Know your guests' preferences before they arrive at your hotel: evidence from TripAdvisor

Know Your Guests' Preferences Before They Arrive at Your Hotel: Evidence from TripAdvisor

Purpose

Toward achieving a better guest experience, the current study uses the word frequency comparison technique to evaluate the types of attributes and services that are used most frequently in guests' five- and one-star reviews on TripAdvisor. The paper also investigates the differences between reviews written by men and women.

Design/methodology/approach

A combined sentiment and text analysis was applied to 329,849 UK hotel reviews from UK TripAdvisor to identify factors that influence customer satisfaction, including those with gender differences.

Findings

Our findings reveal important differences between the male- and female-produced terms. The results show that female travelers pay more attention to the hotel's core products and their comfort compared to male travelers. In terms of food and beverage, men's comments tended to focus on pubs, beer, and certain types of food. In contrast, women's comments were more likely to be related to healthy eating, such as homemade, vegan, and vegetarian foods, as well as fruits and healthy breakfasts. Women also pay more attention to the soft skills of staff such as friendliness, helpfulness, and welcoming messages.

Implications

While core attributes of a hotel stay remain crucial for all guests, disparities exist between the language men and women use to describe them. For core products, women pay more attention to the room's cleanliness, comfort, and features such as bed, pillow, blanket, towel, toiletries, and decoration, whereas men pay more attention to the layout, size, and type of room. Hotels may use gender as a segmentation variable and use these findings in their marketing campaigns.

Originality/value

This is one of the first studies offering insights into the differences between the male and female reactions to and preferences for hotel services at a national level. Following a novel method, this study has listed and ranked attributes and differentiated them based on gender.

Keywords: *Guest experience, Hotel Industry, Gender, TripAdvisor, Online reviews*

Introduction

Nowadays, travelers can easily share their experiences online through the social web (Del Vecchio et al., 2018). Google Maps, TripAdvisor, and online travel agencies provide opportunities for travelers to publicly comment on their experiences (Li et al., 2018). Travelers' reviews describe, reconstruct, and disclose their experiences (Min et al., 2014; Xiang et al., 2015). Reviews can influence customers' decision-making processes, including the choice of accommodation (Pantano et al., 2017). This also facilitates interactive communications between hotels and customers (Buhalis and Law, 2008; Sigala, 2008; Ye et al., 2011), giving a communication triangle between a) recommending travelers, b) prospective guests, and c) service providers. From a business perspective, the information shared online provides invaluable support for decision making. Future customer demand and hotels' financial performance are directly impacted by online reviews (Zhao et al., 2019).

The social web heavily affects the hotel industry (Sigala, 2008; Xiang and Gretzel, 2010). TripAdvisor claims to host 702 million reviews about hotels, restaurants, and other attractions (TripAdvisor, 2019). Expedia also claims 7.5 million verified reviews. Mining this content may help managers gain valuable insights into customer-related problems and better understand their needs, attitudes, and behaviors in the interest of drawing business-related conclusions (Expedia, 2019). This underlines the need to use big data to capture customer reactions to tourism-related services. With the importance of gender differences in tourist behavior, it is also critical for hotel businesses to meet gender-specific needs and successfully differentiate their products (Hao and Har, 2014).

To ensure consumer satisfaction and to improve service quality and memorability, hospitality businesses should follow a customer-centric approach. Hospitality businesses recognize the importance of having a customer-centric approach and valuing their customers' needs and expectations to enhance consumer satisfaction and experiences (Correia et al., 2013; Prayag et al., 2013). The hotel industry offers relatively homogeneous products and services. Understanding guests' preferences and satisfaction criteria is essential for distinguishing businesses from their competitors (Hao and Har, 2014; Rahimi, 2008; Rahimi and Gunlu, 2016). Hotel businesses try to identify and understand their guests' preferences to improve their operations and marketing efforts. However, predicting customers' preferences is often challenging (Li et al., 2015). Customer reviews posted online are powerful tools for helping us understand consumer behaviors and expectations (Gour et al., 2021; Rhee and Yang, 2015).

Social web data analysis has been investigated in previous studies (Ye et al., 2009; Yang et al., 2013; Xiang et al., 2015). For example, Alrawadieh and Law (2019) aimed to categorize the determinants of guest satisfaction in hotels by utilizing web data. Their analysis revealed that the size and quality of the room are the key drivers of traveler satisfaction. Although young European male travelers were the ones who posted the most reviews, the study did not consider the gender differences. The sample represented only ten hotels in one urban city with the highest ratings on TripAdvisor. While previous studies have shown that male and female travelers have different satisfaction criteria and preferences, particularly concerning hotel services and products (Chan and Wong, 2006; Sim et al., 2006), no study has investigated this issue in detail in the context of the UK accommodation industry. In response to Sánchez-Franco et al.'s (2019) call for future research to focus on the differences in hotel experiences between men and women because they cognitively organize and structure their stays differently, we are attempting to address the below question.

- *RQ1: What are the main gender differences in satisfaction criteria for UK hotels?*

Due to the high costs of investments in the hotel industry, it is essential to understand which products and services travelers most appreciate. In selecting and evaluating accommodation, there is a gap between what managers believe is important and what guests say is essential (Lockyer, 2005; Rahimi and Kozak, 2016). When choosing a hotel to stay in, customers consider different layers in a service offering such as the core services (e.g., bed, shower) and auxiliary services (e.g., breakfast service) (Grönroos, 2020). Some services enhance the service offering (e.g., spa) whereas others enable the use of the core service (e.g., check-in, booking). In this regard, it is important to find the types of attributes and services that customers deem important. Previous studies with limited sample sizes have listed and ranked attributes, but the findings cannot be generalized (Callan, 1995; Dolnicar and Otter, 2003; Qu et al., 2000). In particular, there is limited information about what customers perceive to be important in high- and low-level service providers. Hence, our second research question.

- *RQ 2: What types of attributes are mentioned in reviews with high (five stars) or low (one star) ratings?*

Literature Review

Hotel Products and Guests' Experiences

Consumer behavior, experience, and satisfaction have a long history in marketing dating back to the pioneering study by Cardozo in 1965 (Gundersen et al., 1996). Studies of consumer behavior and customer experience are vital elements of the post-purchase period in hotels, assisting managers in understanding the reasons for hotel guest satisfaction. The findings can help organizations understand their future customers' needs better and adjust their services (Rahimi, 2017; Rahimi et al., 2017). Hotel services and products can be divided into levels (Heung, 2000). The hotel room is the first-level core product. The next level includes facilities, support, and augmenting elements. This also includes interactions with other customers and service providers that offer various value-adding products and services (Kotler et al., 2006). Hotels may also be described as a set of attributes including location, room, food and beverages, security, and other services (Dolnicar and Otter, 2003; Qu et al., 2000).

Individuals have different proclivities that influence their reactions to hotel attributes. Qu et al. (2000) reported that frontline employees' service delivery, room quality, value for money, safety and security, business-related services, and other services are the main factors in customers' overall experiences. British travelers consider service, staff, and attitude as the most critical factors in their hotel stays (Callan and Bowman, 2000). Hotel attributes were classified by Shanka and Taylor (2004) into room amenities, service experiences, and other physical facilities. Guests rated the service experience as the most important of the three. Zhou et al. (2014) analyzed customer reviews about four- and five-star hotels and identified 23 specific attributes under six main categories: room, hotel facilities, food, price, location, and employees.

Knutson (1988) surveyed frequent business and leisure travelers to investigate the process by which they choose hotels and why they return to the same hotel. By comparing the six hotels' attributes of room, location, safety/security, service, friendliness, and room rates, Knutson (1988) found that the importance of hotel attributes varies depending on the hotel category and the traveler type. Across all traveler types and hotel segments, the clean/comfortable room category was the most crucial aspect. Choi and Chu (1999) investigated hotel guests' perceptions of service quality across three-, four-, and five-star hotels in Hong Kong. They found that a higher hotel classification level led to a greater importance placed on the relevant attributes, except for security and IDD facilities. For five- and four-star hotels, the most important attributes were service quality and room quality, whereas for three-star hotels, security and room quality were the most important.

Tanford et al. (2012) surveyed hotel guests in the US and found that price, utility, amenity, brand, image, and green policies were the most critical factors of satisfaction. Stringam and Gerdes (2010) found that hotel guests refer to the bed and room features more frequently when giving lower ratings. According to Choi and Chu (2001), room attributes, service quality, and value are the main factors that contribute to hotel guests' satisfaction and their intention to return. In their study, Mattler et al. (2006) found that the room, wellness facilities, friendliness and service, reception, and breakfast and restaurant were the key factors affecting guests' satisfaction with lodging services.

Most of the survey methods used in the research reported above have poor sample quality and low response rates, thus leading to broad and unreliable outcomes. This highlights the need for greater use of big data as a significant driver for creating value for firms and customers, particularly in the hotel industry (Verhoef et al., 2016). Investigating data from user-generated websites, such as TripAdvisor, with data mining methods and applications can offer useful insights into service performances (Bilgihan et al., 2018; Filieri et al., 2021). These procedures may help to identify the ancestors of service satisfiers and dissatisfiers of service providers.

Social Web Data in Hospitality and Tourism

Online social media platforms provide great opportunities to collect wealthy, true, and unsolicited data on travelers' experiences (Alaei et al., 2019). The importance of utilizing social media and data mining tools has been studied extensively (Chu et. al., 2020; Dhiratara et al. 2016; Nusair, 2020). Studying customer ratings has the advantage of being able to determine customer satisfaction. Reviews in textual forms reveal travelers' experiences and show the attributes that are valued greatly by customers (Zhao et al., 2019). Two types of online content analysis were used in previous studies: reviews from professional websites (e.g., TripAdvisor) and social media posts (e.g., Twitter). Most datasets relate to hotels (Banerjee and Chua, 2016; Rhee and Yang, 2015; Zhou et al., 2014; Xiang et al., 2015; Zhou et al., 2014). Zhou et al. (2014) found attributes that influence hotel guests' satisfaction by using reviews posted on Agoda.com for hotels in Hangzhou, China. Using customers' reviews from TripAdvisor, Rhee and Yang (2015) ranked various six hotel attributes (value, location, sleep quality, rooms, cleanliness, and service) for hotels. Banerjee and Chua (2016) also found that traveler rating patterns differ between chain versus independent hotels.

Xiang et al. (2015) applied text analysis to consumer reviews from Expedia to investigate the relationship between hotel customer experiences and satisfaction. Liu et al. (2017) analyzed 412,784 user-generated reviews on TripAdvisor for 10,149 hotels from five Chinese cities and found wide disparities in the relationships between foreign tourists' satisfaction and hotel attributes. Li et al. (2020) studied the impact of hotel attributes on traveler satisfaction by reviewing consumer-generated reviews from TripAdvisor across different cities in China. They found that cleanliness, location, room, service, and value are essential attributes for luxury hotels. Zhou et al. (2014) found 17 attributes influencing customer satisfaction. Previous studies have mainly exploited simple statistical textual analysis (e.g., Berezina et al., 2016). Very few studies used large quantities (typically of more than one million records) of data for their analyses. This highlights the need for greater use of big data as a major driver for creating value for firms and customers, particularly in the hotel and tourism field (Mariani and Borghi, 2021; Mariani et al., 2018).

Gender Differences

Guest satisfaction criteria are largely known. However, factors such as gender may affect guests and result in different preferences for specific hotel attributes (Liu et al., 2017). There can be variations between how female and male guests observe the hotel environment and operations as well as in how they evaluate and retrieve information and make judgments. For example, female guests tend to process information in a more detailed fashion than male guests (Babakus and Yavas 2008; Karatepe et al. 2006). According to Snipes and Thomson (2006), there are differences between male and female customers in evaluating service quality. Female guests rate service quality lower than male guests (Lin et al., 2001; Zeithaml et al., 2006). Thus, female guests seem to place more emphasis on the dependability of service and accuracy of information. Females also weigh affiliation and social interactions with employees more heavily (Noble et al., 2006). On the other hand, males attach more importance to tangible aspects while female guests focus on service quality (Mittal and Kamakura, 2001). Despite the importance of gender differences, hoteliers may assume that all consumers, regardless of their gender, have similar preferences, leading to gender blindness and consumer dissatisfaction (Hao and Har, 2014). Therefore, we see a need to account for gender differences in hospitality marketing research and strategies (Smith & Carmichael, 2006). Furthermore, it is vital to understand both groups as consumers/internet users (Moss et al., 2006) and, more importantly, to capture their views and expectations in their online expressions.

Methodology

TripAdvisor UK data and methods

The raw data used in this study was an extensive random collection of user reviews about UK hotels extracted from the TripAdvisor UK website. Reviews were only taken from a single country to increase the homogeneity of the dataset. This reduces the risk that any gender differences identified are a result of spurious factors, such as international differences in the proportion of women that comment. Only UK reviewers (see below for details) were included for the same reason and because it is known that there are major international differences in cultural styles for writing reviews and for the factors considered important (Hofstede, 1983).

The TripAdvisor.co.uk website was used because it contains reviews in English of UK-based attractions. It maintains a comprehensive list of URLs of webpages containing reviews (and other site webpages) as part of its public sitemap database. All 13,234,039 URLs of pages containing reviews of English destinations were extracted from the sitemap. These URLs contained the identifying URL segment *ShowUserReviews*. A random number generator was utilized to assign a number to each URL the URLs were sorted by these values before selecting one in 20 URLs to give a large final sample. The result is a genuine random sample of TripAdvisor review pages for England. These URLs were crawled using the free web crawler SocSciBot (socscibot.wlv.ac.uk) at a rate of one per second during February 2017. SocSciBot obeys the robots.txt convention to govern web crawling and the one-URL-per-second speed restriction was designed so as not to risk overloading the TripAdvisor web servers (for crawling ethics, see Thelwall and Stuart, 2006).

Information about the reviews crawled with SocSciBot was extracted by a program added to the free big data analysis software Mozdeh (mozdeh.wlv.ac.uk). For each webpage, it extracted the text of each review (most pages contained multiple reviews) as well as the name of the reviewer, their geographic location (if they had entered one), and their review rating (10, 20, 30, 40, or 50). In addition, Mozdeh extracted the type of attraction as categorized by TripAdvisor (e.g., Hotel, Camping, Things to do) and its AA star rating (https://www.tripadvisorsupport.com/hc/en-us/articles/200614057-What-do-star-ratings-for-hotels-mean-), if any (none, one to five stars, including half stars). A review can appear on multiple pages within TripAdvisor and so, before analysis, duplicate reviews (i.e., with the same author name and text) were automatically filtered out. This resulted in 3,067,237 reviews about all attractions.

Each reviewer was assessed by Mozdeh to ensure they were based in the UK. A reviewer was categorized as British if their address contained United Kingdom, UK, Great Britain, England, Scotland, Northern Ireland, Wales, or the names of one of the ten largest UK cities (London, Birmingham, Leeds, Glasgow, Sheffield, Bradford, Edinburgh, Liverpool, Manchester, Bristol). Spaces and commas were employed to avoid false matches, such as New England (USA) instead of England, UK. Some large UK cities have namesakes overseas. Notable examples include Birmingham (Alabama) and London (Ontario). It seemed intuitively likely that the majority of reviews of UK attractions written by people living in cities called London or Birmingham would be British. Nevertheless, to check this, random samples of users matching the above criteria were checked for false matches and none were identified. For example, two-thirds of people giving a London address also specified that they were from the UK and none indicated that they were from a different London. Thus, the error rate for this method seems unlikely to be much higher than 1 in 1000. To focus on hotels, attractions of other types were filtered out, leaving hotels from one to five stars, including half stars for the final analysis.

User genders were needed for some of the analyses. Although users can register their gender in TripAdvisor and this information can be extracted from their homepages on the site, this information is not displayed alongside their review. It would have been possible to crawl the homepages of all reviewers, but this would cause an undesirable substantial increase in the crawling of TripAdvisor (more pages than the original crawl). Instead, an alternative gender detection heuristic was used. This method extracted users' first names and inferred their genders therefrom when evident. The first name of an individual was assumed to be the first text segment of their name, terminating at the first space. Usernames without spaces but employing camel case (i.e., lower case followed by upper case, as in "WordBoundaries") were instead split before the first capital letter after the initial letter. The entire username was assumed to be the first name in the remaining cases. The first names were then checked against dictionaries of words that are overwhelmingly used for males or females to assign a likely gender. For example, all users called Mohammed would be categorized as male, but all users called Susan would be categorized as female.

The gendered list of names was taken from a publicly available source derived from the genders of people submitted to the 2001 US census, covering the many different nationalities present in the USA. Names were included only when 90% of people were either male or female. A UK-based list would have been preferable, but the most extensive evidence-based list was that of the USA. Although this method causes some false matches (e.g., Alison using Ali as a short form name), it seems to be reasonably accurate overall. Its main drawback is that it leaves unclassified about two-thirds of users who choose abstract usernames or have rarer first names or less gender-specific names. The first name list was extended to include "Mr," "Mrs," "Ms," and "Miss."

Sentiment analysis

Sentiment analysis extracts opinions, attitudes, and emotions from online resources (Li et al., 2017). To automatically express the strength and polarity of sentiment expressed in the reviews, the program SentiStrength (Thelwall et al., 2012) was used and adapted to TripAdvisor. SentiStrength employs a lexicon of just under 3000 positive and negative words and word stems, each with an estimated positive or negative term strength. It matches these to the review tests and uses the scores to evaluate the strength of positive and negative sentiment in a review on a scale of 1 (no positive sentiment) to 5 (very strong positive sentiment) and -1 (no negative sentiment) to -5 (very strong negative sentiment). Thus, a review that contains the phrase "I hate the nice room" would get a dual score (-4, 2) due to the words "hate" [-4] and "nice" [2]. These scores are modified by a range of linguistic rules to deal with other ways of expressing or modifying sentiment, such as negation, questions, emoticons, and booster words. These methods, in tandem, enable SentiStrength to estimate the strength of positive and negative sentiment in a text with close to human-level accuracy (Thelwall et al., 2010).

SentiStrength was modified for the TripAdvisor dataset to align it with reviews in three different ways. First, the dual scales were converted to a single overall scale of -4 to 4 by totaling the positive and negative scores. Second, the SentiStrength scores were compared to TripAdvisor ratings after converting the latter to the modified SentiStrength scale as follows: 10 -> -4, 20 -> -2, 30 -> 0, 40 -> 2, 40 -> 4. The machine learning feature of SentiStrength was then used to adjust its sentiment dictionary terms to align with the ratings given by reviewers. The machine learning was conducted on a balanced subset with equal numbers of 10, 20, 30, 40, and 50 reviews. Finally, the term suggestion feature of SentiStrength (Thelwall and Buckley, 2013) was used to suggest additional sentiment terms to add to the dictionary. These terms were manually checked and added. These included, for example, many terms related to dirt and animal infestations. The result of this was a version of SentiStrength that was tailored to detect sentiment in TripAdvisor reviews in a way that aligns with the TripAdvisor rating system and is sensitive to the words used by customers in their reviews.

Word frequency comparisons

A word frequency comparison method was used to compare the proportion of reviews that contain a term between two different groups (male-authored reviews vs. female-authored reviews; reviews with high ratings vs. reviews with low ratings). Any word that occurs in a higher proportion of reviews for one group than for the other potentially associates with either something that is more important for one group than with another or a different communication style.

The word frequency method measures the statistical significance of the difference between the two groups using a 2x2 chi-square test that assesses whether there is a difference between the two groups in the proportion of reviews that contain the given word. The chi-square method is well suited for sentiment analysis studies as a successful feature selector tool (Khan et al., 2016). A chi-square test involves checking whether observed frequencies in one or more categories match expected frequencies. Thus, higher chi-square values indicate a stronger association between a term and one of the groups compared to the other group. In theory, a chi-square score above the critical value (3.84 for a test at the 0.05 level, 6.64 for the 0.01 level, and 10.8 for the 0.001 level) is statistically significant, but because the word frequency method calculates many chi-square values simultaneously, the confidence levels associated with them are not robust (this is the Bonferroni problem; see Perneger, 1998). Therefore, a Benjamini-Hochberg correction was used to safeguard the familywise error rate for each set of chi-square word frequency tests (Benjamini and Hochberg, 1995). The correction factor included all tests with a word frequency high enough to give a statistically significant result.

Plural words were converted into singular before comparing word frequencies to merge words with the same meaning. No further stemming was conducted as it would have risked merging similar terms with different meanings. For example, even the basic Porter (1980) algorithm merges “park” and “parking,” which may be unrelated in the context of hotels. Similarly, topic modeling algorithms (e.g., Wallach, 2006) were not used as these tend to obscure fine-grained distinctions and are not transparent to interpret.

Results

Gender differences in satisfaction criteria for UK hotels

The absence of credible research on male and female approaches toward hotel services was one of the main drivers of this research objective. The results below (refer to Tables 1 and 2) show the differences between male- and female-generated terms among 329,849 downloaded reviews. Through reading comments about the terms and conducting follow-up word association tests on them, the criteria for both male and females were grouped into the following themes: *Core Products (Room-related)*, *Core Products (Food and Beverage-related)*, *Hotel's Facilities*, *Related External Factors*, *Positive Sentiments*, *Negative Sentiments*, *Occasion*, *Female-related Terms*, *Male-related Terms*, *Kids*, *Personal City/Destination*, and *Shop* (Qu et al., 2000; Dolnicar and Otter, 2003; Rahimi and Kozak, 2016; Zhou et al., 2014; Banerjee and Chua, 2016; Rhee and Yang, 2015; Thelwall, 2018).

This provides a dictionary of the terms that are more used by males and females. Statistically significant terms were randomly selected for follow-up word association tests. For example, "Clean" was mainly associated with "room." "Cosy" was mainly associated with "room," "bed," "lounge," and "sofa." "Fresh" was associated with "fruit," "air," "milk," "coffee," "salad," and "tea." "Friendly" was associated with "staff." "Wide" was associated with "room" and "corridors." "Timber" was associated with "wall" and "floor."

Table 1- Themes derived from terms occurring more in comments authored by females than in comments written by males at the 0.1 percent level, using a chi-square test with a Benjamini-Hochberg correction.

Theme	Statistically significant terms within the theme (Female)
Core Products (Room)	Clean, Cosy, Accommodation, Decorated, Comfy, Toiletries, Flower, Hairdryer, Candle, Warm, Towel, Bedding, Fluffy, Robe, Comfortable, Bed, Homely, Decorate, Slipper, Gown, Blanket, Dryer, Pillow, Mirror, Toilet, Sofa,
Core Products (Food and Beverage)	Delicious, Cake, Tea, Yummy, Scone, Homemade, Chocolate, Cocktail, Gluten, Clotted Cream, Sandwiches, Scrummy, Jam, Fruit, Biscuit, Catered, Dessert, Prosecco, Milk, Lunch, Café, Smoothly, Scrumptious, Fresh, Brownie, Pudding, Vegan, Cuppa, Cupcake, Strawberry, Quiche, Tea, Mulled Wine, Ice Cream, Yogurt, Pancake, Veggie, Champagne, Breakfast, Dairy, Coffee, Marshmallow, Panini, Ice, Allergy, Diet, Milkshake, Cheesecake, Teapot, Toffee,
Facilities	Spa, Massage, Facial Treatment, Pool, Therapist, Garden, Jacuzzi, Swim, Hairdresser, Bubbly, Lodge, Singing, Balcony, Tearoom, Lifeguard,
External factors/ Attributes	Helpful, Friendly, Staff, Welcoming, Treat, Wait, Picnic, Peaceful, Weekend, Atmosphere, Entertainment, Beach, Welcome, Waitress, Caravan, Dancing, Voucher, Groupon, Farm, Sunshine, Thoughtful, Vegetarian, Lush, Outdoor, Zoo, Weather, Downfall, Safari, Smelt, Sand,
Positive Sentiments	Lovely, Beautiful, Amazing, Fab, Loved, Gorgeous, Fabulous, Wonderful, Beautifully, Absolutely, Brilliant, Perfect, Fantastic, Defiantly, Extremely, Highly, Cute, Stunning, Spotless, Trilled, Magical, Super, Excited, Incredible, Immaculate, Delightful, Loveliest, Impressed, Incredible, Adore, Hilarious,
Negative Sentiments	Trouble, Disappointment, Weren't, Hadn't, Wasn't, Couldn't, Didn't, Rude, Disgusted, Disappoint,
Occasion	Birthday, Wedding, Christmas, Celebrate, Anniversary, Bridal, Bride, Bridesmaid, Party, Hen, Honeymoon, Groom
Female	Mum, Daughter, Sister, Girl, Mother, She, Girly, Granddaughter, Niece,
Male	Husband, Boyfriend, Hobby, He, Son, Dad, Partner, His, Grandson, Nephew,
Kids	Children, Child, Toddler, Baby, Babies, Pushchair, Playground, Highchair,

Theme	Statistically significant terms within the theme (Female)
Personal	We, Our, Us, Family, X, Xx, Relaxing, Gift, Dog, Xxx, Married, Age, Luxurious, Romantic,
Female related terms	Hair, Mam, Lady, Spoilt, Manicure, Makeup, Lavender, Fiancé, Pamper, Pedicure, Pregnant, Rose,
City / Destination	Cottage,
Shop	-----

Table 2 shows the dictionary of male extracted terms.

Table 2- Themes derived from terms occurring more in comments authored by males than in comments written by females at the 0.1 percent level, using a chi-square test with a Benjamini-Hochberg correction.

Theme	Statistically significant terms within the theme (male)
Core Products (Room)	Adjacent, Larger, Wide, Tidy, Wi-Fi, Timber, Layout, En-Suit, Feature, Refurbishments,
Core Products (Food and Beverage)	Beer, Pint, Pub, Wine, Brewery, Cask, Draught, Finest, Restaurant, Bar, Indian, Cider, Cuisine, Brewed, Thirsty, Dining, Bitter, Michelin, Steak, Italian, Madras, Gastro, Culinary, Doombar, Peperoni, Boozer, Cheers, Cod Fillet, Cooked, Curry, Restaurant, Grill, Gastronomic, Guinness, Vindaloo, Desert, Dinning, Sommelier, Dish, Chinese, Whiskies, Jalfrezi,
Facilities	Service, Parking, Yard, First-Class,
External factors/ Attributes	Quality, Recommendation, Value, Selection, Venue, Price, Railway, Countryside, Summary, Aircraft, Reputation, Historic, Locomotive, Vehicle, Coast, Pricing, Cheapest, Welcome, Enthusiast, Golf, Valley, Museum, Village, Victorian, Greeting, Aviation, Fishing, Cycling, Value-For-Money
Positive Sentiments	Excellent, Good, Superb, Decent, Reasonable, Pleasant, Indeed, Superbly, Terrific, Certainly, Comprehensive, Extensive, Acceptable, Tremendous, Magnificent, Worthwhile, Ample, Outstanding, Sensibly, Impressive, Enjoyable, Exceptional, Excellent, Exemplary, Admirably, Competitively, Friendly, Significant, Remarkably, Worthy, Marvellous
Negative Sentiments	Poor, Avoid, Nonsense, Patronised
Occasion	Stag,
Female	Wife, Girlfriend, Missus, Ex,
Male	Mate, Mr, Lad, Guy,
Kids	----
Personal	Class, Standard, Premier, Style, Premium,
Male related terms	Investment, Council, Match, Football, Engine, Sport, Rugby, Motorcycle, Cricket, League
City / Destination	UK, England, British, Sussex, Europe, Devon, Midlands, France,
Shop	Outlet, Weatherspoon, Hilton, Beefeater

Male and female travelers have different preferences, particularly concerning hotel services and products (Chan and Wong, 2006; Sim et al., 2006). As mentioned before, hotel businesses must meet gender-specific needs/expectations to successfully differentiate their products (Hao and Har, 2014). Table 3 shows theses difference based on the most significant generated terms (***.alpha=0.001, **.alpha=0.01, *alpha=0.05). The percentage shows the word's match in the total reviews.

Table 3 - rank-order of preferences based on gender.

Themes	Female	Male
Core Products (Room)	Cleanness (14.60***) Comfort (8.60***) Features of the Room such as Bed, Pillow, Blanket, Towel, Toiletries And Decoration	Layout 0.50 *** Adjacent (0.50 ***) En-Suite (0.40 ***)
Core Products (Food and Beverage)	Breakfast (16.80***) Fruit (1.70***) Vegetarian (1.60***) Homemade Food (1.30 ***) Vegan (0.30 ***)	Pub (9.30 ***) Beer (5.30 ***) Pint (1.60 ***) Indian (1.80 ***) Italian (1.60 ***) Chinese (0.80 ***) Nigerians (0.10 ***)
Facilities	Pool (2.50 ***) Spa (1.90 ***) Swim (1.00 ***) Hairdresser (0.80 ***) Massage (0.40 ***) Jacuzzi (0.40 ***) Facial Treatment (0.20 ***)	Services (27.50 ***) Parking Facilities (4.90 ***)
External factors/ Attributes	Friendliness (26.60 ***) Helpfulness (12.40 ***) Atmosphere (9.40 ***) Welcoming (6.50 ***) Peaceful (1.00 ***)	Quality of service (9.80 ***) Value for the money (9.70 ***) Recommendation (4.30 ***) Countryside (1.50 ***) Museum (1.40 ***) Golf club (0.60 ***) Fishing (0.20 ***) Cycling (0.20 ***)
Occasion	Birthday (5.10***) Wedding (2.60 ***) Christmas (2.10 ***) Anniversary (1.10 ***) Honeymoon (0.10 ***)	Stag (0.30 ***)

Types of attributes mentioned in reviews with high (5 stars) or low (1 star) ratings

Five-star reviews for hotels were compared to reviews with lower ratings (Table 4). The results revealed the most statistically significant criteria that are important for British travelers when choosing hotels and writing positive reviews about them. The results also show the most positive sentiments used by travelers to explain their experiences. The criteria are above the 0.1 percent level and are grouped into themes. This was followed by reading comments containing the terms and conducting follow-up word association tests on them. Words are listed in the table based on the ascending order of both statistical significance and chi-square.

Table 4 – Themes derived from terms occurring more in comments with high ratings (5 stars) than in comments with other ratings (1-4 stars) at the 0.1 percent level, using a chi-square test with a Benjamini-Hochberg correction.

Theme	Statistically significant terms within the theme
Core Products (Room)	Comfortable, Accommodation, Cosy, Furnished, Fluffy, DVD, En-Suite, Candle
Core Products (Food and Beverage)	Delicious, Fresh, Homemade, Cake, Tastefully, Catered, Yummy, Scrumptious, Chocolate, Finest, Refreshing, Indian, Tastiest, Scrummy, Breakfast, Champagne, Italy, Mouth-Watering, Dietary, Biscuit, Lush, Feast, Icing, Gluten, Vegan, cocktail, flapjack, clotted cream
Facilities	Garden
External factors/ Attributes	Welcoming, Host, Friendly, Spotless, Team, Helpful, Knowledgeable, Home, Atmosphere, Professional, Hesitate, Tour, Treat, Hospitality, Peaceful, Informative, Freshly, Guide, Knowledge, Homely, Countryside, Fun, Passionate, Thoughtfully, Warmly, Farm, Greeted, Recondition, Lake, Compliment, Beach, Weekend, Hospitable, Smiling, Touch, Friendliness, Caring, Staff, Flower, Museum, Kindness, Coast, Professionalism, Wildlife, Fiona, Cathedral, Zoo, Castle
Positive Sentiments	Excellent, Amazing, Wonderful, Greet, Lovely, Perfect, Superb, , Best, Beautiful, Brilliant, Fabulous, Beautifully, Recommended, Absolutely, Perfection, Thoroughly, Gorgeous, Outstanding, Stunning, Delightful, Truly, Fab, Exceptional, Immaculate, Incredible, Pleasure, Faultless, Wow, Well, Awesome, Memorable, Impeccable, Divine, Wonderfully, Exceeded, Unique, Magical, Sublime, Perfectly, Charming, Spacious, Nicest, Magnificent, Fascinating, Breathtakingly, Terrific, Treasure, Genuinely, Impressed, Executional, Amazingly, Brilliantly, Heavenly, Unforgettable, Glorious, Sensational, Thrilled
Negative Sentiments	Fail
Occasion	Birthday, Wedding, Anniversary, Celebrated, Congratulations, Surprise, Honeymoon
Female	Sue, Penny, Sarah, Karen, Jane, Liz, Caroline, Helen, Linda, Lisa, Carol, Emma, Debbie, Sally, Mum
Male	David, Dave, Andy, John, Mike, Paul, Ian, Simon, Richard, Tony, Andrew, Rob, Phil, James, Gary, Kevin, Nick, Martin, Tim, Colin, Sam
Kids	---
Personal	Class, Relaxed, Family, Xx, X, Married, Luxuries, Stylish, Xxx, Jewel
City / Destination/type of accommodation	Cottage, Cornwall, Blackpool, Guesthouse, Devon, Liverpool, York
Shop	---

The results were achieved by selecting one-star reviews out of 329,849 downloaded reviews. The results revealed the most statistically significant criteria for British travelers when rating down an accommodation provider and highlight the negative sentiments they used to explain their experience. The criteria are above the 0.1 percent level and are grouped into themes (Qu et al., 2000; Dolnicar and Otter, 2003; Rahimi and Kozak, 2016). This was followed by reading comments containing the terms and conducting follow-up word association tests on them (Table 5). Words are listed in the table based on the ascending order of both statistical significance and chi-square.

Table 5- Themes derived from terms occurring more in comments with low ratings (1 star) than in comments with other ratings (2-5 stars) at the 0.1 percent level, using a chi-square test with a Benjamini-Hochberg correction.

Theme	Statistically significant terms within the theme
Core Products (Room)	Carpet, Mould, Toilet, Stain, Frozen, Damp, Soggy, Cleaned, Dirt, Dump, Stained, Uncomfortable, Freezing, Dust, Mouldy, Cigarette, Sheet, Curtain, Window, Wallpaper, Matters, Pain
Core Products (Food and Beverage)	Tasteless, Inedible, Burnt, Raw, Kitchen, Smelt, Plague, Smelly, Overcooked, Undercooked, Microwaved, Poisoning, Cutlery, Hygiene, Greasy, Watery, Stank, Uncooked, Reheated, Chicken, Served, Sause, Chef, Water, Chips, Lettuce, Bar, Unappetising, Tomato, Vomiting, Stomach, Burger, Fried, Salty, Flavourless, Cup, Onion
Facilities	Security,
External factors/ Attributes	Dirty, Asked, Manager, Refund, Complain, Waste, Cold, Apology, Refused, Attitude, Overpaid, Ruined, Waiting, Waitress, Unhelpful, Ignored, Reply, Broken, Mistake, Bothered, Angry, Smell, Rudely, Unprofessional, Sick, Waiter, Bother, Wait, Excuse, Care, Apologise, Insult, Unfriendly, Shouted, Shouting, Poorly, Crap, Overprice, Cancelled, Behaviour, Misfortune, Nasty, Compensation, Urine, Unclean, Rudeness, Rudest, Advertised, Payment, Manner, Arguing, Lied, Ignorant, Shabby, Abuse, Reluctant, Robed, Inconvenience, Queue, Drug
Positive Sentiments	Politely,
Negative Sentiments	Worst, Terrible, No, Disgusting, Never, Filthy, Avoid, Appalling, Bad, Worse, Shocking, Horrible, Dreadful, Rubbish, Did Not, won't, Disgrace, Shocked, Disgraceful, Aggressive, Wasn’t, Hadn’t, Arrogant, Badly, Shame, Unacceptable, Barley, Appalled, Wouldn’t, Embarrassed, Disappointing, Unpleasant, Ashamed, Ridiculous, Miserable, Couldn’t, Horrendous, Disaster, Awful, Poor, Pathetic, Hell, Abrupt, Humiliated
Occasion	
Female	Girl, Women
Male	He, Him
Kids	----
Personal	----
City / Destination	----
Shop	Iceland

Discussions and Conclusions

Conclusions

This study aimed to evaluate the types of attributes and services that are used most frequently in guests' reviews on TripAdvisor with five-star ratings as well as those that are less desirable with one-star ratings. Online reviews provide an irreplaceable opportunity for hotel businesses to understand their customers' views and expectations and make better strategic and tactical decisions to create value. Focusing on the service sector, customers expect to be treated as individuals and to be provided with services that they desire, not a standard solution (Gwinner et al., 2005). As it is too expensive to understand all customers' needs/wants, they can be classified into different segments based on other criteria such as their gender (Kim et al., 2018; Cobanoglu et al., 2003). This study used 329,849 hotel reviews from Trip Advisor in the UK to investigate the main gender differences in satisfaction criteria for UK hotels.

Theoretical and Practical Implications

Overall, our study findings are in line with previous studies (Callan and Bowman, 2000; Choi and Chu, 1999; Nasution and Mavondo, 2008; Qu et al., 2000), highlighting the importance of hotel attributes such as clean/comfortable room, convenient location, safety/security, prompt/courteous service, friendliness, and room rates. The analysis reveals statistically significant differences between male and female terms. For core products, women pay more attention to the cleanness, comfort, and features of the room such as bed, pillow, blanket, towel, toiletries, and decoration whereas men pay more attention to the room's layout, size, and type. In general, female travelers pay the most attention to the core products of the hotel and their comfort.

For food and beverages, the main foci of male comments were on the pub, beer, pint, and individual types of food, such as Indian, Italian, Chinese, or Nigerian. In comparison, the main matches in females' comments were related to healthy eating and dietary types, such as vegan, vegetarian, fruit, and healthy breakfast. This suggests that females pay more attention to healthy eating than males. In terms of facilities, female-associated terms included Spa, Massage, Facial Treatment, Pool, Jacuzzi, Swim, and Hairdresser while male terms were more focused on services and parking facilities. This is in line with Choi and Chu (2001). For external factors/hotel attributes, males write more about the Quality of service, Recommendation from friends, Value for the money, and nearby facilities such as Countryside, Golf club, Fishing, Cycling, and Museum. Females write more about friendliness, helpfulness, and welcoming messages from staff. Females seemed to be mainly looking for a peaceful place with an entertaining atmosphere. Most previous studies have highlighted the role and importance of staff performance and friendliness in customers' overall satisfaction (Matzler et al., 2006; Qu, Ryan, and Chu, 2000). Our results have shown that these factors are more important for female travelers.

Considering occasions, females mentioned a range of terms that are important for them (i.e., Birthday, Wedding, Christmas, Anniversary, Honeymoon) while the only occasion term associated with male comments was Stag. For family-related terms, females used many, including Children, Child, Toddler, Baby, Babies, Pushchair, Playground, and Highchair while males generally used none. In line with Hao and Har (2014), safety and location are more important considerations for female customers than males. Females are more likely to use warm, positive sentiment terms such as Lovely, Beautiful, Amazing, Fab, Loved, Gorgeous, Fabulous, Wonderful, Beautifully, Absolutely, Brilliant, Perfect, and Fantastic to express their experiences. In contrast, males use more technical or weaker positive language such as Excellent, Good, Superb, Decent, Reasonable, Pleasant, Indeed, Superbly, Terrific, and Certainly. The negative sentiments of Disappointment and Disgusted are more used by females while males use terms such as Poor and Nonsense.

The types of attributes mentioned in reviews with high (5 stars) or low (1 star) ratings were also investigated. For core products, comfort, room attributes, and facilities are the main terms. Likewise, Mccartney and Ge (2016) found that the quality of service and facilities weighed most heavily in the minds of travelers searching for hotels. The terms included fresh, homemade, finest, tasteful, and mouth-watering for the food and beverage theme. Other essential words were vegan and vegetarian, which shows the importance of vegetarian options (Waters, 2018). Given that only reviews from the UK were analyzed, some terms in this theme were geographically specific. For example, the term of Devon was mentioned by travelers and is associated with cream, tea, scones, clotted cream, and jam (Devonshire tea and Cornish cream tea).

Other external factors and hotel attributes play critical roles in high rating reviews. These include hotel attributes such as the garden, atmosphere, or staff helpfulness. These are in line with Zaman et al.'s (2016) study about the six hotel selection criteria—location, sleep quality, comfort and equipment, service, value for money, and cleanliness—as the main factors influencing the decision-making process of travelers. The terms that were associated with staff in five-star reviews were Knowledge, Greeted, Hospitable, Thoughtfully, Friendliness, Professionalism, Caring, Kindness, Friendly, Welcoming, Team, Knowledgeable, Hospitality, Treat, Informative, Passionate, Fun, and Warmly. Many reviewers also mentioned the location of the hotel and being close to the Cathedral Museum, Castle, Zoo, and Beach. The only negative sentiment among all five-star comments was *fail*. The most mentioned destinations were Cottage, Cornwall, Blackpool, Guesthouse, Devon, Liverpool, and York. For the one-star reviews, the terms that emerged for the core products were Mould, Frozen, Damp, Dirt, Uncomfortable, Dust, and Stained.

No terms related to occasions, children, personal, and city/destination. No positive sentiment was detected. This suggests that very dissatisfied customers usually ignore any positive aspect of the services they receive (Gelbrich and Roschk, 2011; Tronvoll, 2011). Previous studies focused on hotel choice factors (Kim and Park, 2017; Kim and Perdue, 2013; Yavas and Babakus, 2005). The above findings are in line with previous studies highlighting the hotel attributes that are crucial when hotel guests select a hotel and how different groups select certain attributes. Our study provides specific practical implications for hotel managers.

Limitations

This study is not free of limitations. First, our study is restricted to hotels in the UK with English language comments, but the results may differ in other countries and in other languages. Only men and women were examined, their genders inferred from their usernames. Further studies can investee the customer experience and gender preferences in different contexts such as Airbnb, restaurants, spa facilities, or destinations. This research suffers from the inherent limitations of sentiment analysis. Sentiment vocabularies may have affected the results. The current research does not consider other segmentation variables such as traveler nationality or traveler type (e.g., business vs. leisure). Comparative research that examines the satisfiers of domestic business travelers vs. international leisure travelers would be valuable for hoteliers. Business and leisure travelers might also have different preferences based on the purposes of their trips.

References

Alrawadieh, Z., & Law, R. (2019). Determinants of hotel guests' satisfaction from the perspective of online hotel reviewers. *International Journal of Culture, Tourism and Hospitality Research, 13* (1), 84-97.

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: capitalizing on big data. *Journal of Travel Research*, *58*(2), 175-191.

Alexander, M. G., & Wood, W. (2000). Women, men and positive emotions: A social role interpretation. In A. Fischer (Ed.), Gender and emotion: Social psychological perspectives (pp. 189-210). Cambridge: Cambridge University Press.

Baggio, R. (2016), "Improving tourism statistics: merging official records with big data", in Fuchs, M., Höpken, W. and Lexhagen, M. (Eds), Big Data and Business Intelligence in the Travel and Tourism Industry, Mid Sweden University, Östersund, pp. 89-92.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological), 57(1), 289-300.

Berezina, K., Bilgihan, A., Cobanoglu, C. and Okumus, F. (2016), "Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews", Journal of Hospitality Marketing & Management, Vol. 25 No. 1 pp. 1-24.

Bilgihan, A., Seo, S., and Choi, J. (2018), "Identifying restaurant satisfiers and dissatisfiers: Suggestions from online reviews", Journal of Hospitality Marketing & Management, Vol. 27, No. 5, 601-625.

Bowen, J. T., & Chen, S. L. (2001). The relationship between customer loyalty and customer satisfaction. International journal of contemporary hospitality management, 13(5), 213-217.

Bryson, S., Kenwright, D., Cox, M., Ellsworth, D. and Haimes, R. (1999), "Visually exploring gigabyte datasets in real time", Communications of the ACM, Vol. 42 No. 8, pp. 83-90.

Buhalis, D. and Foerste, M. (2015), "SoCoMo marketing for travel and tourism: empowering cocreation of value", Journal of Destination Marketing and Management, Vol. 4 No. 3, pp. 151-161.

Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. Tourism management, 29(4), 609-623.

Callan, R.J. (1995). Hotel classification and grading schemes, a paradigm of utilisation and user characteristics. *International Journal of Hospitality Management*, 14, 271-284.

Carter, P.L. (2016), "Where are the enslaved?: tripadvisor and the narrative landscapes of Southern plantation museums", Journal of Heritage Tourism, Vol. 11 No. 3, pp. 235-249.

Chan, E. S., & Wong, S. C. (2006). Hotel selection: When price is not the issue. Journal of Vacation Marketing, 12(2), 142-159.

Chen, M., Mao, S. and Liu, Y. (2014), "Big data: a survey", Mobile Networks and Applications, Vol. 19 No. 2, pp. 171-209.

Choi, T. Y., & Chu, R. (2001). Determinants of hotel guests' satisfaction and repeat patronage in the Hong Kong hotel industry. International Journal of Hospitality Management, 20(3), 277-297.

Chu, R. K., & Choi, T. (2000). An importance-performance analysis of hotel selection factors in the Hong Kong hotel industry: A comparison of business and leisure travelers. Tourism Management, 21(4), 363–377.

Chu, S.-C., Deng, T. and Cheng, H. (2020). The role of social media advertising in hospitality, tourism and travel: a literature review and research agenda, International Journal of Contemporary Hospitality Management, Vol. 32 No. 11, pp. 3419-3438. https://doi.org/10.1108/IJCHM-05-2020-0480

Correia, A., Kozak, M. and Ferradeira, J. (2013), “From tourist motivations to tourist satisfaction. International”, International Journal of Culture, Tourism and Hospitality Research, Vol. 7 No. 4, pp. 411-424.

d’Amore, M., Baggio, R. and Valdani, E. (2015), “A practical approach to big data in tourism: a low-cost raspberry pi cluster”, In Tussyadiah, I. and Inversini, A. (Eds), Information and Communication Technologies in Tourism 2015 (Proceedings of the International Conference in Lugano, Switzerland, February 3-6), Springer, Berlin, pp. 169-181.

Del Vecchio, P., Mele, G., Ndou, V., & Secundo, G. (2018). Creating value from social big data: Implications for smart tourism destinations. *Information Processing & Management*, *54*(5), 847-860.

Dolnicar, S. and Ring, A. (2014), “Tourism marketing research: past, present and future”, Annals of Tourism Research, Vol. 47, pp. 31-47.

Dolnicar, S., & Otter, T. (2003). Which hotel attributes matter? A review of previous and a framework for future research.

Expedia (2019) *Fast Facts.* Available at: https://newsroom.expedia.com/fast-facts(Accessed: 17/01/19).

Filieri, R., Acikgoz, F., Ndou, V. and Dwivedi, Y. (2021). Is TripAdvisor still relevant? The influence of review credibility, review usefulness, and ease of use on consumers’ continuance intention. *International Journal of Contemporary Hospitality Management*, Vol. 33 No. 1, pp. 199-223. https://doi.org/10.1108/IJCHM-05-2020-0402

Fischer, A., & Manstead, A. (2000). The relation between gender and emotion in different cultures. In A. Fischer (Ed.), Gender and emotion: Social psycholgical perspectives (pp. 71-94). Cambridge: Cambridge University Press.

García-Palomares, J.C., Gutiérrez, J. and Mínguez, C. (2015), “Identification of tourist hot spots based on social networks: a comparative analysis of European metropolises using photo-sharing services and GIS”, Applied Geography, Vol. 63, pp. 408-417.

Gelbrich, K., & Roschk, H. (2011). A meta-analysis of organizational complaint handling and customer responses. *Journal of Service Research*, *14*(1), 24-43.

Gerard, G., Osinga, E.C., Lavie, D. and Scott, B.A. (2016), “Big data and data science methods for management research”, Academy of Management Journal, Vol. 59 No. 5, pp. 1493-1507.

Gretzel, U., Sigala, M., Xiang, Z. and Koo, C. (2015), “Smart tourism: foundations and developments”, Electronic Markets, Vol. 25 No. 3, pp. 179-188.

Gour, A., Aggarwal, S. and Erdem, M. (2021). Reading between the lines: analyzing online reviews by using a multi-method Web-analytics approach. *International Journal of Contemporary Hospitality Management*, 33 (2), pp. 490-512.

Grönroos, C. (2020). Service Marketing Research Priorities. *Journal of Services Marketing. 34*(3), 291–298

Gundersen, M. G., Heide, M., & Olsson, U. H. (1996). Hotel guest satisfaction among business travelers: what are the important factors?. *Cornell hotel and restaurant administration quarterly*, *37*(2), 72-81.

Gwinner, K.P., Bitner, M.J., Brown, S.W., & Kumar, A. (2005). Service customization through employee adaptiveness. Journal of Service Research, 8(2), 131 –148.

Hao, J. S. C., & Har, C. O. S. (2014). A study of preferences of business female travelers on the selection of accommodation. *Procedia-Social and Behavioral Sciences*, *144*, 176-186.

Heung, V. C. (2000). Satisfaction levels of mainland Chinese travelers with Hong Kong hotel services. *International Journal of Contemporary Hospitality Management*, *12*(5), 308-315.

Hofstede, G. (1983). National cultures in four dimensions: A research-based theory of cultural differences among nations. International Studies of Management & Organization, 13(1-2), 46-74.

Jackson, S. (2016), “Prediction, explanation and big(ger) data: a middle way to measuring and modelling the perceived success of a volunteer tourism sustainability campaign based on ‘nudging”, Current Issues in Tourism, Vol. 19 No. 7, pp. 643-658.

Jansz, J. (2000). Masculine identity and restrictive emotionality. In A. Fischer (Ed.), Gender and emotion: Social psycholgical perspectives (pp. 166-186). Cambridge: Cambridge University Press.

Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, *74*(7), 2561-2573.

Kandampully, J., & Suhartanto, D. (2003). The role of customer satisfaction and image in gaining customer loyalty in the hotel industry. Journal of Hospitality & Leisure Marketing, 10(1-2), 3-25.

Khan, F. H., Qamar, U., & Bashir, S. (2016). Senti-CS: Building a lexical resource for sentiment analysis using subjective feature selection and normalized Chi-Square-based feature weight generation. *Expert Systems*, *33*(5), 489-500.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. Business Horizons, 54, 241-251.

Kim, D., & Park, B. J. R. (2017). The moderating role of context in the effects of choice attributes on hotel choice: A discrete choice experiment. Tourism Management, 63, 439–451.

Kim, D., & Perdue, R. R. (2013). The effects of cognitive, affective, and sensory attributes on hotel choice. International Journal of Hospitality Management, 35, 246–257.

Kim, J., Kim, P. B., & Kim, J. E. (2018). Impacts of temporal and gender difference on hotel selection process. *Journal of Hospitality Marketing & Management*, 1-22.

Kotler, P., Bowen, J. T., Makens, J. C., & Baloglu, S. (2006). Marketing for hospitality and tourism.

Li, G., Law, R., Vu, H. Q., Rong, J., & Zhao, X. R. (2015). Identifying emerging hotel preferences using emerging pattern mining technique. Tourism management, 46, 311-321.

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.

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301-323.

Li, Z., Fan, Y., Jiang, B., Lei, T., & Liu, W. (2019). A survey on sentiment analysis and opinion mining for social multimedia. *Multimedia Tools and Applications*, *78*(6), 6939-6967.

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. Tourism Management, 29 (3), 458–468.

Lu, W. and Stepchenkova, S. (2015), “User-generated content as a research mode in tourism and hospitality applications: topics, methods, and software”, Journal of Hospitality Marketing and Management, Vol. 24 No. 2, pp. 119-154.

Mariani, M., Baggio, R., Fuchs, M., & Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, *30*(12), 3514-3554.

Mariani, M. and Borghi, M. (2021). Customers’ evaluation of mechanical artificial intelligence in hospitality services: a study using online reviews analytics. International Journal of

Contemporary Hospitality Management, Vol. 33 No. 11, pp. 3956-3976. https://doi.org/10.1108/IJCHM-06-2020-0622.

Mayer-Schönberger, V. and Cukier, K. (2013), Big Data: A Revolution That Will Transform How We Live, Work, and Think, Houghton Mifflin Harcourt, New York, NY.

Mccartney, G., & Ge, Z. (2016). An examination of the pull and push factors influencing hotel selection by Chinese outbound travelers. International Journal of Tourism Sciences, 16(1–2), 41– 59.

McCleary, K. W., Weaver, P. A., & Hutchinson, J. C. (1993). Hotel selection factors as they relate to business travel situations. Journal of Travel Research, 32(2), 42–48.

McCleary, K. W., Weaver, P. A., & Lan, L. (1994). Gender-based differences in business travelers' lodging preferences. The Cornell Hotel and Restaurant Administration Quarterly, 35(2), 51–58.

Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q. and Vespignani, A. (2013), "The twitter of babel: mapping world languages through microblogging platforms", PLoS ONE, Vol. 8 No. 4, doi: 10.1371/journal.pone.0061981.

Moss, G., Gunn, R., & Heller, J. (2006). Some men like it black, some women like it pink: consumer implications of differences in male and female website design. *Journal of Consumer behaviour*, *5*(4), 328-341.

Nusair, K. (2020), "Developing a comprehensive life cycle framework for social media research in hospitality and tourism: A bibliometric method 2002-2018", International Journal of Contemporary Hospitality Management, Vol. 32 No. 3, pp. 1041-1066. https://doi.org/10.1108/IJCHM-09-2019-0777

Paldino, S., Bojic, I., Sobolevsky, S., Ratti, C. and González, M.C. (2015), "Urban magnetism through the lens of geo-tagged photography", EPJ Data Science, Vol. 4 No. 1, pp. 1-17.

Pantano, E., Priporas, C. V., & Stylos, N. (2017). 'You will like it!'using open data to predict tourists' response to a tourist attraction. *Tourism Management*, *60*, 430-438.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. British Medical Journal, 316(7139), 1236.

Porter, M. F. (1980). An algorithm for suffix stripping. Program, 14(3), 130-137.

Prayag, G., Hosany, S. and Odeh, K. (2013), "The role of tourists' emotional experiences and satisfaction in understanding behavioral intentions", Journal of Destination Marketing and Management, Vol. 2 No. 2, pp. 118-127.

Qu, H., Ryan, B., & Chu, R. (2000). The importance of hotel attributes in contributing to travelers' satisfaction in the Hong Kong hotel industry. *Journal of Quality Assurance in Hospitality & Tourism*, *1*(3), 65-83.

Rahimi, R. & Kozak, M. (2017). Impact of Customer Relationship Management on Customer Satisfaction: The Case of a Budget Hotel Chain. Journal of Travel & Tourism Marketing, 34(1), 40-51.

Rahimi, R. (2008). Feasibility study of application and implimantation of customer relationship management (CRM) in hotel industry.Master Thesis. Lulea University.

Rahimi, R. (2017). Customer relationship management (people, process and technology) and organisational culture in hotels: Which traits matter?. *International Journal of Contemporary Hospitality Management*, *29*(5), 1380-1402.

Rahimi, R., & Gunlu, E. (2016). Implementing Customer Relationship Management (CRM) in hotel industry from organisational culture perspective. International Journal of Contemporary Hospitality Management. 28 (1). 89 – 112.

Rahimi, R., & Kozak, M. (2017). Impact of Customer Relationship Management on Customer Satisfaction: The Case of a Budget Hotel Chain. Journal of Travel & Tourism Marketing, 34(1), 40-51.

Rahimi, R., Köseoglu, M. A., Ersoy, A. B., & Okumus, F. (2017). Customer relationship management research in tourism and hospitality: a state-of-the-art. *Tourism Review*, *72*(2), 209-220.

Rondan-Cataluña, F. J., & Rosa-Diaz, I. M. (2014). Segmenting hotel clients by pricing variables and value for money. Current Issues in Tourism, 17(1), 60–71.

Saleh, F., & Ryan, C. (1992). Client perceptions of hotels: A multi-attribute approach. Tourism Management, 13(2), 163–168.

Sánchez-Franco, M. J., Navarro-García, A., & Rondán-Cataluña, F. J. (2019). A naive Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services. *Journal of Business Research*, *101*, 499-506.

Sigala, M. (2008). Web 2.0, social marketing strategies and distribution channels for city destinations: enhancing the participatory role of travelers and exploiting their collective intelligence. In M. Gascó-Hernández, T. Torres-Coronas (Eds.), Information communication technologies and city marketing: digital opportunities for cities around the world . IDEA Publishing.

Sigala, M., Lockwood, A., & Jones, P. (2001). Strategic implementation and IT: Gaining competitive advantage from the hotel reservations process. International Journal of Contemporary Hospitality Management, 13 (7), 364–371.

Sim, J., Mak, B., & Jones, D. (2006). A model of customer satisfaction and retention for hotels. Journal of Quality Assurance in Hospitality & Tourism, 7(3), 1-23.

Smith, W. W. & Carmicheal, B. A. (2006). Domestic business travel in Canada with a focus on the female market. Journal of travel and tourism marketing, 21 (1), 65 – 76.

Stoppard, J. M., & Gunn Gruchy, C. D. (1993). Gender, context, and expression of positive emotion. Personality and Social Psychology Bulletin, 19(2), 143-150.

Stringam, B. B., & Gerdes, J. (2010). An analysis of wordof-mouse ratings and guest comments of online hotel distribution sites. Journal of Hospitality Marketing & Management, 19(7), 773–796.

Stylos, N., Bellou, V., Andronikidis, A., & Vassiliadis, C. A. (2017). Linking the dots among destination images, place attachment, and revisit intentions: A study among British and Russian tourists. *Tourism Management*, *60*, 15-29.

Supak, S., Brothers, G., Bohnenstiehl, D. and Devine, H. (2015), “Geospatial analytics for federally managed tourism destinations and their demand markets”, Journal of Destination Marketing and Management, Vol. 4 No. 3, pp. 173-186.

Teeroovengadum, V. (2020), “Service quality dimensions as predictors of customer satisfaction and loyalty in the banking industry: moderating effects of gender”, European Business Review.

Thelwall, M. & Stuart, D. (2006). Web crawling ethics revisited: Cost, privacy and denial of service. Journal of the American Society for Information Science and Technology, 57(13), 1771-1779.

Thelwall, M. (2016). Sentiment analysis for small and big data. *The SAGE handbook of online research methods*, 344-355.

Thelwall, M. (2018). Can museums find male or female audiences online with YouTube?. Aslib Journal of Information Management, 70(5), 481-497.

Thelwall, M., & Buckley, K. (2013). Topic-based sentiment analysis for the Social Web: The role of mood and issue-related words. Journal of the American Society for Information Science and Technology, 64(8), 1608–1617.

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social Web, Journal of the American Society for Information Science and Technology, 63(1), 163-173.

Thelwall, M., Buckley, K., Paltoglou, G. Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. Journal of the American Society for Information Science and Technology, 61(12), 2544–2558.

TripAdvisor(2019) *Press Releases*. Available at: https://tripadvisor.mediaroom.com/UK (Accessed: 17/01/19).

Tronvoll, B. (2011). Negative emotions and their effect on customer complaint behaviour. *Journal of Service Management*, *22*(1), 111-134.

Verhoef, P.C., Kooge, E. and Walk, N. (2016), Creating Value with Big Data Analytics: Making Smarter Marketing Decisions, Routledge, London.

Vogel, D. L., R.Wester, S., Heesacker, M., & Madon, S. (2003). Confirming gender stereotypes: A social role perspective. Sex Roles, 48(11-12), 519-528.

Wallach, H. M. (2006). Topic modeling: beyond bag-of-words. In Proceedings of the 23rd international conference on Machine learning (pp. 977-984). New York, NY: ACM Press.

Waters, J. (2018). A model of the dynamics of household vegetarian and vegan rates in the UK. *Appetite*, *127*, 364-372.

Wood, S.A., Guerry, A.D., Silver, J.M. and Lacayo, M. (2013), "Using social media to quantify naturebased tourism and recreation", Scientific Reports, Vol. 3 No. 1, doi: 10.1038/srep02976.

Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction?. International Journal of Hospitality Management, 44, 120-130.

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.

Yavas, U., & Babakus, E. (2005). Dimensions of hotel choice criteria: Congruence between business and leisure travelers. International Journal of Hospitality Management, 24, 359–367.

Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. Computers in Human Behavior, 27(2), 634-639.

Zammuner, V. L. (2000). Men's and women's lay theories of emotion. In A. Fischer (Ed.), Gender and emotion: Social psychological perspectives (pp. 48-70). Cambridge: Cambridge University Press.

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.

Zhou, L., Ye, S., Pearce, P. L., & Wu, M.-Y. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. International Journal of Hospitality Management, 38, 1e10.