Examining Airbnb guest satisfaction tendencies: a text mining approach

Mariana Cavique (correspondent author), msbcs@iscte-iul.pt, Iscte – Instituto Universitário de Lisboa, and Universidade Europeia, Portugal. ORCID: 0000-0001-8489-8644

Mariana Cavique is a lecturer of the Tourism course at Universidade Europeia, Portugal, and PhD student in Information Science and Technology at Iscte, Portugal. She holds a master's degree in Electrotechnical Engineering from Instituto Superior Técnico, Universidade de Lisboa, Portugal, and her research interests are in text mining applied to tourism.

Ricardo Ribeiro, **ricardo.ribeiro@iscte-iul.pt**, Iscte – Instituto Universitário de Lisboa, Portugal & INESC-ID Lisboa, Portugal. ORCID: 0000-0002-2058-693X

Ricardo Ribeiro holds a PhD (2011) in Information Systems and Computer Engineering from Instituto Superior Técnico, Universidade de Lisboa, Portugal. He is an Associate Professor at the Department of Information Science and Technology of Iscte and a researcher at INESC-ID, Portugal, working in speech and language processing. His current research interests focus on high-level information extraction from unrestricted text or speech and improving machine-learning techniques using domain-related information. He is the editor of a book on the computational processing of Portuguese.

Fernando Batista, **fernando.batista@iscte-iul.pt**, Iscte – Instituto Universitário de Lisboa, Portugal & INESC-ID Lisboa, Portugal. ORCID: 0000-0002-1075-0177

Fernando Batista received his PhD (2011) in Computer Science and Engineering from Instituto Superior Técnico, Universidade de Lisboa, Portugal. He is currently an Associate Professor at Iscte – Instituto Universitário de Lisboa, and researcher at INESC-ID Lisboa, Portugal. His current research focuses on written and spoken language processing, Machine Learning, and Text Mining for social media. He is member of the organization team of LxMLS - the Lisbon Machine Learning Summer School (from 2016 to 2022), Senior Member of the IEEE, and member of the ISCA Speech.

Antónia Correia, acorreia@ualg.pt, CEFAGE, University of Algarve, Portugal. ORCID: 0000-0002-6707-8289

She holds a Ph.D. from University of Algarve in tourism demand (Portugal) and an aggregation from University of Algarve; all relating to tourism and hospitality management. Her research interests are in consumer behavior and tourism economics. She is part of the editorial board of several journals in tourism, she has publications in top tier tourism journals.

Acknowledgements

This work was supported by national funds through FCT, Fundação para a Ciência e a Tecnologia, under projects UIDB/50021/2020 and UID/ECO/04007/2021.

Declarations of interest: none

Abstract:

Given Airbnb's changes since its inception and the dynamism of customer preferences, a study that sheds light on how customer satisfaction is evolving is relevant. An automated method is proposed for identifying these satisfaction tendencies at a large scale. This study follows a text mining approach to analyse 590,070 reviews posted between 2010 and 2019 on the Airbnb platform in Lisbon. Topic Modelling is employed in order to identify the main topics discussed in the reviews, and Sentiment Analysis to understand the topics that compose guest's satisfaction in the context of Airbnb services. Three major topics are extracted from Airbnb reviews: 'host's service', 'physical aspects', and 'location'. Although a positivity bias in guest reviews is confirmed, the satisfaction level seems to be decreasing over the years. The results also reveal that 'physical aspects' is the predominant topic when considering the negative guest reviews. This research considers big data the base to create knowledge, data spanning over the years, offering consistency to the research.

Keywords: Airbnb, online reviews, Topic Modelling, Sentiment Analysis, satisfaction, hospitality

1. Introduction

The rapid growth of Web 2.0 offered the means for an exponential increase of online user-generated content in the form of text, video, audio, and image. Among all the content generated by users, online travel reviews have wildly proliferated (Baka, 2016) and have been recognized as a rich and credible data source for tourism studies. This type of data influences consumer purchasing decisions and provides information about tourist experience and satisfaction (H. Li et al., 2013; Lu & Stepchenkova, 2012; Zhou et al., 2014).

In the context of peer-to-peer (P2P) accommodation, online reviews play an even more important role than for conventional hotels because online user-generated content is often the dominant or sole communication channel for customers. Most of the hosts of P2P accommodations are micro-entrepreneurs that do not announce themselves on other promotional channels (e.g., TV, radio, outdoors) as do conventional hotels.

In light of the perception of online reviews value, deriving information about the guest experience and satisfaction from guests reviews have attracted increasing research attention (Belarmino et al., 2019; Celata et al., 2020; Cheng & Jin, 2019; Ding et al., 2020; Lawani et al., 2019; Luo & Tang, 2019; Serrano et al., 2020; Tussyadiah & Zach, 2017; Zhang, 2019b, 2019a; Zhu et al., 2020), as an alternative to traditional methods such as questionnaire surveys (J. Li et al., 2019; Priporas et al., 2017) and interviews (Sthapit & Jiménez-Barreto, 2018).

A significant advantage of using online reviews is that researchers can monitor customers' changing perceptions of experience by analysing online reviews with timestamp data. Nonetheless, the literature examining the time effect on P2P accommodation reviews (Ding et al., 2020) is still scarce. This study addresses this gap, exploring online reviews in Airbnb, the leading provider in the P2P accommodation segment.

Responding to Zhang's (2019b) call to examine Airbnb reviews in different countries and the call for time-based research (Correia & Kozak, 2021), this study analyses online reviews over ten years in Lisbon, the capital of Portugal. Although rarely studied in the literature, this country has experienced in the last decades a strong tourism growth in general, evolving to an important tourist destination in Europe and in the world. This rapid and sustained growth extends to the local accommodation, mostly offered by Airbnb (NOVASCHOOLOFBUSINESS, 2016). This research analyses the guest reviews in Lisbon, which besides being the capital of Portugal, is also the city with the most significant number of Airbnb properties in the country (NOVASCHOOLOFBUSINESS, 2016).

This study aims to evaluate guest satisfaction by discovering the aspects and sentiments present in text reviews and its distribution through time. The specific objectives are 1) to identify the dominant topics and sentiments present in text reviews; 2) to evaluate the evolution of the topics and sentiments in a ten years' time span; 3) to identify the guest sentiment towards each identified topic. This study employs Topic Modelling for discovering the topics that best characterize the information in the reviews and Sentiment Analysis to determine whether the reviews are positives, negatives, or neutrals, being this a proxy of satisfaction.

The contributions of this research to the body of knowledge are two-fold. First, we combine two text mining methods and apply these to a scope of location not considered in other studies, giving significant insights to improve Airbnb service. Second, we investigate customer satisfaction over a ten-year period, offering consistency to the results and allowing the identification of trends.

The structure of this paper is as follows. Section 2 reviews the existing literature on the subject. Section 3 explains the dataset used, while Section 4 presents the methodology applied in this study. In Section 5, experimental results and analysis are presented. Finally, Section 6 discusses the conclusions, implications, and limitations of this work.

2. Literature review

Customer satisfaction is one of the main topics in hospitality research since it directly influences return intention and recommendation and, consequently, business success (Flint et al., 2011; Wilkins et al., 2010). Although there is not a consensual definition of customer satisfaction, many of the suggested definitions conceptualize it as the emotional response to the evaluation of the perceived difference between expectations and final result of the service (Millán & Esteban, 2004). Many attempts have been made to measure and understand it, mainly because identifying the service factors that determine tourist satisfaction can assist in decision making processes for service quality improvement. Previous studies have employed surveys and interviews, such as Mohsin and Lockyer (2010) that assess the service quality perception of customers of luxury hotels, or Priporas et al. (2017) that investigate customers' perceptions of the service quality facets of Airbnb accommodation.

Presently, one of the most prevalent forms used by tourists to express satisfaction with a service is the provision of online reviews/ratings. This data source has been considered a reliable source of information by customers (Litvin et al., 2008) since it reflects real user experiences and showed an impact on purchasing intention. The experimental study of Mauri and Minazzi (2013) shows a positive correlation between hotel purchasing intention and expectations of the customers and the valence of the review. The value of online reviews to better understand the hospitality market is also corroborated by Kim and Park (2017), showing that online reviews rating is a more significant predictor than traditional feedback for explaining hotel performance metrics.

Due to the recognized importance of online reviews, an increasing number of studies have been using them as the primary data to investigate the major Airbnb service attributes sought by guests. The advances in text mining approaches have been a major contribution to the analysis of this type of unstructured data. Studies exploring key content and themes from online reviews applied different methods such as Topic Modelling (Celata et al., 2020; Ding et al., 2020; Zhang, 2019a), agglomerative hierarchical cluster analysis with word co-occurrence networks (Tussyadiah & Zach, 2017), or Leximancer concept-mapping algorithm (see Smith and Humphreys (2006) for a detailed explanation on Leximancer's algorithm) (Cheng & Jin, 2019). Although the order of importance of each Airbnb attribute varies between studies, analysis based on guest reviews tends to focus on a similar collection of themes relating to the relationship with the host (Celata et al., 2020; Cheng & Jin, 2019; Ding et al., 2020; Tussyadiah & Zach, 2017; Zhang, 2019a), physical aspects of the property (Celata et al., 2020; Cheng & Jin, 2019; Ding et al., 2020; Tussyadiah & Zach, 2017; Zhang, 2019a) and the convenience of the location (Celata et al., 2020; Cheng & Jin, 2019; Ding et al., 2020; Tussyadiah & Zach, 2017; Zhang, 2019a).

In order to better understand how the different attributes found in the text contribute to Airbnb guest satisfaction, a limited number of studies explore the association between the attributes found and rating and sentiment scores (Cheng & Jin, 2019; Luo & Tang, 2019; Tussyadiah & Zach, 2017). Those studies assume the sentiments expressed in text and rating scores being a proxy of customer satisfaction. The present study is based on the same assumption, since the sentiment identified in the text reflects the customer emotional state in relation to the service. We combine Topic Modelling and Sentiment Analysis approaches, taking advantage of the potential of both methodologies, allowing not only to revisit the topics of guest satisfaction but also to understand whether guests are talking positively or negatively about those topics.

Despite the history of reviews being available online, studies analysing Airbnb guest reviews throughout the time are still scarce (Ding et al., 2020). Ding et al. (2020) beyond examining the differences between international and Malaysian Airbnb users when commenting on Airbnb accommodation experience, present topics proportion over time between 2014 and 2019. The interpretation of guests' expectations trends in the study of Ding *et al.* (2020) is hampered by the higher number of topics and, consequently, by the distinction of topics with very semantic proximity. In this sense, a macro analysis that involves a smaller number of topics would clarify the trends of the topics expressed by guests.

Given Airbnb's changes since its inception, such as the proliferation of multi-unit hosts and the professionalization of the platform (Dogru et al., 2020), as well as the dynamic of the consumers' preferences, analysing the evolution of guests sentiment throughout the time is relevant to clarify how customer satisfaction is evolving. This study addresses this gap, presenting a Sentiment Analysis over a ten-year period.

3. Dataset

The dataset used in this paper is obtained from the *Inside Airbnb* website (http://insideairbnb.com/), an independent and non-commercial initiative that collects public information from Airbnb's website to allow public analysis, discussion, and community benefit. The large open datasets provided by this website have been used extensively in research, and particularly for studying Airbnb reviews (Cheng & Jin, 2019; Lawani et al., 2019; Serrano et al., 2020; Tussyadiah & Zach, 2017; Zhang, 2019b; Zhu et al., 2020).

Inside Airbnb periodically releases new versions of the datasets for many cities worldwide, including information about the accommodations, the availability calendar for 365 days in the future, and the reviews for each property. This study uses a dataset compiled on April 29, 2020, concerning the reviews written by the guests for each accommodation in Lisbon. However, in order to avoid the interference of the COVID-19 pandemic effect on Airbnb activity (Hu & Lee, 2020), the time span selected is between July 2010 and December 2019. Our study is restricted to reviews in English, which corresponds to 60% of the total reviews. Non-English reviews were detected and removed, as well as automatic reviews in the form of '… This is an automated posting', generated for example when a booking is cancelled. In total, 590,070 reviews were selected for our analysis, with an average length of 54 words (terms), varying between 1 word and 1,008 words.

Our dataset includes four variables: property Id, date in which the review was written, reviewer Id, and the review text itself.

4. Methodology

Figure 1 illustrates the adopted research process. The general framework consists of creating a Topic model (LDA Model) with the reviews posted between July 2010 and December 2019 and then applying the model on aggregated documents. In parallel, the sentiment of each review is discovered.

[Figure 1 near here]

In the following sub-sections, we present the methodologies employed in this study: Topic Modelling, which allows the detection of the topics discussed in the reviews, and Sentiment Analysis, which enables the understanding of the positivity/negativity expressed in reviews.

4.1. Topic Modelling

Topic Modelling is a natural language processing and text mining method that can be applied to a collection of text documents for discovering the latent abstract 'topics' occurring in those documents in an unsupervised way. This technique has increasingly been used in research because it is an automated process that enables the analysis of much larger text collections than would be feasible to process manually. Topic Modelling has been widely used in the tourism and hospitality sector to extract important tourist product/service attributes from online reviews (Guo et al., 2017; Taecharungroj & Mathayomchan, 2019).

For extracting the main attributes of guest satisfaction expressed in reviews, we use Latent Dirichlet Allocation (LDA). LDA generates 'topics' as lists of words based on the likelihood of words to occur within documents and infers mixtures of these topics in each document. LDA is the most common method for Topic Modelling for its better probability/statistical foundation that overcomes the overfitting problem of other methods (Blei et al., 2003).

As usually done in text mining, due to the large number of non-informative words and characters in online texts, the text must be pre-processed in order to reduce the noisy content and to improve the effectiveness and efficiency of extracting the important attributes. The pre-processing of the texts of the reviews consisted in the following sequence of steps: (1) remove non-alphabetic characters; (2) convert all text into lowercase letters; (3) normalize contractions (e.g., ‘isn’t’ to ‘is not’); (4) consider only nouns and verbs after assigning parts of speech to each word (POS tagging); (5) lemmatize (i.e., reduce tokens to their root form, such as ‘enjoyed’ into ‘enjoy’); (6) remove of pre-defined stop-words (common words with little meaning, such as articles and auxiliary verbs) and Airbnb domain-specific stop words (‘apartment’, ‘place’, ‘stay’ and ‘location’); (7) remove of low frequency words (terms that are not present in at least 1.5% of the documents); (8) replace the most common host's names with ‘@host’; (9) correct some spelling mistakes such as ‘barrio’ instead of ‘bairro’; (10) translate ‘Lisboa’ (the name of the city written in Portuguese) by ‘Lisbon’. The implementation of the text pre-processing tasks was performed using regular expressions and by applying a set of Python Natural Language Processing modules, such as: *langdetect* for language detection; *Natural Language Toolkit* (Bird et al., 2009) for assigning parts of speech, performing lemmatization and removing stop-words; and *scikit-learn* (Pedregosa et al., 2011) for converting the collection of text documents to their matrix representations. For data manipulation, such as filtering or data aggregation, we have used the *pandas* library (McKinney & others, 2010), a high-level data manipulation tool.

In the creation of the LDA model, each review is considered an independent document. The number of topics to be discovered in the text, a required parameter for LDA, was estimated through an iterative process. After experimenting different numbers of topics (between 2 and 10), we selected a 3-topic model solution considering its topic's semantic coherence. Topic coherence measures have been proposed in the literature to evaluate the interpretability of topics learned from a statistical topic model. Such measures, based on word similarity of the top words of a topic, are grounded on a reference corpus (Röder et al., 2015). In our experiments, we consider the C_v measure designed for LDA (Röder et al., 2015), which achieved the maximum value for 3 topics. This solution was also validated qualitatively by the authors considering the clearness and interpretability of the topics, and its proximity to tourism literature. The LDA method was implemented with the *gensim* module (Rehurek & Sojka, 2011), a Python library designed specifically for Topic Modelling.

In order to observe the evolution of topics over time, all the reviews written in the same month were aggregated in one document (time-level aggregation), leading to a dataset composed by 111 documents used for inference.

4.2. Sentiment Analysis

Sentiment Analysis has become a popular automated process of examining a text's polarity, in general, categorized into two main approaches: the machine-learning and the rule-based approaches. The machine learning approach is usually performed using sentiment models previously trained on labelled textual datasets. This approach requires time and effort to label the texts manually and relies heavily on the training data. The rule-based approach does not require human-labelled documents since it uses word lexicons where each word or sequence of words is associated with a specific sentiment score. Machine-learning approach is considered a domain-dependent task, where models trained for a given domain do not perform well in other domains. The use of rule-based approaches is often reported in the literature because of their simplicity and because adapting sentiment lexicons for certain domains is a relatively simple process. This approach proved to be efficient, sufficiently accurate, and has been increasingly used for analysing social media data (Kumaresh et al., 2019).

In this study, Sentiment Analysis is performed using VADER (Valence Aware Dictionary for sEntiment Reasoning), a rule-based Sentiment Analysis tool that is specifically adapted to detect sentiments expressed in social media (Hutto & Gilbert, 2014). VADER is based on lists of words associated with valence scores for sentiment intensity, validated by humans. More positive words have higher positive ratings and more negative ones have lower negative ratings. For computing the sentiment polarity, VADER combines a valence-based lexicon attuned explicitly to social media contexts, along with five general rules for changing sentiment intensity: (1) Punctuation, namely the exclamation point, and (2) Capitalization, that increase the magnitude of the sentiment intensity; (3) Degree modifiers (e.g., adverbs such as 'extremely' or 'marginally') that can either increase or decrease the intensity; (4) The conjunction 'but', and (5) Negation, which changes the sentiment polarity. VADER produces a compound score standardized to range between -1 and 1. A score near 1 means the review is strongly positive, -1 means the review is very negative, and scores between –0.05 and +0.05 are considered neutral, according to Hutto and Gilbert (2014). This work considers the previously mentioned thresholds, and each review is classified as positive, neutral, or negative. In order to provide a comprehensive evolution of sentiment over time, the number of reviews for each polarity is taken into account for each month.

As illustrated in Figure 1, the topics were also inferred by sentiment polarity, enabling the understanding of the positivity/negativity expressed regarding those topics. In this process, all the reviews with the same sentiment polarity were aggregated in one document (sentiment-level aggregation), leading to a dataset composed by 3 documents used for inference.

4.3. Trends Study for Older and Newer Properties

In order to understand whether topic and sentiment changes over time were driven by newer properties or were valid for all properties, we performed an additional experiment comparing reviews trends from properties listed before and after 2015.

The March 2015 snapshot (the oldest snapshot available online) was defined as the benchmark for distinguishing older from newer properties, as in 2014, when legal regimes for local accommodation were created, the number of Airbnb accommodations significantly increased, with repercussions on the number of reviews. The information of the properties listed on the mentioned snapshot was collected from Tom Slee's blog database (Slee, n.d.).

The reviews were divided into three groups: reviews written about properties that were listed on the Airbnb website in March 2015 (older properties), reviews from properties that were not present in the chosen snapshot and did not have reviews before the snapshot (newer properties), and reviews from properties that were not present in the chosen snapshot but had some reviews written before the snapshot (older properties). The latter group refers to the properties that were removed from Airbnb website before March 2015 or that have not been captured during the data collection by Tom Slee (Slee, n.d.). In total, 2,067 old properties and 16,274 new properties were considered.

5. Results and discussion

This section presents our findings. First, we analyse the topics achieved by the Topic Modelling approach. Then, we analyse the overall sentiment trends over time by performing Sentiment Analysis of each one of the reviews. Finally, by combining both Topic Modelling and Sentiment Analysis, we analyse the content's overall sentiment related to each topic.

5.1. Topic Modelling results

As previously mentioned, we have adopted a topic model solution with three topics for analysing the guest reviews. Figure 2 was created using the pyLDAvis tool (Sievert & Shirley, 2015) and represents the topics in an inter-topic distance map of two dimensions. Each circle in the image represents a topic. The area of the circle is proportional to the frequency of each topic over the entire corpus, and the distance between circles indicates the similarity between topics. The smaller the distance, the more words the topics have in common. In Figure 2, these representations are complemented by the lists of the most frequent terms for each topic, along with the frequency of occurrence.

[Figure 2 near here]

Topic 1 covers a variety of concepts related to the satisfaction with the service provided by the host. The prevalence of the terms ‘host’ and ‘@host’, which represents the first name of the host according to our pre-processing method, reveals the focus of these topics in the characteristics of the host and the need to give a face to the host (e.g., ‘*Joao is a great host and helped us out a lot*’, ‘*Sergio is a wonderful host, very helpful and lives in the neighborhood so he has great suggestions for your stay*’). Reviews highlight the feeling of being at home while staying in an Airbnb accommodation, the hospitality and care of the host (e.g., ‘*It's the small details that count and Antonio's details are the story of a great stay in Lisbon.*’), as well as the helpfulness of the host providing recommendations for places to eat and go, and his availability to communicate with guests. This topic shows expressions of appraisal by the stay, particularly by the apartment, host, and location (e.g., ‘recommend stay’, ‘recommend apartment’, ‘recommend place’, ‘love stay’, ‘thank host’, ‘enjoy stay’). These results are consistent with previous studies (Cheng & Jin, 2019; Ding et al., 2020) that verified that the concepts within the themes ‘location’ and ‘host’ greatly influence Airbnb users’ recommendations.

Topic 2 describes the physical aspects of the property, including spaces (e.g., ‘room’, ‘bedroom’, ‘kitchen’, ‘bathroom’, ‘view’), facilities (e.g., ‘shower’, ‘balcony’, ‘machine’, ‘window’, ‘floor’, ‘door’, ‘bed’), and functionality/preservation of the property (e.g., ‘use’, ‘work’, ‘problem’, ‘issue’). The quality of sleep (e.g., ‘sleep’, ‘night’, ‘bed’, ‘noise’) is also considered important in this topic. The term ‘host’ is the most common word of Topic 1 but it is not an exclusive term of this topic. Although in a lower frequency, this concept is also present in Topic 2, related to the host's responsiveness in solving the issues/problems of the property.

Topic 3, labelled as ‘Location’, focus on the convenience of the location of the property, both by the proximity to points of interest (e.g., restaurants, shops, supermarkets) and the convenience of transportation (access to public transportation (e.g., train station, metro, bus, tram) and walking distances). The expression ‘everything need’ is exclusive of this topic, referring to the offers of the neighbourhood.

The service of the host, as well as the physical aspects of the property and the convenience of the location are consensual in the literature as important attributes in guest’s satisfaction. These attributes were found as the main dimensions on reviews of Airbnb accommodations in other countries such as the USA (Luo & Tang, 2019; Tussyadiah & Zach, 2017; Zhang, 2019b, 2019a), Australia (Cheng & Jin, 2019), Malaysia (Ding et al., 2020) and Italy (Celata et al., 2020). These results reinforce that these three aspects are cross-cutting primary attributes influencing guests experience and satisfaction.

In addition, we have performed an inference of topics over time, represented in Figure 3, where the reviews trends from older and newer properties are illustrated. The analysis revealed (for both groups) a constant decrease of the proportion of the topic ‘Host’s service’ since 2016 compensated with an increase of the proportion of the topic related to the physical aspects.

[Figure 3 near here]

Regarding Topic 1 terms, '@host' is the one with the greatest decrease, showing that the hosts names have been less mentioned on reviews. These results lead us to conclude that guests may have had less direct contact with the host during accommodation and that the service is becoming less humanized.

Our findings complements the analysis reported by Ding et al. (2020), showing that the topic related to the more humanized side of service has a downward trend, and that can be explained by the professionalization of the Airbnb platform (Dogru et al., 2020), which delivers a more standardized service due to its growth. However, further research is required to explain this pattern in the primary concerns of guests.

5.2. *Sentiment Analysis results*

Sentiment Analysis shows that most reviews (98%) are positive, with 1% of the reviews being negative and the other 1% being neutral. Our results confirm the bias towards positivity, also perceived in previous studies on Airbnb reviews (Bridges & Vásquez, 2018; Zervas et al., 2018; Zhang, 2019b). Although this positivity bias is verified, in general, in travel review websites, positive ratings tend to skew higher on Airbnb than other established platforms (Zervas et al., 2018). Some of the reasons that may explain the greater positivism or the lack of negativity in a review made on Airbnb are as follows: guest expectations being more realistic in the case of peer-to-peer accommodations since the property description is written by an individual rather than a corporate marketing effort (Yannopoulou, 2013); and, guests trying to avoid to harm the reputation of the host when negative experiences with rentals are not serious and hosts showing a strong client orientation (Bulchand-Gidumal & Melián-González, 2020).

In order to understand the tendencies of Airbnb reviews sentiment, a time-based view is presented in Figure 4. Although the local accommodation figure was introduced into the Portuguese legal system in 2008, since 2014, when its own legal regimes were created, the number of Airbnb accommodations significantly increased, with repercussions on the number of reviews.

[Figure 4 near here]

The number of reviews represented in Figure 4 reflects the seasonality in tourism in general and similarly in Airbnb's operation. Seasonality in Lisbon derives essentially from the natural seasons (e.g., summer and winter), and so the minimum number of reviews is verified between January-February, and the maximum number of reviews is verified in September.

Although the number of negative and neutral reviews is very low in percentage, the time-based view clearly suggests an increasing tendency in the percentage of negative and neutral reviews.

We have performed an additional experiment, represented in Figure 5, comparing the reviews from properties listed before and after 2015, and we have verified that the two groups of reviews have the same decreasing pattern of positive sentiment as well as the increasing pattern of negative and neutral sentiment, leading us to conclude that the changing trend was not driven by newer properties.

[Figure 5 near here]

Since 92% of the reviews are from guests reviewing the service in Lisbon for the first time, this analysis is a good representation of the sentiments of Airbnb guests at different times. These findings are consistent with the empirical study of Guttentag (2020), which suggests that later adopters of Airbnb (the ones that follow the first individuals to adopt Airbnb) exhibit lower degrees of satisfaction. Guttentag' study is based on innovation diffusion concepts (Rogers, 2003) and on online surveys, covering the process of different individuals deciding to adopt Airbnb at different times. Guttentag (2020) argues that Airbnb is recognized as a product evolving over time, with a concerted effort to provide a more reliable and professional (i.e., hotel-like) service. So earlier adopters accepted a lodging product that was less hotel-like in nature than that which exists today. The findings of Guttentag (2020) suggest that later adopters show a greater preference towards hotel-style Airbnb features (e.g., entire home rentals, Instant Book, Airbnb Plus, and Superhosts) and, beyond that, exhibit lower expectations and satisfaction towards Airbnb. Our results, obtained through an automated process, corroborates that more recent Airbnb adopters exhibit lower levels of satisfaction.

5.3. Distribution of topics by sentiment polarity

Although non-positive reviews represent a small percentage of total reviews, it was assumed that a closer examination of the topics by polarity would give some insights about the causes of dissatisfaction in the Airbnb experience, which can help service improvement. Figure 6 represents the distribution of the topics by sentiment polarity, which shows that non-positive reviews are mainly related to the physical aspects of the property. These results are in accordance with Villeneuve and O'Brien (2020) results that show that Airbnb reviews with indoor environmental quality related terms had a statistically significantly worse overall sentiment score than those without.

[Figure 6 near here]

A closer look at the text of negative reviews with the presence of Topic 2 (71% of the negative reviews text), reveals that the causes of dissatisfaction are related to the lack of comfort of the house (e.g., noisy, dirty, smelly, cold) and the lack of accuracy of the facilities (e.g., 'broken machine', 'different from pictures'). Those attributes are also found in the study of Villeneuve and O'Brien (2020), which explores top reported causes for indoor environmental quality complaints on Airbnb in Canada, pointing out as main causes of dissatisfaction the lack of acoustic comfort, thermal comfort, indoor air quality, and visual (light) comfort. While in the text of negative reviews in which the Topic 1 prevail, the hosts' lack of communication, flexibility, and presence is criticized. In a smaller proportion, Topic 3 (location) is also found in negative reviews, where the causes of dissatisfaction are related to the steep locations and lack of parking spaces, characteristic of the historical centre of Lisbon.

The text of positive reviews with the presence of the topics related to service (41% of the positive reviews text) shows appreciation for the hospitality and helpfulness of the host. Positive reviews where Topic 2 is the predominant topic show that Airbnb guests appreciate a well-equipped and cleaned house, as well as unexpected amenities (e.g., 'touch', 'wine'), as already verified by Cheng and Jin (2019) in Sydney reviews. The text of positive reviews in which Topic 3 is predominant focuses on the convenience of location and transportation, which is consistently reported in the literature as an attribute valued by guests (Celata et al., 2020;

Cheng & Jin, 2019; Ding et al., 2020; Luo & Tang, 2019; Tussyadiah & Zach, 2017; Zhang, 2019b, 2019a), as it defines how the tourist will explore the destination.

Figure 7 shows examples of positive and negative reviews for each one of the topics. In these examples, we have highlighted the words in the text that are the most representative of the corresponding topic.

[Figure 7 near here]

6. Conclusions

This study provides insights into the primary concerns of the guests about the Airbnb accommodation in Lisbon, using text mining approaches, in line with the growing recognition of the importance of using online reviews to understand tourist experience and satisfaction better.

From a methodological perspective, this study presents an automatic method based on unsolicited and easily accessible guest generated content that allows us to better understand the evolution of guest satisfaction, in alternative to time consuming traditional methods, such as surveys used by Guttentag (2020). This paper demonstrates the suitability of using Latent Dirichlet Allocation and Sentiment Analysis for extracting meaningful information from large, unstructured human-authored text data. Combining these two text mining approaches made it possible to identify the most prevalent guest satisfaction factors and analyse the time effect. The methodology proposed opens new lines of investigation that can be pursued for understanding the life cycle of Airbnb.

Based on a Topic Modelling approach, the most predominant topics were captured: ‘host’s service’, ‘physical aspects’, and ‘location’. These results corroborate the existing literature, supporting that these three aspects are cross-cutting primary Airbnb guests' concerns and

reinforcing the convergence in the Airbnb experience (Brochado et al., 2017). Furthermore, our study shows a downward trend of the topic related to host's service, bringing significant implications.
The Sentiment Analysis results confirm a positive bias in guest reviews (98% positive reviews); however, the analysis of the time effect shows that guest satisfaction follows a decreasing pattern in a ten years' time span.

By combining Topic Modelling and Sentiment Analysis, it was possible to provide a comprehensive picture of the positively and negatively perceived dimensions of the Airbnb experience. The results reveal that 'physical aspects' is the predominant topic in the negative guest reviews.

6.1. Research implications

The results from this study are relevant for Airbnb and the wider hospitality sector because they provide an understanding of how guests evaluate their experience on Airbnb. This study confirms through Sentiment Analysis of guest reviews that more recent Airbnb adopters exhibit lower degrees of satisfaction. As explained by Guttentag (2020), these results can be interpreted in the light of the evolution of Airbnb as a brand, which nowadays promotes a more reliable and professional (i.e., hotel-like) service, and consequently attracts guests less accepting of uncertainty and lodging products less hotel-like in nature. From a managerial standpoint, the findings of this study suggest that Airbnb practitioners must manage guests' expectations better since customer satisfaction is decreasing over time.

Our Topic Modelling results showing that the host's service topic has been in decline raise other hypotheses and relevant research questions: Is the decline of this topic a sign that hosts are delivering a less humanized service? Is Airbnb service tending to be more standardized due to the increasing professionalization of the hosts? Since the original value proposition of

Airbnb is supported on authentic local experiences and social interaction, the decline of the humanized service has significant implications. It not only corroborates that Airbnb is losing its sharing economy ethos (Demir & Emekli, 2021), but also suggests that Airbnb practitioners should take actions in order to distinguishes Airbnb from conventional forms of accommodation.

This research brings to light the urgency to give back to Airbnb the initial characteristics that justified the increasing growth of this lodging mean - the humanized service or feel at home away from home (Dias et al., 2015). Airbnb is competing with hotels but while hotels seem to be adopting the sharing economy attributes and experiences to compete more effectively, Airbnb seems to be closer to the traditional hotel approach, which can result in a non-competitive strategy.

6.2. Limitations and future work

Despite the contributions mentioned above, this study still has some limitations that should be addressed in further research. First, this study applies LDA, a popular unsupervised Topic Modelling approach to efficiently discover the topics and weights from a large text corpus. However, standard LDA does not allow the use of supervised labels to incorporate independent expert knowledge into the learning process. Future research could employ a semi-supervised approach, where a sample of documents (reviews) is manually read and tagged, with the topics observed by independent expert labellers, to guide the model learning. Such a semi-supervised approach could improve document classifications and align the resulting model with human expectations for Topic Modelling and extraction. Second, our study is restricted to reviews written in English. It would be valuable to incorporate data in a bilingual or multilingual context for cross-cultural analysis. Third, the current COVID-19 pandemic is undoubtedly impacting tourists' accommodation decisions and satisfaction criteria. Thus, future comparative studies between the reviews written before and after the beginning of the pandemic would be a useful next step forward that could demonstrate a greater seek for hygiene and safety. Finally, our dataset was collected from *Inside Airbnb* website, which offers direct download data reportedly collected from Airbnb's website. Although widely used in academic research, this open dataset lacks investigations regarding the quality of the data. Alsudais (2021) examines the dataset and explains an issue of added incorrect data. Findings indicate that reviews were incorrectly added to some listings due to a new feature implemented by Airbnb (Airbnb experiences). However, the low number of IDs with incorrect data suggests that it may not currently be severe enough to affect published papers using the *Inside Airbnb* dataset.

References

Alsudais, A. (2021). Incorrect Data in the Widely Used Inside Airbnb Dataset. *Decision Support Systems*, *141*, 113453. https://doi.org/10.1016/j.dss.2020.113453

Baka, V. (2016). The becoming of user-generated reviews: Looking at the past to understand the future of managing reputation in the travel sector. *Tourism Management*, *53*, 148–162. https://doi.org/10.1016/j.tourman.2015.09.004

Belarmino, A., Whalen, E., Koh, Y., & Bowen, J. T. (2019). Comparing guests' key attributes of peer-to-peer accommodations and hotels: mixed-methods approach. *Current Issues in Tourism*, *22*(1), 1–7. https://doi.org/10.1080/13683500.2017.1293623

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc."

Bridges, J., & Vásquez, C. (2018). If nearly all Airbnb reviews are positive, does that make them meaningless? *Current Issues in Tourism*, *21*(18), 2057–2075. https://doi.org/10.1080/13683500.2016.1267113

Brochado, A., Troilo, M., & Shah, A. (2017). Airbnb customer experience: Evidence of convergence across three countries. *Annals of Tourism Research*, *63*, 210–212. https://doi.org/10.1016/j.annals.2017.01.001

Bulchand-Gidumal, J., & Melián-González, S. (2020). Why are ratings so high in the sharing economy? Evidence based on guest perspectives. *Current Issues in Tourism*, *23*(10), 1248–1260. https://doi.org/10.1080/13683500.2019.1602597

Celata, F., Capineri, C., & Romano, A. (2020). A room with a (re)view. Short-term rentals, digital reputation and the uneven spatiality of platform-mediated tourism. *Geoforum*, *112*(June 2020), 129–138. https://doi.org/10.1016/j.geoforum.2020.04.007

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

Correia, A., & Kozak, M. (2021). Past, present and future: trends in tourism research. *Current Issues in Tourism*, 1–16. https://doi.org/10.1080/13683500.2021.1918069

Demir, E., & Emekli, G. (2021). Is Airbnb no longer a sharing economy platform? Evidence from Europe's top 10 Airbnb destinations. *Anatolia*, *32*(3), 470–488. https://doi.org/10.1080/13032917.2021.1890626

Dias, J. A., Correia, A., & López, F. J. M. (2015). The meaning of rental second homes and places: the owners' perspectives. *Tourism Geographies*, *17*(2), 244–261. https://doi.org/10.1080/14616688.2014.959992

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*, 102676.

https://doi.org/10.1016/j.ijhm.2020.102676

Dogru, T., Mody, M., Suess, C., Line, N., & Bonn, M. (2020). Airbnb 2.0: Is it a sharing economy platform or a lodging corporation? *Tourism Management*, *78*, 0–17. https://doi.org/10.1016/j.tourman.2019.104049

Flint, D. J., Blocker, C. P., & Boutin, P. J. (2011). Customer value anticipation, customer satisfaction and loyalty: An empirical examination. *Industrial Marketing Management*, *40*(2), 219–230. https://doi.org/10.1016/j.indmarman.2010.06.034

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, *59*, 467–483. https://doi.org/10.1016/j.tourman.2016.09.009

Guttentag, D., & Smith, S. L. J. (2020). The diffusion of Airbnb: a comparative look at earlier adopters, later adopters, and non-adopters. *Current Issues in Tourism*, 1–20. https://doi.org/10.1080/13683500.2020.1782855

Hu, M. R., & Lee, A. D. (2020). Airbnb, COVID-19 Risk and Lockdowns: Global Evidence. *SSRN Electronic Journal*, *December*. https://doi.org/10.2139/ssrn.3589141

Hutto, C. J., & Gilbert, E. (2014). VADER : A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International AAAI Conference on Weblogs and Social Media*, 216–225.

Kim, W. G., & Park, S. A. (2017). Social media review rating versus traditional customer satisfaction. *International Journal of Contemporary Hospitality Management*, *29*(2), 784–802. https://doi.org/10.1108/IJCHM-11-2015-0627

Kumaresh, N., Bonta, V., & Janardhan, N. (2019). A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis. *Asian Journal of Computer Science and*

Technology, *8*(S2), 1–6. https://doi.org/10.51983/ajcst-2019.8.S2.2037

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

Li, H., Ye, Q., & Law, R. (2013). Determinants of Customer Satisfaction in the Hotel Industry: An Application of Online Review Analysis. *Asia Pacific Journal of Tourism Research*, *18*(7), 784–802. https://doi.org/10.1080/10941665.2012.708351

Li, J., Hudson, S., & So, K. K. F. (2019). Exploring the customer experience with Airbnb. *International Journal of Culture, Tourism, and Hospitality Research*, *13*(4), 410–429. https://doi.org/10.1108/IJCTHR-10-2018-0148

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, *29*(3), 458–468. https://doi.org/10.1016/j.tourman.2007.05.011

Lu, W., & Stepchenkova, S. (2012). Ecotourism experiences reported online: Classification of satisfaction attributes. *Tourism Management*, *33*(3), 702–712. https://doi.org/10.1016/j.tourman.2011.08.003

Luo, Y., & Tang, R. (Liang). (2019). Understanding hidden dimensions in textual reviews on Airbnb: An application of modified latent aspect rating analysis (LARA). *International Journal of Hospitality Management*, *80*, 144–154. https://doi.org/10.1016/j.ijhm.2019.02.008

Mauri, A. G., & Minazzi, R. (2013). Web reviews influence on expectations and purchasing intentions of hotel potential customers. *International Journal of Hospitality*

Management, *34*(1), 99–107. https://doi.org/10.1016/j.ijhm.2013.02.012

McKinney, W., & others. (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*, *445*, 51–56.

Millán, Á., & Esteban, Á. (2004). Development of a multiple-item scale for measuring customer satisfaction in travel agencies services. *Tourism Management*, *25*(5), 533–546. https://doi.org/10.1016/j.tourman.2003.07.002

Mohsin, A., & Lockyer, T. (2010). Customer perceptions of service quality in luxury hotels in New Delhi, India: An exploratory study. *International Journal of Contemporary Hospitality Management*, *22*(2), 160–173. https://doi.org/10.1108/09596111011018160

NOVASCHOOLOFBUSINESS. (2016). *O Alojamento Local em Portugal - qual o fenómeno?* https://www.hoteis-portugal.pt/docs/content/twgQoT7vP6TAxRGMJSCG3w7BLEOyRVd5.pdf

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Priporas, C. V., Stylos, N., Rahimi, R., & Vedanthachari, L. N. (2017). Unraveling the diverse nature of service quality in a sharing economy: A social exchange theory perspective of Airbnb accommodation. *International Journal of Contemporary Hospitality Management*, *29*(9), 2279–2301. https://doi.org/10.1108/IJCHM-08-2016-0420

Rehurek, R., & Sojka, P. (2011). Gensim--python framework for vector space modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, *3*(2).

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *WSDM 2015 - Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, 399–408. https://doi.org/10.1145/2684822.2685324

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

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*, 1–19. https://doi.org/10.1080/09669582.2020.1838529

Sievert, C., & Shirley, K. (2015). *LDAvis: A method for visualizing and interpreting topics*. 63–70. https://doi.org/10.3115/v1/w14-3110

Slee, T. (n.d.). *Tom Slee's personal blog and online repository*. http://tomslee.net/category/airbnb-data

Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods*, *38*(2), 262–279. https://doi.org/10.3758/BF03192778

Sthapit, E., & Jiménez-Barreto, J. (2018). Exploring tourists' memorable hospitality experiences: An Airbnb perspective. *Tourism Management Perspectives*, *28*, 83–92. https://doi.org/10.1016/j.tmp.2018.08.006

Taecharungroj, V., & Mathayomchan, B. (2019). Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. *Tourism Management*, *75*, 550–568. https://doi.org/10.1016/j.tourman.2019.06.020

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

Villeneuve, H., & O'Brien, W. (2020). Listen to the guests: Text-mining Airbnb reviews to explore indoor environmental quality. *Building and Environment*, *169*, 106555. https://doi.org/10.1016/j.buildenv.2019.106555

Wilkins, H., Merrilees, B., & Herington, C. (2010). The determinants of loyalty in hotels. *Journal of Hospitality Marketing and Management*, *19*(1), 1–21. https://doi.org/10.1080/19368620903327626

Yannopoulou, N. (2013). User-Generated Brands and Social Media: Couchsurfing and Airbnb. *Contemporary Management Research*, *9*(1), 85–90. https://doi.org/10.7903/cmr.11116

Zervas, G., Proserpio, D., & Byers, J. (2018). A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average. *SSRN Electronic Journal*, 1–22. https://doi.org/10.2139/ssrn.2554500

Zhang, J. (2019a). Listening to the Consumer: Exploring Review Topics on Airbnb and Their Impact on Listing Performance. *Journal of Marketing Theory and Practice*, *27*(4), 371–389. https://doi.org/10.1080/10696679.2019.1644953

Zhang, J. (2019b). 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

Zhou, L., Ye, S., Pearce, P. L., & Wu, M. Y. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. *International Journal of Hospitality Management*, *38*, 1–10. https://doi.org/10.1016/j.ijhm.2013.12.004

Zhu, L., Lin, Y., & Cheng, M. (2020). Sentiment and guest satisfaction with peer-to-peer accommodation: When are online ratings more trustworthy? *International Journal of Hospitality Management*, *86*, 102369. https://doi.org/10.1016/j.ijhm.2019.102369

Figures

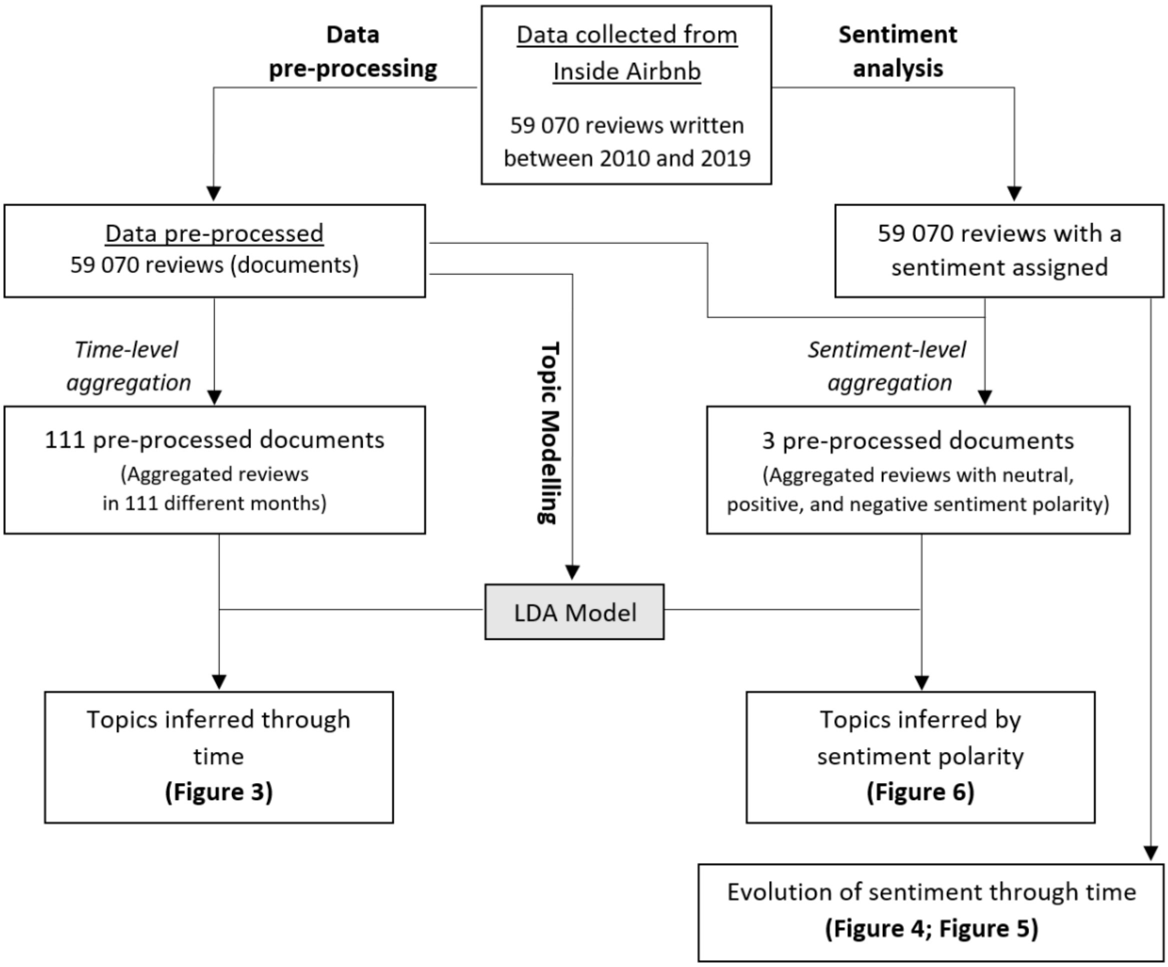


Fig. 1- Illustration of the Proposed Method.

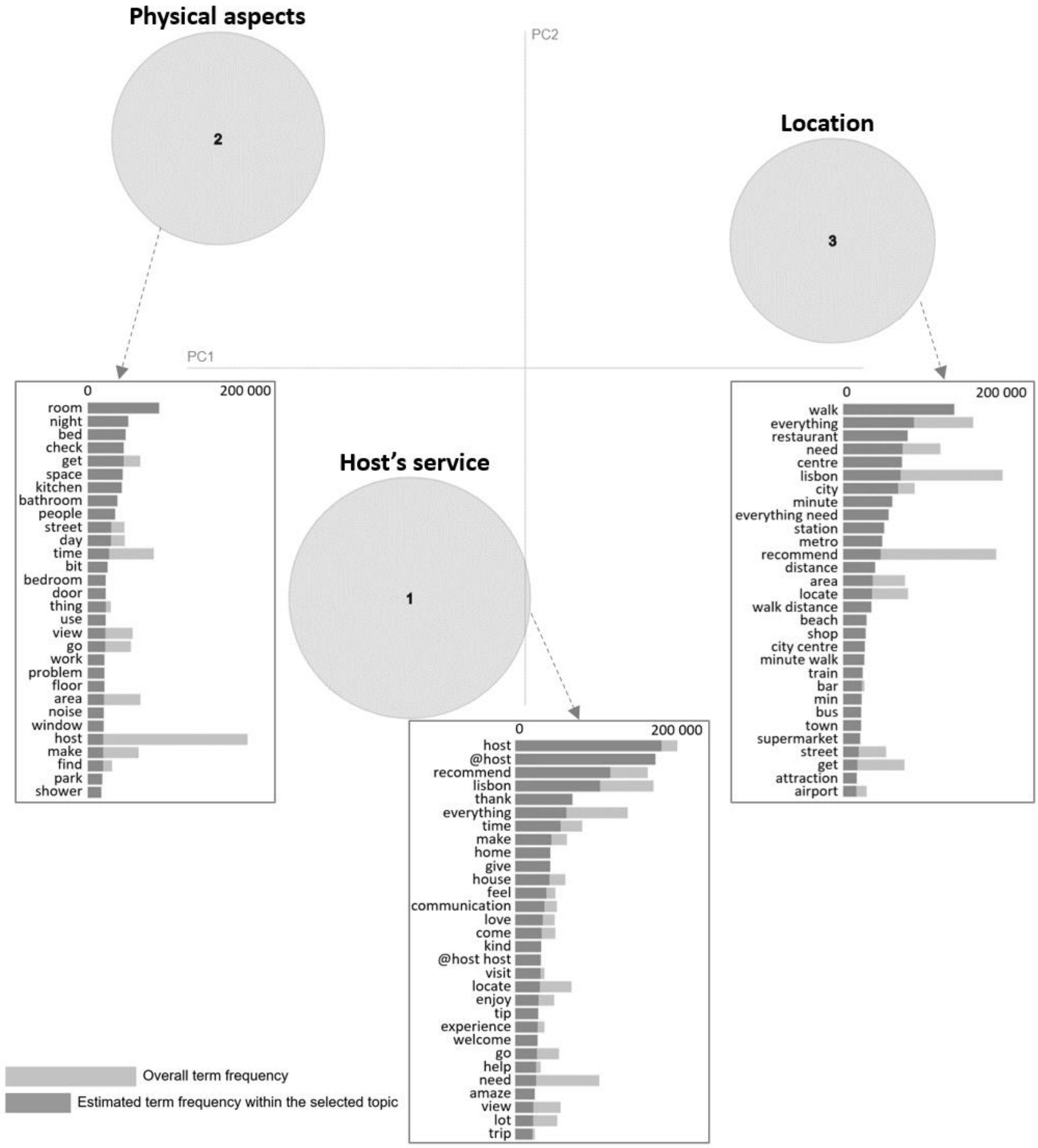


Fig. 2 Representation of the topics as an inter-topic distance map of two dimensions, and histograms with the 30 most relevant terms in each topic

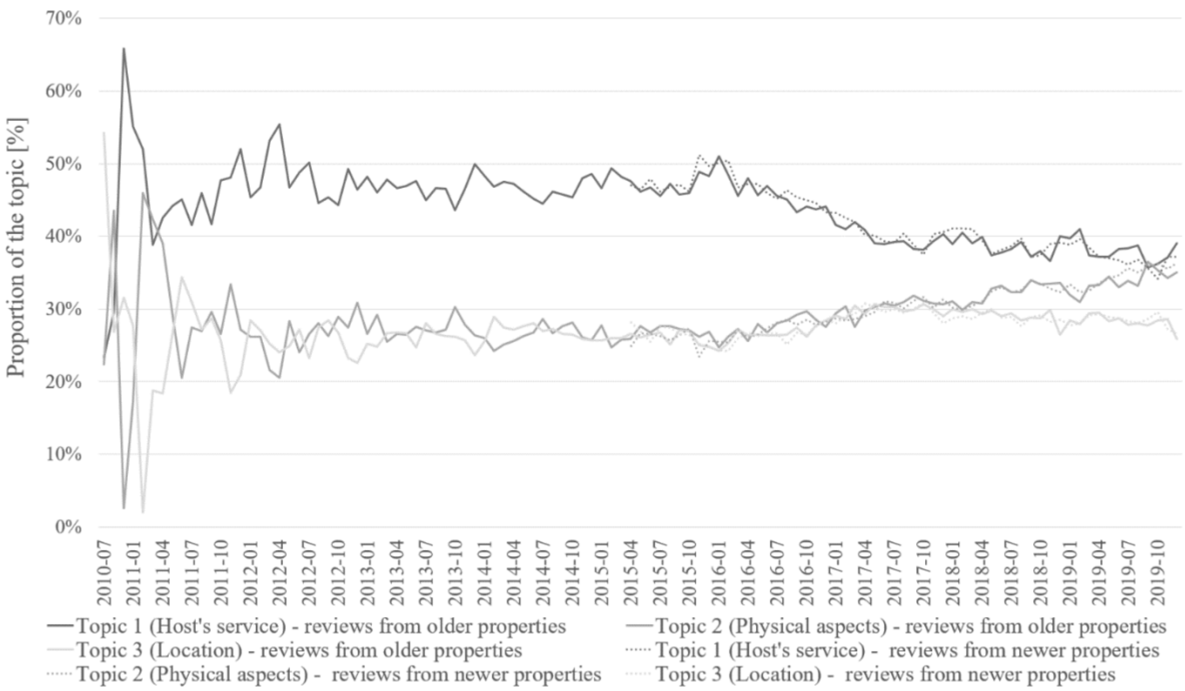


Fig. 3 Topics monthly evolution

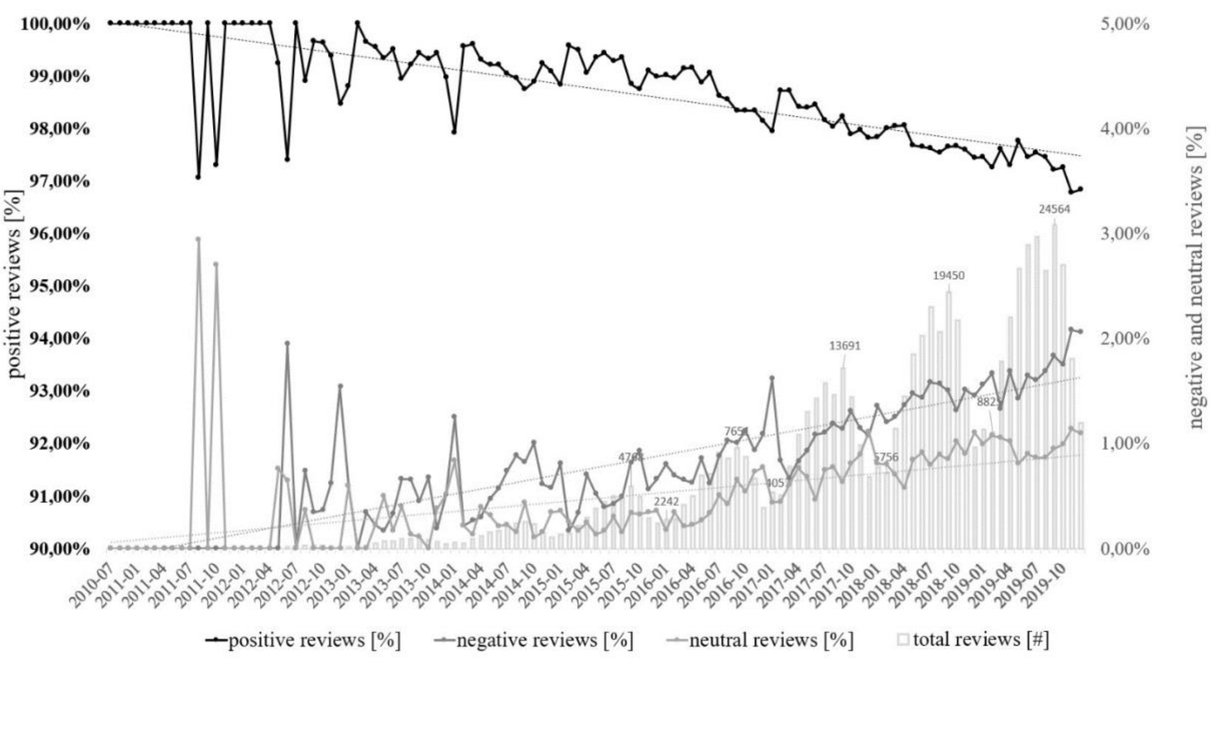
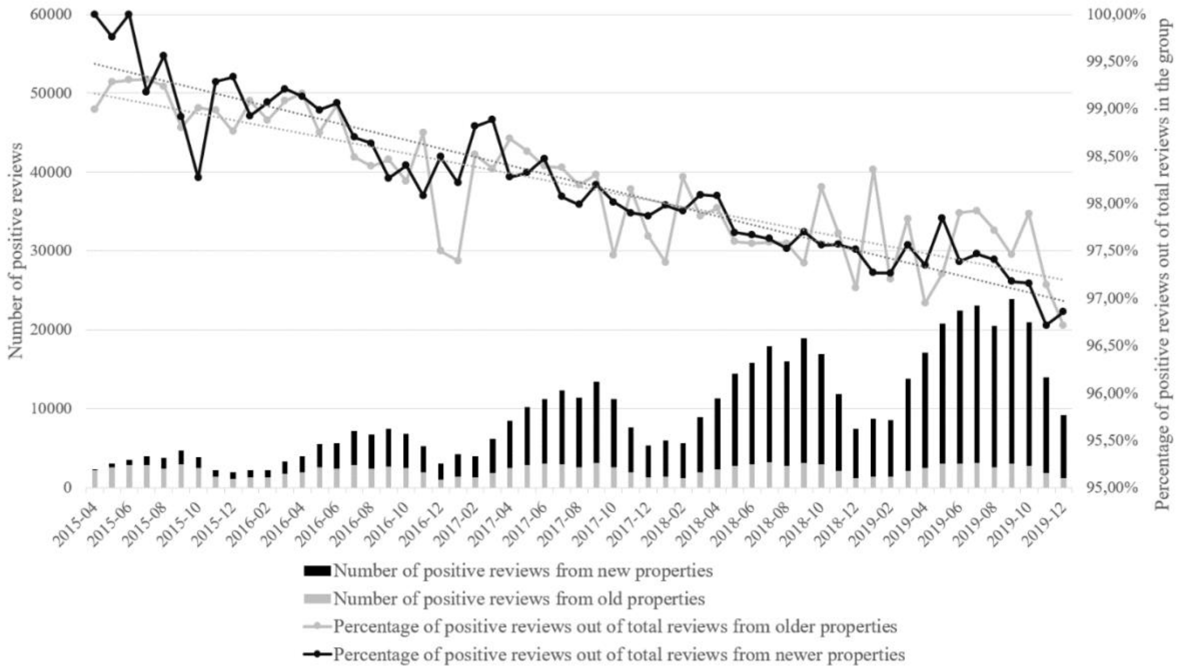
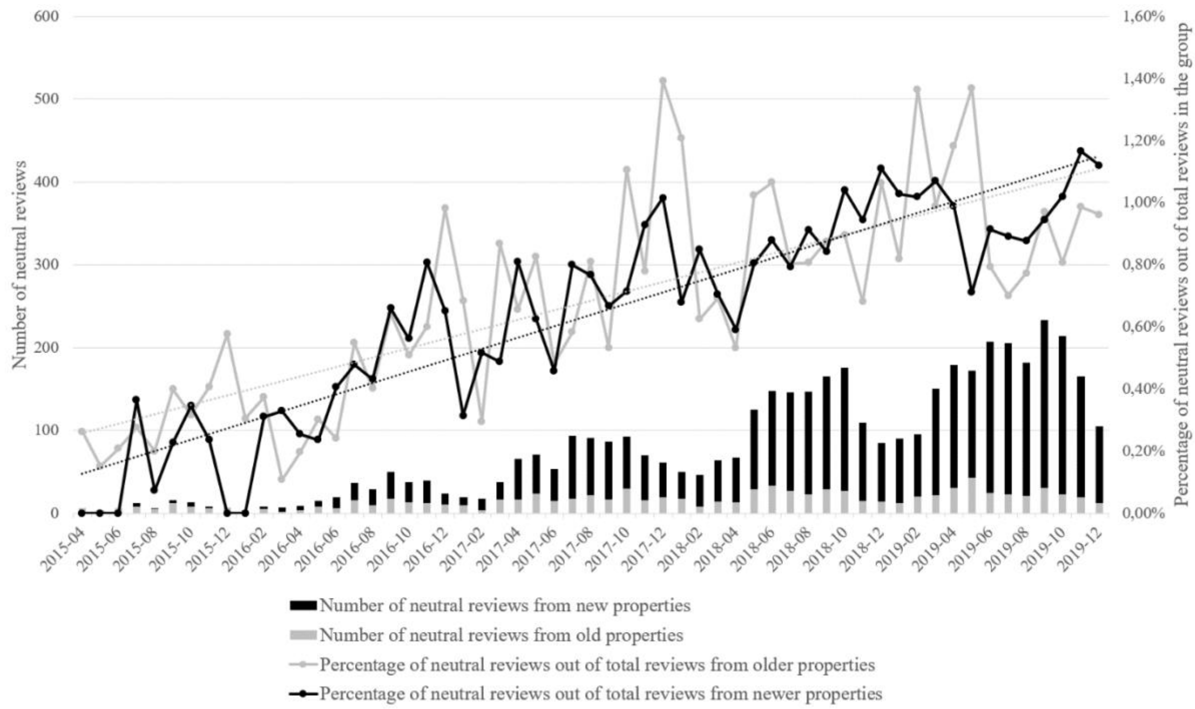


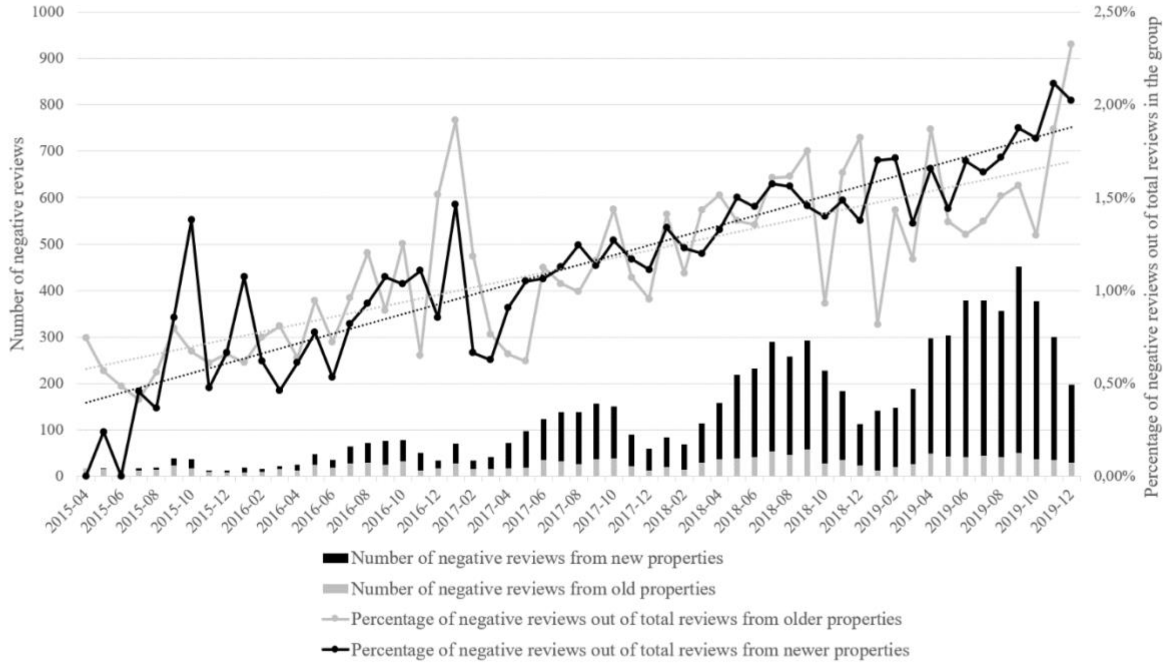
Fig. 4 Monthly evolution of the sentiment perceived in reviews



(a) positive sentiment



(b) neutral sentiment



(c) negative sentiment

Fig. 5 Monthly evolution of the sentiment perceived in reviews from older and newer properties

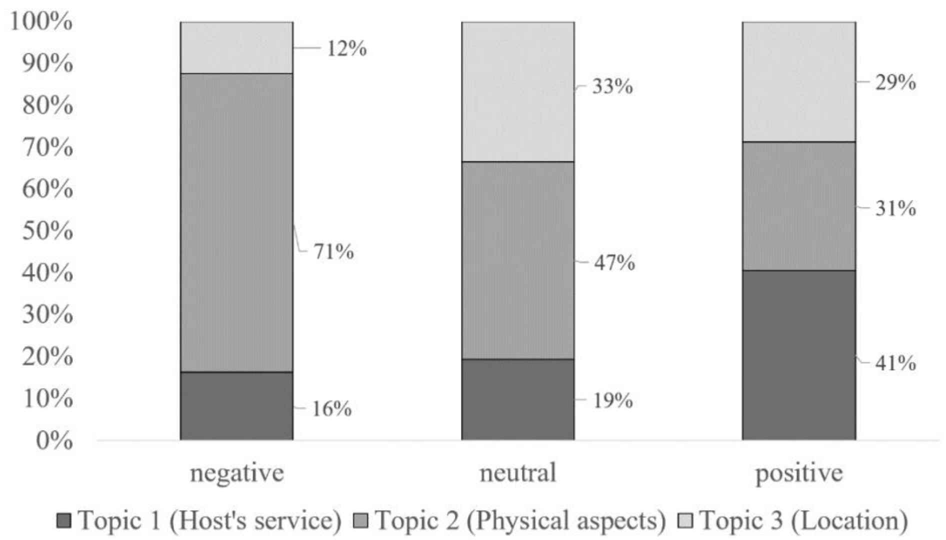
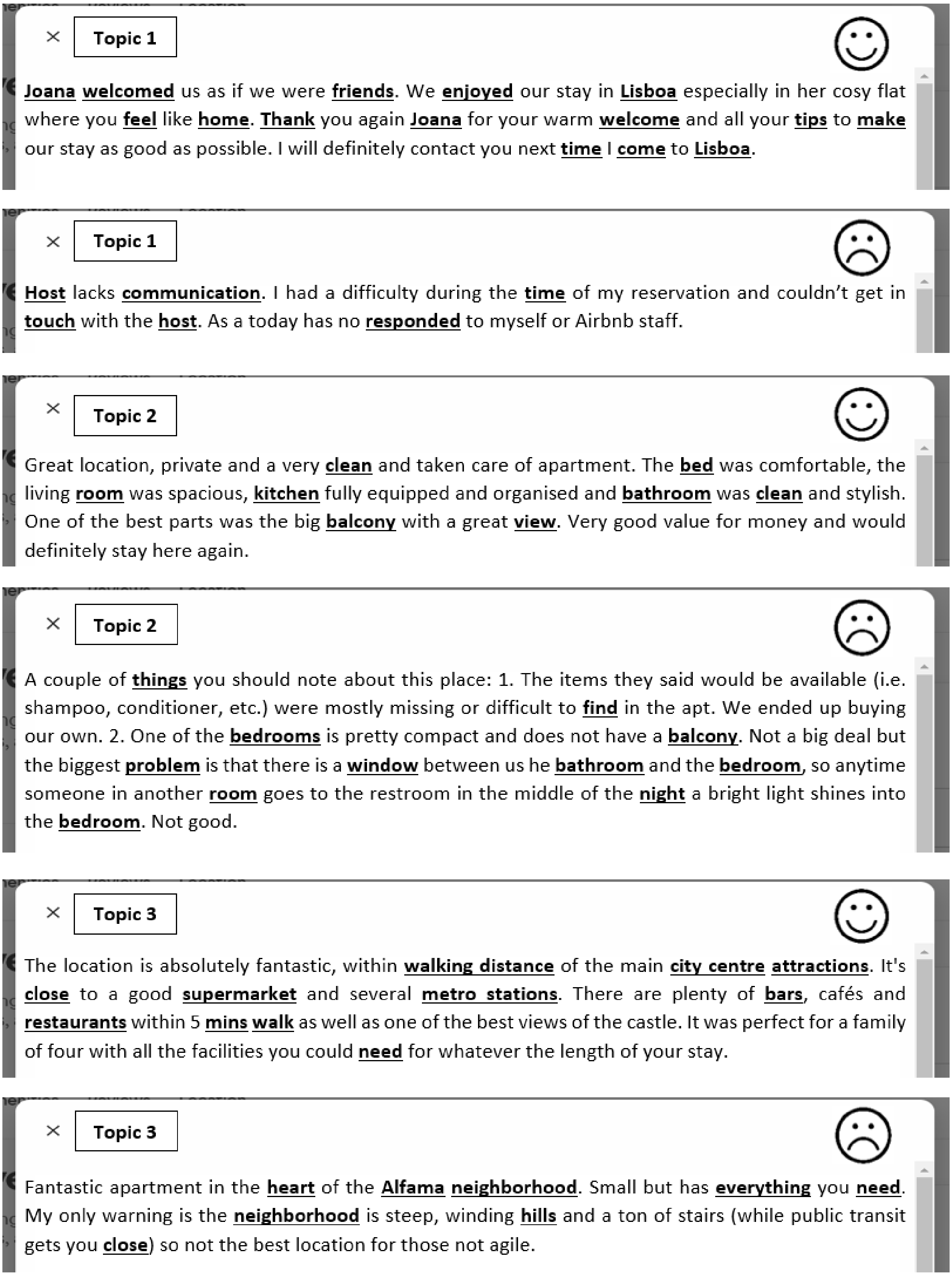


Fig. 6 Proportion of the topics per sentiment polarity



Topic 1

Joana **welcomed** us as if we were **friends**. We **enjoyed** our stay in **Lisboa** especially in her cosy flat where you **feel** like **home**. **Thank** you again **Joana** for your warm **welcome** and all your **tips** to **make** our stay as good as possible. I will definitely contact you next **time** I **come** to **Lisboa**.

Topic 1

Host lacks **communication**. I had a difficulty during the **time** of my reservation and couldn't get in **touch** with the **host**. As a today has no **responded** to myself or Airbnb staff.

Topic 2

Great location, private and a very **clean** and taken care of apartment. The **bed** was comfortable, the living **room** was spacious, **kitchen** fully equipped and organised and **bathroom** was **clean** and stylish. One of the best parts was the big **balcony** with a great **view**. Very good value for money and would definitely stay here again.

Topic 2

A couple of **things** you should note about this place: 1. The items they said would be available (i.e. shampoo, conditioner, etc.) were mostly missing or difficult to **find** in the apt. We ended up buying our own. 2. One of the **bedrooms** is pretty compact and does not have a **balcony**. Not a big deal but the biggest **problem** is that there is a **window** between us he **bathroom** and the **bedroom**, so anytime someone in another **room** goes to the restroom in the middle of the **night** a bright light shines into the **bedroom**. Not good.

Topic 3

The location is absolutely fantastic, within **walking distance** of the main **city centre** **attractions**. It's **close** to a good **supermarket** and several **metro stations**. There are plenty of **bars**, cafés and **restaurants** within 5 **mins** **walk** as well as one of the best views of the castle. It was perfect for a family of four with all the facilities you could **need** for whatever the length of your stay.

Topic 3

Fantastic apartment in the **heart** of the **Alfama** **neighborhood**. Small but has **everything** you **need**. My only warning is the **neighborhood** is steep, winding **hills** and a ton of stairs (while public transit gets you **close**) so not the best location for those not agile.

Fig. 7 Examples of positive and negative reviews for each topic