



# Exploring Sources of Satisfaction and Dissatisfaction in Airbnb Accommodation Using Unsupervised and Supervised Topic Modeling

Kai Ding<sup>1</sup>, Wei Chong Choo<sup>1</sup>, Keng Yap Ng<sup>2</sup>, Siew Imm Ng<sup>1</sup> and Pu Song<sup>3\*</sup>

<sup>1</sup> Department of Management and Marketing, Faculty of Economics and Management, Universiti Putra Malaysia, Seri Kembangan, Malaysia, <sup>2</sup> Department of Software Engineering and Information System, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Seri Kembangan, Malaysia, <sup>3</sup> Department of Preschool Education, Guiyang Preschool Education College, Guiyang, China

## OPEN ACCESS

### Edited by:

Jose Ramon Saura,  
Rey Juan Carlos University, Spain

### Reviewed by:

Marco Ieva,  
University of Parma, Italy  
Vera Gelashvili,  
Rey Juan Carlos University, Spain  
Juan Gabriel Martínez Navalón,  
Rey Juan Carlos University, Spain

### \*Correspondence:

Pu Song  
songpuedu@gmail.com

### Specialty section:

This article was submitted to  
Organizational Psychology,  
a section of the journal  
Frontiers in Psychology

Received: 27 January 2021

Accepted: 17 March 2021

Published: 21 April 2021

### Citation:

Ding K, Choo WC, Ng KY, Ng SI and Song P (2021) Exploring Sources of Satisfaction and Dissatisfaction in Airbnb Accommodation Using Unsupervised and Supervised Topic Modeling. *Front. Psychol.* 12:659481. doi: 10.3389/fpsyg.2021.659481

This study aims to examine key attributes affecting Airbnb users' satisfaction and dissatisfaction through the analysis of online reviews. A corpus that comprises 59,766 Airbnb reviews from 27,980 listings located in 12 different cities is analyzed by using both Latent Dirichlet Allocation (LDA) and supervised LDA (sLDA) approach. Unlike previous LDA based Airbnb studies, this study examines positive and negative Airbnb reviews separately, and results reveal the heterogeneity of satisfaction and dissatisfaction attributes in Airbnb accommodation. In particular, the emergence of the topic "guest conflicts" in this study leads to a new direction in future sharing economy accommodation research, which is to study the interactions of different guests in a highly shared environment. The results of topic distribution analysis show that in different types of Airbnb properties, Airbnb users attach different importance to the same service attributes. The topic correlation analysis reveals that home like experience and help from the host are associated with Airbnb users' revisit intention. We determine attributes that have the strongest predictive power to Airbnb users' satisfaction and dissatisfaction through the sLDA analysis, which provides valuable managerial insights into priority setting when developing strategies to increase Airbnb users' satisfaction. Methodologically, this study contributes by illustrating how to employ novel approaches to transform social media data into useful knowledge about customer satisfaction, and the findings can provide valuable managerial implications for Airbnb practitioners.

**Keywords:** customer satisfaction, sharing economy, Airbnb, text mining, supervised topic modeling, big data, user-generated content

## INTRODUCTION

With the advancement of internet technology and the development of Web 2.0 application, the tourism and hospitality industry has experienced a significant change (Litvin et al., 2008). One of the significant changes is that an increasing number of customers are actively acquiring and sharing information on products and services on online social platforms (Podnar and Javernik, 2012) instead of solely relying on information provided by service providers. Information shared by customers on social media platforms is also regarded as user-generated content (UGC). On the

internet, various types of UGC are generated, and customer reviews are considered to be the most influential UGC, which establishes a communication channel among different customers, as well as between service providers and customers (Casaló et al., 2015). For customers, online customer reviews are revealed to be an essential information source in decision-making (Shengli and Fan, 2019), a statistic indicates that almost 95% of travelers read online reviews before purchasing a travel product (Ady and Quadri-Felitti, 2015). For service providers, customer reviews reflect customers' needs and expectations, enabling novel approaches to increase customer satisfaction (Phillips et al., 2015).

In recent years, increasing attention has been paid to the importance of evaluating social media content to improve the quality of the customer experience and identify service recovery solutions, especially in the hospitality industry. Particularly, with the emergence of different text analytics, various studies have been conducted to understand customer satisfaction in the hospitality industry. For instance, Berezina et al. (2015) examined factors causing hotel customers' satisfaction and dissatisfaction by applying text network analysis. Kim et al. (2016) conducted a content analysis of online hotel reviews to identify the main drivers of hotel customer satisfaction and dissatisfaction. Xu and Li (2016) identified the determinants of customer satisfaction and dissatisfaction toward hotels through latent semantic analysis. Additionally, Padma and Ahn (2020) applied word frequency analysis to determine major themes of the luxury hotel's service quality related to hotel guest satisfaction and dissatisfaction in Malaysia. The previous studies have demonstrated the usefulness of extracting meaningful information regarding customer satisfaction from customer reviews in the hospitality industry. Some studies (e.g., Berezina et al., 2015; Xu and Li, 2016) found that the antecedents of customer satisfaction and dissatisfaction are not identical, suggesting that these two constructs can be examined separately to support strategy development for improving customer satisfaction.

Despite the wide application of text analytics in academic research, existing research on UGC in the accommodation sector of the sharing economy is probably still in its infancy (Sutherland and Kiatkawsin, 2020). Hence, this study aims to explore the sources of satisfaction and dissatisfaction in accommodations listed in a sharing economy platform, namely Airbnb, through a systematic analysis of UGC. The sharing economy is a newly developed business model based on consumption and sharing of goods and services among strangers (Gossen and Scholl, 2016). This business model has become the most rapidly developing business trends in history, with over 24 billion US dollars invested in the venture capital finding since 2010 (Wallenstein and Shelat, 2017). Airbnb is the leading example of the sharing economy in the accommodation sector (Liu and Mattila, 2017). Although both traditional hotels and Airbnb are mainly providing accommodation services, Airbnb users are often found to pursue different lodging experiences than customer choosing traditional hotels (Yi et al., 2019), such as living like a local and the opportunity to interact with the host and local community (Tussyadiah and Pesonen, 2016). Due to the distinctive preferences of Airbnb users, the results regarding

customer satisfaction attributes identified from traditional hotels may not be compatible with the Airbnb context (Luo and Tang, 2019; Sutherland and Kiatkawsin, 2020).

Many studies have been conducted to understand the satisfaction of Airbnb users through the analysis of customer reviews, but previous studies (e.g., Möhlmann, 2015; Tussyadiah and Zach, 2017) only focused on satisfaction attributes and neglected dissatisfaction attributes, which leads to the assumption that the negative performance of those satisfaction attributes could be the source of Airbnb user dissatisfaction. According to Herzberg's two-factor theory (Herzberg, 1966), satisfaction and dissatisfaction are two independent continuums instead of two opposite extremes. In addition, the negative and positive performance of the same attribute would result in asymmetric impacts on customer satisfaction in the hospitality industry (Bi et al., 2019). Gu and Ryan (2008) also suggest that attributes result in customer dissatisfaction are not identical to those that lead to customer satisfaction. Therefore, in order to help Airbnb practitioners develop more effective strategies to improve user satisfaction, it is crucial to understand both guest satisfaction and dissatisfaction by examining UGC on social media (Kim et al., 2016). Kano et al. (1984) highlight the non-linear nature of customer satisfaction attributes, suggesting that service attributes contribute unequally to customer satisfaction. In order to rank the importance of service attributes in customer reviews, previous Airbnb studies (Ju et al., 2019; Luo and Tang, 2019; Ding et al., 2020) tend to be based on the emphasis level of those attributes. However, attributes with a higher emphasis level in customer reviews do not have the same level of influence on customer satisfaction (Xu, 2020a). Therefore, it is necessary to conduct further studies to determine the relative importance of identified attributes to customer satisfaction, which can help Airbnb practitioners to set priorities when developing strategies to increase customer satisfaction. Besides, this study also intends to extend previous studies by comparing Airbnb users' preferences when staying in four different types of Airbnb properties (e.g., an entire property, a hotel room, a shared room, and a private room), as Xu (2020b) found that Airbnb users staying in different types of Airbnb properties put emphasis on different aspects of the lodging experience. The findings can provide valuable suggestions for hosts to adopt more targeted strategies to ensure a satisfactory lodging experience for Airbnb users.

Given the research gap, our study aims to answer the following set of questions:

- (1) How to extract both satisfaction and dissatisfaction attributes from customer reviews in Airbnb accommodation?
- (2) How to determine the relative importance of attributes extracted from customer reviews to Airbnb users' satisfaction and dissatisfaction?
- (3) How do Airbnb users' review emphasis on satisfaction attributes and dissatisfaction attributes differ when they stay in different types of Airbnb properties?

To answer these questions, Latent Dirichlet Allocation (LDA), an unsupervised topic modeling technique, is employed to identify attributes that cause Airbnb user satisfaction and dissatisfaction.

The results of LDA can provide proportional distributions of extracted topics in different documents, enabling us to examine how Airbnb user's perception of service attributes varies in different types of Airbnb properties. Lastly, supervised LDA (sLDA), a supervised topic modeling technique, is used to determine the relative importance of service attributes that are associated with user satisfaction and dissatisfaction.

This study makes the following theoretical contributions. First, from the perspective of methodology, this is the first study introducing the sLDA into customer satisfaction research in the lodging industry, providing a more direct solution to conduct topic regression analysis. Second, this study contributes to existing literature regarding Airbnb service improvement by enhancing our understanding of attributes resulting in both user satisfaction and dissatisfaction. As for Airbnb practitioners, this study provides valuable insights into required priority actions to increase the level of user satisfaction and specific service improvement suggestions for different types of Airbnb properties.

The rest of the paper is organized as follows. Section Literature review presents the relevant studies on Airbnb satisfaction and textual data analysis techniques. Section Methodology describes research methods and the process of data analysis. In section Results, the data and the results are analyzed. Section Conclusion concludes this study.

## LITERATURE REVIEW

### Airbnb Satisfaction Attributes

To increase customer satisfaction, one of the essential tasks is to identify attributes that are most related to customers' needs and expectations (Yang et al., 2011). Different sources of data have been used to identify attributes affecting the satisfaction of Airbnb users. One stream of studies used traditional survey methods (Möhlmann, 2015; Tussyadiah, 2016; Priporas et al., 2017; Lee and Kim, 2018; Sthapit et al., 2019) while another stream of studies adopted online reviews as the data source

to examine Airbnb users' satisfaction (Tussyadiah and Zach, 2017; Ju et al., 2019). The major findings of these studies are summarized in **Table 1**.

Many other studies have also been conducted to identify key attributes from Airbnb user experience that could drive user satisfaction, and the findings of these studies can serve as useful references to compare with the results of the present study. Tussyadiah and Zach (2015) analyzed 41,560 reviews collected from Portland and Oregon, US, and identified three major attributes ("location," "host," and "property") that frequently appeared in the Airbnb user reviews. Brochado et al. (2017) examined Airbnb user experience by analyzing 1,776 reviews collected from three countries (India, Portugal, and the USA), and eight themes were identified in this study, including "stay," "host," "place," "location," "apartment," "room," "city," and "home." Besides, Cheng and Jin (2019) identified influential attributes on Airbnb users' lodging experiences by analyzing 170,123 Airbnb reviews from Sydney, Australia, and three key influential attributes were identified in this study, including "location," "amenities," and "host." Luo and Tang (2019) analyzed 250,439 Airbnb reviews collected from Los Angeles and identified five latent aspects associated with the lodging experience of Airbnb users, including "communication," "experience," "location," "product/service," and "value."

Although online reviews have been used in previous studies to identify attributes that Airbnb users are concerned about, these studies failed to distinguish between the service attributes that determine Airbnb users' satisfaction and dissatisfaction. Moreover, most of the previous studies mainly focused on attributes that influence Airbnb users' satisfaction, but there is little research on dissatisfaction attributes. Customer dissatisfaction is an apparent reality in all the service industry. Previous studies that examine the impact of service attributes on customer satisfaction (Chowdhary and Prakash, 2005; Chen et al., 2014; Kim et al., 2016; Xu and Li, 2016; Park et al., 2020) indicate that it is insufficient to only identify one dimension of satisfaction. The assumption of one-dimensional concept is

**TABLE 1** | Summary of Airbnb satisfaction attributes.

References	Methodology	Key attributes	Country
Möhlmann (2015)	Survey with partial least squares technique	Utility, trust, cost savings, and familiarity.	Germany
Tussyadiah (2016)	Survey with an exploratory factor analysis	Enjoyment factors, monetary benefits (value), and accommodation amenities.	US
Priporas et al. (2017)	Survey with partial least squares technique	Service quality attributes from the questionnaire developed by Akbaba (2006).	Thailand
Tussyadiah and Zach (2017)	Text mining and regression analysis	location and feeling welcome	USA
Lee and Kim (2018)	Survey with structural equation modeling	Hedonic and utilitarian values.	USA
Sthapit et al. (2019)	Web-based survey with confirmatory factor analysis	Functional value and social value.	Italy
Ju et al. (2019)	Text mining and exploratory factor analysis	Facility service quality, visually appealing, room/house, comfortable bed, helpful host, and friendly host.	USA and Canada

that an individual service attribute can generate both satisfaction and dissatisfaction (Oliver, 1980). This assumption has been questioned by early customer satisfaction studies suggesting that the absence of service attributes which generate satisfaction might not result in dissatisfaction (Kano et al., 1984; Cadotte and Turgeon, 1988). In addition, according to Herzberg's two-factor theory (Herzberg, 1966) that has been widely applied in customer satisfaction studies in the hospitality industry (Berezina et al., 2015; Kim et al., 2016; Tontini et al., 2017; Gerdt et al., 2019), certain attributes contributing to satisfaction do not create dissatisfaction and vice versa. More specifically, Herzberg (1966) categorized satisfaction into two distinctive groups, namely, motivation factors (satisfiers), and hygiene factors (dissatisfiers). The author suggests that missing of hygiene factors can lead to dissatisfaction, but their presence does not contribute to the enhancement of satisfaction as those attributes are perceived as guaranteed features. The realization of motivational factors can lead to the increase of customer satisfaction, but their existence does not necessarily generate satisfaction. Therefore, satisfaction and dissatisfaction are not a two-phase continuum, with one increasing and the other decreasing (Berezina et al., 2015).

Based on the Herzberg's two-factor theory, this study attempted to identify both satisfaction and dissatisfaction attributes through the analysis of Airbnb online reviews. Online reviews have been widely used as a valuable source of information to understand customer satisfaction in the hospitality industry (Guo et al., 2017). These reviews provide descriptions about customers' consumption experience and show their opinions toward provided services. Comparing to using customer ratings that are based on numeric figures to measure customer satisfaction (Felbermayr and Nanopoulos, 2016), customer reviews are more informative by presenting customer satisfaction and dissatisfaction in a more detailed manner (Xu and Li, 2016). In addition, in some ways, adopting the customer review based approach outperforms conventional methods, such as case studies, and survey approaches that have been heavily applied in customer satisfaction research. First, due to the open structure of customer reviews, these spontaneously generated UGC can provide more authentic reports of customers' experiences compared to surveys and case studies (Lucini et al., 2020). Second, the setting of survey questions is usually based on previous studies (Guo et al., 2017), which is not conducive to identifying novel subjects not considered in previous questionnaires, such as customer preferences in a newly developed business.

## Sentiment Analysis in Consumer Research

In the business research, sentiment analysis refers to the process of identifying different emotions (positive, negative, and neutral) toward a product or service in the text using computer-aided sentiment detection tools (Nasukawa and Yi, 2003), and these emotions were found to be strongly related to customer satisfaction. Kumar and Zymbler (2019) applied a machine learning-based sentiment analysis approach to assess customer satisfaction from airline reviews. Additionally, Zhu et al. (2020) examined the relationship between sentiment and

guest satisfaction in Airbnb accommodation, adopted a lexicon-based method to classify words in each review as positive and negative, and represented guest satisfaction using accumulated online ratings. The results of this study revealed that Airbnb listings are more likely to receive higher rating scores when their reviews have a higher degree of positive sentiment. Geetha et al. (2017) examined the relationship between online hotel review sentiment and customer rating score that is used to represent the customer's satisfaction level in this study. This study concluded that the change of customer rating is influenced by reviewing sentiment polarity, and the strength of influence of review sentiment polarity on the variation of customer ratings is different in budget and premium hotels, with a value of 44 and 21%, respectively.

These studies demonstrated that sentiment polarity can be a good indicator of customers' satisfaction, and this concept has also been implemented in previous studies (e.g., Xu et al., 2017; Wang et al., 2018). After drawing insights from previous studies and considering that customer rating for individual comment is not available on the Airbnb platform, the present study used the sentiment polarity of each review as a proxy for the satisfaction level of Airbnb users; specifically, positive reviews indicate satisfaction while negative reviews indicate dissatisfaction.

In terms of techniques used to conduct sentiment analysis, lexicon-based methods based on a pre-defined list of lexical features have been widely applied in social media research (Choi et al., 2020). Considering that implementing lexicon-based methods are less time-consuming and convenient to implement, and many lexicons developed by researchers in the text mining field can be directly applied to the analysis, hence we decided to use a lexicon-based method to conduct sentiment analysis.

## Topic Modeling

Topic modeling is a machine learning-based text mining approach, which aims to reveal themes from text documents (Blei, 2012). Although human coding remains the gold standard for textual content analysis (Short et al., 2010), analyzing a large volume of customer online data is far beyond human processing capacity (Kumar and Zymbler, 2019). Topic modeling provides an automatic solution to analyze those unstructured big data without requiring manual coding, thus reducing time and costs, and also human bias. Comparing to the word-frequency based text mining approach applied in previous studies (Ju et al., 2019; Padma and Ahn, 2020), topic modeling can provide more meaningful results, as frequency-based approaches ignore the context and the underlying relationship between the words (Ahmad and Laroche, 2015).

LDA is one of the popular topic modeling techniques. The underlining assumption of the LDA model is that the words of each document result from a mixture of topics, and each topic is a distribution over the vocabulary. The outcome of LDA includes a certain number of topics represented by a list of words with a high probability of co-occurrence. Two statistical outputs are also provided in LDA, including the estimated probability of different topics within each document and the probability that a word is used to represent a topic. LDA has been widely used in consumer research to understand customers' perceptions of



service. For instance, Sanchez-Franco et al. (2019) used LDA to examine the relationship between different aspects of hospitality services and service quality. Ibrahim and Wang (2019) employed LDA to extract the customers' major topics of concern to support the online retailing service improvement. Besides, LDA also provides a good solution to understand customer's satisfaction in the service-based industry. For instance, Guo et al. (2017) used LDA to identify the key dimensions of hotel customer satisfaction and analyzed 266,533 online customer reviews collected from TripAdvisor. The findings of this study revealed 19 important dimensions regarding customer satisfaction. Lucini et al. (2020) used LDA to explore dimensions of airline customers by analyzing 55,000 online customer reviews. Additionally, Bi et al. (2019) demonstrated the suitability of using LDA to extract customer satisfaction dimensions from online reviews. This study extracted 18 meaningful customer satisfaction dimensions from 25,314 electronic product reviews.

Except for LDA, there are also some other topic modeling techniques, such as structural topic model (STM) developed by Roberts et al. (2014), and dynamic topic model (DTM) developed by Blei and Lafferty (2006). Although all these topic modeling tools also can extract latent topics from the text, these tools have different focuses and require additional variables to be added in the application. For example, the emphasis of using STM is to examine the impact of pre-defined covariates on the changes of topic prevalence (Korfiatis et al., 2019). As for DTM, it is often applied to identify changes in customer perception toward certain products over time (Ha et al., 2017). Despite the availability of different topic modeling techniques, the selection of appropriate tools should be based on which of these models best fit the research objectives (Saura, 2020). After evaluating different topic modeling techniques and referring to previous studies (Guo et al., 2017; Lucini et al., 2020), we found that LDA is more suitable for this study based on its simplicity in application and its confirmed effectiveness in customer satisfaction studies.

Although LDA is capable of identifying attributes related to customer's satisfaction from a large textual document, the results from the LDA model cannot clearly differentiate satisfaction and dissatisfaction attributes. This is because the allocation of words to each topic in the LDA model does not consider the sentiment polarity of each document, causing many topics to contain both positive and negative words. To overcome this limitation, the present study conducts LDA analysis on negative reviews and positive reviews separately, as positive topics indicating customer satisfaction have a higher probability to appear in positive reviews and vice versa (Berezina et al., 2015; Xu and Li, 2016; Hu et al., 2019). Besides, reviews with negative emotions are found to be more authentic (Chen, 2020). Therefore, a separate analysis of positive and negative reviews can shed light on the attributes that customers truly care about.

sLDA is another topic modeling approach implemented in the present study. sLDA is an extension of LDA, in which an additional response variable is modeled to determine the latent topics with the best predictive power (Blei and McAuliffe, 2007). Comparing with other topic modeling tools (e.g., STM and DTM), sLDA outperforms these tools by providing a more direct solution to conduct topic regression analysis (Blei and McAuliffe,

2007). To the best of our knowledge, sLDA has not been applied in hospitality research. In other contexts, Chai (2019) applied sLDA to analyze employee satisfaction dataset that contains both textual feedbacks and numeric ratings. The author demonstrated the effectiveness of using sLDA to extract employees' satisfaction factors that can predict satisfaction ratings. Blei and McAuliffe (2007) employed sLDA to predict movie ratings and compared the results with the LDA-based linear regression and lasso, revealing that sLDA achieved better predictive performance. In addition to sLDA, labeled LDA ( $L$ -LDA) is another joint topic model (Flaherty et al., 2005) while it fundamentally differs from the sLDA (Blei and McAuliffe, 2007). Kim and Kang (2018) employed  $L$ -LDA to determine the positive and negative attributes extracted from cosmetic reviews. In this study, customer ratings were incorporated as the document label, and the findings revealed words associated with different ratings.

For the present study, sLDA is applied to determine topics with predictive power to different levels of Airbnb user satisfaction that were represented by sentiment scores. Hence, the attributes that are most associated with Airbnb users' satisfaction and dissatisfaction can be identified. Another benefit of applying sLDA is that using more than one topic modeling techniques can generate a more meaningful analysis (Williams and Betak, 2019).

The present study differs from the previous Airbnb online reviews studies by dividing reviews into the positive and negative group. Hence, satisfaction factors and dissatisfaction factors can be clearly distinguished, and the following comparison analysis can provide additional insights into how Airbnb users' satisfaction can be increased. Besides, this study explored Airbnb reviews and their associated sentiment scores using a novel method to conduct further analysis. The joint analysis of online reviews and sentiment scores can provide us with the relative importance of different satisfaction and dissatisfaction attributes.

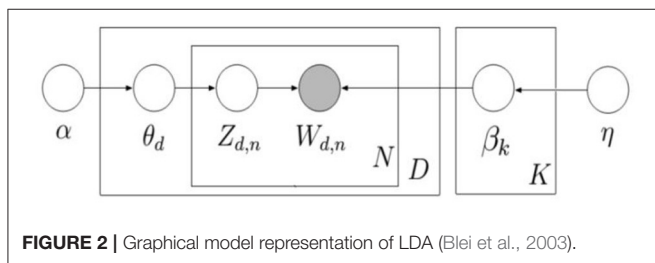
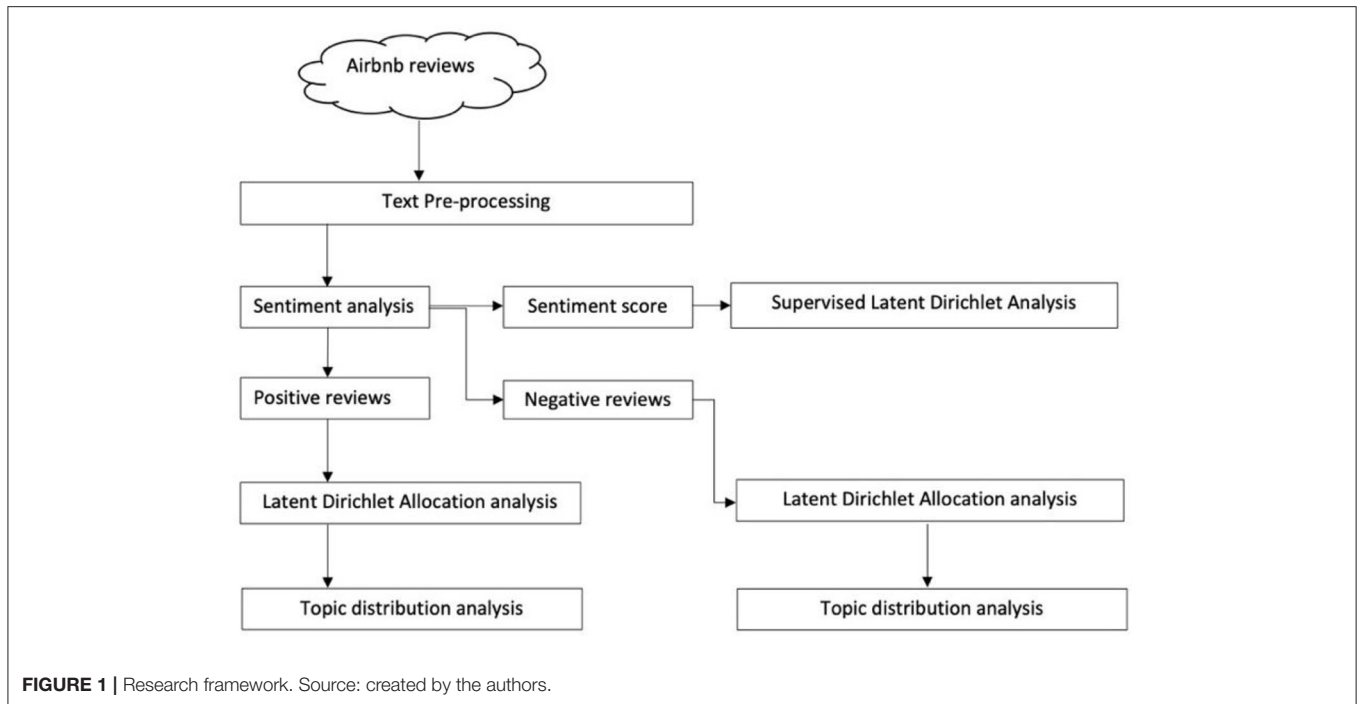
## METHODOLOGY

### Research Design Framework

The main workflow of the present study is presented in **Figure 1**, including the following steps: (1) collecting Airbnb review dataset; (2) pre-processing text to reduce corpus dimensions and unnecessary noise; (3) conducting sentiment analysis and re-organize the corpus; (4) extracting topics from positive and negative reviews separately using LDA; (5) examining the topic distributions in different types of Airbnb properties; (6) identifying topics that are highly associated with Airbnb user satisfaction and dissatisfaction by referring to the statistical results of sLDA.

### LDA and sLDA Method

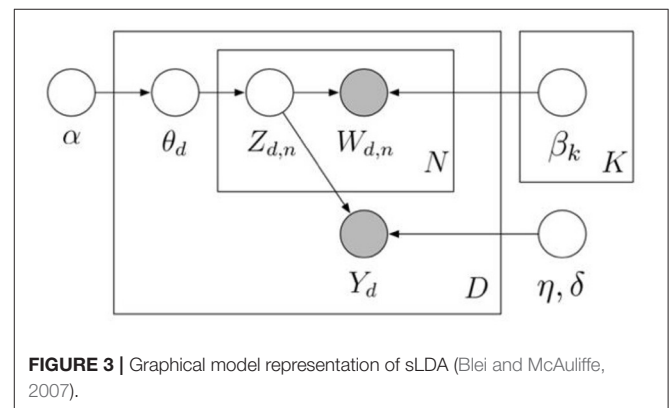
Each Airbnb review is considered to be a document, and a corpus consists of  $N$  documents represented in  $d = (D_1, D_2, \dots, D_N)$ . Each document is composed of  $m$  words represented in  $D = \{w_1, w_2, \dots, w_m\}$ . The LDA plates are illustrated in **Figure 2**.  $K$  represents the total number of topics;  $\beta_k$  denotes topics distribution over different  $K$ ;  $N$  refers to the number of words in a document;  $D$  represents the total number of the document;  $Z_{d,n}$  (per-word topic assignment) and  $W_{dn}$  (per-word topic



proportions) are word-level parameters;  $\theta_d$  (per-document topic proportions) denotes the document-level parameter, which is an essential metric for comparing topic distributions in different documents. Each node in the plate represents a random variable. Specifically, the node with shadow represents the observed words in the corpus, and the unshaded nodes denote the hidden variables. The directed edges indicate a dependence between the corresponding variables. For instance,  $Z_{d,n}$  is conditionally dependent on  $\theta_d$ , and  $W_{d,n}$  is conditionally dependent on  $Z_{d,n}$ . Besides,  $\alpha$  and  $\eta$  are two Dirichlet parameters. In this paper, the R package “topicmodel” using Gibbs sampling is selected to perform the LDA computation.

The plate of the sLDA model is exhibited in **Figure 3**. In addition to adding the response variable ( $Y_d$ ) to the sLDA, there are many similarities between the sLDA and LDA. In this study, the sentiment score of each document is used as the response variable. The sentiment score ( $y_n$ ) of each review is modeled as:

$$y_n = \lambda_1 \bar{z}_{1n} + \lambda_2 \bar{z}_{2n} + \dots + \lambda_K \bar{z}_{Kn} + \epsilon_n$$



where  $\lambda_K$  stands for the regression coefficients of the model;  $\bar{z}_{kn}$ ,  $k = 1, \dots, K$  indicates the estimated proportion of each topic in document  $n$ . This regression model is fitted without an intercept, as the components of  $\bar{z}_n$  always sum to one (Blei and McAuliffe, 2007). Regarding notations,  $\eta$  and  $\sigma^2$  are pre-set constants for the normal distribution, explaining how learned topics affect the supervision (Blei and McAuliffe, 2007). Consistent with Chai (2019), the sLDA analysis was conducted in this study using the sample code in the “lda” package.

### Data Collection

The Airbnb dataset for this study was acquired from the Inside Airbnb website where the researcher can obtain publicly accessible Airbnb data for the purpose of research, and this website also provided data for several published Airbnb related studies (e.g., Horn and Merante, 2017; Cheng and Jin, 2019).

We collected Airbnb online review data from 12 cities in 11 different countries from January 2012 to January 2020, with a total of 1.2 million reviews. Those cities include Beijing, Hong Kong, Tokyo, Cape Town, Paris, London, Seattle, Ottawa, Mexico City, Buenos Aires, Santiago, and Rio. The reason for selecting 12 different cities is for the purpose of improving the generalizability of the findings (Sutherland and Kiatkawsin, 2020), which is differentiated from most of previous Airbnb studies that only focus on an individual country (e.g., Cheng and Jin, 2019; Ding et al., 2020). Airbnb has expanded its business to more than 220 countries (Airbnb, 2017), and it is particularly popular for international tourists (Guttentag et al., 2018). The research results with higher generality can help Airbnb and its hosts understand some common expectations of Airbnb users and also support them to provide satisfactory services to guests from different countries. Based on the available data, we selected some well-known capitals and popular international cities from countries on different continents, where Airbnb is widely used by tourists.

## Data Pre-processing

The data pre-processing was conducted using R programming. First, all the non-English reviews were filtered using the “cld2” package. Considering that short reviews may produce an undesirable outcome because LDA is built on co-occurrence of words, reviews < 50 words were excluded (Lee and Yu, 2018). The “TM” package was used to remove numbers, punctuations, and extra white space. All the words were converted to the lower case. Besides, we chose to only lemmatize all the words in the corpus (i.e., removing the inflectional ending of a word) instead of stemming words, which could result in the loss of information and poor interpretability (Hagen, 2018). Then, common stop words such as “the,” “a,” and “and” were removed, and a list of customized stop words such as city names, duplicated reviews, and system-generated phrases were also removed. Many compound words were converted as single tokens, such as “Wi Fi” to “wif.” Moreover, words appearing < 1% of the corpus were removed to reduce noisy words.

## Sentiment Analysis

After data pre-processing, sentiment analysis was performed using the R package “SentimentR” (Rinker, 2019), with zero indicating a neutral sentiment and both polarities reflecting negative and positive sentiments. Compared with other lexicon-based sentiment analysis techniques that can only classify the

review as positive, neutral, and negative, one advantage of “sentimentr” is that its numeric output can be directly used to perform further regression analysis. The numeric output of sentiment analysis is used as the response variable in the sLDA model. Besides, the dictionary of “sentimentr” is incorporated with valence shifters, such as “not” and “can’t,” improving the accuracy of measuring sentiment.

To ensure a balanced sample with an equal number of positive and negative reviews, we first identified the number of negative reviews in each city and then randomly selected an equal number of positive reviews in the same city. The sample summary is exhibited in **Table 2**.

The comparison of frequent terms in positive and negative reviews is illustrated in **Figure 4**, where the horizontal axis and the vertical axis indicate the proportion of words appeared in negative and positive comments, respectively; the words close to the red line have similar frequencies in both types of reviews, such as “location,” “respond,” and “size,” and most of them are neutral words. In positive reviews, Airbnb users often compliment host behavior using words “thoughtful,” “knowledgeable,” “warmly,” and “attentive,” which rarely appeared in negative reviews. On the contrary, Airbnb hosts are often described as “rude,” “unprofessional,” and “terrible” in the negative reviews. Additionally, Airbnb users use more general words to express their feelings in positive reviews, such as “excellent,” “nice,” and “fantastic.” By contrast, negative reviews contain more words related to a specific aspect of the Airbnb lodging experience, such as “dirty,” “stain,” “smell,” “bug,” and “stink” that describe the sanitary condition of the room. Moreover, “refund” appeared significantly more and with comparatively high frequency in negative reviews. The general information regarding factors leading to Airbnb user’s satisfaction and dissatisfaction is exhibited in the term frequency map, which can serve as a reference to compare with the topic modeling results.

## RESULTS

### Topic Number Identification

In line with the research of Lee and Yu (2018) and Roque et al. (2019), the optimal topic number for LDA analysis was determined using the R package “ldatuning” (Murzintcev, 2015). Besides, a range of the optimal topic number can be obtained with the “ldatuning” package by running four different methods developed by Griffiths and Steyvers (2004), Cao et al. (2009),

**TABLE 2** | Descriptive statistics of review samples.

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	Total
No. of listing	297	762	1,621	3,535	5,794	8,276	11,067	13,685	2,943	27,980
No. of reviews	360	943	2,026	4,334	7,233	10,288	13,991	17,469	3,122	59,766
Entire home/apt	266	654	1,414	3,037	4,905	7,246	10,091	13,264	2,467	43,344
Private room	91	282	593	16	2,209	2,852	3,599	3,756	604	14,002
Shared room	3	3	13	1,247	78	76	100	170	15	1,705
Hotel room	0	4	6	34	41	114	201	279	36	715





**TABLE 3** | Topic label.

Type of reviews	Topic no.	Topic name	Literatures
Positive reviews	Topic 1	Sleep disturbance	Y
	Topic 2	Help from hosts	Y
	Topic 3	Public transportation	Y
	Topic 4	Amenities	Y
	Topic 5	Location	Y
	Topic 6	Check in/out	Y
	Topic 7	View	Y
	Topic 8	Neighborhood environment	Y
	Topic 9	Room size	Y
	Topic 10	Home-like experience	Y
	Topic 11	Easy access to desired places	Y
	Topic 12	Room experiences	Y
	Topic 13	Communication	Y
	Topic 14	Revisit intention	Y
Negative reviews	Topic 1	Property issues	Y
	Topic 2	Unmatched descriptions	Y
	Topic 3	Room temperature	N
	Topic 4	Kitchen experiences	Y
	Topic 5	Noise	Y
	Topic 6	Location	Y
	Topic 7	Easy access to desired places	Y
	Topic 8	Booking and refund	Y
	Topic 9	Hosts' unpleasant behavior	Y
	Topic 10	Home-like experience	Y
	Topic 11	Door lock/key	Y
	Topic 12	Bathroom problems	Y
	Topic 13	Poor room maintenance	Y
	Topic 14	Guest conflicts	N
Topic 15	Dirtiness and smell	Y	
Topic 16	Communication	Y	
Topic 17	Bed size/condition	Y	
Topic 18	Checking in/out	Y	

dimension “convenience” to group those topics. In **Table 4**, service attributes extracted from positive and negative reviews are presented under the corresponding dimensions. In general, Airbnb users discussed significantly more on tangibility related attributes. More specifically, in the positive reviews, topics related to external environment were mentioned more often, while in the negative reviews, Airbnb users talked more about the functional failure of internal facilities. Except for the dimension of convenience, the topics under the remaining dimensions are

all related to the host, which indicates the important role of hosts in Airbnb user experience.

In positive reviews, two topics are associated with the external environment: “neighborhood environment” and “views.” The other three attributes are associated with accommodation experiences: “amenities,” “room size,” and “sleeping disturbance.” In negative reviews, topics related to tangibility are mainly about different issues or problems encountered by Airbnb users. As for topic “property issues,” there are only general descriptions regarding property-related issues without mentioning a specific aspect. “Room temperature” includes Airbnb users’ complaints about uncomfortable room temperature. Particularly, Airbnb users often highlighted that windows failed to regulate the room temperature. Concerning “kitchen experiences,” Airbnb users often complained about the insufficiency or poor condition of provided supplies for the cooking purpose. Besides, the topic “bathroom problems” involves Airbnb users’ complaints regarding the cleanliness of the bathroom, as well as shower water temperature and pressure. “Door lock/key” includes safety concerns of Airbnb users and inconvenience in collecting the keys.

Three topics are under the dimension of reliability, and all these topics are extracted from negative reviews. The topic “unmatched description” refers to Airbnb users’ dissatisfaction that listing description does not match the real situation. “Booking and refund” involves many Airbnb users’ frustrations toward the booking cancellation and also some refund issues, such as delayed refund or incorrect amount of refund. As for “hosts’ unpleasant behavior,” Airbnb users often complained about the irresponsible behaviors of hosts, such as failing to provide timely and effective assistance to solve the problems, inconsistent with the topic “help for hosts” from the dimension of assurance.

Under the dimension of responsiveness, the topic “communication” is extracted from both positive and negative reviews. In positive reviews, Airbnb users often highlighted the prompt reply from hosts. Besides, it can be observed from the representative reviews of this topic that the communication was mainly related to the check-in process, and Airbnb users also expressed their appreciations to clear instructions provided by the host. By contrast, “communication” extracted from negative reviews is related to Airbnb users’ complaints that they are not able to reach the host or receive a timely reply from the host.

Under the dimension of empathy, “home-like experience” was extracted from both positive and negative reviews. The representative reviews of this topic reveal that Airbnb hosts provided Airbnb users home-like experiences by serving guests with great attitudes and paying close attention to their needs. In addition, Airbnb users often express gratitude to hosts for providing thoughtful daily items (e.g., soap, shampoo) and different services (e.g., breakfast, laundry). “Check-in/out” in positive reviews refers to that Airbnb hosts are able to provide flexible check-in and out services on certain occasions, such as flight delay and changing schedules. By contrast, “checking in/out” in negative reviews is often connected with Airbnb users’ dissatisfaction that they are not able to check-in at a scheduled time and receive a flexible check-in or out services from the host.

**TABLE 4** | Topic classification.

Review type	Dimension	Tangibility	Reliability	Responsiveness	Assurance	Empathy	Convenience
Positive reviews		Amenities		Communication	Help from hosts	Home-like experience	Public transportation
		View				Check in/out	Location
		Neighborhood environment					Easy access to desired places
		Room size					
		Sleep disturbance					
		Room experiences					
Positive reviews		Property issues	Unmatched descriptions	Communication		Home-like experience	Location
		Room temperature	Booking and refund			Check in/out	Easy access to desired places
		Kitchen experiences	Hosts' unpleasant behavior				
		Noise					
		Door lock/key					
		Bathroom problems					
		Poor room maintenance					
		Dirtiness and smell					
		Bed size/condition					

Under the dimension of convenience, “location” was extracted from both positive and negative reviews, presenting Airbnb users’ general descriptions about the locations; words such as “perfect,” “excellent,” “nice” were often used. Another common topic “easy access to desired places” indicates that Airbnb users care about the distance between the property and their point-of-interest, such as restaurants, shops, public transportation, and tourist attractions. Additionally, “public transportation” emerged as an individual topic in positive reviews, suggesting that satisfied Airbnb users emphasize more on transport-related convenience.

The comparative topic analysis reveals the positive and negative performance of the same attributes, demonstrating that some attributes can be the sources of both Airbnb users’ satisfaction and dissatisfaction. Besides, three topics associated with Airbnb users’ favorable lodging experience were extracted from negative reviews, and one topic associated with Airbnb users’ complaints was extracted from positive reviews, supporting the coexistence of customer satisfaction and dissatisfaction (Chen et al., 2014).

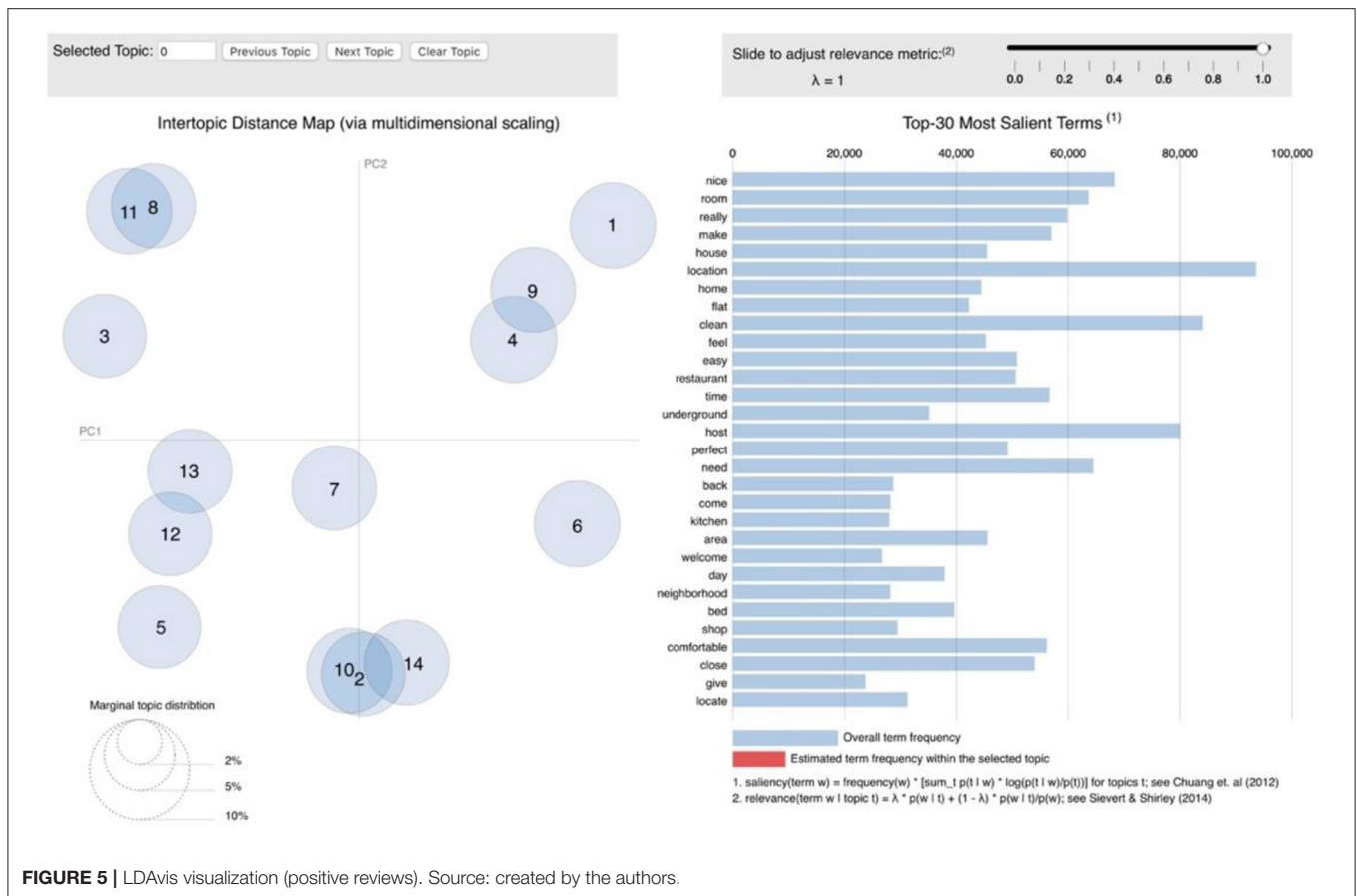
## Topic Correlation Analysis

The topic models were visualized using the R package “LDavis” (Sievert and Shirley, 2014). “LDavis” is a web-based interactive graphical interface in which researchers can observe the relationship between different topics generated by the LDA model and the distribution of top topic words by clicking

on the corresponding topic. In this study, “LDavis” is mainly used to reveal the relationship between different topics. In **Figures 5, 6**, the size of the topic circle indicates the proportional prevalence of the topic. Besides, it can be observed that the topics extracted from both positive and negative reviews shared similar proportions in the present study. The closer two topics are, the more likely they are to be discussed together in the same comment.

According to **Figure 5**, topic 4 (amenities) and topic 9 (room size) are closely connected because both topics contain many words related to room characteristics. Airbnb users often commented on topic 10 (home-like experience), topic 2 (help from hosts), and topic 14 (revisit intention) simultaneously. This topic cluster indicates that facilitating home-like experience and providing thoughtful help could induce Airbnb users’ revisit intention. Additionally, topic 11 (easy access to desired places) and topic 8 (neighborhood environment) are closely related as both topics are related to the location of the property. Besides, the connection between topic 13 (communication) and topic 12 (room experience) is mainly because the word “host” is often mentioned in both topics.

**Figure 6** indicates that topic 8 (booking and refund), topic 9 (host’s unpleasant behavior), topic 16 (communication), and topic 18 (check-in/out) formed a cluster, and all these four topics contain Airbnb users’ complaints about host regarding different issues, such as booking canceled by the host, bad attitude, and



no response from the host during the check-in period. Besides, topic 1 (property issues) and topic 2 (unmatched descriptions) are both related to the property. The connection between Topic 3 (room temperature) and topic 6 (location) could be explained by the possibility of room temperature being uncomfortable due to geographic location. Although there is no overlap between topic 12 (bathroom problems) and topic 14 (guest conflicts), they are close to each other; consequently, guest conflicts could arise in the bathroom-related issues. The slight overlapping between topic 4 (kitchen experience) and topic 13 (poor room maintenance) indicates that poor maintenance of certain areas could result in the poor cooking experience.

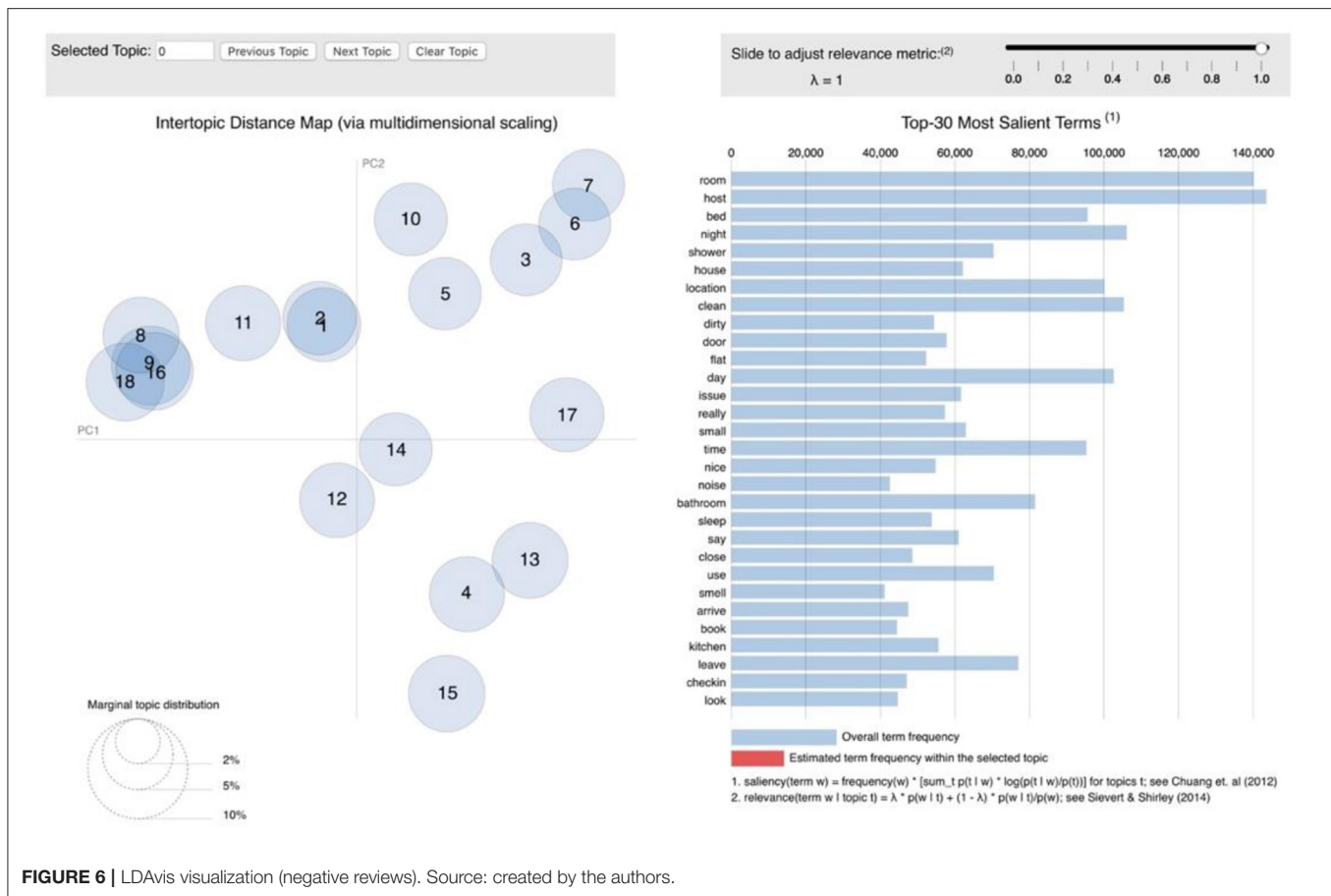
### Topic Distribution Analysis

The distributions of topics extracted from both positive and negative reviews among four types of Airbnb properties were analyzed. Besides, the ANOVA test was conducted on the topic distributions, revealing that all topics extracted from both positive and negative reviews have a significant difference ( $p < 0.05$ ) among these four types of Airbnb properties.

Figure 7 illustrates that Airbnb users who lived in an entire property are more likely to comment on location-related topics, such as easy access to desired places, neighborhood environment, and surrounding views; room size is much less discussed. It is noticeable that Airbnb users who chose a private room comment

significantly more on the home-like experience and emphasize comparatively more public transportation. Those who lived in a shared room often write about the help from the host and room size and pay less attention to the neighborhood environment. As for the hotel room, Airbnb users comment much less on home-like experience and revisit intention but write significantly more on room size.

As presented in Figure 8, Airbnb users who live in an entire property complain more about bathroom problems, the sanitary condition of the room, poor kitchen experience, poor room maintenance, property issues, and uncomfortable room temperature while they are much less likely to have conflicts with other guests. Besides, those who live in a hotel room tend to complain more about issues related to booking and refund, unpleasant check-in and out experiences; they comment much less on home-like experience. However, those who chose a private room write slightly more on home-like experience. Airbnb users from the shared room and private room write significantly more on conflicts with other guests; this may be explained that Airbnb users who stay in these two types of accommodation have more opportunities to interact with other guests, and many conflicts arise from people's different hygiene habits. Additionally, those who stay in a shared room complain more about the size and condition of the bad that cause poor sleeping experience and the host's unpleasant behavior



while they discuss much less on bathroom problems and property issues.

## sLDA Analysis

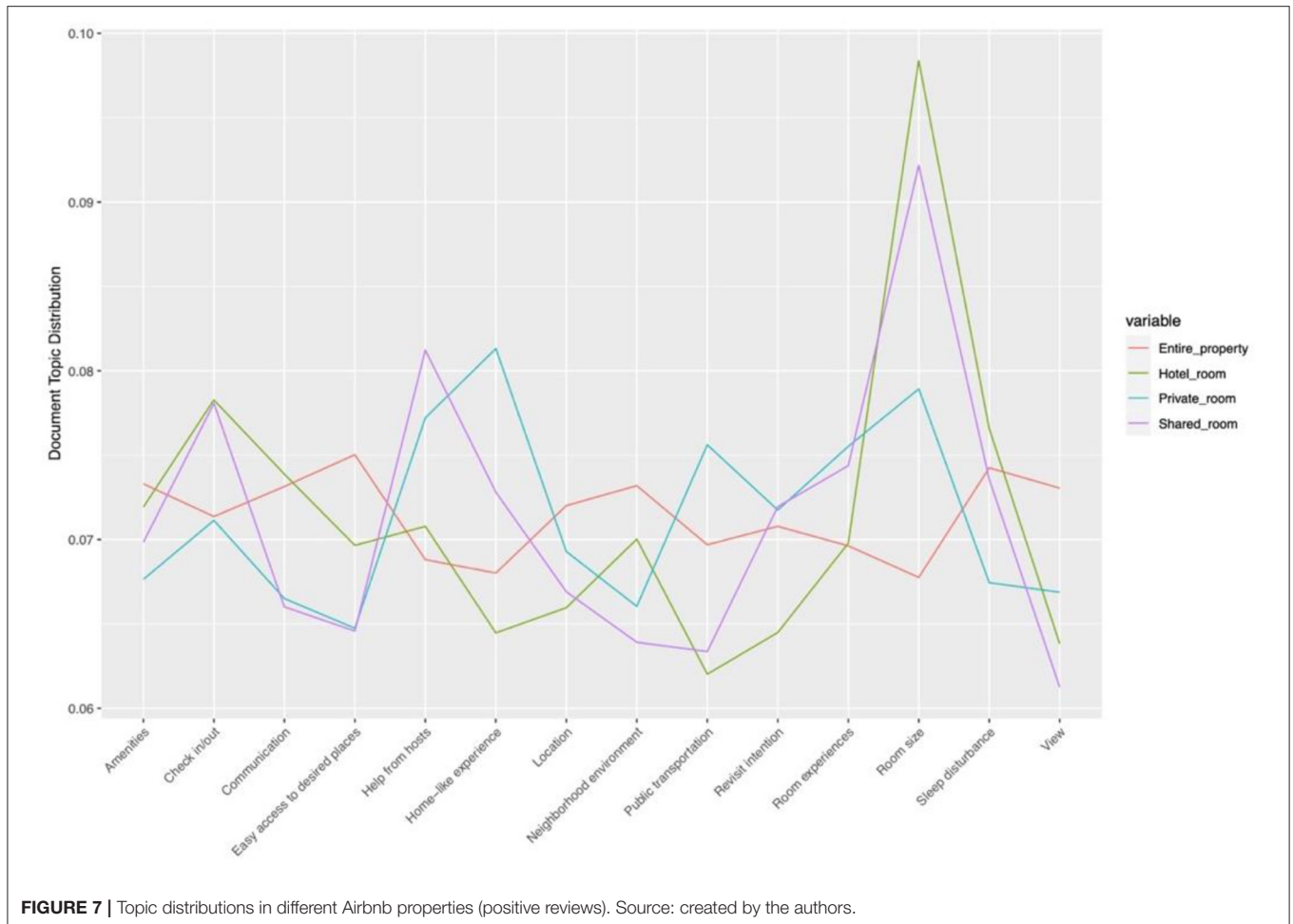
sLDA was implemented using the “lda” package. The parameter setting is:  $\alpha = 1$ ,  $\eta = 0.1$ ,  $\sigma^2 = 0.25$ . A suitable topic number was determined by conducting test-run with different topic numbers ranging from 1 to 40. Both quantitative and qualitative criteria were used to make the final decision. The quantitative criterion refers to the perplexity of the topic model. Perplexity based on probability was used to measure how well a topic model will fit a sample (Hagen, 2018); a lower perplexity score indicates better performance. The result of the perplexity metric in **Figure 9** indicates that the increase in topic number can generate a better perplexity result. However, it can be observed by examining the qualitative results that many overlapping topics in models exhibited lower perplexity score, and the logical connections of top words of some topics were not interpretable. This was because topic models with good performance in perplexity were demonstrated to be less interpretable (Chang et al., 2009). Consequently, the human judgment approach was used for the present study to evaluate the coherence and interpretability of the model. After the performance of different models was evaluated, a 20-topic solution that achieved the best semantically coherent results in the range of 1 to 40 topics was selected; it

is a common topic solution when conducting topic modeling analysis in the hospitality industry (e.g., Guo et al., 2017; Hu et al., 2019). Ten top words associated with each topic are presented in **Appendix C**. The topic labeling process is identical to LDA. Except “guest conflicts,” the remaining topics or relevant concepts that have appeared in previous LDA analyses are also reflected in sLDA results. The topic of “guest conflicts” does not appear in the sLDA results is because that the number of words related to this topic is not significant enough to emerge as an individual topic; moreover, this topic is only strongly associated with certain types of Airbnb properties; hence, it will not affect the overall evaluation.

The sLDA analysis mainly focuses on evaluating the relationship between different topics and the review polarity score that indicates the satisfaction level of Airbnb users. **Table 5** presents the estimated regression coefficient of each attribute, where a negative value indicates dissatisfaction while a positive value indicates satisfaction. Looking at **Table 5**, it is apparent that dirty environment of the accommodation (topic 12) is the most likely driver of dissatisfaction. At the same time, it can be seen that the host’s timely reply (topic 9) and warm hospitality (topic 2) are the most important drivers of satisfaction.

It can be observed that nine out of the 20 topics have a positive value of the coefficient and the other 11 topics have a negative value of the coefficient. As for statistical significance, except for





topic 5, the remaining topics are significantly related to review polarity score. Topic 5 is at the 5.97% level, suggesting that topic 5 is not statistically relevant to Airbnb users' satisfaction. Topic 9 has the highest coefficient value of 0.884, followed by topic 2 with a coefficient value of 0.767, demonstrating that the improvement of these two aspects is most likely to increase Airbnb users' satisfaction. Regarding topics with a negative value of the coefficient, topic 12 has the lowest coefficient value of  $-0.4567$ ; thus, the improvement of the sanitary condition of the room should be prioritized to prevent the Airbnb users from being dissatisfied. Another two leading causes of dissatisfaction are topic 1 and 20, with coefficient values of  $-0.318$  and  $-0.297$  respectively.

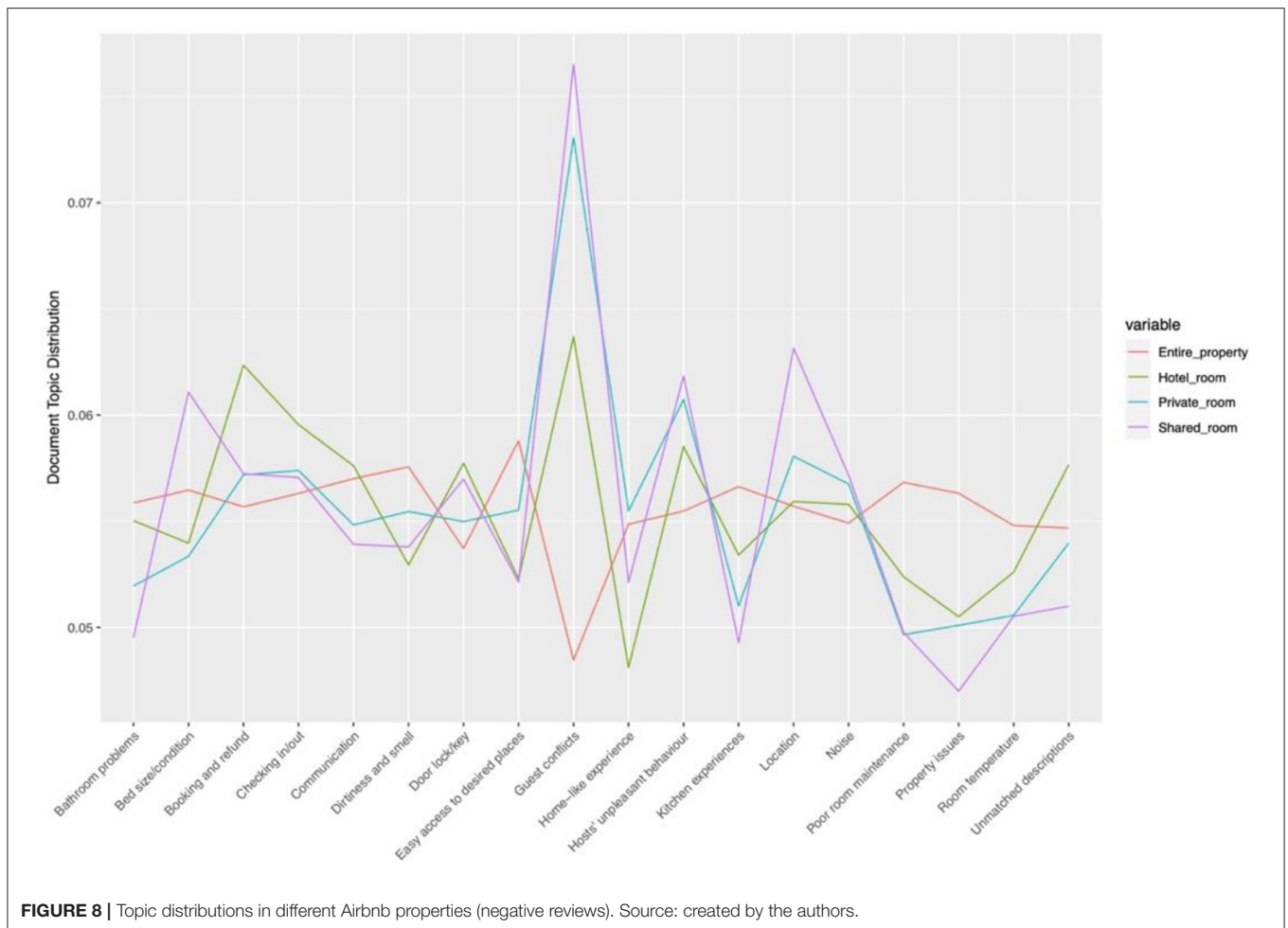
In **Figure 10**, the thickness of the topic bar represents the  $t$ -value, which indicates the statistical significance of the corresponding topic. The gradation of the topic color represents the estimated value, the lighter the color denotes a higher estimated score, and the darker color indicates a lower estimate score. Among topics with a positive coefficient, the higher the estimated coefficient of a topic on the x-axis, the higher the probability that satisfied Airbnb users talk about it. Among topics with negative coefficients, the lower the coefficient, the higher the probability that dissatisfied Airbnb users talk about it.

## CONCLUSION

### Discussion

In this study, LDA is employed to identify sources of Airbnb users' satisfaction and dissatisfaction. Different from previous Airbnb customer review studies that disregard the balance of positive and negative reviews in samples, this study used samples consisting of an equal number of positive and negative reviews. LDA was employed to extract latent topics from positive and negative reviews separately, contributing to differentiating satisfaction and dissatisfaction attributes. The results revealed 14 topics from positive reviews and 18 topics from negative reviews, supporting that the sources of dissatisfaction for customers are more diverse compared to the sources of satisfaction (Xu and Li, 2016). Except for "guest conflicts" and "room temperature," the remaining topics identified in this study have also been reported in previous Airbnb studies. However, previous studies mainly reported the positive performance of certain attributes such as "communication" and "kitchen experience," while overlooking the negative performance. Thus, this study provided a more comprehensive understanding of Airbnb service attributes.

As revealed by comparing LDA results from positive reviews and negative reviews, attributes under the dimension of



**FIGURE 8 |** Topic distributions in different Airbnb properties (negative reviews). Source: created by the authors.

tangibility are the major sources of Airbnb users' dissatisfaction. Those attributes associated with dysfunctional household equipment and poor room conditions are often discussed in negative reviews. However, the positive performance of those attributes in positive reviews was not often mentioned by Airbnb users, demonstrating that internal facilities related tangible attributes do not play a crucial role in formulating a satisfied lodging experience; this finding is consistent with Priporas et al. (2017). The attributes under the dimension of convenience are the major source of user satisfaction, consistent with Lee and Kim's (2018) finding that hedonic value is strongly associated with Airbnb users' satisfaction. However, in order to avoid Airbnb user dissatisfaction, it is important for hosts to provide an accurate description of the location of the property, because we found that complaints about unmatched descriptions are mainly related to the location. As for topics related to Airbnb hosts, both satisfied and dissatisfied Airbnb users emphasized interactions with Airbnb hosts, indicating that the behaviors of Airbnb hosts can be a driver of Airbnb user satisfaction and dissatisfaction. The relevant topics from positive reviews present the compliments to Airbnb hosts' positive behaviors, such as hosts' timely responses and assistance. Although a

positive topic related to Airbnb hosts can also be observed in negative reviews, the remaining topics are all about complaints regarding Airbnb hosts, such as bad attitudes, late responses, and irresponsible behaviors.

In this study, sLDA is applied to determine the relative importance of attributes to Airbnb users' satisfaction and dissatisfaction. The statistical results of sLDA model reveal the different strength of service attributes to Airbnb users' satisfaction and dissatisfaction. More specifically, maintaining positive communication has the strongest predictive power to Airbnb user satisfaction. This could be related to that positive interpersonal communication with guest can help to foster initial trust in hosts (Gruber, 2020), which has a positive relationship with Airbnb user satisfaction (Liang et al., 2018). In contrast, poor room conditions such as dirty or smelly environment are the main driver of Airbnb user dissatisfaction. The topic "bed size and condition" is tested to be not significantly related to Airbnb users' satisfaction.

Furthermore, this study also extends previous studies by examining topic distributions in four types of Airbnb properties, contributing to revealing Airbnb users' different emphasis when they stay in different types of Airbnb properties. The findings

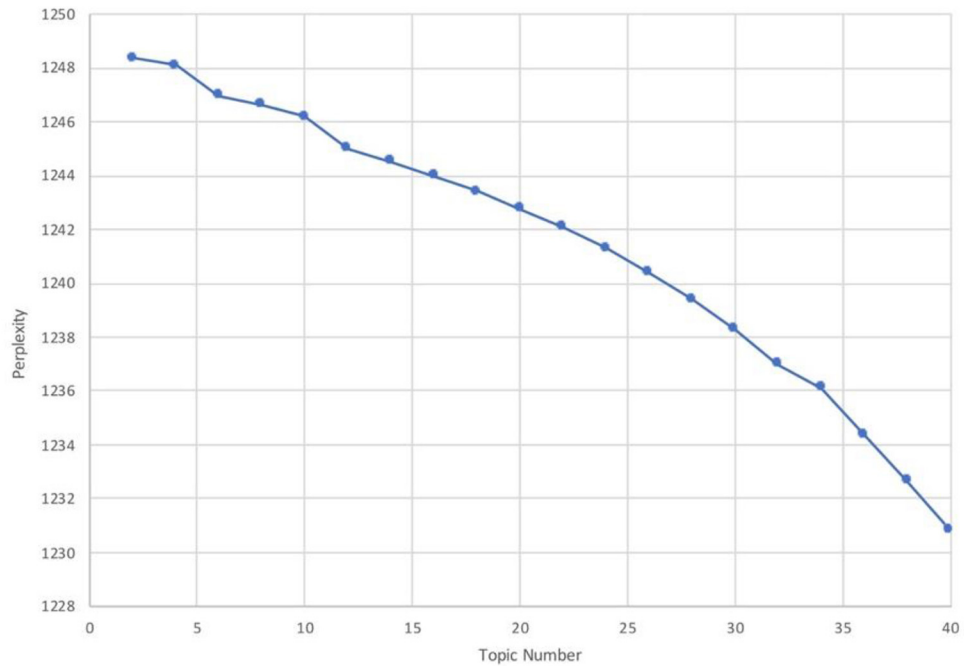


FIGURE 9 | Perplexity scores of different topic solutions. Source: created by the authors.

TABLE 5 | sLDA statistical summary.

Topic no.	Coefficient	Estimate std. error	t-value	Pr(> t )
Topic 1	-0.318495	0.007493	-42.508	< 2e-16***
Topic 2	0.767366	0.006205	123.673	< 2e-16***
Topic 3	-0.244690	0.007951	-30.774	< 2e-16***
Topic 4	-0.174047	0.008800	-19.778	< 2e-16***
Topic 5	-0.016144	0.008573	-1.883	0.0597 .
Topic 6	0.513106	0.006926	74.086	< 2e-16***
Topic 7	-0.051490	0.008101	-6.356	2.08e-10***
Topic 8	0.476621	0.008196	58.153	< 2e-16***
Topic 9	0.884127	0.006776	130.470	< 2e-16***
Topic 10	-0.083269	0.008994	-9.258	< 2e-16***
Topic 11	-0.195362	0.008295	-23.551	< 2e-16***
Topic 12	-0.456692	0.006571	-69.500	< 2e-16***
Topic 13	0.074721	0.008926	8.371	< 2e-16***
Topic 14	0.324420	0.007145	45.405	< 2e-16***
Topic 15	-0.251838	0.008823	-28.544	< 2e-16***
Topic 16	0.061481	0.008435	7.289	3.17e-13***
Topic 17	0.459367	0.006850	67.059	< 2e-16***
Topic 18	0.406608	0.006090	66.769	< 2e-16***
Topic 19	-0.033446	0.007920	-4.223	2.42e-05***
Topic 20	-0.296729	0.006919	-42.887	< 2e-16***

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1. Residual standard error: 0.1432 on 59,746 degrees of freedom, Multiple R-squared: 0.735. Adjusted R-squared: 0.7349. F-statistic: 8,285 on 20 and 59,746 DF, p-value: < 2.2e-16.

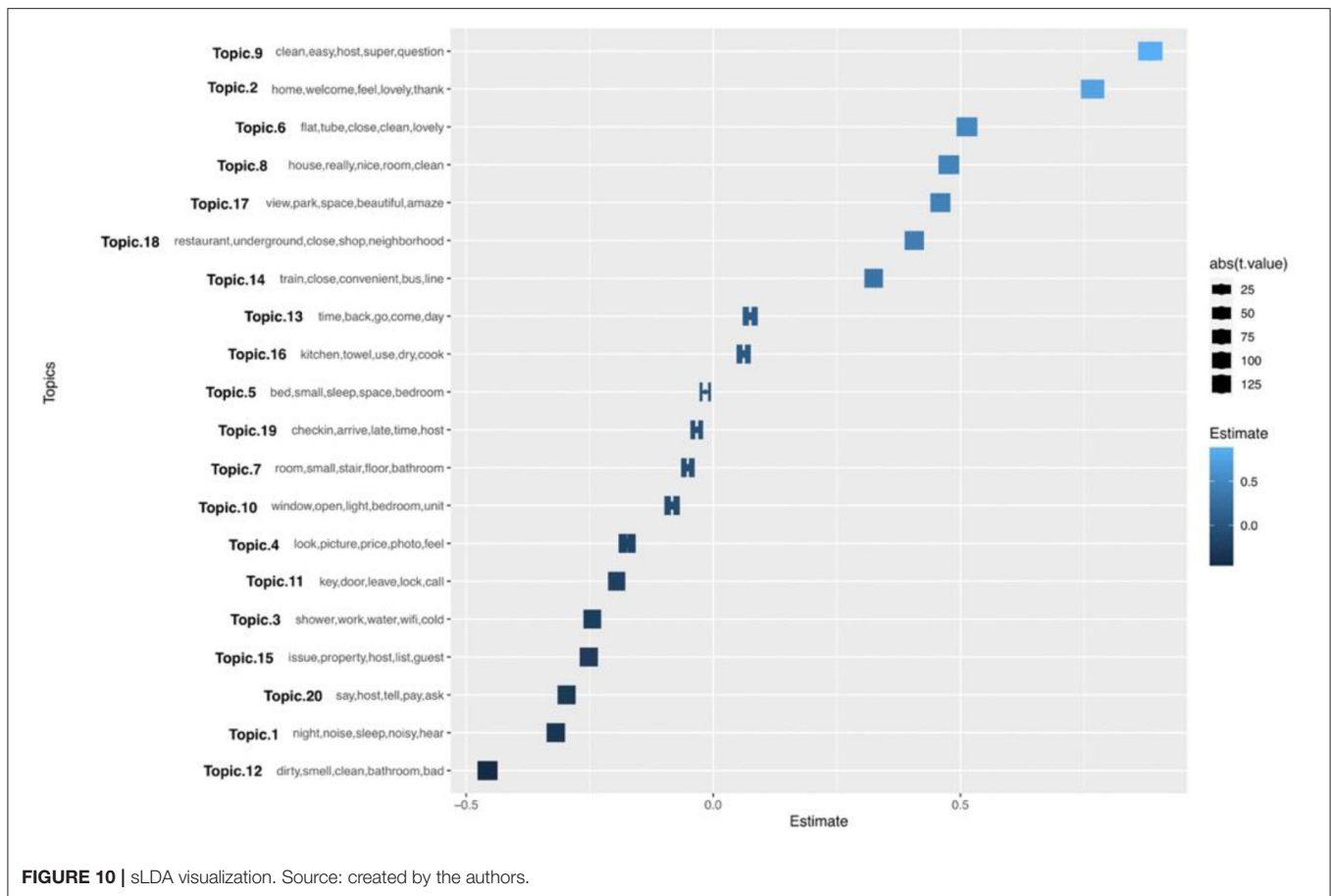


FIGURE 10 | sLDA visualization. Source: created by the authors.

show that Airbnb users who booked the entire property cared more about and convenience for exploring local communities, while those who stayed in the hotel room and shared room were more likely to encounter conflicts with other guests.

## Theoretical Implications

This study makes the following contributions to the literature. This is the first study to look separately at the attributes of customer satisfaction and dissatisfaction in the accommodation sector of the sharing economy. The findings support the hypotheses of Herzberg's two-factor theory in terms of the different nature of satisfaction and dissatisfaction attributes. Some findings of this study can provide a basis for the future development of questionnaires used for measuring Airbnb user satisfaction. For instance, two topics ("room temperature" and "guest conflicts") that had not been reported in previous studies were identified. In particular, the emergence of "guest conflict" in this study can provide a valuable implication for future studies to examine interactions between different guests in a highly shared environment. This study also supports the existence of hybrid factors that can be both satisfiers and dissatisfiers (Kano et al., 1984). This study identified service attributes that can be categorized into a hybrid category, such as "communication" and "check in/out."

This study extends the existing Airbnb studies that largely focus on the overall evaluation of Airbnb user experience (Cheng and Jin, 2019; Ju et al., 2019; Luo and Tang, 2019; Sutherland and Kiatkawsin, 2020) by comparing the emphasis level of service attributes discussed by users staying in four types of Airbnb properties. The results revealed Airbnb users' different expectations toward different types of Airbnb properties. The topic correlation analysis revealed the potential associations among different topics. In particular, the findings suggest the close connection between "home-like experience," "help from hosts" and "revisit intention," which can provide insights for future Airbnb loyalty studies.

From the perspective of methodology, LDA has been extensively applied in existing customer review studies while less attention has been paid to employing sLDA. This is the first study that employed sLDA to examine customer reviews in the lodging industry, and the results demonstrated the effectiveness of using sLDA to obtain insights from customer reviews into how different attributes affect customer satisfaction. Therefore, future studies can follow the same procedures of the present study by incorporating different outcome variables to identify topics with stronger predictive power to a specific variable.



## Practical Implications

This study has many valuable implications for Airbnb managerial practice. First, it revealed the heterogeneity of satisfaction and dissatisfaction attributes in Airbnb accommodation. The results can help Airbnb practitioners understand what aspects should be given more attention in order to increase guest satisfaction and avoid putting too much effort into the aspects that are not very valued by Airbnb users. For instance, Airbnb hosts need to emphasize on providing outstanding customer services in order to increase Airbnb user satisfaction. Besides, even though significant improvement in dissatisfiers (e.g., tangible features of the room) may not guarantee the increase of guest satisfaction, Airbnb hosts still have to maintain these aspects at an acceptable performance level to avoid guest dissatisfaction. As a successful business must not only create satisfaction, but also avoid dissatisfaction (Kim et al., 2016).

Second, the results from analyzing the topic distributions in four types of Airbnb properties suggest that Airbnb hosts renting shared and private rooms should develop measures in advance to avoid potential conflicts between different guests, as the topic “guest conflicts” appeared significantly more in these two types of Airbnb property. Third, the results of this study provide Airbnb practitioners valuable insights into priority setting when dealing with satisfaction and dissatisfaction attributes. For satisfiers, Airbnb hosts should give priority to improving communication experience of the guest, such as providing quick responses, and preparing useful pre-arrival instructions. On the other hand, hosts must ensure the cleanliness and hygiene of the room, especially the bathroom when taking measures to avoid dissatisfaction.

Forth, we also recommend Airbnb management to provide necessary training to Airbnb hosts regarding how to engage with Airbnb users in different situations and cater for their needs, considering that Airbnb hosts are regarded as the key contributor to Airbnb users’ evaluation. Different from traditional hotel staff, they are generally trained to provide standardized service to guests. Some Airbnb hosts may not have been engaged in the service industry, resulting in their lack of service awareness and making them unable to facilitate a satisfactory lodging experience to Airbnb users. Last, the identified attributes that reveal customers’ expectations can be used as useful references for Airbnb and hosts to create social media content that is of value to customers, which can contribute to the building of trust and confidence in the platform (Martínez-Navalón et al., 2019).

## REFERENCES

- Ady, M., and Quadri-Felitti, D. (2015). *Consumer Research Identifies How to Present Travel Review Content for More Bookings*. Available online at: <http://www.trustyou.com/press/study-reveals-travelers-preferred-summarized-review-content-full-text-reviews> (accessed September 14, 2020).
- Ahmad, S. N., and Laroche, M. (2015). How do expressed emotions affect the helpfulness of a product review? Evidence from reviews using latent semantic analysis. *Int. J. Electron. Commer.* 20, 76–111. doi: 10.1080/10864415.2016.1061471

## Limitations

Although using the lexicon-based approach to conduct sentiment analysis can achieve sufficiently accurate results for academic research (Pang and Lee, 2008), the sentiment lexicon used in this study is not exclusively designed for the lodging industry, resulting in inaccurate sentiment scoring to certain opinion words. Hence, it is encouraged for future studies to use the sentiment lexicon with context emphasizing the lodging industry to achieve better results. Besides, this study only examined how Airbnb users’ perceptions to different service attributes vary in different types of Airbnb property, making it insufficient to explain Airbnb users’ behaviors. Future studies can examine Airbnb users’ perceptions by looking at some other perspectives, such as Airbnb users’ cultural backgrounds, the purpose of stay, and price ranges of the property. Furthermore, the inclusion of data from 12 different countries for this study can improve the generalizability of the results. However, some drivers of Airbnb users’ satisfaction or dissatisfaction in certain countries could be overlooked. Consequently, future studies could examine Airbnb reviews in a region or country that has not been explored to identify some context-specific service attributes.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: <http://insideairbnb.com/get-the-data.html>.

## AUTHOR CONTRIBUTIONS

KD: conceptualization, methodology, original draft preparation, software, and revision writing. WC: validation, reviewing, and editing. KN: software, data curation, and text mining. SN: investigation, reviewing, and editing. PS: revision writing, visualization, and editing. All authors have read and agreed to the submitted version of the manuscript.

## ACKNOWLEDGMENTS

We thank SpeedProp Global Sdn Bhd for its scholarship support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.659481/full#supplementary-material>

- Airbnb (2017). *About us*. Available online at: <https://www.airbnb.com.au/about/about-us> (accessed March 1, 2021).
- Akbaba, A. (2006). Measuring service quality in the hotel industry: a study in a business hotel in Turkey. *Int. J. Hosp. Manag.* 25, 170–192. doi: 10.1016/j.ijhm.2005.08.006
- Arun, R., Suresh, V., Madhavan, C. V., and Murthy, M. N. (2010). On finding the natural number of topics with latent dirichlet allocation: some observations. *Springer Berlin Heidelberg* 6118, 391–402. doi: 10.1007/978-3-642-13657-3\_43
- Berezina, K., Bilgihan, A., Cobanoglu, C., and Okumus, F. (2015). Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews. *J. Hosp. Mark. Manag.* 25, 1–24. doi: 10.1080/19368623.2015.983631

- Bi, J. W., Liu, Y., Fan, Z. P., and Cambria, E. (2019). Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model. *Int. J. Prod. Res.* 57, 7068–7088. doi: 10.1080/00207543.2019.1574989
- Blei, D. (2012). Probabilistic topic models. *Commun. ACM* 55:77. doi: 10.1145/2133806.2133826
- Blei, D.M., and McAuliffe, J. D. (2007). “Supervised topic models,” in: *Advances in Neural Information Processing Systems 20 - Proceedings of the 2007 Conference* (Vancouver, BC: Curran Associates Inc).
- Blei, D. M., and Lafferty, J. D. (2006). “Dynamic topic models,” in *Proceedings of the 23rd International Conference on Machine Learning* (New York, NY: ACM). doi: 10.1145/1143844.1143859
- Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.* 3, 993–1022. Available online at: <https://www.scopus.com/record/display.uri?eid=2-s2.0-0141607824&origin=inward>
- Brochado, A., Troilo, M., and Aditya, S. (2017). Airbnb customer experience: evidence of convergence across three countries. *Ann. Tour. Res.* 63, 210–221. doi: 10.1016/j.annals.2017.01.001
- Cadotte, E. R., and Turgeon, N. (1988). Dissatisfiers and satisfiers: Suggestions for consumer com-plaints and compliments. *J. Consumer Satisfact. Dissatisfact. Complain. Behav.* 1, 74–79.
- Cao, J., Xia, T., Li, J., Zhang, Y., and Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing* 72, 1775–1781. doi: 10.1016/j.neucom.2008.06.011
- Casaló, L. V., Flavián, C., Guinalú, M., and Ekinci, Y. (2015). Avoiding the dark side of positive online consumer reviews: enhancing reviews’ usefulness for high risk-averse travelers. *J. Bus. Res.* 68, 1829–1835. doi: 10.1016/j.jbusres.2015.01.010
- Chai, C. (2019). Text mining in survey data. *Survey Pract.* 12, 1–14. doi: 10.29115/SP-2018-0035
- Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., and Blei, D. M. (2009). Reading tea leaves: how humans interpret topic models. *Adv. Neural Inf. Process Syst.* 288–296. Available online at: <https://www.scopus.com/record/display.uri?eid=2-s2.0-84863381525&origin=inward#>
- Chen, A., Lu, Y., Gupta, S., and Qi, X. (2014). Can customer satisfaction and dissatisfaction coexist? An issue of telecommunication service in China. *J. Inf. Technol.* 29, 237–252. doi: 10.1057/jit.2013.26
- Chen, M. J. (2020). Examining the influence of emotional expressions in online consumer reviews on perceived helpfulness. *Inf. Process Manag.* 57:102266. doi: 10.1016/j.ipm.2020.102266
- Cheng, M., and Jin, X. (2019). What do Airbnb users care about? An analysis of online review comments. *Int. J. Hosp. Manag.* 76, 58–70. doi: 10.1016/j.ijhm.2018.04.004
- Choi, J., Yoon, J., Chung, J., Coh, B. Y., and Lee, J. M. (2020). Social media analytics and business intelligence research: a systematic review. *Inf. Process. Manag.* 57:102279. doi: 10.1016/j.ipm.2020.102279
- Chowdhary, N., and Prakash, M. (2005). Service quality: Revisiting the two factors theory. *J. Serv. Res.* 5, 61–75. Available online at: <https://www.scopus.com/record/display.uri?eid=2-s2.0-34548030171&origin=inward#>
- Deveaud, R., SanJuan, E., and Bellot, P. (2014). Accurate and effective latent concept modelling for *Ad Hoc* information retrieval. *Doc. Num.* 17, 61–84. doi: 10.3166/dn.17.1.61-84
- Ding, K., Choo, W. C., Ng, K. Y., and Ng, S. I. (2020). Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation. *Int. J. Hosp. Manag.* 91:102676. doi: 10.1016/j.ijhm.2020.102676
- Felbermayr, A., and Nanopoulos, A. (2016). The role of emotions for the perceived usefulness in online customer reviews. *J. Interact. Mark.* 36, 60–76. doi: 10.1016/j.intmar.2016.05.004
- Flaherty, P., Giaeffer, G., Kumm, J., Jordan, M. I., and Arkin, A. P. (2005). A latent variable model for chemogenomic profiling. *Bioinformatics.* 21, 3286–3293. doi: 10.1093/bioinformatics/bti515
- Geetha, M., Singha, P., and Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels - an empirical analysis. *Tour. Manag.* 61, 43–54. doi: 10.1016/j.tourman.2016.12.022
- Gerdt, S. O., Wagner, E., and Schewe, G. (2019). The relationship between sustainability and customer satisfaction in hospitality: an explorative investigation using eWOM as a data source. *Tour. Manag.* 74, 155–172. doi: 10.1016/j.tourman.2019.02.010
- Gossen, M., and Scholl, G. (2016). The sharing economy. *Ä-kologisches Wirtschaften-Fachzeitschrift* 31:41. doi: 10.14512/OEW310141
- Griffiths, T. L., and Steyvers, M. (2004). Finding scientific topics. *Proc. Natl. Acad. Sci. U.S.A.* 101, 5228–5235. doi: 10.1073/pnas.0307752101
- Gruber, S. (2020). Personal trust and system trust in the sharing economy: a comparison of community- and platform-based models. *Front. Psychol.* 11:e581299. doi: 10.3389/fpsyg.2020.581299
- Gu, H., and Ryan, C. (2008). Chinese clientele at Chinese hotels: preferences and satisfactions. *Int. J. Hosp. Manag.* 27, 337–345. doi: 10.1016/j.ijhm.2007.10.005
- Guo, Y., Barnes, S. J., and Jia, Q. (2017). Mining meaning from online ratings and reviews: tourist satisfaction analysis using latent Dirichlet allocation. *Tour. Manag.* 59, 467–483. doi: 10.1016/j.tourman.2016.09.009
- Guttentag, D., Smith, S., Potwarka, L., and Havitz, M. (2018). Why tourists choose airbnb: a motivation-based segmentation study. *J. Travel Res.* 57, 342–359. doi: 10.1177/0047287517696980
- Ha, T., Beijnon, B., Kim, S., Lee, S., and Kim, J. H. (2017). Examining user perceptions of smartwatch through dynamic topic modeling. *Telemat. Inform.* 34, 1262–1273. doi: 10.1016/j.tele.2017.05.011
- Hagen, L. (2018). Content analysis of e-petitions with topic modeling: how to train and evaluate LDA models? *Inf. Process Manag.* 54, 1292–1307. doi: 10.1016/j.ipm.2018.05.006
- Herzberg, F. (1966). *Work and the Nature of Man*. Cleveland, OH: World Publishing.
- Horn, K., and Merante, M. (2017). Is home sharing driving up rents? Evidence from Airbnb in Boston. *J. Hous Econ.* 38, 14–24. doi: 10.1016/j.jhe.2017.08.002
- Hu, N., Zhang, T., Gao, B., and Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tour. Manag.* 72, 417–426. doi: 10.1016/j.tourman.2019.01.002
- Ibrahim, N. F., and Wang, X. (2019). A text analytics approach for online retailing service improvement: evidence from twitter. *Decis. Support Syst.* 121, 37–50. doi: 10.1016/j.dss.2019.03.002
- Ju, Y., Back, K., Choi, Y., and Lee, J. (2019). Exploring Airbnb service quality attributes and their asymmetric effects on customer satisfaction. *Int. J. Hosp. Manag.* 77, 342–352. doi: 10.1016/j.ijhm.2018.07.014
- Kano, N., Seraku, N., Takahashi, F., and Tsuji, S. (1984). Attractive quality and must-be quality. *Hinshitsu: Quality. J. Japan. Soc. Q. Control.* 14, 39–48.
- Kim, B., Kim, S., and Heo, C. Y. (2016). Analysis of satisfiers and dissatisfiers in online hotel reviews on social media. *Int. J. Contemp. Hosp. Manag.* 28, 1915–1936. doi: 10.1108/IJCHM-04-2015-0177
- Kim, S. G., and Kang, J. (2018). Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews. *Inf. Process Manag.* 54, 938–957. doi: 10.1016/j.ipm.2018.06.003
- Knutson, B., Stevens, P., Wullaert, C., Patton, M., and Yokoyama, F. (1990). LODGSERV: a service quality index for the lodging industry. *J. Hosp. Tour. Res.* 14, 227–284. doi: 10.1177/109634809001400230
- Korfiatis, N., Stamolampros, P., Kourouthanassis, P., and Sagiadinos, V. (2019). Measuring service quality from unstructured data: a topic modeling application on airline passengers’ online reviews. *Expert Syst. Appl.* 116, 472–486. doi: 10.1016/j.eswa.2018.09.037
- Kumar, S., and Zymbler, M. (2019). A machine learning approach to analyze customer satisfaction from airline tweets. *J. Big Data.* 6:62. doi: 10.1186/s40537-019-0224-1
- Lee, K., and Yu, C. (2018). Assessment of airport service quality: a complementary approach to measure perceived service quality based on Google reviews. *J. Air Transp.* 71, 28–44. doi: 10.1016/j.jairtraman.2018.05.004
- Lee, S., and Kim, D. Y. (2018). The effect of hedonic and utilitarian values on satisfaction and loyalty of Airbnb users’. *Int. J. Contemp. Hosp. Manag.* 30, 1332–1351. doi: 10.1108/IJCHM-09-2016-0504
- Liang, L. J., Choi, H. C., and Joppe, M. (2018). Exploring the relationship between satisfaction, trust and switching intention, repurchase intention in the context of Airbnb. *Int. J. Hosp. Manag.* 69, 41–48. doi: 10.1016/j.ijhm.2017.10.015
- Litvin, S., Goldsmith, R., and Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tour. Manag.* 29, 458–468. doi: 10.1016/j.tourman.2007.05.011

- Liu, S. Q., and Mattila, A. S. (2017). Airbnb: online targeted advertising, sense of power, and consumer decisions. *Int. J. Hosp. Manag.* 60, 33–41. doi: 10.1016/j.ijhm.2016.09.012
- Lucini, F. R., Tonetto, L. M., Fogliatto, F. S., and Anzanello, M. J. (2020). Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. *J. Air Transp.* 83:101760. doi: 10.1016/j.jairtraman.2019.101760
- Luo, Y., and Tang, R. (2019). Understanding hidden dimensions in textual reviews on Airbnb: An application of modified latent aspect rating analysis (LARA). *Int. J. Hosp. Manag.* 80, 144–154. doi: 10.1016/j.ijhm.2019.02.008
- Martínez-Navalón, J. G., Gelashvili, V., and Debasa, F. (2019). The impact of restaurant social media on environmental sustainability: An empirical study. *Sustainability* 11, 1–24. doi: 10.3390/su11216105
- Möhlmann, M. (2015). Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again. *J. Consum. Behav.* 14, 193–207. doi: 10.1002/cb.1512
- Murzintcev, N. (2015). *Package Ldatuning: Tuning of the LDA Models Parameters*. Available online at: <http://cran.r-project.org/pub/R/web/packages/ldatuning/ldatuning.pdf> (accessed September, 11, 2020).
- Nasukawa, T., and Yi, J. (2003). “Sentiment analysis: capturing favorability using natural language processing,” in *Proceedings of the 2nd International Conference on Knowledge Capture* (Sanibel Island: ACM). doi: 10.1145/945645.945658
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction. *Decisions. J. Mark.* 17, 460–469. doi: 10.1177/002224378001700405
- Padma, P., and Ahn, J. (2020). Guest satisfaction and dissatisfaction in luxury hotels: an application of big data. *Int. J. Hosp. Manag.* 84:102318. doi: 10.1016/j.ijhm.2019.102318
- Pang, B., and Lee, L. (2008). Opinion mining and sentiment analysis. *Found. Trends Inf. Ret.* 2, 1–135. doi: 10.1561/9781601981516
- Park, S., Lee, J., and Nicolau, J. L. (2020). Understanding the dynamics of the quality of airline service attributes : satisfiers and dissatisfiers. *Tour. Manag.* 81:104163. doi: 10.1016/j.tourman.2020.104163
- Phillips, P., Zigan, K., Santos Silva, M., and Schegg, R. (2015). The interactive effects of online reviews on the determinants of Swiss hotel performance: a neural network analysis. *Tour. Manag.* 50, 130–141. doi: 10.1016/j.tourman.2015.01.028
- Podnar, K., and Javernik, P. (2012). The effect of word of mouth on consumers’ attitudes toward products and their purchase probability. *J. Promot. Manag.* 18, 145–168. doi: 10.1080/10496491.2012.668426
- Priporas, C., Stylos, N., Vedanthachari, L., and Santiwatana, P. (2017). Service quality, satisfaction, and customer loyalty in Airbnb accommodation in Thailand. *Int. J. Tour. Res.* 19, 693–704. doi: 10.1002/jtr.2141
- Rinker, T. (2019). *Sentimentr. R package version 2.7.1*. Available online at: <https://CRAN.R-project.org/package=sentiment> (accessed September 21, 2020).
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., et al. (2014). Structural topic models for open-ended survey responses. *Am. J. Polit. Sci.* 58, 1064–1082. doi: 10.1111/ajps.12103
- Roque, C., Lourenço Cardoso, J., Connell, T., Schermers, G., and Weber, R. (2019). Topic analysis of road safety inspections using latent dirichlet allocation: a case study of roadside safety in Irish main roads. *Anal. Prev.* 131, 336–349. doi: 10.1016/j.aap.2019.07.021
- Sanchez-Franco, M. J., Cepeda-Carrion, G., and Roldán, J. L. (2019). Understanding relationship quality in hospitality services. *Internet Res.* 29, 478–503. doi: 10.1108/IntR-12-2017-0531
- Saura, J. R. (2020). Using data sciences in digital marketing: framework, methods, and performance metrics. *J. Innov. Know.* 11, 92–102. doi: 10.1016/j.jik.2020.08.001
- Shengli, L., and Fan, L. (2019). The interaction effects of online reviews and free samples on consumers’ downloads: an empirical analysis. *Inf. Process Manag.* 56:102071. doi: 10.1016/j.ipm.2019.102071
- Short, J. C., Broberg, J. C., Coglisier, C. C., and Brigham, K. H. (2010). Construct validation using computer-aided text analysis (CATA): An illustration using entrepreneurial orientation. *Organ. Res. Methods* 13, 320–347. doi: 10.1177/1094428109335949
- Sievert, C., and Shirley, K. (2014). “Ldavis: a method for visualizing and interpreting topics,” in *Workshop on Interactive Language Learning, Visualization, and Interfaces* (Baltimore, MD). doi: 10.3115/v1/W14-3110
- Staphit, E., Chiappa, G. D., Coudounaris, D. N., and Bjork, P. (2019). Determinants of the continuance intention of Airbnb users: consumption values, co-creation, information overload and satisfaction. *Tour. Rev.* 75, 511–531. doi: 10.1108/TR-03-2019-0111
- Sutherland, I., and Kiatkawsin, K. (2020). Determinants of guest experience in airbnb: a topic modeling approach using LDA. *Sustainability* 12:3402. doi: 10.3390/su12083402
- Tontini, G., Bento, G., dos, S., Milbratz, T. C., Volles, B. K., and Ferrari, D. (2017). Exploring the nonlinear impact of critical incidents on customers’ general evaluation of hospitality services. *Int. J. Hosp. Manag.* 66, 106–116. doi: 10.1016/j.ijhm.2017.07.011
- Tussyadiah, I. P. (2016). Factors of satisfaction and intention to use peer-to-peer accommodation. *Int. J. Hosp. Manag.* 55, 70–80. doi: 10.1016/j.ijhm.2016.03.005
- Tussyadiah, I. P., and Pesonen, J. (2016). Impacts of peer-to-peer accommodation use on travel patterns. *J. Travel Res.* 55, 1022–1040. doi: 10.1177/0047287515608505
- Tussyadiah, I. P., and Zach, F. (2017). Identifying salient attributes of peer-to-peer accommodation experience. *J. Travel Tour. Mark.* 30, 636–652. doi: 10.1080/10548408.2016.1209153
- Tussyadiah, I. P., and Zach, F. J. (2015). Hotels vs. Peer-to-Peer accommodation rentals: text analytics of consumer reviews in Portland, Oregon. *SSRN Electron. J.* doi: 10.2139/ssrn.2594985
- Wallenstein, J., and Shelat, U. (2017). *Hopping Aboard the Sharing Economy*. Available online at: <http://www.bcg.com/publications/2017/strategy-accelerating-growth-consumer-products-hopping-aboard-sharing-economy.aspx> (accessed August 20, 2020).
- Wang, Y., Lu, X., and Tan, Y. (2018). Impact of product attributes on customer satisfaction: an analysis of online reviews for washing machines. *Electron. Commer. Res. Appl.* 29, 1–11. doi: 10.1016/j.elerap.2018.03.003
- Williams, T., and Betak, J. (2019). A comparison of LSA and LDA for the analysis of railroad accident text. *Proc. Comp. Sci.* 11, 11–15. doi: 10.5383/JUSPN.11.01.002
- Xu, X. (2020a). Examining an asymmetric effect between online customer reviews emphasis and overall satisfaction determinants. *J. Bus. Res.* 106, 196–210. doi: 10.1016/j.jbusres.2018.07.022
- Xu, X. (2020b). How do consumers in the sharing economy value satisfaction? Evidence from online reviews. *Decis. Support Syst.* 128:113162. doi: 10.1016/j.dss.2019.113162
- Xu, X., and Li, Y. (2016). The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: a text mining approach. *Int. J. Hosp. Manag.* 55, 57–69. doi: 10.1016/j.ijhm.2016.03.003
- Xu, X., Wang, X., Li, Y., and Haghghi, M. (2017). Business intelligence in online customer textual reviews: understanding consumer perceptions and influential factors. *Int. J. Inf. Manag.* 37, 673–683. doi: 10.1016/j.ijinfomgt.2017.06.004
- Yang, C. C., Jou, Y. T., and Cheng, L. Y. (2011). Using integrated quality assessment for hotel service quality. *Qual. Quant.* 45, 349–364. doi: 10.1007/s11135-009-9301-4
- Yi, J., Yuan, G., and Yoo, C. (2019). The effect of the perceived risk on the adoption of the sharing economy in the tourism industry: the case of Airbnb. *Inf. Process Manag.* 57:102108. doi: 10.1016/j.ipm.2019.102108
- Zhu, L., Lin, Y., and Cheng, M. (2020). Sentiment and guest satisfaction with peer-to-peer accommodation: when are online ratings more trustworthy? *Int. J. Hosp. Manag.* 86:102369. doi: 10.1016/j.ijhm.2019.102369

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Ding, Choo, Ng, Ng and Song. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.