies seem to be aligned, it becomes noticeable that the type of hotel, geography and type of traveller influences customer perceptions and therefore, hotel ratings.

When it comes to tangible attributes, studies demonstrate that hotel facilities, as well as food and beverage, exert a very important influence over customer experience and satisfaction, with those attributes alternating ranking positions according to the geography and samples of each study (Almeida and Pelissari, 2019; Annisa & Surjandari, 2019; Chaves et al. (2012); Ban et al., 2019, Egresi, Puiu, Zotic & Alexandru, 2020; Gunasekar, & Sudhakar, 2019 and Li et al., 2020). Moreover, results from these studies demonstrate that room quality can directly sway profitability and guests' willingness to recommend the accommodation, besides leading to customer grievances and unfavourable reviews.

Many studies also refer to location as a key attribute of satisfaction. This feature presents a significant impact on online rates and textual reviews as presented by Anagnostopoulou, Buhalis, Kountouri, Manousakis and Tsekrekos (2019), Chaves et al. (2012), Gunasekar and Sudhakar (2019) and Li et al. (2020). Outcomes of these studies demonstrated that hotel location is very significant when considering profitability and volume of bookings, which means that hoteliers should be aware that unchangeable attributes have also a considerable influence on business performance. Finally, online textual reviews can prove the quality of tangible and intangible services, which, in turn, must be the hotelier's priorities, particularly when the main purpose of the business is to promote the well-being of their guests (Li et al. 2020; Xu, 2019).

3. Methodology

3.1 Sampling and data collection

In order to accomplish the tasks proposed by this research, online reviews of hotels located in the Algarve and written by guests who reside in the UK, France and Portugal need to be collected from Tripadvisor. This study opted to work with this platform because it stands out as the leading global advice website dedicated to tourist products and services with more than 300 million members and 500 million reviews of hotels and other tourism-related featured items (Liang, Liu & Wang, 2019; Zhao et al., 2019). Also, the website allows hotel practitioners to identify hotel characteristics closely connected to customer overall satisfaction at lower costs compared to traditional approaches (Litvin, 2019).

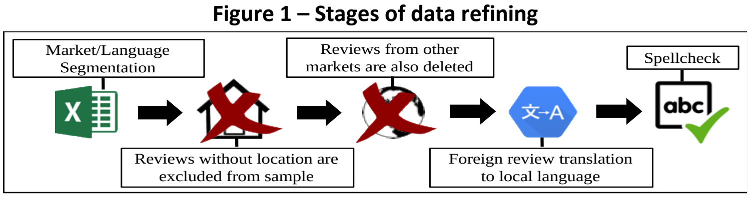
Before starting the review collection process, it was necessary to check if all 170 hotels listed by Registo Nacional de Turismo (2019), were available on Tripadvisor's website. All hotels were searched on the review platform, according to their names, address, postal code and the official website. A total of 161 hotels were considered eligible to have their reviews collected. Among the excluded sample, four hotels had duplicated entries, two establishments had not been reviewed, one accommodation had only one review, one hotel had been converted into a hostel and one unit was not found.

The second stage consisted of extracting the reviews of the selected hotels. First, each hotel unit has been given a singular ID number. Second, each hotel website link on Tripadvisor was included in an Excel spreadsheet and a bot built in C# has been used as a web scraper to collect online reviews posted from January 2019 to December 2019. A total of 15,585 ratings and text reviews written in three different languages were obtained from the platform. 10,744 reviews were written in English, 2,716 in Portuguese and finally 2,125 reviews were registered in French. Also, username and country of residence related to each rating and text review were collected.

3.2 Data refining

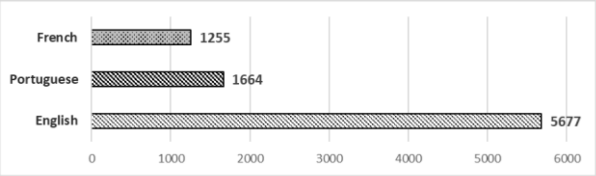
Data refining is an essential process when analysing a large volume of text. This extensive procedure consists mainly of converting unstructured text data into structured forms (Ban, Choi, Choi, Lee & Kim, 2019). Consequently, a five-stage process shown in Figure 1 is conducted with the purpose of promoting as accurate data analysis as possible. First, the text reviews are segmented by language in an Excel spreadsheet to facilitate further market segmentation. Second, reviews missing the city or country of residence are excluded from the sample, since it precludes market segmentation. Next, all reviews in which users are not from the target markets are also excluded leading to a sample of 8,596 valid reviews. Then, with the help of an online translator, all reviews written in the foreign languages covered by this study were translated to their original language and segmented accordingly (for example, a review posted by a French resident in English is translated into French and segmented as French market). Finally, with the assistance of an online spell checker, all reviews are revised in order to be correctly interpreted by further analysis software.

Figure 1 – Stages of data refining



As aforementioned, the United Kingdom is Algarve's most representative foreign market, representing almost 40% of the total overnight guests in the area. Its representation is reflected in the total number of reviews (5,677), which surpasses the number of Portuguese (1,664) and French (1,255) reviews combined. Figure 2 presents the final amount of valid reviews for each of the selected markets.

Figure 2 – Number of valid reviews per market



Although the number of valid reviews for each language differs, previous studies, such as García-Pablos, Duca, Cuadros, Linaza and Marchetti (2016), Antonio et. Al (2018), Phillips et al. (2020), also work with different amounts of reviews per language, with goals that include the verification of sentiment strength from different languages towards accommodations or destinations. Furthermore, all valid reviews correspond to the same time period and region, therefore avoiding bias when it comes to the sentiment analysis of the reviewers. For this reason, this study takes into consideration all valid reviews from each language in order to obtain results for the chosen region and timeline.

3.3 Sentiment analysis

Sentiment analysis - also known as opinion mining - is the field of study responsible for analysing individuals' reviews, opinions, sentiments and emotions towards products and services. This type of investigation mainly focuses on positive or negative sentiments expressed on people's opinions (Liu, 2012). It is also described as a computational study that investigates texts, which include someone's opinion about a specific business, such as hotel reviews (Çalı & Balaman, 2019).

It is also through the conduction of opinion mining that a more precise interpretation of text reviews is assured since this method aims to classify texts according to their sentiment orientation. This categorization can be binary (positive, negative), ternary (positive, neutral, negative) or to the context of reviews, as for instance, thumbs up or thumbs down, approval or disapproval (Çalı & Balaman, 2019; Duan, Yu, Cao & Levy, 2016; Han, Mankad, Gavirneni & Verma, 2016).

Travellers' online reviews contain objective and subjective statements about hotels and travel experiences. Consequently, the sentiment analysis' main aim is to deal with the subjective expressions utilized to describe emotions, opinions and feelings (Çalı & Balaman, 2019; Liu, 2012). Moreover, sentiment analysis is proven to be an important tool to quantify guest's opinions in the textual component of reviews (Antonio et al. 2018). Thereby, as a means to reach the goals proposed by this research, a lexicon-based approach sentiment analysis is adopted to identify the polarity of each selected online review. By doing so, it will be possible to uncover and classify any relevant semantic and emotional information regarding the targeted markets selected by this study.

3.4 Data analysis

Throughout the years, Tripadvisor's database has become noticeable in the hospitality industry due to the fact that its extraction and handling eases the detection of business strengths and weaknesses, benchmarking, measurement of customer overall satisfaction, hotels image improvement and market visibility (Lima & Viana, 2017). The platform is also widely used by scholars in pursuit of useful and strategic information regarding the hotel industry. Ahani et al. (2019a, 2019c), Ahani, Nilashi, Ibrahim, Sanzogni and Weaven (2019b), Han et al. (2016), Hu and Chen (2016), Li, Law, Vu, Rong and Zhao (2015), Litvin (2019), Yi, Li and Jai (2018), Yu, Li and Jai (2017), Zhao, Xu and Wang (2019) are recent examples of studies that made use of Tripadvisor as a primary source of information to identify customer overall satisfaction, preferences and segment markets to provide new insights as well as managerial and theoretical implications. Hence, the online textual reviews analysed in this study are open and unstructured data acquired from Tripadvisor. Even though the data has been previously refined, some of the challenges faced by this research are the impossibility to employ direct analysis methods, such as standard statistical or econometric procedures, and also to deal with a substantial number of reviews, which can be very time consuming and lead to biased results if done manually (Han et al., 2016; Limberger, Meira, Añaña & Sohn, 2016; Loo & Leung, 2018).

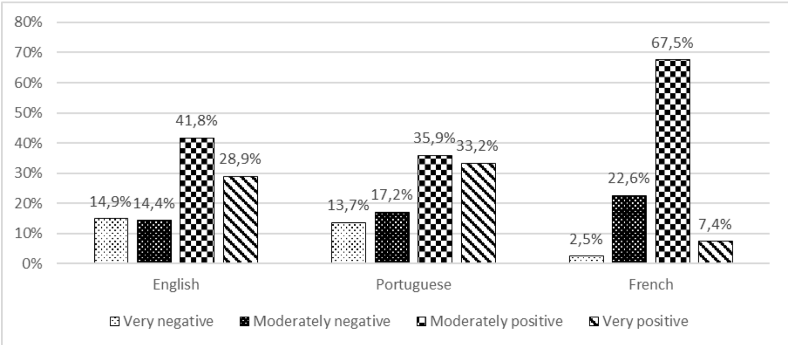
For this reason and as pointed out by recent studies such as Ahani et al. (2019a); Ban et al. (2019); Çalı and Balaman, (2019); Francesco and Roberta, (2019) and Oliveira et al. (2020), an appropriate software devoted to analysing large quantity of data is required. Therefore NVivo 12, a tool developed to assist researchers in qualitative data analysis, was chosen to assist the following content analysis of this study. The chosen tool is then used for the following procedures: first, all refined data has been imported into NVivo 12 separately; after, three different projects were created. One for each market and local language (English, Portuguese and French); next, a general sentence-level sentiment analysis is carried out to detect the sentiment polarity in each sentence of the text reviews. By doing so, this study aims to identify the overall sentiment of each market towards their hotel experience in the Algarve. In this first inquiry, the total amount of reviews and all their content were taken into consideration. Then, a frequency analysis is conducted in order to rank the 30 most frequent nouns related to hotel attributes. As a result, it was possible to uncover which hotel attributes are mostly discussed by each market in the online reviews posted on Tripadvisor. Also, frequency analysis uncovered common codes, also known as themes, which helped to define which hotel attributes could be compared afterwards. Finally, an aspect level sentiment analysis was conducted to identify the sentiment polarity of opinions expressed on seven different hotel attributes, namely: room, pool, staff, breakfast, restaurant, beach and view. Finally, a comparison between all three markets is performed.

4. Results

4.1 Comparison of overall market sentiment analysis

As shown in Figure 3, the results obtained from the general sentence-level sentiment analysis present a binary categorization of the sentiment found in the total number of reviews of each market. The automatic process carried by NVivo 12 applies a scoring system where each word containing sentiment has a pre-defined score and it is classified on a Likert-type sentiment scale under the following labels: very negative, moderately negative, moderately positive and very positive. It is possible to visualize that English and Portuguese reviews followed a similar pattern, while French reviews remained centred in the two moderate groups. Still, all three markets wrote more positive sentences than negative ones in their reviews, which validates Bayer and Emir, (2017) as well as Tontini, Dos Santos Bento, Milbratz, Volles and Ferrari (2017)'s findings that travellers mainly write reviews to share their positive experiences instead of the negative ones. The English, Portuguese and French markets presented respectively 71%, 69%, 75% of absolute positive sentiment towards their hotel experience in the Algarve, which represents an average of 71.6%. In terms of overall satisfaction, results show that, if on the one hand, the local market shows a higher percentage of very positive sentiment, on the other hand, it is behind the two other markets in terms of general positive sentiment. Furthermore, although presenting a very low percentage of very positive sentiment, the French residents' reviews display three-quarters of positive sentiment.

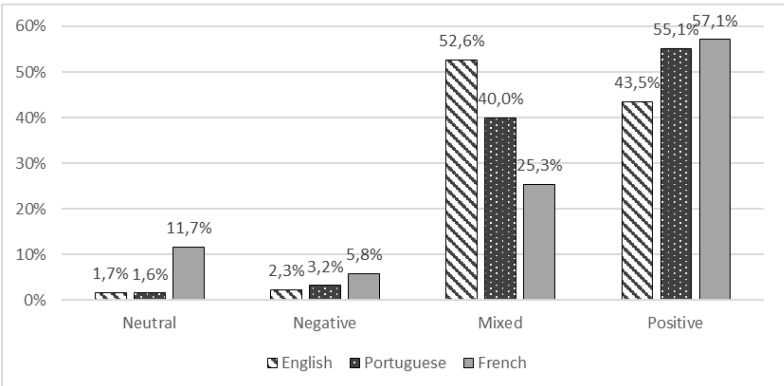
Figure 3 – Overall sentiment of the reviews



When conducting the sentiment analysis of the unstructured data, NVivo searches for expressions of sentiment in the source material, however, it is important to highlight that the tool only can analyse one language at a time and does not recognize sarcasm, double negatives, slangs, dialect variations, idioms or ambiguity. Still, the software provides a handy overview when it comes to the analysis of a large volume of data in a short period of time.

Regarding the total number of sentiment coding references in relation to the total number of reviews detected in each language, it was verified that slightly over 10% of French reviews were classified as neutral. In other words, these coding references were considered to have no sentiment at all. In turn, English and Portuguese text reviews had less than 2% of neutral coding references detected. Also, French reviews presented more negative coding references (5.8%), which represents more negative mentions than the two other markets summed. Concerning mixed references, when you have both positive and negative sentiments identified in the same text passage, results show that more than half of the English reviews (52.6%) mention pros and cons about their hotel experience on a single Tripadvisor review. This pattern of writing reviews was followed by the Portuguese (40%) and only by a quarter of the French posts. With regard to positive coding references, these were mainly detected on French reviews (57.1%), followed by Portuguese (55.1%) and English samples (43.5%). Figure 4 shows the comparison among all three markets.

Figure 4 – Overall sentiment of coding references



4.2 Frequency analysis

As previously performed by Antonio et al. (2018), Ban et al. (2019), Gerdt, Wagner and Schewe (2019), as well as Khoo-Lattimore and Ekiz (2014) a frequency analysis was conducted in order to uncover the 30 most mentioned words associated with hotel experience in each language. This type of analysis is essential, especially at the beginning of the textual data analysis, since it can assist in identifying which hotel attributes are mostly discussed and assessed in the reviews, besides revealing significant results about the interaction between what is written in all three different languages.

Taking this into account, unrelated words, verbs, adjectives, articles, conjunctions, prepositions, pronouns and others were excluded from the query. Only nouns related to hotel attributes were maintained. Also, the word hotel, which was the most frequent noun in the reviews was removed since it is not a hotel attribute, but the hotel unit itself. As shown in Table 2, after conducting the word search query in each one of the languages, it was found that room, staff and pool are respectively the three most mentioned words, regardless of the language used. Similar results are found in Khoo-Lattimore and Ekiz (2014), which found room, staff and food as the most recurrent themes. Also, six terms amongst the top ten most mentioned words are the same in all languages. Apart from the three words aforementioned, breakfast, restaurant and beach figure amongst the ten most-cited hotel attributes.

The outcomes of this query support the selection of hotel attributes that are going to have their sentiment detected and contrasted among all selected markets in order to answer whether satisfaction and dissatisfaction attributes are similar among the British, French and Portuguese markets.

Table 2 – Frequency analysis results

	English		Portuguese		French	
1	**room**	2.21%	**quarto**	2.09%	**chambre**	2.15%
2	**staff**	1.56%	**funcionários**	1.81%	**personnel**	1.31%
3	**pool**	1.39%	**piscina**	1.32%	**piscines**	1.22%
4	**food**	0.96%	**café da manhã**	1.28%	**club**	0.83%
5	**breakfast**	0.92%	**praia**	0.99%	**restaurant**	0.77%
6	**bars**	0.92%	**localização**	0.83%	**plage**	0.70%
7	cleaning	0.89%	qualidade	0.80%	vue	0.68%
8	restaurant	0.88%	serviços	0.73%	petit-déjeuner	0.66%
9	days	0.85%	restaurante	0.65%	séjours	0.65%
10	beach	0.77%	estrelas	0.51%	salle	0.57%

4.3 Automatic and manual theme coding

After conducting the word search query, an automatic coding process is executed on NVivo to identify the most relevant themes in each one of the three languages. The unstructured data are analysed by NVivo 12 with the assistance of a language pack. Themes are then automatically detected by analysing the content and the sentence structure within it. In doing so, the significance is assigned to some themes over others based on how frequently they occur in the reviews.

The themes found by the tool are combined into groups and the results are presented as a code for each broad idea, with child codes attached to each main theme. All the relevant content is coded to the theme codes that are created. The results are summarized in a code matrix which shows the codes for each broad idea, and the number of coding references from each source. Still, human supervision is necessary since text analytics is a complex process and manual coding is always going to be more accurate owing to the fact that human sentiments, such as sarcasm, slangs, idioms or ambiguity cannot be identified by the tool.

During the human checking process, it was verified that the French and Brazilian Portuguese expressions for breakfast, respectively 'petit-déjeuner' and 'café da manhã' were broken into different themes. When separate, the word 'petit' means small and 'déjeuner' means lunch. Hence, human supervision was necessary to identify if customers were evaluating breakfast, lunch or saying that something was small. The same applies to the expression 'café da manhã'. It was necessary to check if comments were related to breakfast (café da manhã), coffee (café) or morning (manhã). Also, similar themes were merged. For example, 'hôtel' and 'l'hôtel' (the hotel) were unified among French themes. 'Pool' and 'pool area' were merged among English themes, since their comments were related to the same hotel attribute.

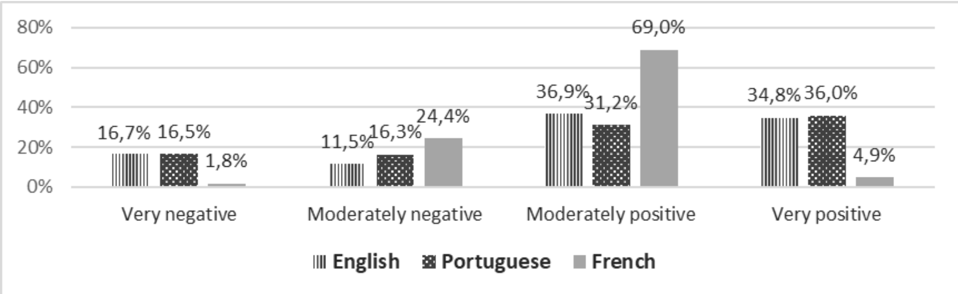
Finally, taking into account the theme coding results and the frequency word outcomes, six hotel attributes are chosen to have their sentiment compared: staff; breakfast; restaurant; pool; room and beach. Besides figuring among the most ten frequent words in all languages, these expressions are also among the most mentioned detected themes. Furthermore, sentiment towards the theme 'hotel' is also checked in the interest of verifying how hotels in the Algarve are evaluated by the selected markets.

4.4 Sentiment analysis towards selected hotel attributes

After concluding the previous process, each hotel attribute is individually submitted to a sentiment analysis performed by NVivo 12. Next, the outcomes are compared between all markets to verify if satisfaction and dissatisfaction attributes are similar among them. However, a prior sentiment analysis towards the theme hotel is conducted with the purpose of verifying the selected markets' general opinion about the topic.

As shown in Figure 5, Portuguese reviews show the highest overall level of dissatisfaction over hotels (32.8%), followed by English (28.2%) and French reviews (26.2%). With respect to positive sentiment, French reviews present almost three quarters (73.9%) of approving text content, whilst only 4.9% were detected as very positive, which is almost seven times weaker in positive sentiment when compared to English reviews (34.8%) and Portuguese reviews (36.0%). According to previous results shown in Figure 4, French reviews sentiments are significantly centred in moderate sentiment levels, whilst English and Portuguese reviews are similarly distributed. Still, as shown in Figure 5, considerable sentiment differences are observed amongst all languages.

Figure 5 – Sentiment towards hotel



With regard to the verified hotel attributes, other dissimilarities are observed as presented in Figure 6. When both moderately positive and very positive sentiment are summed up, reviews written in French stand out among all six selected attributes. That leads to the conclusion that French reviews are in general more positive than the two other analysed languages. Nevertheless, a significant small amount of very positive sentiment is found in French reviews. Staff is the hotel attribute in which the highest percentage of very positive sentiment was found (11%). Breakfast is the feature with the highest rate of moderately positive sentiment (93.8%) and absolute positive sentiment in French reviews (93.8%), while room is the most criticized. Not only approximately 30 percent of negative sentiment towards hotel room is detected, but also the highest amount of very negative sentiment (4.3%).

With respect to English reviews, the highest absolute value of positive sentiment is found towards beach (81.3%), followed by staff (73.7%) and restaurant (73.5%). When it comes to very positive sentiment, staff is the most praised hotel feature among English residents (40.2%). At the other end, as observed in the French reviews, room presents the highest negative sentiment rate, both on an absolute (33%) and very negative sentiment rate (20.9%). Finally, Portuguese reviews present the highest level of very positive sentiment towards one single attribute: beach (50.6%), followed by breakfast (40.7%) and pool (37.4%). Beach also corresponds to the highest absolute positive rate of Portuguese reviews (77.1%). In regard to wholly negative sentiment, Portuguese reviews also demonstrated most of their dissatisfaction towards room (34.1%), followed by staff (33.1%), which was also the attribute where most of the very negative reviews were focused on (16.7%). Figure 6 shows the sentiment comparison between all three languages towards each hotel feature.

Figure 6 – Sentiment towards hotel attributes

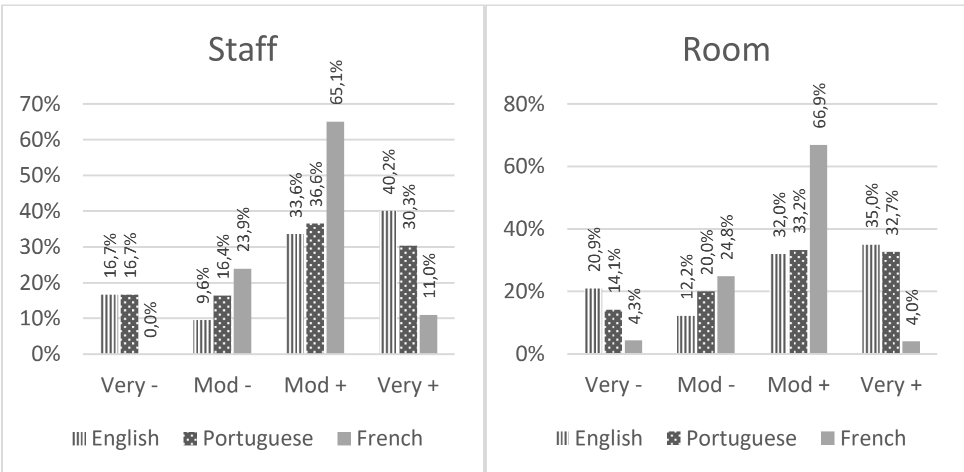
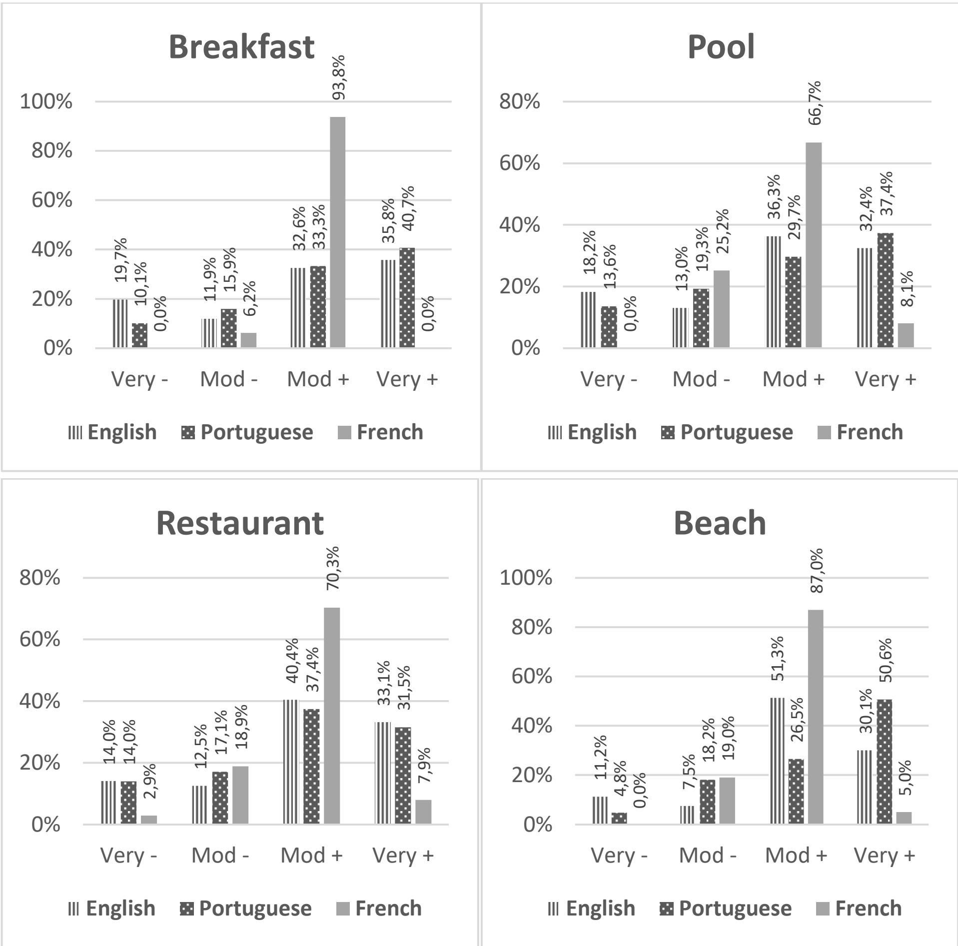


Figure 6 – Sentiment towards hotel attributes



5. Discussion

In line with Ahani et al. (2019b), Antonio et al. (2018), Francesco and Roberta (2019), Mariani and Predvoditeleva (2019), as well as Philips et al. (2020), the outcomes of this study identified different perceptions and opinions from the selected markets when posting online reviews on Tripadvisor. Although some similarities amongst satisfaction and dissatisfaction towards hotel attributes have been observed, in general, English, French and Portuguese online reviews share different opinions about their hotel stay experience in the Algarve. Hence, findings validate the fact that opinions expressed on online reviews might differ depending on guest's cultural background as well as geographic distance.

The results showed the English reviews manifested the highest rates of very positive sentiment towards staff (40.2%), breakfast (35.8%) and room (35%), respectively. Portuguese reviews demonstrated very positive sentiment towards beach (50.6%), breakfast (40.7%) and pool (37.4%). In turn, French reviews had their higher positive sentiment rates towards staff (11%), pool (8.1%) and restaurant (7.9%) accordingly. Despite the fact that staff, breakfast and pool are very positively evaluated by two of the three different markets, these attributes do not present the same rate of satisfaction, which indicates different emphasis and perceptions towards hotel attributes.

If on the one hand, the staff is highly appreciated by British residents, accumulating the highest rate of very positive sentiment (40.2%) and the second-highest rate of absolute positive sentiment (73.7%) (beach is the first one - 81.3%), on the other hand, the same feature mustered the second-highest rate of whole negative sentiment in Portuguese reviews (33%) (room presented the highest whole negative sentiment rate – 34.1%). That not only demonstrates different market opinions but also denotes that hoteliers must make some adjustments to better satisfy their local market in that specific attribute.

Furthermore, a different pattern regarding French reviews is noticed. Even though absolute positive or negative sentiment rates did not present significant fluctuation in comparison with English and Portuguese reviews, a very low rate of strong sentiment is detected. These results may indicate that French residents are satisfied with all hotel attributes, however, there is still room to surpass their expectations. Results also show that, when moderately positive and very positive reviews are summed up, French reviews present a higher rate of absolute positive sentiment than English and Portuguese reviews towards all attributes. Still, in line with Antonio et al. (2018), English reviews presented more sentiment strength than Portuguese reviews.

6. Conclusion and implications

The central objective of this study was to verify how distinct markets evaluate hotels in the Algarve through the analysis of online reviews, in order to identify if satisfaction and dissatisfaction attributes are similar among some of the main markets of overnight stay tourists in the region. After the analysis of 8,596 online reviews related to 161 Portuguese hotel units posted on Tripadvisor between January 1st and December 31st 2019, the results show that, although some similarities amongst satisfaction and dissatisfaction with regard to hotel attributes have been observed, opinions towards online reviews might differ depending on the geographic distance or cultural differences.

According to the general outcomes of this study, hoteliers should prioritize room quality, since this specific hotel attribute is the most mentioned in all languages and gathers the highest average rate of absolute negative reviews (32.1%), which represents one-third of all negative sentiment in all the observed languages. Moreover, special attention should be paid to pool maintenance, quality and condition as this feature has an average rate of 29.8% of whole negative sentiment besides being the third most mentioned hotel feature by all three markets. Ranked third on this list and mainly criticized by the local market is staff (27.7%). Apart from being the second most mentioned word in all languages, such results reinforce that staff issues also need to be addressed by hotels, such as providing an adequate number of personnel and training, besides assuring empathy and efficiency to provide clients with a high-quality service.

Regarding the average rate of absolute positive reviews, findings show that hotels located next to the beach can take advantage of this feature when promoting their business, since this attribute presents an average rate of 83.5%. Additionally, the results suggest that breakfast, which presented the second highest satisfaction rate among all markets (78.7%), can contribute to increase customer satisfaction, especially in the French market, which demonstrated a high approval of this feature. Finally, restaurant was the third-highest average satisfaction rate and hoteliers can exploit this feature to attract customers from the selected markets. As pointed out by Khoo-Lattimore and Ekiz (2014), positive feedback from guest provides guidance to the maintenance of ongoing excellence in customer satisfaction and may also result in favourable electronic word-of-mouth and recommendations.

This study also consolidates the fact that a text analytics software can be a useful tool to assist with word frequency queries and sentiment analysis, however, human supervision is strictly necessary to minimize bias and improve the accuracy of the results. Furthermore, this research confirms that sentiment analysis is a useful tool to quantify the textual reviews posted by guests in foreign languages, extending the findings of Antonio et al. (2018). In closing, hotels can allocate their limited market budget more accurately when their target markets opinions towards the products and services are comprehended and recognized. Hence, data mining and text analysis is proven to be an efficient tool in order to assist hotel businesses to improve their offerings.

This study contributes to the current hospitality management and marketing literature by extending the current research on the use of online reviews, especially regarding the analysis of reviews written in different languages. Also, this study reinforces Philips et al. (2020)'s statement that online travellers' opinions are multifaceted constructs displaying varying patterns of rating behaviour, but also reveals significant dissimilarities among travellers based on geographic distances. In addition, such as Antonio et. Al (2018), Francesco and Roberta (2019), Mariani and Predvoditeleva (2019) and Phillips et al. (2020), the outcomes demonstrate that travellers from different nations place dissimilar emphasis on hotel attributes, as sentiment strength towards hotel attributes differs in each language.

In line with Ahani et al. (2019a; 2019b), Antonio et al. (2018), Francesco and Roberta (2019), Mariani and Predvoditeleva (2019) and Philips et al. (2020), the outcomes of the present study can assist hoteliers with up-to-date practical implications to better understand the predilection of residents from three different European countries, for the purpose of developing efficient marketing strategies or better allocate their market budget. In practice, the method performed by this study demonstrates that English, French and Portuguese residents place different emphasis on hotel attributes. Hence, the findings of this study can assist hoteliers to have a better understanding of the preferences of each segment, which, in turn, can help managers to make important operational and service improvements. By doing so, hoteliers can receive better reviews from satisfied customers and increase their profit. The hotel industry can also use this method as a benchmarking tool and analyse their competitors' performance and make strategic marketing decisions to stand out in such a competitive market.

Still, this study presents limitations that must be considered. First of all, Tripadvisor was the only platform examined by this study. Additionally, the sample size has been limited since only reviews from British, French and Portuguese residents about hotels located in the Algarve and during the year of 2019 were considered, thus the results of this study cannot be generalised. Future studies can validate the aforementioned results by applying different data mining techniques or comparing findings with data obtained from other platforms. Furthermore, the findings of this study can be extended to other countries, periods, different hotel attributes or even uncover which reasons lead to satisfaction and dissatisfaction rates of each market towards the selected hotel attributes.

Funding

This paper is financed by National Funds provided by FCT-Foundation for Science and Technology through project UIDB/04470/2020.

References

Anagnostopoulou, S., Buhalis, D., Kountouri, I., Manousakis, E., & Tsekrekos, A. (2019). The impact of online reputation on hotel profitability. *International Journal of Contemporary Hospitality Management, 32*(1), 20-39. https://doi.org/10.1108/IJCHM-03-2019-0247

Ahani, A., Nilashi, M., & Ibrahim, O. (2019a). Travellers segmentation and choice prediction through online reviews: The case of Wellington's hotels in New Zealand. *Journal of Soft Computing and Decision Support Systems*, *6*(5), 23-30. https://jscdss.com/index.php/files/article/view/209/pdf_249

Ahani, A., Nilashi, M., Ibrahim, O., Sanzogni, L., & Weaven, S. (2019b). Market segmentation and travel choice prediction in spa hotels through Tripadvisor's online reviews. *International Journal of Hospitality Management*, *80*, 52-77. https://doi.org/10.1016/j.ijhm.2019.01.003

Ahani, A., Nilashi, M., Yadegaridehkordi, E., Sanzogni, L., Tarik, A., Knox, K., Samad, S., & Ibrahim, O. (2019c). Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels. *Journal of Retailing and Consumer Services*, *51*, 331-343. https://doi.org/10.1016/j.jretconser.2019.06.014

Almeida, G., & Pelissari, A. (2019). Customer Satisfaction based on the attributes of accommodation services. *Revista Brasileira de Pesquisa em Turismo, 13*(2), 32-53. https://doi.org/10.7784/rbtur.v13i2.1516

Annisa, R., & Surjandari, I. (2019). Opinion mining on Mandalika hotel reviews using latent dirichlet allocation. *Procedia Computer Science, 161*, 739-746. https://doi.org/10.1016/j.procs.2019.11.178

Antonio, N., De Almeida, A., Nunes, L., Batista, F., & Ribeiro, R. (2018). Hotel online reviews: different languages, different opinions. *Information Technology & Tourism*, *18*(1-4), 157-185. https://doi.org/10.1007/s40558-018-0107-x

Ban, H., Choi, H., Choi, E., Lee, S., & Kim, H. (2019). Investigating key attributes in experience and satisfaction of hotel customer using online review data. *Sustainability*, *11*(23), 65-70. https://doi.org/10.3390/su11236570

Bayer, R., & Emir, O. (2017). Examination of urban hotels' online review site performance with a comparative approach based on Tripadvisor and Booking. *Journal of Business Research Turk*, *9*(3), 72-85. http://doi.org/10.20491/isarder.2017.288

Çalı, S., & Balaman, Ş. (2019). Improved decisions for marketing, supply and purchasing: Mining big data through an integration of sentiment analysis and intuitionistic fuzzy multi criteria assessment. *Computers & Industrial Engineering*, *129*, 315-332. https://doi.org/10.1016/j.cie.2019.01.051

Chaves, M., Gomes, R., & Pedron, C. (2012). Analysing reviews in the Web 2.0: Small and medium hotels in Portugal. *Tourism Management, 33*(5), 1286-1287. https://doi.org/10.1016/j.tourman.2011.11.007

Duan, W., Yu, Y., Cao, Q., & Levy, S. (2016). Exploring the impact of social media on hotel service performance: A sentimental analysis approach. *Cornell Hospitality Quarterly*, *57*(3), 282-296. https://doi.org/10.1177/1938965515620483

Egresi, I., Puiu, V., Zotic, V., & Alexandru, D. (2020). Attributes that contribute to guest satisfaction: A comparative study of reviews posted on booking. com and on Airbnb's platform. *Acta Geobalcanica, 6*(1), 7-17. https://doi.org/10.18509/AGB.2020.01

Francesco, G., & Roberta, G. (2019). Cross-country analysis of perception and emphasis of hotel attributes. *Tourism Management*, *74*, 24-42. https://doi.org/10.1016/j.tourman.2019.02.01

García-Pablos, A., Duca, A., Cuadros, M., Linaza, M., & Marchetti, A. (2016). Correlating languages and sentiment analysis on the basis of text-based reviews. *Information and Communication Technologies in Tourism*, *2*(5), 565-577. https://doi.org/10.1007/978-3-319-28231-2_41

Gerdt, S., Wagner, E., & Schewe, G. (2019). The relationship between sustainability and customer satisfaction in hospitality: An explorative investigation using eWOM as a data source. *Tourism Management*, *74*, 155-172. https://doi.org/10.1016/j.tourman.2019.02.010

Gunasekar, S., & Sudhakar, S. (2019). Does hotel attributes impact customer satisfaction: A sentiment analysis of online reviews. *Journal of Global Scholars of Marketing Science, 29*(2), 180-195. https://doi.org/10.1080/21639159.2019.1577155

Han, H., Mankad, S., Gavirneni, N., & Verma, R. (2016). What guests really think of your hotel: Text analytics of online customer reviews. *Cornell Hospitality Report*, *16*(2), 3-17. https://scholarship.sha.cornell.edu/chrreports/4/

Hu, Y., & Chen, K. (2016). Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings. *International Journal of Information Management*, *36*(6), 929-944. https://doi.org/10.1016/j.ijinfomgt.2016.06.003

Instituto Nacional de Estatística (2019, March 24). *Estatísticas do Turismo 2018*. Instituto Nacional de Estatística, I. P. https://www.ine.pt/xurl/pub/358629548

Khoo-Lattimore, C., & Ekiz, E. (2014). Power in praise: Exploring online compliments on luxury hotels in Malaysia. *Tourism and Hospitality Research*, *14*(3), 152-159. https://doi.org/10.1177/1467358414539970

Li, G., Law, R., Vu, H., Rong, J., & Zhao, X. (2015). Identifying emerging hotel preferences using emerging pattern mining technique. *Tourism Management*, *46*, 311-321. https://doi.org/10.1016/j.tourman.2014.06.015

Li, H., Liu, Y., Tan, C. W., & Hu, F. (2020). Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews. *International Journal of Contemporary Hospitality Management, 32*(5), 1713-1735. https://doi.org/10.1108/IJCHM-06-2019-0581

Liang, X., Liu, P., & Wang, Z. (2019). Hotel selection utilizing online reviews: a novel decision support model based on sentiment analysis and DL-VIKOR method. *Technological and Economic Development of Economy, 25*(6), 1139-1161. https://doi.org/10.3846/tede.2019.10766

Lima, S., & Viana, F. (2017). Criterios competitivos en operaciones de servicios de hotelería en el contexto de online travel reviews. *Estudios y Perspectivas en Turismo*, *26*(3), 585-606. https://dialnet.unirioja.es/descarga/articulo/6328656.pdf

Limberger, P., Meira, J., Añaña, E., & Sohn, A. (2016). A qualidade dos serviços na hotelaria: um estudo com base nas on-line travel reviews (OTRS). *Turismo-Visão e Ação*, *18*(3), 690-714. http://www.redalyc.org/articulo.oa?id=261056061012

Litvin, S. (2019). Hofstede, cultural differences, and Tripadvisor hotel reviews. *International Journal of Tourism Research*, *21*(5), 712-717. https://doi.org/10.1002/jtr.2298

Loo, P., & Leung, R. (2018). A service failure framework of hotels in Taiwan: Adaptation of 7Ps marketing mix elements. *Journal of Vacation Marketing*, *24*(1), 79-100. https://doi.org/10.1177/1356766716682555

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, *5*(1), 1-167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Mariani, M., & Predvoditeleva, M. (2019). How do online reviewers' cultural traits and perceived experience influence hotel online ratings?

International Journal of Contemporary Hospitality Management, *31*(12), 4543-4573. https://doi.org/10.1108/IJCHM-11-2018-0927

Oliveira, A., Renda, A. & Correia, M. (2020). Online reviews: A pathway to improve hotel management. *Dos Algarves: A Multidisciplinary e-Journal*, *36*, 108-126. https://doi.org/10.18089/DAMeJ.2020.36.7

Phillips, P., Antonio, N., De Almeida, A., & Nunes, L. (2020). The influence of geographic and psychic distance on online hotel ratings. *Journal of Travel Research*, *59*(4), 722-741. https://doi.org/10.1177/0047287519858400

PORDATA (2019, March 25). *Alojamentos turísticos: total e por tipo de estabelecimento*. Available at: https://www.pordata.pt/Portugal/Alojamentos+tur%c3%adsticos+total+e+por+tipo+de+estabelecimento-2562-211752

Registo Nacional de Turismo (2019, April 9). *Consulta ao registo - Empreendimentos Turísticos*. Available at: https://rnt.turismodeportugal.pt/RNET/Registos.ConsultaRegisto.aspx?Origem=CP&MostraFiltro=True

Tontini, G., Dos Santos Bento, G., Milbratz, T., Volles, B., & Ferrari, D. (2017). Exploring the nonlinear impact of critical incidents on customers' general evaluation of hospitality services. *International Journal of Hospitality Management*, *66*, 106-116. https://doi.org/10.1016/j.ijhm.2017.07.011

Travel BI (2020, April 6). *Turismo em Portugal | 2018*. Available at: https://travelbi.turismodeportugal.pt/ptpt/Documents/Turismo%20em%20Portugal/turismo-em-portugal-2018.pdf

World Tourism Organization (2019, March 25). *International Tourism Highlights, 2019 Edition*. UNWTO. DOI: 10.18111/9789284421152

World Tourism Organization (2020, September 19). *Impact of COVID-19 on global tourism made clear as UNWTO counts the cost of standstill*. UNWTO. https://webunwto.s3.eu-west-1.amazonaws.com/s3fs-public/2020-07/200728-barometer-en.pdf

Xu, X. (2020). How do consumers in the sharing economy value sharing? Evidence from online reviews. *Decision Support Systems, 128*, 113-162. https://doi.org/10.1016/j.dss.2019.113162

Xu, X. (2019). Examining the relevance of online customer textual reviews on hotels' product and service attributes. *Journal of Hospitality & Tourism Research, 43*(1), 141-163. https://doi.org/10.1177/1096348018764573

Xu, X. (2018). Does traveler satisfaction differ in various travel group compositions? *International Journal of Contemporary Hospitality Management*, *30*(3), 1663-1685. https://doi.org/10.1108/IJCHM-03-2017-0171

Xu, X., Wang, X., Li, Y., & Haghighi, M. (2017). Business intelligence in online customer textual reviews: Understanding consumer perceptions and influential factors. *International Journal of Information Management*, *37*(6), 673-683. https://doi.org/10.1016/j.ijinfomgt.2017.06.004

Yi, S., Li, X., & Jai, T. (2018). Hotel guests' perception of best green practices: A content analysis of online reviews. *Tourism and Hospitality Research*, *18*(2), 191-202. https://doi.org/10.1177/1467358416637251

Ying, S., Chan, J., & Qi, X. (2020), "Why are Chinese and North American guests satisfied or dissatisfied with hotels? An application of big data analysis". *International Journal of Contemporary Hospitality Management, 32*(10), 3249-3269. https://doi.org/10.1108/IJCHM-02-2020-0129

Yu, Y., Li, X., & Jai, T. (2017). The impact of green experience on customer satisfaction: Evidence from Tripadvisor. *International Journal of Contemporary Hospitality Management*, *29*(5), 1340-1361. https://doi.org/10.1108/IJCHM-07-2015-0371

Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, *76*, 111-121. https://doi.org/10.1016/j.ijhm.2018.03.017

Zhou, L., Ye, S., Pearce, P., & Wu, M. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. *International Journal of Hospitality*