Quality 4.0: Big data analytics to explore service quality attributes and their relation to user sentiment in digital home-sharing platforms

Natalia Amat-Lefort[1], Federico Barravecchia[2], Luca Mastrogiacomo[3]

[1]*n.amat.lefort@liacs.leidenuniv.nl* [2]*federico.barravecchia@polito.it* [3]*luca.mastrogiacomo@polito.it*

[1]Universitat Internacional de Catalunya, Barcelona, Spain
[1]Science Based Business (SBB), Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, Netherlands
[2,3] Politecnico di Torino, DIGEP (Department of Management and Production Engineering), Torino, Italy

Abstract

The fourth industrial revolution has presented significant challenges and opportunities for quality management. Quality 4.0 has been introduced as a response to the need to adapt to technology innovations and update traditional quality models. In this context, one of the most important factors for adopting Quality 4.0 is the leveraging of big data to collect insights and quality perceptions from clients. Therefore, user reviews have emerged as a valuable source of information, which can be analysed with machine learning procedures to uncover latent quality dimensions.

This study applies a combination of text mining techniques (topic modelling and sentiment analysis) to analyse Airbnb reviews, identifying service quality attributes and assessing their relation to the users' sentiment. A total of 2,735,437 reviews written by Airbnb guests in four European cities are analysed, aiming to obtain a better understanding of the preferences and service quality perspectives of Airbnb users. First, topic modelling is applied to find the most relevant service attributes mentioned in the reviews. Then, sentiment analysis is used to assess the positiveness/ negativeness of the users' feedback. By analysing the correlation between service attributes and sentiment score, we find that four attributes show a significant positive relation to the guest's sentiment. These are Apartment views, Host tips and advice, Location and Host friendliness. On the other hand, the following attributes are negatively correlated with user sentiment: Sleep disturbance, Website responsiveness, Thermal management and Hygiene issues. The results of this study and the attributes identified can serve as a reference to extend previous quality assessment scales designed for peer-to-peer accommodation services. Additionally, this paper responds to the need to update quality management frameworks by integrating data-driven methodologies to identify quality attributes.

Keywords: Quality 4.0; Service quality; Big data analytics; Text mining; Topic modelling;

Sentiment analysis; Sharing economy

1. Introduction

More than 2 million people stay on Airbnb every night, on average (Airbnb, 2021a). Being one of the most popular sharing economy businesses, Airbnb is a peer-to-peer online platform that allows hosts to offer their spare rooms or properties to guests, as an alternative to hotels. In fact, Airbnb has disrupted the hospitality industry (Guttentag, 2015), due to its impact on the market share of traditional accommodation options such as hotels (Zervas et al., 2017). In such a fluctuating industry, service quality is essential to achieve competitive advantage (Chen & Chen, 2014). Indeed, given that service quality has a strong positive association with customer satisfaction, it is widely acknowledged as a key success factor for businesses (Saha & Theingi, 2009; Shemwell et al., 1998).

In order to manage quality in the era of the fourth industrial revolution, it is necessary to adapt to technology innovations and update traditional quality models (Saihi et al., 2021). As a response to these technological challenges and opportunities, a new paradigm of quality management has emerged, known as Quality 4.0 (Dias et al., 2021; Zonnenshain & Kenett, 2020). Among the most relevant requirements for implementing Quality 4.0 is understanding how to use big data to collect insights and quality perceptions from customers (Antony et al., 2021). Thanks to machine learning methods, user reviews have become a valuable source of information for businesses (Barravecchia et al., 2021). More specifically, several studies have shown that user-generated content such as customer reviews can provide valuable insights about service quality attributes (Barravecchia et al., 2020; Brochado et al., 2019; Chakrabarti et al., 2018; Ding et al., 2020; James et al., 2017; Ju et al., 2019; Korfiatis et al., 2019; Mastrogiacomo et al., 2021; Palese & Usai, 2018; Su & Teng, 2018).

In the case of Airbnb, over 250 million reviews have been written by hosts and guests from all over the world (Airbnb, 2021b), creating an extensive source of information. Analysing the large amounts of unstructured text that can be found in online reviews is far beyond manual processing capacity. On the contrary, text mining offers an efficient way to extract actionable insights from large-scale databases of reviews and ratings (Hu & Liu, 2004). Due to their unstructured nature, textual reviews contain much more detailed insights of each guest’s personal experience than traditional surveys (Tsao et al., 2020). According to Sparks & Browning (2010), listening to customer feedback from user reviews is essential to understand how service quality can be improved in the hospitality industry.

Considering the advantages of user-generated content as a source of information, text mining has become increasingly popular in quality research regarding Airbnb services. Initially, researchers used customer feedback obtained from surveys to explore Airbnb service quality (Ju et al., 2019; Priporas et al., 2017). More recently, Ding et al. (2020) applied topic modelling to extract service quality attributes from 242,020 Airbnb reviews in Malaysia, identifying four new service attributes. Furthermore, other studies based on Airbnb reviews have obtained significant insights about customer experience (M. Cheng & Jin, 2019; Joseph & Varghese, 2019), user preferences (Serrano et al., 2020) or consumer behaviour (Lee et al., 2019). The results of these studies confirm that text mining can be useful to examine Airbnb reviews and contribute to improving previous survey-based assessment instruments.

Existing research on Airbnb reviews has focused mainly on identifying popular review topics, based on the assumption that these topics reflect desirable service attributes (Ju et al., 2019). However, there is very limited literature discussing which attributes have a positive (or negative) relation to the sentiment expressed by the user. Considering that each review contains a narrative of the guest's personal experience (Lee et al., 2019), negative comments are more likely to mention the attributes that did not fulfil their expectations (W. Zhang et al., 2012). On the other hand, positive reviews with a high sentiment score are more prone to mention the attributes that left the user satisfied. Consequently, by analysing the relationship between the attributes mentioned in each review and its sentiment score, it is possible to infer which attributes should be avoided and which ones should be pursued.

The current work aims to shed light on this issue, by identifying positive and negative service attributes in a dataset of 2,735,437 Airbnb reviews. To this end, this study combines two text mining techniques: sentiment analysis and topic modelling. More specifically, a rule-based model known as VADER (Valence Aware Dictionary for Sentiment Reasoning) is used to assign a sentiment score to each review. The score calculated by the model indicates both the polarity (positive or negative) and the intensity (strength) of the emotions in each review. Once the sentiment analysis has been performed, the service attributes mentioned in each review are identified by applying the Structural Topic Model (STM), which is a variant of the Latent Dirichlet Allocation (LDA) technique. This topic modelling method identifies the underlying topics and calculates an estimation of each topic's weight in each review. Finally, by estimating the correlation of each attribute with the sentiment score, we can infer which ones are positive or negative, i.e. (desirable or undesirable service quality attributes).

In terms of the study's geographical scope, reviews from four European countries have been included: France, Italy, Spain and Portugal. According to the GLOBE project (Global Leadership and Organisational Behaviour Effectiveness), these countries belong to the "Latin Europe" cluster. A study by Gupta et al. (2002) validated these clusters. Even though these countries are highly ranked among the most popular tourist destinations in the world (World Tourism Organization, 2020), previous Airbnb research has not analysed them. According to data published by Airbnb (2019), the European countries where Airbnb had the highest direct economic impact were France ($10.8 billion), Spain ($6.9 billion), Italy ($6.4 billion), UK ($5.6 billion), and Portugal ($2.3 billion). Thus, this study focuses on Airbnb in these Latin European countries, with the aim of contributing to generate cross-regional insights of peer-to-peer accommodation services.

Therefore, the main research objectives of this study are (1) to *identify the main service quality attributes mentioned in Airbnb reviews in Latin Europe* and (2) *determine which attributes have a positive (or negative) correlation with the user's sentiment*. The current work contributes to the quality management literature by extracting novel quality insights from Airbnb users' reviews in Europe using text mining techniques. This study also contributes to the literature by proposing a new approach to analyse user-generated content, combining topic modelling and sentiment analysis to detect positive and negative service quality attributes.

The rest of the paper is organized as follows. First, the literature review (Section 2) describes the antecedents of service quality assessment in the shared accommodation industry, as well as the existing text mining research on Airbnb reviews. The methodological approach is explained in Section 3, while Section 4 presents the results obtained on the analysis of Airbnb data. Finally, Section 5 discusses the most relevant conceptual and practical implications of the study, as well as the limitations and opportunities for future research.

2. Literature review

2.1. Service quality assessment in the shared accommodation sector

From the viewpoint of business administration, service quality can be measured as the discrepancy between the customer's expectations of a service and the actual experience of that service (Parasuraman et al., 1988). Since this concept was introduced, numerous studies have found a positive relationship between service quality and customer satisfaction (H. Lee et al., 2000; Saha & Theingi, 2009; Spreng & Mackoy, 1996) and loyalty (Bloemer et al., 1999; de Ruyter et al., 1998; Sivadas & Baker-Prewitt, 2000) across different sectors. Moreover, it has been proven that service quality is also linked to overall profitability of service companies (Chang & Chen, 1998; Zeithaml, 2000). Consequently, service quality is key to increase competitive advantage in a business, due to its positive impact on relevant success factors for companies.

On the other hand, the rise of shared accommodation platforms has generated a lot of research interest, considering its significant impact on the hotel industry (Zervas et al., 2017). Several studies have analysed the motivations of shared accommodation users (Guttentag et al., 2018; So et al., 2018), as well as other aspects such as the determinants of their trust (Hawlitschek et al., 2016; Mittendorf, 2018; Tussyadiah & Park, 2018), loyalty (Lalicic & Weismayer, 2018), repurchase intentions (Liang et al., 2018), and satisfaction (Möhlmann, 2015; Tussyadiah, 2016).

An emerging field of study has focused on studying the key drivers of service quality in the shared accommodation sector. Priporas et al. (2017) investigated the service quality perceptions of Airbnb users in Phuket (Thailand), through a survey based on measurement tool developed by Akbaba (2006). Their findings pointed out that guests value the convenience of shared accommodation services, as well as the hospitality provided by Airbnb hosts.

Following that line of research, Ju et al. (2019) carried out a survey based on Airbnb users in the US and Canada, aiming to identify the dimensions of service quality attributes in this sector. They initially used text mining to extract service attributes, which were later used to design their survey and propose a service quality assessment instrument for shared accommodation services. The study found that Airbnb's main service quality attributes are associated with the website, host, and facility, and that each one of them can have distinct impacts on customer satisfaction.

One common aspect of the previously mentioned studies is that their conclusions were mainly based on a survey that was completed by Airbnb users. Until recently, there were not many ways to access information about consumer opinions, other than questionnaires and interviews. Nevertheless, due to the rise of digital platforms that collect feedback from users, user-generated content has emerged as a new source of information.

2.2. Quality 4.0 and user-generated content

Over the course of history, revolutions have occurred when disruptive technologies and new paradigms initiate a process of profound change in society and the economy (Schwab, 2016). Nowadays, industries worldwide are experiencing a fourth industrial revolution due to the widespread use of digital technologies. This environment presents several challenges and opportunities for quality management, which needs to transform to adapt to the new Quality 4.0 paradigm (Saihi et al., 2021). This new concept of quality emphasizes the importance of incorporating technology innovations and data-driven solutions into quality management (Zonnenshain & Kenett, 2020). According to a recent study by Antony et al. (2021), understanding how to efficiently handle big data in quality management is the most relevant factor for a successful adoption of Quality 4.0, regardless of the size and type of company.

One of the many ways big data can have an impact on organizations is by helping them detect quality determinants. Historically, perceived service quality was assessed through interviews and surveys (DeVellis, 2016). Even though these approaches are widely accepted, they are often costly and time-consuming (Barravecchia et al., 2021). In this context, one of the opportunities generated by the Industry 4.0 is that it has become much easier for consumers to share their opinions about a specific product or service through online review platforms, company websites and social media (Palese & Usai, 2018). This type of user-generated content (ratings and reviews) can be a source of valuable information about consumer satisfaction and perceived service quality (Gu & Ye, 2014; Hargreaves, 2015). Using text mining tools, this content can be analysed to identify quality determinants, noting that critical service features are frequently mentioned in user reviews (Barravecchia et al., 2021). Moreover, as opposed to traditional surveys, user comments and ratings are not affected by laboratory effects (Yacouel & Fleischer, 2012), which makes them an excellent tool to capture the customer's actual thoughts.

2.3. Comparison of text mining methods applied to Airbnb's user-gerated content

A recent stream of literature has focused on exploring Airbnb reviews using text mining approaches, in order to get a better understanding of the users' opinions and behaviour. Tussyadiah & Zach (2017) presented one of the first studies in this field, applying term-frequency and word co-occurrence networks in Airbnb reviews to identify the service attributes sought by guests from Portland (Oregon). The results indicated that the most common attributes in user reviews were associated with the listing's location, host (service and hospitality), and property (facilities and atmosphere).

Afterwards, other studies combined similar word-frequency based techniques with traditional surveys. In particular, Ju et al. (2019) used first text mining to find service quality attributes, followed by a survey to examine the effects of these attributes on customer satisfaction. Ranjbari et al. (2020) also analysed user reviews to extract the service attributes and conducted a survey to calculate an Importance-Performance Analysis (IPA) matrix. On the other hand, Cheng & Jin (2019) combined term-frequency with sentiment analysis, calculating the probability of an Airbnb user experience concept being mentioned in a positive and negative context. They found that "noise" was the main cause of negative sentiments in Sydney.

Even though term-frequency based approaches can be useful to extract latent topics from unstructured textual data, they have certain limitations. When word frequency counts are used to extract meaning from texts, they assume that each word has only one meaning (S. Lee et al., 2010) . However, words can have different meanings depending on how they are combined with other words. Additionally, word frequency methods require the definition of each topic by a list of manually selected keywords, making it difficult to analyse the trend of a particular topic (Ding et al., 2020).

In contrast, other text mining methods such as topic modelling and clustering can analyse groups of words together, instead of counting them separately (M. W. Berry & Castellanos, 2008). In this way, these techniques can consider the variations in each word's meaning across different contexts. In terms of clustering methods, most of the previous studies used techniques based on word co-occurrence networks. For instance, Lee et al. (2019), applied hierarchical clustering to assess the evolution over a 5-year time span of the attributes that influence Airbnb customer experience in London. They found that there were basic attributes which were considered important throughout the years (such as cleanliness and location), but certain attributes only appeared in particular years and seasons, such as "Wi-Fi" or "Quietness".

Although conventional clustering methods overcome some of the limitations of term-frequency based techniques, topic modelling presents several advantages over them. Clustering is based on the idea of classifying documents into different groups based on a specific similarity measure (M. W. Berry & Castellanos, 2008). With clustering, every document (reviews, in our case) appears only in one group (such as topics or attributes). In topic modelling, each topic is defined by a group of words and each document is assigned a probability of being associated to a topic (Jelodar et al., 2019). Hence, when applied to the analysis of customer reviews, topic modelling can capture the fact that the same review could mention several different topics.

One of the most widely used topic models is Latent Dirichlet Allocation (LDA), which has also been used to analyse Airbnb reviews and compare the extracted topics with hotel reviews (J. Zhang, 2019). LDA is used as an unsupervised probabilistic topic modelling technique, which assumes that each document (or review) is composed of a set of topics (Blei, 2012). Research has proven that LDA can effectively detect the main topics in social media text and explore their evolution over time (Wang et al., 2012). Furthermore, LDA does not assume that the documents should have a particular structure or follow any specific grammar rules, which makes it particularly appropriate for big data analysis of online reviews (Feldman & Sanger, 2006)

Despite the fact that LDA has many advantages as a text mining technique, it does not offer the possibility of including document metadata into the topic model. However, an extension of LDA known as the structural topic model (STM) can overcome this limitation (Roberts et al., 2019). As opposed to LDA, STM allows researchers to incorporate arbitrary information about each document (such as the date of the review or the nationality of the reviewer) into the model. A study by Ding et al. (2020) used STM to extract key service quality attributes from a database of Airbnb reviews posted in Malaysia during a five-year period. They identified 22 topics related to service quality and found certain differences between the perspectives of Malaysian and international users. Their results confirmed the effectiveness of STM as a method to identify service quality attributes from Airbnb user reviews.

After reviewing the features and limitations of each text mining technique, the method selected for this study was STM. Additionally, this study combines topic modelling with sentiment analysis to assess the relation of each service attribute to the user's sentiment. Hence, the current study identifies positive and negative service attributes (in terms of their correlation with sentiment score) from Airbnb reviews in France, Italy, Spain and Portugal. This represents a methodological advance, given that very few studies have analysed both the topics and the sentiment of reviews, particularly in Airbnb services.

Moreover, the literature about European Airbnb users is very scarce, as most of the previous studies focused on guests from the US, Australia or Asia (see Table 1). For this reason, this study aims to fulfil the aforementioned gap in the literature. Table 1 presents an overview of the main features and methods applied by previous text mining literature that analysed Airbnb reviews to identify service attributes, also highlighting the main differences among the studies.

					Text mining techniques		
Study	**N. of reviews**	**Time span**	**Geographical scope**	**Survey**	**Word frequency/ co-occurrence**	**Topic modelling**	**Sentiment analysis**
(Tussyadiah & Zach, 2017)	41,560	5 years	US		X		
(Ju et al., 2019)	16,430	3 days	US	X	X		
(Lee et al., 2019)	169,666	5 years	UK		X		
(J. Zhang, 2019)	1,026,988	6 years	US		X	X	
(M. Cheng & Jin, 2019)	181,263	N/A*	Australia		X		X
(Ranjbari et al., 2020)	112,138	1 year	Australia	X	X		
(Ding et al., 2020)	242,020	5 years	Malaysia			X	
This study	2,735,437	10 years	France, Italy, Spain, Portugal			X	X

Table 1. Main features of previous text mining literature that analysed Airbnb reviews to identify service attributes. *N/A: Not available, unspecified.

3. Methodology

In order to identify the service quality attributes of Airbnb and gather the necessary data to estimate their correlation with sentiment score, this study has applied a seven-step methodology: (i) dataset extraction; (ii) data pre-processing; (iii) sentiment analysis; (iv) identification of the optimal number of topics; (v) topic extraction (vi) mapping topics onto service quality attributes; (vii) labelling. Figure 1 shows an overview of the methodology applied.

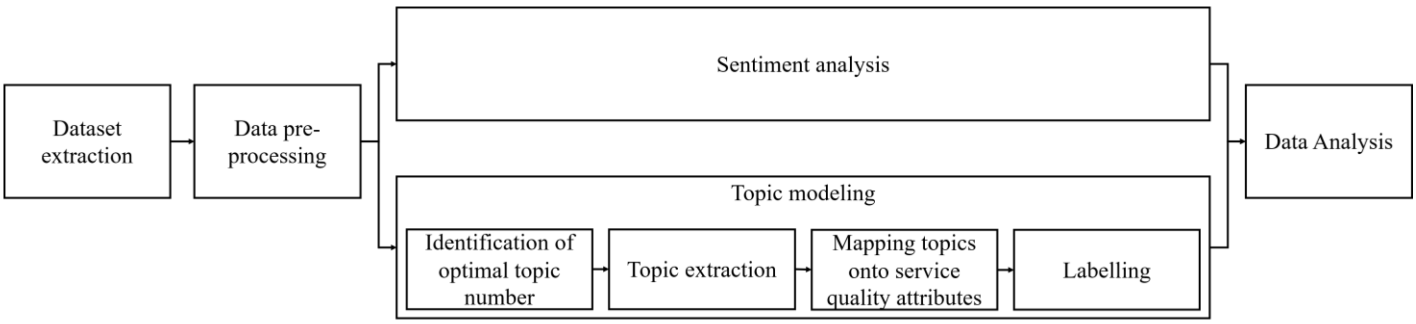

Figure 1. Overview of the methodology.

3.1. Dataset extraction

Data was extracted from the InsideAirbnb website, which is an independent, non-commercial initiative that provides public access to Airbnb data (InsideAirbnb, 2021), including the reviews for each listing.

In total, 2,735,437 reviews were collected from 100,454 listings in Paris, Rome, Barcelona and Lisbon as of July 2020 (review dates ranged from 2010 to 2020). These cities were selected because they are the most popular tourist destinations (Euromonitor, 2019) in the countries defined in the study's geographical scope (France, Italy, Spain and Portugal).

3.2. Data pre-processing

In order to prepare the data for the text mining process, it is necessary to conduct several pre-processing procedures (Feinerer et al., 2008). First, reviews were filtered by language using the texcat package in R (Hornik et al., 2013), in order to exclude non-English reviews. Afterwards, the following processes were carried out, following the approaches recommended in previous studies (Barravecchia et al., 2020; Ding et al., 2020; Mastrogiacomo et al., 2021): (1) Text normalization (conversion of all text into lowercase letters to reduce ambiguity), (2) Text tokenization (separating each sentence into words), (3) Text stemming to remove word inflections or derivations, leaving only the word's stem (e.g. The words "visited" and "visiting" are reduced to their stem "visit"), (4) Removal of the punctuation, numbers and common English stop words (e.g. "and", "the"), (5) Removal of the words shorter than 2 or longer than 15 characters, as well as the reviews with less than 10 words; (6) Removal of words with an extremely low frequency (appearing in less than 20 documents) which would not be representative for the analysis, (7) Removal of the words generally not significant to identify topical content (such as: "another", "review", "made", "did", "done", etc.) and (8) N-gram analysis to find frequently co-occurring words and transform them into unigrams (e.g., "customer service" to "customerservice").

3.3. Sentiment analysis

Sentiment analysis is defined as "the computational study of people's opinions, attitudes and emotions toward an entity" (Medhat et al., 2014, p. 1093). In the case of customer reviews, sentiment analysis can be used to classify, measure and monitor the user's emotions towards a particular product or service.

There is a wide variety of tools that can be used to perform sentiment analysis, each one of them presenting particular characteristics. In order to compare these tools, Ribeiro et al. (2016) developed a benchmark named SentiBench. One of the tools that appeared as a highly ranked sentiment analysis model was VADER (Valence Aware Dictionary for sEntiment Reasoning).

For the current study, VADER was applied to calculate the sentiment score of each review on a scale of -1 (most negative) to +1 (most positive). VADER is a rule-based model for sentiment analysis, which has been specifically adapted to provide a highly accurate sentiment assessment in social media text. In fact, it has been proven that VADER can even outperform individual human raters (F1 Classification Accuracy = 0.96, as opposed to 0.84 from humans) when analysing social media content (Hutto & Gilbert, 2014).

Another sentiment analysis tool known as LIWC (Linguistic Inquiry and Word Count) is also among the few that have been validated by humans (Pennebaker et al., n.d.), but VADER presents certain advantages over it. In particular, VADER is more accurate in social media contexts because it takes into account the characteristics of the text that can affect the perceived sentiment intensity (Hutto & Gilbert, 2014). For instance, it can capture the sentiment expressed by emoticons, punctuation (such as exclamation points), capitalization (i.e., ALL-CAPS to emphasize sentiment), acronyms, slang, etc.

Considering these advantages, VADER was selected to conduct the sentiment analysis, assigning a score between -1 and +1 to each review.

3.4. Identification of the optimal number of topics and topic extraction

In the context of natural language processing, topic models are probabilistic models used to discover latent topics in a collection of documents (Lafferty & Blei, 2009). Aside from identifying the underlying topics in a text corpus, topic modelling associates a set of keywords to each topic and estimates the distribution of topic weights for each document (Blei, 2012).

For the current study, the Structural Topic Model (STM) was used to extract the latent topics from a dataset of Airbnb customer reviews. STM is an extended version of previous widely accepted probabilistic topic models such as Latent Dirichlet Allocation (LDA) (Blei, 2012). The main difference between STM and previous models is that it can incorporate document-level covariate information into the topic model. Including these covariates can improve the model's inference and interpretability, and they can have an impact on the topics' content and/or prevalence (Roberts et al., 2019).

The optimal number of topics (T) is an essential parameter for topic modelling techniques. One of the methods that can be used to identify T is by applying the topic modelling algorithm iteratively (with different numbers of topics) to find the iteration with the best performance (Wallach et al., 2009). In order to evaluate the performance of each iteration, several metrics can be analysed, such as the semantic coherence, held-out likelihood or exclusivity (Roberts et al., 2019). The current study used held-out likelihood, which refers to probability of unseen held-out documents given a trained model (Wallach et al., 2009). On average, better models will be associated to a higher held-out likelihood.

Figure 2 presents the values obtained when calculating the held-out likelihood as a function of T (from 7 to 100). Taking into account the held-out likelihood calculation, the current study selected the number of topics that resulted in the best overall performance. This approach is similar to the one applied in previous topic modelling studies (Ding et al., 2020; Korfiatis et al., 2019). As it can be observed in Figure 2, the highest held-out likelihood was obtained when the value of T was 65. Therefore, 65 was identified as the optimal number of topics.

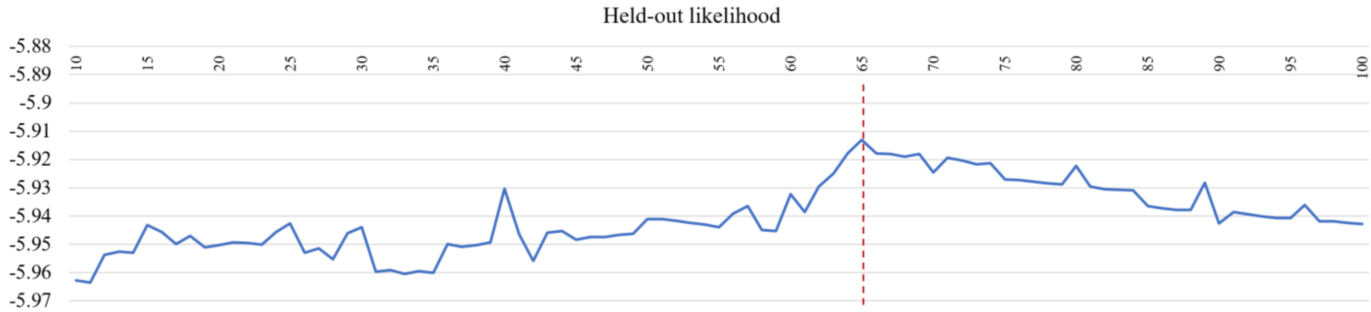

Figure 2. Held-out likelihood for varying numbers of topics.

However, not all of the topics identified in the corpus corresponded to service quality attributes, which is why it was necessary to define a criterion to map these topics to service quality attributes. The next section describes how the mapping was conducted.

3.5. Mapping topics onto service quality attributes

As a result of applying STM to the pre-processed text corpus, the topics were successfully extracted. In order to identify the topics related to service quality attributes, the results were compared with the Airbnb service quality dimensions defined by Ju et al. (2019). Their scale was selected as a reference because it was the first survey-based approach specifically designed to assess service quality in Airbnb. Table 2 shows the mapping of topics onto their corresponding service quality attributes.

The topics that could not be assigned to these categories were not considered as service quality attributes. More specifically, the following topics could not be matched with any service quality dimensions: Price, Suggestions to other users, References to other reviews, Length of stay. Furthermore, other topics only reflected the user's overall satisfaction with the service but did not specify any particular service attribute (e.g., "Great stay! We had a good time, will come back soon"). These topics were also not retained for the analysis. Similarly, topics that referred to automated or null reviews were not taken into account. Finally, several topics related to reviews that were written in more than one language (aside from English) were also discarded. As a result, out of the 65 topics identified in the corpus, 37 were mapped onto service quality attributes.

3.6. Labelling

Through topic modelling, each review was assigned a set of probabilities (topic weights) of being associated to each one of the topics identified in the corpus, with values between 0 and 1. Therefore, topic modelling can capture the fact that the same review could mention several different topics, and it can also reflect the relevance of each topic in a particular review. A high topic weight of "Topic X" reflects a greater probability that a particular review is related to "Topic X". In a hypothetical example, the following review (with keywords marked in bold) obtains a topic weight of 55% for the "Proximity to restaurants” attribute, 17% for “Cleanliness”, 28% for “Host friendliness” and 0% for “Proximity to public transports” (assuming these were the only topics in the model):

> *“The apartment was absolutely beautiful and very **clean**! It was close to lots of **restaurants**, so there were many options to enjoy delicious **local cuisine** nearby. There was a nice pizza **restaurant** across the street. Our **host** was very **friendly** and **helpful** before and during our stay. We highly recommend this apartment!”*

However, topic labels are not provided by the STM. The process of assigning appropriate labels to each topic derived from the model is known as topic labelling. In order to provide a label for each attribute identified through the topic modelling process, the reviews with the highest topic weight for each attribute were used as a reference. Additionally, the most relevant keywords associated to each attribute were also used as a reference for the labelling process. Table 2 shows the top 5 keywords for each attribute, which are the words with the highest probability within each topic. Highest probability words are very useful to select an appropriate label for each topic (Roberts et al., 2019).

In the case of attributes identified in previous Airbnb studies, the label source indicates the study from which the name of the attribute was extracted. Six of the attributes were labelled manually by the authors through group discussion and evaluation, given that they could not be found in previous literature. Lastly, the reviews with the highest weight for each topic were analysed to confirm the adequacy of their respective labels.

Figure 3 illustrates (with examples) the process of extracting topics, labelling them and mapping them onto service quality dimensions, in order to identify service quality attributes.

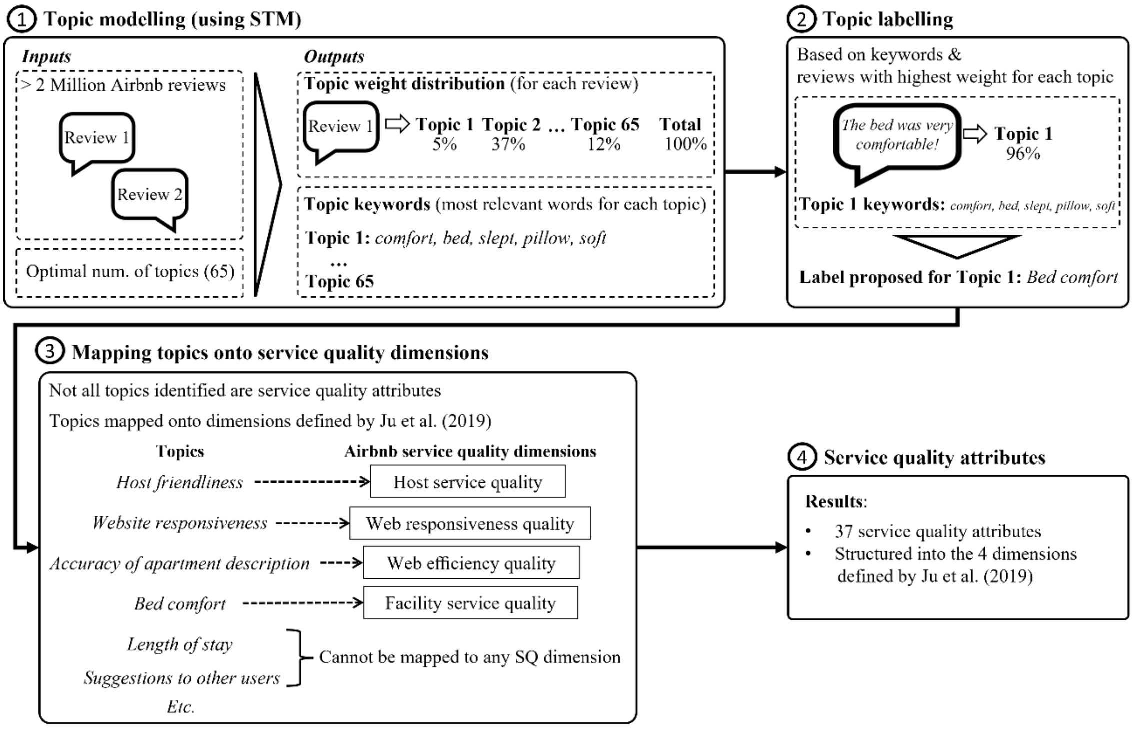

Figure 3. Process to identify service quality attributes from topic modelling results.

4. Data analysis and results

4.1. Airbnb service quality attributes

As a result of mapping the topic modelling outcome onto the Airbnb service quality dimensions specified by Ju et al. (2019), several quality attributes were identified. Table 2 shows the service quality attributes found in the dataset of Airbnb reviews that was analysed. Each one of these attributes was related to one of the following dimensions: Host service quality, Web responsiveness quality, Web efficiency quality and Facility service quality.

Host service quality refers to the quality of the interaction between the host and the guest. Within this dimension, several topics were identified as service attributes, including Host friendliness, Host tips and advice, Host responsiveness and Check-in/ Check-out flexibility.

Web responsiveness quality indicates the platform's ability to respond to the user's inquiries, as well as to compensate guests when there is a problem. One of the topics (which was associated to keywords such as "person", "guest", "ask", "detail", "care" and "contact) was matched with this dimension. Therefore, the service attribute linked to this topic was labelled "Website responsiveness".

The dimension of Web efficiency measures the quality of the website's information and its organization. Two of the topics identified were mapped onto this dimension, considering that they were related to the accuracy of the information found on the website (such as the apartment's description and pictures).

Finally, Facility service quality refers to the quality of the place where the service is provided: the apartment, the room, the neighbourhood, etc. This dimension was the one that was matched with the highest number of topics, which were divided in four categories: apartment features (i.e. Cleanliness, Bed comfort), Equipment (i.e. Apartment furniture, Internet connection), Location (i.e. Proximity to restaurants, Distance from city centre) and Other attributes (Family friendliness).

Dimensions (Ju et al., 2019)	**Id.**	**Service attributes identified**	**Topic label source**	**Relevant keywords (stemmed for text pre-processing purposes)**
Host service quality	1	Host tips and advice	Adapted from Ju et al. (2019)	visit, place, see, tip, short
	2	Host welcome	Adapted from Tussyadiah & Zach (2017)	welcom, inform, warm, met, gave
	3	Host friendliness	Adapted from Ranjbari et al. (2020)	help, friend, kind, pleasant, thought
	4	Host kindness	Adapted from Ju et al. (2019)	host, autom, accommod, offer, wonder
	5	Feeling at home	Adapted from J. Zhang (2019)	home, made, make, feel, trip
	6	Communication with host	Adapted from Sutherland & Kiatkawsin (2020)	easi, communic, get, check, checkin
	7	Host responsiveness	J. Zhang (2019)	quick, respons, alway, question, respond
	8	Late check-in	Adapted from J. Zhang (2019)	late, wait, reserv, flight, delay
	9	Check-in/check-out flexibility	Ranjbari et al. (2020)	check, even, earli, leav, accommod
Web responsiveness quality	10	Website responsiveness	Parasuraman et al. (1988)	person, guest, ask, detail, care
Web efficiency quality	11	Accuracy of apartment description	Adapted from Ranjbari et al. (2020)	expect, photo, descript, accur, smaller
	12	Accuracy of apartment pictures	Labeled by authors	just, look, pictur, exact, describ
Facility service quality		**Apartment features**		
	13	Apartment safety	Ranjbari et al. (2020)	build, door, safe, felt, outsid
	14	Room aesthetics	Cheng et al. (2019)	room, bathroom, bedroom, kitchen, live
	15	Cleanliness	Ranjbari et al. (2020)	clean, good, flat, hous, near
	16	Bed comfort	Adapted from J. Zhang (2019)	comfort, bed, slept, pillow, soft
	17	Apartment dimensions	Labeled by authors	small, size, unit, shower, cloth
	18	Sleeping capacity	Sutherland & Kiatkawsin (2020)	space, access, plenti, spot, coupl
	19	Views	Ranjbari et al. (2020)	beauti, amaz, stay, view, absolut
	20	Relax/ unwind areas in apartment	Adapted from Sutherland & Kiatkawsin (2020)	day, long, relax, sightse, end

	21	Ease of access (stairs)	Labeled by authors	floor, stair, top, elev, luggag
	22	Food and drinks	Adapted from J. Zhang (2019)	breakfast, food, coffe, even, wine
	23	Thermal management (water/ heating issues)	Sutherland & Kiatkawsin (2020)	work, problem, shower, water, hot
	24	Hygiene issues (uncleanliness)	Labeled by authors	toilet, bad, smell, unfortun, seem
		Equipment		
	25	Room equipment	Adapted from J. Zhang (2019)	provid, use, kitchen, includ, cook
	26	Internet connection	Adapted from Lee et al. (2019)	wifi, connect, internet, watch, view
	27	Apartment furniture	Ranjbari et al. (2020)	well, quiet, spacious, equip, decor
		Location characteristics		
	28	Location (overall)	Tussyadiah & Zach (2017)	locat, apart, stay, recommend, love
	29	Neighbourhood features	Lawani et al. (2019)	neighborhood, huge, plus, far, favorit
	30	Accessibility	Sutherland & Kiatkawsin (2020)	also, apart, around, near, easy
	31	Proximity to attractions	Ranjbari et al. (2020)	street, away, step, market, block
	32	Distance from city centre	Labeled by authors	citi, can, center, centr, park
	33	Proximity to public transports	Adapted from Ranjbari et al. (2020)	close, metro, station, conveni, nearbi
	34	Proximity to restaurants	Adapted from J. Zhang (2019)	restaur, area, lot, shop, bar
	35	Reachability from the airport	Labeled by authors	airport, bus, train, take, stop
	36	Sleep disturbance (night noise)	Sutherland & Kiatkawsin (2020)	night, bit, window, sleep, nois
		Other attributes		
	37	Family friendliness	Sutherland & Kiatkawsin (2020)	famili, kid, adult, children, apart

Table 2. Mapping and labeling of identified topics onto service quality attributes

4.2. Correlation of service attributes with user sentiment

Identifying Airbnb's service quality attributes was the first research objective of the current study, together with defining which attributes have a positive (or negative) correlation with the user's sentiment.

In order to estimate the relation of each quality attribute to user sentiment, the Pearson's correlations between each review's sentiment score and the weights of each topic for that particular review were calculated. This was possible because STM provided a distribution of topic weights for each review, and the sentiment analysis assigned a sentiment score to each review. With the purpose of determining which correlations were statistically significant, their p-values and the levels of significance specified in the Table of critical values were considered (Weathington et al., 2012).

Out of the 37 service attributes detected through topic modelling, eight of them presented a statistically significant correlation with the sentiment score. Their returned p-values were < 0.001, confirming their significance. As it can be observed in the correlation matrix (Figure 4), half of these attributes were positively correlated with the sentiment score (Views, Host tips and advice, Location, Host friendliness). However, the other attributes showed a negative correlation with the sentiment (Sleep disturbance, Website responsiveness, Thermal management, Hygiene issues). The attribute with the greatest negative correlation with sentiment score was Hygiene issues ($r = -0.58$), followed by Thermal management (water and heating issues, $r = -0.29$). On the contrary, Views was the attribute with the most significant positive correlation with sentiment score ($r = 0.27$), followed by Host tips and advice ($r = 0.25$).

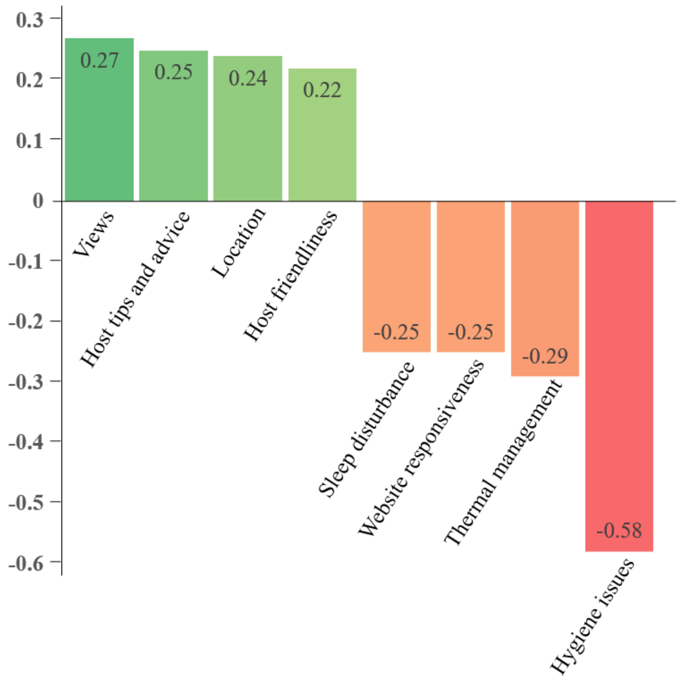

Figure 4. Pearson's correlations between service attributes and sentiment score.

Figure 5 shows (graphically, in the form of scatter plots) the relationship between these service attributes and the sentiment score.

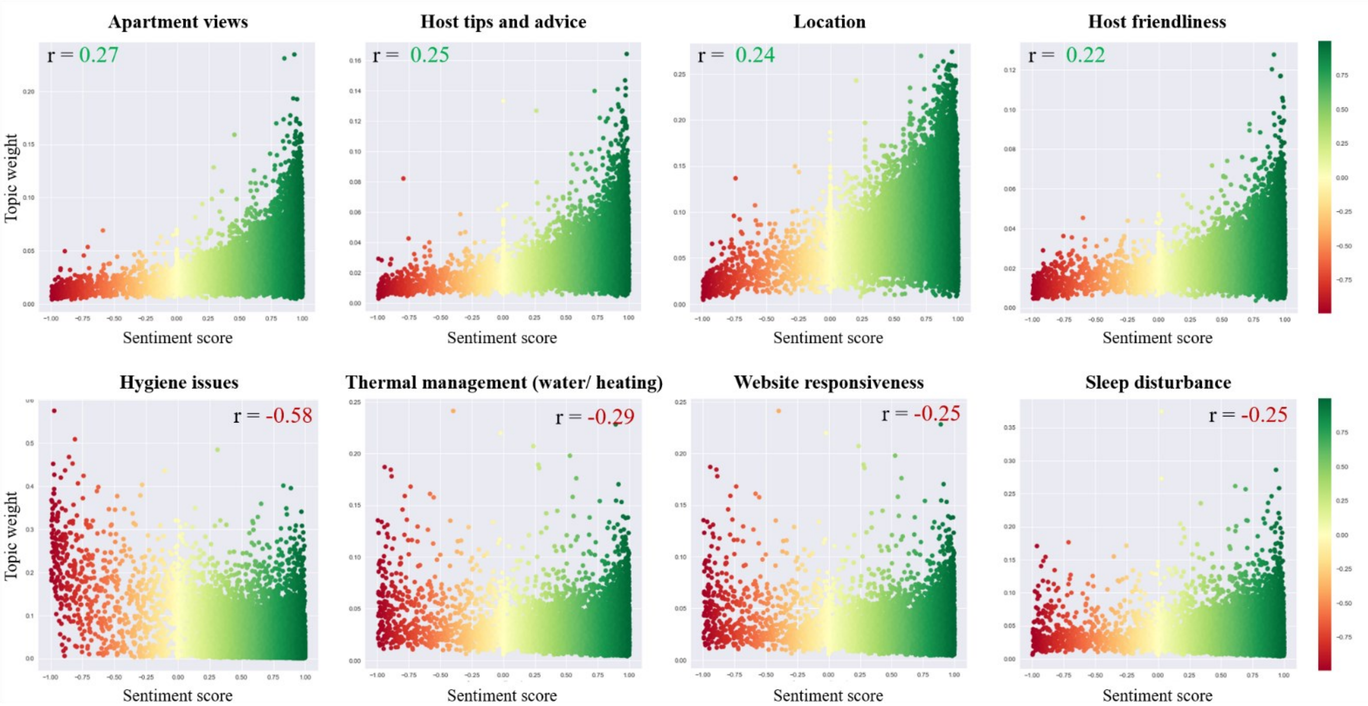

Figure 5. Scatter plots showing the relationship between service attributes and sentiment score. First row: positive service attributes; second row: negative service attributes.

The vertical axis in the scatter plots represents the topic weight (which reflects the importance of an attribute in a particular review), whereas the horizontal axis represents the sentiment score assigned to each review (ranging from -1 to 1). As it can be observed in the first row of Figure 5, in the four topics identified as positive service attributes, higher topic weights are related to higher sentiment scores. On the contrary, in the case of negative service attributes (second row of Figure 5), higher topic weights are associated to lower sentiment scores. However, the trend in negative attributes is not as strong as in the positive ones, given that the dataset presented a positive emotion bias, as it is often the case in user-generated content (2019).

5. Conclusions and discussion

5.1. Key findings

This study used a data-driven approach (combining topic modelling and sentiment analysis) to identify Airbnb service quality attributes, based on reviews written by users in four European countries. First, topic modelling (STM) was used to detect the main topics mentioned in the reviews. In order to determine which topics were related to service quality attributes, they were mapped onto the service quality dimensions proposed by Ju et al. (2019). As a result, 37 attributes related to service quality were identified (see Table 2), 6 of which had not been detected in previous Airbnb studies. In total, 9 attributes were linked to the interaction with the host, 3 were about the interaction with the online platform and 25 were associated to the facilities (apartment features, apartment equipment and location).

Additionally, sentiment analysis was used to find which service attributes were positively or negatively correlated with the user's sentiment. The following attributes were found to have a significant positive relationship with the guest's sentiment: Apartment views, Host tips and advice, Location, Host friendliness. On the other hand, these attributes were shown to be negatively related to user sentiment: Sleep disturbance, Website responsiveness, Thermal management, Hygiene issues. These insights provide a better understanding of the needs and quality perceptions of digital home-sharing platform users. Furthermore, the results of this study and the methodology proposed can be used as a reference to extend existing quality assessment frameworks.

5.2. Theoretical and practical implications

From a theoretical point of view, this study contributes to the Quality 4.0 literature, providing an example of how quality management research can obtain insights from customers by using Big data analytics. According to Saihi et al. (2021), most of the existing Quality 4.0 literature focuses on describing its potential benefits but requires further detail in terms of real applications through practical examples and case studies. On the other hand, Saihi et al. (2021) noted that the majority of studies emphasized the potential advantages of using advanced technology for quality control, while very few focused on other quality management tools and methodologies. Therefore, the current study aims to contribute to rethinking how quality research can adapt to the digital era and leverage technology to obtain customer insights. Furthermore, aside from validating the Airbnb service quality dimensions proposed by Ju et al. (2019), the results of this study can also be used to extend existing Airbnb service quality questionnaires. More specifically, 4 of the attributes that were identified within the Host service quality dimensions had not been included in the scale by Ju et al. (2019), such as "Check-in/check-out flexibility". Regarding the Facility service quality dimension, 19 new attributes were detected, such as "Internet connection" or "Room equipment".

From a practical point of view, this study verifies the adequacy of STM as a technique to obtain insights from user-generated content. It also confirmed that STM is a technique that can be used to perform service quality research, which has also been applied successfully in previous Airbnb studies (Ding et al., 2020). Additionally, the current work presents a data-driven approach to analyse user-generated content, by combining STM and sentiment analysis to detect positive and negative service quality attributes. Finally, this research also contributes to the literature of service quality in peer-to-peer accommodation services by identifying the quality attributes perceived by European users of Airbnb. The results provide valuable insights regarding the preferences of Airbnb users in Latin Europe, given that previous studies had focused on users from the US, UK, Australia and Southeast Asia. Therefore, the results obtained in the current work can serve as a reference for future comparative studies among Airbnb users at different tourist destinations.

5.3. Limitations and future research

This study comes with certain limitations, which at the same time present several opportunities for future research. First, the focus of this article was Airbnb, which is the most relevant platform in the home-sharing sector, but there are many other platforms that could also be examined in different industries. Second, although the dataset included reviews from users in different countries, this study did not assess the potential differences among guest preferences regarding service attributes in these countries. Third, due to space limitations, the current study did not analyse the evolution in time of the service attributes that were identified.

References

Airbnb. (2021a). *Airbnb Trust & Safety - Your safety is our priority*. https://www.airbnb.com/trust

Airbnb. (2021b). *New Data: The Airbnb Advantage*. https://news.airbnb.com/new-data-the-airbnb-advantage/

Airbnb Estimated Direct Economic Impact Exceeds $100 Billion in One Year. (n.d.). Retrieved February 6, 2021, from https://news.airbnb.com/airbnb-estimated-direct-economic-impact-exceeds-100-billion-in-one-year/

Akbaba, A. (2006). Measuring service quality in the hotel industry: A study in a business hotel in Turkey. *International Journal of Hospitality Management*, *25*(2), 170–192. https://doi.org/10.1016/j.ijhm.2005.08.006

Antony, J., Sony, M., McDermott, O., Jayaraman, R., & Flynn, D. (2021). An exploration of organizational readiness factors for Quality 4.0: an intercontinental study and future research directions. *International Journal of Quality & Reliability Management*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/IJQRM-10-2021-0357

Barravecchia, F., Mastrogiacomo, L., & Franceschini, F. (2020). Categorizing Quality Determinants in Mining User-Generated Contents. *Sustainability*, *12*(23), 9944. https://doi.org/10.3390/su12239944

Barravecchia, F., Mastrogiacomo, L., & Franceschini, F. (2021). Digital voice-of-customer processing by topic modelling algorithms: insights to validate empirical results. *International Journal of Quality and Reliability Management*. https://doi.org/10.1108/IJQRM-07-2021-0217/FULL/XML

Berry, L. L., Parasuraman, A., & Zeithaml, V. A. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*(1), 12–40.

Berry, M. W., & Castellanos, M. (2008). Survey of text mining II: Clustering, classification, and retrieval. *Survey of Text Mining II: Clustering, Classification, and Retrieval*, 1–240. https://doi.org/10.1007/978-1-84800-046-9

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Bloemer, J., de Ruyter, K., & Wetzels, M. (1999). Linking perceived service quality and service loyalty: a multi-dimensional perspective. *European Journal of Marketing*, *33*(11/12), 1082–1106. https://doi.org/10.1108/03090569910292285

Brochado, A., Rita, P., Oliveira, C., & Oliveira, F. (2019). Airline passengers' perceptions of service quality: themes in online reviews. *International Journal of Contemporary Hospitality Management*, *31*(2), 855–873. https://doi.org/10.1108/IJCHM-09-2017-0572

Chakrabarti, S., Trehan, D., & Makhija, M. (2018). Assessment of service quality using text mining – evidence from private sector banks in India. *International Journal of Bank Marketing*, *36*(4), 594–615. https://doi.org/10.1108/IJBM-04-2017-0070

Chang, T. Z., & Chen, S. J. (1998). Market orientation, service quality and business profitability: A conceptual model and empirical evidence. *Journal of Services Marketing*, *12*(4), 246–264. https://doi.org/10.1108/08876049810226937

Chen, W. J., & Chen, M. L. (2014). Factors Affecting the Hotel's Service Quality: Relationship Marketing and Corporate Image. *Journal of Hospitality Marketing and Management*, *23*(1), 77–96. https://doi.org/10.1080/19368623.2013.766581

Cheng, M., & Jin, X. (2019). What do Airbnb users care about? An analysis of online review comments. *International Journal of Hospitality Management*, *76*, 58–70. https://doi.org/10.1016/j.ijhm.2018.04.004

Cheng, X., Fu, S., Sun, J., Bilgihan, A., & Okumus, F. (2019). An investigation on online reviews in sharing economy driven hospitality platforms: A viewpoint of trust. *Tourism Management*, *71*, 366–377. https://doi.org/10.1016/j.tourman.2018.10.020

de Ruyter, K., Wetzels, M., & Bloemer, J. (1998). On the relationship between perceived service quality, service loyalty and switching costs. *International Journal of Service Industry Management*, *9*(5), 436–453. https://doi.org/10.1108/09564239810238848

DeVellis, R. F. (2016). Scale Development Theory and Applications (Fourth Edition). *SAGE Publication*.

Dias, A. M., Carvalho, A. M., & Sampaio, P. (2021). Quality 4.0: literature review analysis, definition and impacts of the digital transformation process on quality. *International Journal of Quality and Reliability Management*. https://doi.org/10.1108/IJQRM-07-2021-0247/FULL/XML

Ding, K., Choo, W. C., Ng, K. Y., & Ng, S. I. (2020). Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation. *International Journal of Hospitality Management*, *91*(August), 102676. https://doi.org/10.1016/j.ijhm.2020.102676

Euromonitor. (2019). *Top 100 City Destinations: 2019 Edition | Euromonitor*. https://go.euromonitor.com/white-paper-travel-2019-100-cities.html

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

Feldman, R., & Sanger, J. (2006). The Text Mining Handbook. In *The Text Mining Handbook*. Cambridge University Press. https://doi.org/10.1017/cbo9780511546914

Gu, B., & Ye, Q. (2014). First Step in Social Media: Measuring the Influence of Online Management Responses on Customer Satisfaction. *Production and Operations Management*, *23*(4), 570–582. https://doi.org/10.1111/poms.12043

Gupta, V., Hanges, P. J., & Dorfman, P. (2002). Cultural clusters: Methodology and findings. *Journal of World Business*, *37*(1), 11–15. https://doi.org/10.1016/S1090-9516(01)00070-0

Guttentag, D. (2015). Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*, *18*(12), 1192–1217. https://doi.org/10.1080/13683500.2013.827159

Guttentag, D., Smith, S., Potwarka, L., & Havitz, M. (2018). Why Tourists Choose Airbnb: A Motivation-Based Segmentation Study. *Journal of Travel Research*, *57*(3), 342–359. https://doi.org/10.1177/0047287517696980

Hargreaves, C. A. (2015). Analysis of Hotel Guest Satisfaction Ratings and Reviews: An Application in Singapore. In *American Journal of Marketing Research* (Vol. 1, Issue 4). http://www.aiscience.org/journal/ajmrhttp://creativecommons.org/licenses/by-nc/4.0/

Hawlitschek, F., Teubner, T., & Weinhardt, C. (2016). Trust in the Sharing Economy. *Die Unternehmung*, *70*(1), 26–44. https://doi.org/10.5771/0042-059x-2016-1-26

Hornik, K., Mair, P., Rauch, J., Geiger, W., Buchta, C., & Feinerer, I. (2013). The textcat package for n-gram based text categorization in R. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v052.i06

Hu, M., & Liu, B. (2004). Mining opinion features in customer reviews. *Proceedings of the National Conference on Artificial Intelligence*, 755–760.

Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 8, Issue 1). http://sentic.net/

InsideAirbnb. (2021). *Inside Airbnb. Adding Data to the Debate.* http://insideairbnb.com/

James, T. L., Villacis Calderon, E. D., & Cook, D. F. (2017). Exploring patient perceptions of healthcare service quality through analysis of unstructured feedback. *Expert Systems with Applications*, *71*, 479–492. https://doi.org/10.1016/j.eswa.2016.11.004

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, *78*(11), 15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Joseph, G., & Varghese, V. (2019). Analyzing Airbnb customer experience feedback using text mining. In *Big Data and Innovation in Tourism, Travel, and Hospitality: Managerial Approaches, Techniques, and Applications* (pp. 147–162). Springer Singapore. https://doi.org/10.1007/978-981-13-6339-9_10

Ju, Y., Back, K. J., Choi, Y., & Lee, J. S. (2019). Exploring Airbnb service quality attributes and their asymmetric effects on customer satisfaction. *International Journal of Hospitality Management*, *77*, 342–352. https://doi.org/10.1016/j.ijhm.2018.07.014

Korfiatis, N., Stamolampros, P., Kourouthanassis, P., & Sagiadinos, V. (2019). Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews. *Expert Systems with Applications*, *116*, 472–486. https://doi.org/10.1016/j.eswa.2018.09.037

Lafferty, J., & Blei, D. (2009). *Topic Models*. https://doi.org/10.1201/9781420059458.ch4

Lalicic, L., & Weismayer, C. (2018). A model of tourists' loyalty: the case of Airbnb. *Journal of Hospitality and Tourism Technology*, *9*(1), 78–90. https://doi.org/10.1108/JHTT-02-2017-0020

Lawani, A., Reed, M. R., Mark, T., & Zheng, Y. (2019). Reviews and price on online platforms: Evidence from sentiment analysis of Airbnb reviews in Boston. *Regional Science and Urban Economics*, *75*, 22–34. https://doi.org/10.1016/j.regsciurbeco.2018.11.003

Lee, C. K. H., Tse, Y. K., Zhang, M., & Ma, J. (2019). Analysing online reviews to investigate customer behaviour in the sharing economy: The case of Airbnb. *Information Technology and People*, *33*(3), 945–961. https://doi.org/10.1108/ITP-10-2018-0475

Lee, H., Lee, Y., & Yoo, D. (2000). The determinants of perceived service quality and its relationship with satisfaction. In *Journal of Services Marketing* (Vol. 14, Issue 3, pp. 217–231). MCB UP Ltd. https://doi.org/10.1108/08876040010327220

Lee, S., Song, J., & Kim, Y. (2010). An empirical comparison of four text mining methods. *Journal of Computer Information Systems*, *51*(1), 1–10. https://doi.org/10.1080/08874417.2010.11645444

Liang, L. J., Choi, H. C., & Joppe, M. (2018). Exploring the relationship between satisfaction, trust and switching intention, repurchase intention in the context of Airbnb. *International Journal of Hospitality Management*, *69*, 41–48. https://doi.org/10.1016/j.ijhm.2017.10.015

Mastrogiacomo, L., Barravecchia, F., Franceschini, F., & Marimon, F. (2021). Mining quality determinants of Product-Service Systems from User-Generated Contents. *Quality Engineering*.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Mittendorf, C. (2018). Collaborative consumption: the role of familiarity and trust among Millennials. *Journal of Consumer Marketing*, *35*(4), 377–391. https://doi.org/10.1108/JCM-12-2016-2040

Möhlmann, M. (2015). Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again. *Journal of Consumer Behaviour*, *14*(3), 193–207. https://doi.org/10.1002/cb.1512

Palese, B., & Usai, A. (2018). The relative importance of service quality dimensions in E-commerce experiences. *International Journal of Information Management*, *40*, 132–140. https://doi.org/10.1016/j.ijinfomgt.2018.02.001

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (n.d.). *Linguistic Inquiry and Word Count: LIWC2015 Operator's Manual*. Retrieved February 10, 2021, from www.LIWC.net

Priporas, C.-V., Stylos, N., Vedanthachari, L. N., & Santiwatana, P. (2017). Service quality, satisfaction, and customer loyalty in Airbnb accommodation in Thailand. *International Journal of Tourism Research*, *19*(6), 693–704. https://doi.org/10.1002/jtr.2141

Ranjbari, M., Shams Esfandabadi, Z., & Scagnelli, S. D. (2020). A big data approach to map the service quality of short-stay accommodation sharing. *International Journal of*

Contemporary Hospitality Management, *32*(8), 2575–2592. https://doi.org/10.1108/IJCHM-02-2020-0097

Ribeiro, F. N., Araújo, M., Gonçalves, P., André Gonçalves, M., & Benevenuto, F. (2016). SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science*, *5*(1), 1–29. https://doi.org/10.1140/epjds/s13688-016-0085-1

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(1), 1–40. https://doi.org/10.18637/jss.v091.i02

Saha, G. C., & Theingi. (2009). Service quality, satisfaction, and behavioural intentions: A study of low-cost airline carriers in Thailand. *Managing Service Quality*, *19*(3), 350–372. https://doi.org/10.1108/09604520910955348

Saihi, A., Awad, M., & Ben-Daya, M. (2021). Quality 4.0: leveraging Industry 4.0 technologies to improve quality management practices – a systematic review. *International Journal of Quality & Reliability Management*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/IJQRM-09-2021-0305

Schwab, K. (2016). The Fourth Industrial Revolution - Klaus Schwab - Google Books. In *the Fourth Industrial Revolution*.

Serrano, L., Ariza-Montes, A., Nader, M., Sianes, A., & Law, R. (2020). Exploring preferences and sustainable attitudes of Airbnb green users in the review comments and ratings: a text mining approach. *Journal of Sustainable Tourism*. https://doi.org/10.1080/09669582.2020.1838529

Shemwell, D. J., Yavas, U., & Bilgin, Z. (1998). Customer-service provider relationships: An empirical test of a model of service quality, satisfaction and relationship-oriented outcomes. *International Journal of Service Industry Management*, *9*(2), 155–168. https://doi.org/10.1108/09564239810210505

Sivadas, E., & Baker-Prewitt, J. L. (2000). An examination of the relationship between service quality, customer satisfaction, and store loyalty. In *International Journal of Retail & Distribution Management* (Vol. 28, Issue 2, pp. 73–82). MCB UP Ltd. https://doi.org/10.1108/09590550010315223

So, K. K. F., Oh, H., & Min, S. (2018). Motivations and constraints of Airbnb consumers: Findings from a mixed-methods approach. *Tourism Management*, *67*, 224–236. https://doi.org/10.1016/j.tourman.2018.01.009

Sparks, B. A., & Browning, V. (2010). Complaining in cyberspace: The motives and forms of hotel guests' complaints online. *Journal of Hospitality Marketing and Management*, *19*(7), 797–818. https://doi.org/10.1080/19368623.2010.508010

Spreng, R. A., & Mackoy, R. D. (1996). An empirical examination of a model of perceived service quality and satisfaction. *Journal of Retailing*, *72*(2), 201–214. https://doi.org/10.1016/S0022-4359(96)90014-7

Su, Y., & Teng, W. (2018). Contemplating museums' service failure: Extracting the service quality dimensions of museums from negative on-line reviews. *Tourism Management*, *69*, 214–222. https://doi.org/10.1016/j.tourman.2018.06.020

Sutherland, I., & Kiatkawsin, K. (2020). Determinants of guest experience in Airbnb: A topic modeling approach using LDA. *Sustainability (Switzerland)*, *12*(8). https://doi.org/10.3390/SU12083402

Tsao, H. Y., Chen, M. Y., Campbell, C., & Sands, S. (2020). Estimating numerical scale ratings from text-based service reviews. *Journal of Service Management*, *31*(2), 187–202. https://doi.org/10.1108/JOSM-06-2019-0167

Tussyadiah, I. P. (2016). Factors of satisfaction and intention to use peer-to-peer accommodation. *International Journal of Hospitality Management*, *55*, 70–80. https://doi.org/10.1016/j.ijhm.2016.03.005

Tussyadiah, I. P., & Park, S. (2018). When guests trust hosts for their words: Host description and trust in sharing economy. *Tourism Management*, *67*, 261–272. https://doi.org/10.1016/j.tourman.2018.02.002

Tussyadiah, I. P., & Zach, F. (2017). Identifying salient attributes of peer-to-peer accommodation experience. *Journal of Travel and Tourism Marketing*, *34*(5), 636–652. https://doi.org/10.1080/10548408.2016.1209153

Wallach, H. M., Murray, I., Salakhutdinov, R., & Mimno, D. (2009). Evaluation methods for topic models. *Proceedings of the 26th International Conference On Machine Learning, ICML 2009*, *4*, 1105–1112.

Wang, Y., Agichtein, E., & Benzi, M. (2012). TM-LDA: Efficient online modeling of latent topic transitions in social media. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 123–131. https://doi.org/10.1145/2339530.2339552

Weathington, B. L., Cunningham, C. J. L., & Pittenger, D. J. (2012). Appendix B: Statistical Tables. In *Understanding Business Research* (pp. 435–483). Wiley. https://doi.org/10.1002/9781118342978.app2

World Tourism Organization. (2020). UNWTO World Tourism Barometer and Statistical Annex, December 2020. In *UNWTO World Tourism Barometer* (Vol. 18, Issue 7). World Tourism Organization (UNWTO). https://doi.org/10.18111/wtobarometereng.2020.18.1.7

Yacouel, N., & Fleischer, A. (2012). The Role of Cybermediaries in Reputation Building and Price Premiums in the Online Hotel Market. *Journal of Travel Research*, *51*(2), 219–226. https://doi.org/10.1177/0047287511400611

Zeithaml, V. A. (2000). Service Quality, Profitability, and the Economic Worth of Customers: What We Know and What We Need to Learn. *Journal of the Academy of Marketing Science*, *28*(1), 67–85. https://doi.org/10.1177/0092070300281007

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *Journal of Marketing Research*, *54*(5), 687–705. https://doi.org/10.1509/jmr.15.0204

Zhang, J. (2019). What's yours is mine: exploring customer voice on Airbnb using text-mining approaches. *Journal of Consumer Marketing*, *36*(5), 655–665. https://doi.org/10.1108/JCM-02-2018-2581

Zhang, W., Xu, H., & Wan, W. (2012). Weakness Finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. *Expert Systems with Applications*, *39*(11), 10283–10291. https://doi.org/10.1016/j.eswa.2012.02.166

Zonnenshain, A., & Kenett, R. S. (2020). Quality 4.0—the challenging future of quality engineering. *Https://Doi.Org/10.1080/08982112.2019.1706744*, *32*(4), 614–626. https://doi.org/10.1080/08982112.2019.1706744