ENHANCED KNOWLEDGE IN SCIENCES AND TECHNOLOGY
e-ISSN: 2773-6385

EKST

Vol. 6 No. 1 (2026) 385-396
https://publisher.uthm.edu.my/periodicals/index.php/ekst

Sentiment and Text Analysis of Airbnb Reviews to Understand Traveller Preferences in Tokyo and Singapore

Ang Zhan Xiang[1], Khuneswari P Gopal Pillay[1*]

[1] *Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, UTHM Kampus Cawangan Pagoh, Hab Pendidikan Tinggi Pagoh, KM 1, Jalan Panchor, 84600 Pagoh, Muar, Johor, MALAYSIA.*

*Corresponding Author: khuneswari@uthm.edu.my
DOI: https://doi.org/10.30880/ekst.2026.06.01.045

Article Info

Received: 4 January 2026
Accepted: 20 January 2026
Available online: 9 July 2026

Keywords

Word Cloud, Sentiment Analysis, Topic Modelling, Airbnb Review

Abstract

Airbnb named by the name of airbed, breakfast and bed has become a major element of the sharing economy that transforms urban tourism by the means of peer-to-peer accommodation and user-created reviews. In this study, the text mining and natural language processing methods are utilised to analyse and compare the experience of travellers who stayed at Airbnb in Tokyo and Singapore. Airbnb reviews of both cities were gathered and pre-processed according to the standard NLP procedures. The word cloud visualisation based on TF-IDF was implemented to find the terms that were often discussed, whereas a lexicon-based method of sentiment analysis was adopted to determine the overall sentimental distributions. Latent Dirichlet Allocation (LDA) was then used to derive major themes that affect the satisfaction of the guests. The findings display that both cities show a predominantly positive sentiment with the average sentiment score of about 0.93 in Tokyo and 0.91 in Singapore though the neutral sentiments can be more hardly differentiated because of the linguistic vagueness and the imbalance of classes. Three themes were identified as quality of service, convenience of location and room condition. The guests of Tokyo were more concerned about being able to access transportation and being close to it, whereas the guests of Singapore were more concerned about cleanliness, hospitality, and fast check-in. Although the results were limited by the methodological constraints based on the granularity of sentiments, the results indicate the usefulness of TF-IDF and LDA in providing practical recommendations to Airbnb hosts and tourism policymakers.

1. Introduction

Short-term rentals, or STRs, as hosted by application companies like Airbnb, have impacted the hospitality sector greatly by introducing flexibility in accommodation and encouraging community-based travelling [1]. STR systems have brought in innovative types of micro-entrepreneurship and personalised travel experiences, which have sparked discussions around affordability of housing, regenerating neighbourhoods and regulatory governance. Digital transformation, changing consumer demands and mixed policy reactions have influenced their development, especially in urban markets in Tokyo and Malaysia, which attempt to achieve a balance between innovation and stability in the community. The guest reviews have therefore proved to be among the most useful in the tourism research and provide intimate insights to the travels, service quality and tourist attraction popularity [2][3]. Compared to numeric star rating, textual reviews offer more detailed accounts that display the concerns, preferences and expectations of the travellers which help the hosts to know which parts of the services require improvement [4][5].

To dissect these unstructured texts, structured natural language processing (NLP) pipeline of TF -IDF, lexicon-based sentiment analysis and Latent Dirichlet Allocation (LDA) were used. TF -IDF extracts and weights notable keywords by removing common yet non-informative words and highlighting context-dependent words. It has been shown to be an efficient feature-extraction tool in tourism analytics to reveal prevailing review patterns [6]. Sentiment analysis was performed through a lexicon-based method of categorising the emotional polarity into positive, negative and neutral. It is favoured due to its transparency and appropriateness to multilingual or low-resource datasets and has been used in e-reputation management and service assessment [7]. Nevertheless, it is subject to contextual vagueness and domain language. The recurring themes that were identified in medium-length Airbnb reviews were then subjected to LDA topic modelling, a method commonly utilised in tourism research because of its interpretability and policy applicability, although more complex models exist including BERT or GPT [8].

Despite the significant level of analytical rigour that these NLP tools provide, there were still some problems, such as the ambiguity of guest comments, cross-cultural language difference and limitations of lexicon-based sentiment dictionaries. Although sentiment analysis has gained popularity in tourism research, relatively few studies have applied lexicon-based approaches to Airbnb guest reviews in Asian urban contexts, particularly in highly regulated markets such as Tokyo where the regulatory framework and cultural demands do not resemble the Western market [7][9]. Existing literature continues to focus predominantly on Western short-term rental (STR) markets or examines tourism sentiment in aggregate, often overlooking the sociocultural and regulatory specificities of cities such as Tokyo and Singapore. This disconnect hinders the creation of cultural sensitive policy frameworks and the understanding of host expectations of guest in these rare markets. This lack of contextualised analysis limits the development of culturally sensitive policy frameworks and constrains understanding of host–guest dynamics in regulated Asian STR markets.

Tokyo is an example of a fast-growing STR setting, which is conditioned by the development of tourism and tough local policies. In 2017, 28.7 million tourists came to Japan, which is three times more than the number of tourists in five years [10] and subsequently legalised STRs to boost economic growth [11]. The Tokyo STR market is expanding despite the strict rules relating to zoning and operational limitations [12] with hosts adjusting to the regulatory demands. Singapore, however, has one of the most rigid STR systems in Asia which bans rentals less than three months in personal accommodation [13]. However, the constant need to find low-cost, genuine and native accommodation continues [14] and creates rich but under-analysed textual review information.

To fill these gaps, the study applied an integrated computational framework combining TF-IDF–based text mining, lexicon-based sentiment analysis and Latent Dirichlet Allocation (LDA) topic modelling on Airbnb guest reviews in Tokyo and Singapore to identify key insights into customer satisfaction and traveller preferences through topic modelling and word cloud analysis of Airbnb guest reviews and to examine the sentiments and experiences of travellers staying in Airbnb accommodations in Tokyo and Singapore using sentiment analysis and probabilistic topic modelling techniques. Besides that, practical and data-driven recommendations are aimed to provide for Airbnb hosts and policymakers to support service quality enhancement and informed decision-making in the governance of short-term rentals in both cities. The combination of these computational methods with cross-city comparison enables the work to provide a more context-sensitive and data-driven insight into the experiences of guests in Asian regulated markets of STR. The results are meant to guide hosts towards improvement of their services and guide policymakers towards formulation of more responsive and culturally sensitive policies on the governance of STR.

2. Research Methodology

2.1 Data Sources

The data applied in the present research was accessed from Inside Airbnb ranging from year 2011 to 2024. This open platform gives detailed and anonymised datasets on Airbnb listings and reviews of guests in large cities around the world [15]. The information behind the Tokyo and Singapore Airbnb reviews was specifically taken to suit the research scope. The Tokyo and Singapore datasets contained 755,482 and 3,7962 data points, respectively. The dataset consists of attributes such as listing ID, review date, reviewer ID and comments. Only the review text and corresponding metadata needed for text analysis were extracted for downstream processing. Inside Airbnb data is curated for independent research purposes and updated regularly to reflect market changes, ensuring the validity and usefulness of the dataset for academic purposes. Table 1 summarises the attributes of the Tokyo and Singapore Airbnb Review dataset

Table 1 Attributes of the Tokyo and Singapore Airbnb Review dataset

Feature	Type	Description
Listing_ID	Integer	A unique identifier for each Airbnb listing associated with the review.
ID	Integer	A unique identifier for each review entry.
Date	Date (DD-MM-YYYY)	The date when the reviewer posted the review.
Reviewer_ID	Integer	A unique identifier for the reviewer who wrote the review.
Reviewer_Name	String	The name or username of the reviewer.
Comments	String	The full text content of the guest's review describes their experience

2.2 Data Preprocessing

Prior to analysis, Airbnb reviews from Tokyo and Singapore were pre-processed to ensure data quality and consistency [16]. Reviews with missing text were removed, followed by standard cleaning procedures, including lowercase, removal of punctuation, numbers and stop words such as included platform and location terms such as 'airbnb', 'bnb', 'tokyo', 'singapore', host or personal names like 'Jerome' and 'dave', web-formatting artifacts such as 'nbsp', 'br', multilingual functional words and high-frequency generic words that did not contribute meaningful semantic or sentiment information like 'us', 'we', 'it' and 'u' [17]. English translation of non-English reviews was done in order to establish a single corpus. The word forms were standardised through tokenisation and lemmatisation to work out the dimensionality. Quality checks were done in order to reduce noise and maintain semantic structure to aid dependable sentiment analysis and topic modelling [18]. Online accommodation reviews have a high positive sentiment bias in the literature because guests tend to write positive reviews and social norms prevent them from being very negative [3]. In turn, sentiment distributions within Airbnb data are inherently skewed and positive reviews are the largest group compared to neutral and negative [5]. The imbalance was observed in this study and defined as a natural feature of peer-to-peer accommodation platforms and not as a methodological constraint. As the aim of this project was to investigate real-world sentiment patterns of travellers and not to create a balanced classifier, the initial distribution of reviews was retained. In the case of manual validation, the sample of 100 reviews per city was selected randomly so that the existing structural occurrence of sentiments could be representative and a possible systematic error in misclassification could be identified [8][17].

2.3 Word Cloud

The first visualisation of the text-mining was a word cloud that was used to identify the most commonly used words in Airbnb guest reviews. TF-IDF weighting was applied to emphasise informative terms and suppress common words, as defined in equation (1):

$$\boldsymbol{TF-IDF(t,d) = TF(t,d) \times \log\left(\frac{N}{DF(t)}\right)} \quad (1)$$

TF-IDF mitigates the effect of repeatedly used words and conspicuous ones. Such a strategy is popular in text mining and tourism analytics [19].

2.4 Sentiment Analysis

The sentiment analysis method that was adopted was lexicon based to examine the emotional tone of guest reviews. Polarity and subjectivity scores were computed using predefined sentiment lexicons, as defined in equation (2) [20]:

$$\boldsymbol{Polarity} \ \epsilon \ [-\mathbf{1}, \mathbf{1}], \boldsymbol{Subjectivity} \ \epsilon \ [\mathbf{0}, \mathbf{1}] \tag{2}$$

The technique is effective in terms of clear categorization of the reviews as either negative, neutral and positive. Lexicon-based sentiment analysis is commonly applied in tourism studies [21].

2.5 Topic Modelling

Latent Dirichlet Allocation (LDA) was applied to identify latent themes in Airbnb guest reviews. LDA models each document as a mixture of topics, with word probabilities defined by equation (3) [22]:

$$\boldsymbol{P(w|d) = \sum_{k=1}^{K} P(w|z=k) \cdot P(z=k|d)} \tag{3}$$

where K is the number of topics, $P(w|z=k)$ is the probability of word w given topic k, $P(z=k|d)$ is the probability of topic k in document d. The optimal number of topics was determined using coherence scores. LDA remains widely used due to its interpretability in tourism research [8].

2.6 Model Evaluation

To ensure the reliability of the applied text analysis methods, sentiment analysis and topic modelling were evaluated using quantitative metrics and qualitative validation techniques. Lexicon-based sentiment analysis tools, namely VADER and TextBlob were employed to classify guest reviews into positive, neutral and negative categories [23]. Although these methods are rule-based and do not involve supervised learning, their performance was assessed using conventional classification metrics through manual validation. A random sample of 100 reviews from the Tokyo and Singapore Airbnb Review dataset was manually labelled and used as ground truth. Since the number of reviews was extremely high, it was not possible to manually annotate the whole dataset. True to the conventional sentiment analysis practice, a random sampling of reviews was thus employed in the human validation of the polarity outputs [17][20]. Random sampling guaranteed that the chosen reviews lacked the bias to a particular sentiment category, property type or time and gave a representative representation of the entire corpus [17]. A sample size of 100 reviews per city was chosen because previous large scale tourism and social media research typically use small, manually inspected subsets to check the consistency of sentiments and reveal systematic error patterns [20]. This scale is adequate to indicate the prevailing errors and concordance with the human judgement, but is not infeasible in detecting the quality examination in a diligent manner. The predicted sentiment labels were compared against the human annotations, and performance was evaluated using accuracy, precision, recall and F1-score, defined by equation (4), (5), (6) and (7):

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \tag{4}$$

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \tag{5}$$

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP=FN}} \tag{6}$$

$$\boldsymbol{F1\ Score} = \mathbf{2} \times \frac{\boldsymbol{Precision \times Recall}}{\boldsymbol{Precision+Recall}} \tag{7}$$

where *TP*, *FP*, *TN* and *FN* denote true positives, false positives, true negatives, and false negatives, respectively. A confusion matrix was constructed to visualise classification performance and identify misclassification patterns, such as neutral reviews being incorrectly labelled as positive due to generic sentiment terms such as “good” or “nice” used out of context [20]. The confusion matrix structure is shown in Table 2

Table 2 Outcome of confusion matrix for the sentiment analysis

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positive	False Negative
Actual Negative	False Positive	True Negative

arising from preprocessing or lexicon misalignment [21].

For topic modelling, the effectiveness of Latent Dirichlet Allocation (LDA) was evaluated using topic coherence scores, which measure the semantic consistency of high-probability words within each topic [24]. Coherence was computed based on the co-occurrence frequency of top-ranked topic words across the review corpus. A simplified coherence formulation between topic words $\boldsymbol{w_i}$ and $\boldsymbol{w_j}$ is given by equation (8):

$$\boldsymbol{Coherence(w_i, w_j) = log \frac{D(w_i, w_j) + \epsilon}{D(w_j)}} \quad (8)$$

where $\boldsymbol{D(w_i, w_j)}$ denotes the number of documents containing both $\boldsymbol{w_i}$ and $\boldsymbol{w_j}$, $\boldsymbol{D(w_j)}$ denotes the number of documents containing $\boldsymbol{w_j}$, and $\boldsymbol{\epsilon}$ is a small constant to prevent division by zero. The coherence scores were used to guide topic selection and ensure the interpretability of extracted topics.

2.7 Overall Flowchart

Fig. 1 is a flow chart that summarises all the steps involved in the study. In the flow chart, data of Airbnb reviews from Tokyo and Singapore were analysed using the same methodology workflow to be comparable. Starting with data preprocessing which includes the elimination of miss entries and text normalisation such as lowercasing, elimination of punctuation, stop-word filtration, tokenisation and lemmatisation of texts in a dataset [16].

Non-English reviews that were in Japanese, Arabic, Tamil and German were translated into English in Google translate to form a homogenous corpus of languages to be analysed. After preprocessing, word cloud visualisations were created with raw term frequencies and weighting of TF-IDF to identify the common terms used [19]. Sentiment analysis was then undertaken to classify the individual reviews into categories of emotional valence such as positive, negative and neutral based on their score of polarity and subjectivity [21]. Finally, topic modelling was conducted using LDA which was used to determine the major themes through probabilistic topic distributions of the corpus [22].

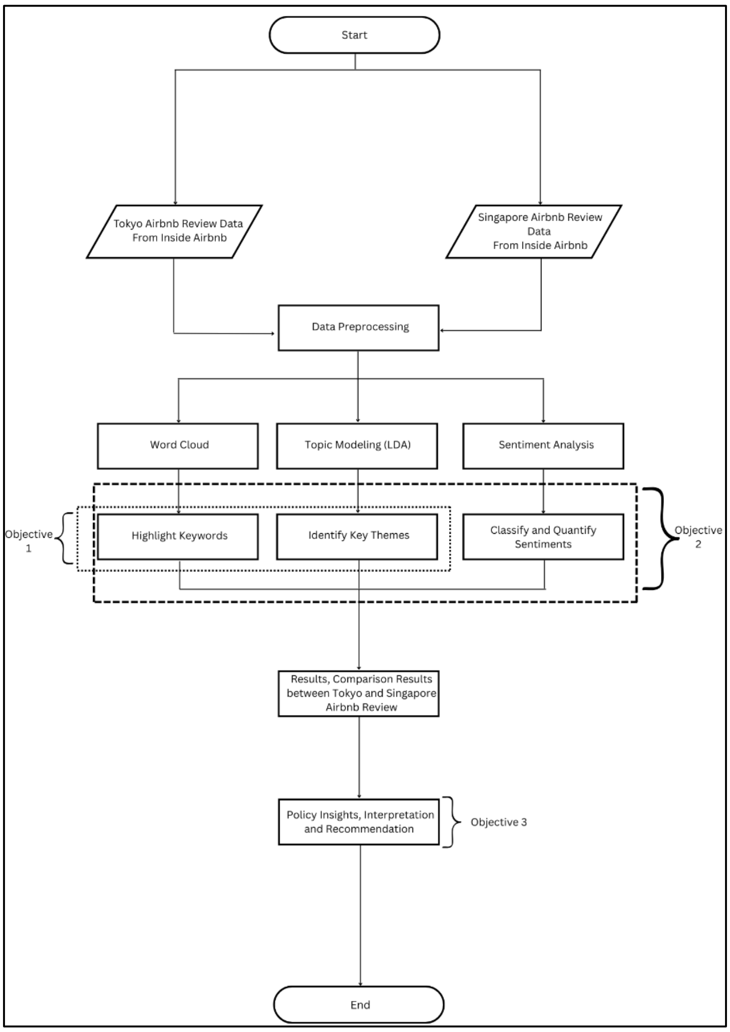

Fig. 1 Flow Chart

3. Results and Discussion

3.1 Data Preprocessing

Data preprocessing was conducted to ensure that the Tokyo and Singapore Airbnb review datasets were suitable for text mining, sentiment analysis and topic modelling. Initially, the datasets contained 755,482 and 37,962 reviews, respectively. Due to the presence of multilingual content and noisy or low-quality entries, preprocessing was necessary. The datasets were loaded using the pandas library with error-handling parameters to exclude corrupted records. Language detection was performed using langdetect, and non-English reviews were translated into English via the deep-translator library. Text normalization included the removal of HTML tags and non-textual symbols, lowercasing, tokenization, stop-word removal and lemmatization using NLTK. Reviews with fewer than three words, generic expressions, or only symbols were excluded. After preprocessing, 443,235 Tokyo reviews and 31,557 Singapore reviews remained.

3.2 Data Descriptive

Based on Fig. 2 and Fig. 3, the data showed distinct growth and decline patterns that correspond to tourism trends from year 2011 to 2024. Although both cities showed a significant decline in the number of reviews in 2020 and 2021, they managed to recover and reached their peak numbers in 2024. Each dataset has a peak number of reviews of 42.3% of their total reviews in Tokyo and 19% of their total reviews in Singapore.

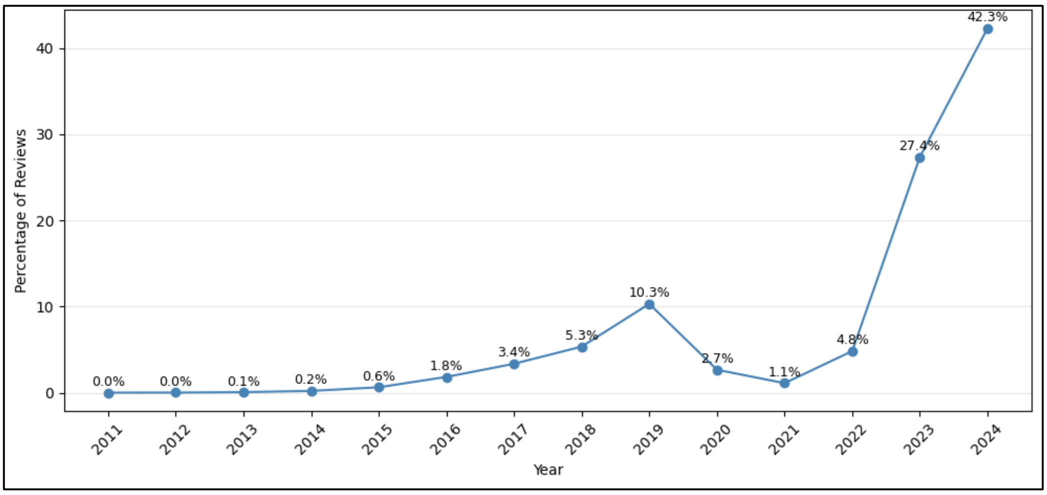

Fig. 2 Yearly Percentage Distribution of Tokyo Airbnb Review Dataset

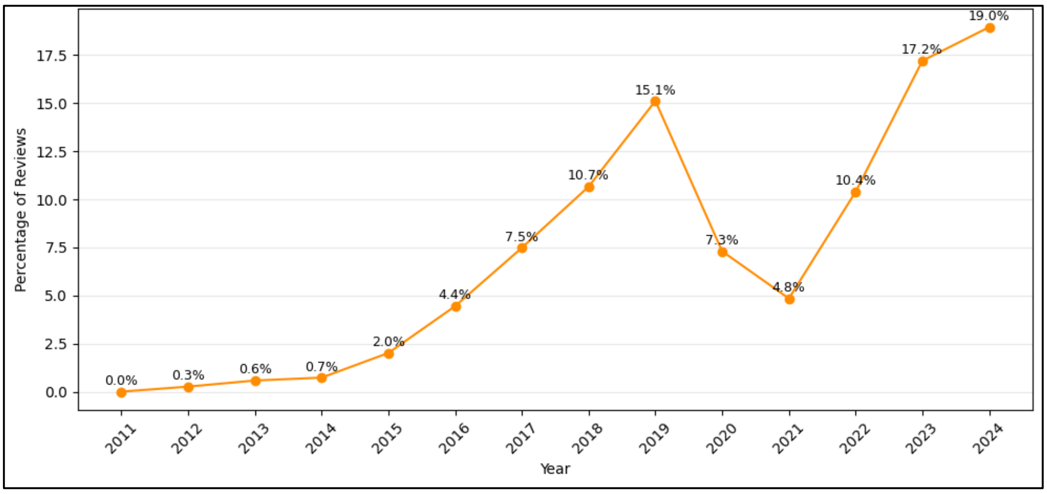

Fig. 3 Yearly Percentage Distribution of Singapore Airbnb Review Dataset

3.3 Word Cloud

(a) (b)

Fig. 4 Word cloud of Airbnb review dataset (a) Tokyo;(b) Singapore

According to Fig. 4, the word clouds of Tokyo and Singapore revealed the fact that positive words prevail in the review of guests in both cities, which points to rather positive Airbnb experiences. The repetitive words such as clean, good, nice, host, recommend, indicate the emphasis on the ideas of cleanliness, location and general quality of stay as the key to the happiness of guests in every place. Yet, one evident contextual disparity lies in the description of what people say about their stay. The word cloud of Tokyo is also rich in words related to mobility such as station, train, walk, minute and distance meaning the importance of the transport accessibility to the visitors. This is very logical since Tokyo is so crowded and trains constitute a very big percentage of its transportation system such that being near a station is a mound of convenience. The constant repetition of such words like easy and close only underline that smooth movement is one of the main elements of satisfaction there. Singapore, however, focuses more on service and comfort conditions which include staff, friendly, helpful, comfortable, room, implying that individuals are sensitive to the human touch and the quality of the living space. The fact that the words such as responsive, check and recommend are present show that positive reviews there are enhanced by the presence of good communication and easy processes. Overall, both cities are highly rated in terms of the satisfaction of their guests, but the word clouds indicate city-specific priorities. Tokyo guests are not dependent on hospitality, comfort, and service quality, but rather on the convenience and accessibility of the services. These insights contribute to the understanding of why it is important to align the offerings of Airbnb to local expectations, and they blend well with the sentiment and topic modelling outcomes.

3.4 Sentiment Analysis

As Fig. 5 shows, Tokyo has 421,613 positive, 13,158 neutral and 8,464 negative reviews. Whereas there were 29,530, 1,126 and 901 positive, neutral and negative reviews respectively in Singapore. These findings show a majority of Airbnb. customers of the two cities reported positive feedback of their stay experiences, indicating that general customer perception towards Airbnb in the two markets is very positive.

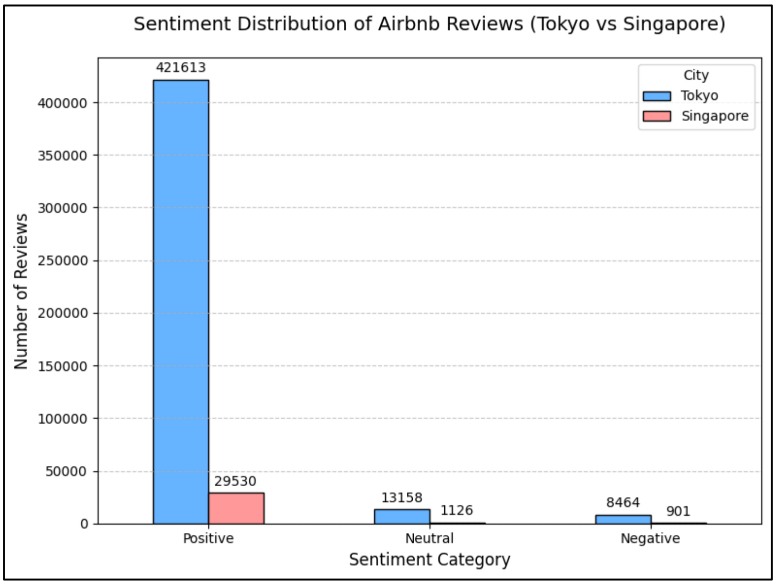

Fig. 5 Sentiment Analysis of Tokyo and Singapore Airbnb Reviews

According to Table 3 and 4, the model sentiment analysis of Tokyo and Singapore Airbnb reviews show that there is consistent. good model performance, especially at the positive sentiments' classification. The Tokyo model accomplished a better result. the overall accuracy (0.93) in comparison to Singapore (0.91) which is slightly higher in classification consistency. In both datasets, the model showed accuracy of high degree, and the values of Tokyo and Singapore were 0.96 and 0.94, respectively, and remember, with the values of 0.99 and 0.97 of Tokyo and Singapore, respectively, of positive sentiments which implies a good estimation of strongly expressed opinions. Whereas negative sentiment classification was acceptable in both conditions, Singapore was perfect in precision but slightly less recall (0.67) indicating. some false categorisations of negative reviews. Conversely, both models were not very effective at recognizing neutral. sentiment with close precision and recall, which is shown by the confusion matrices. This pattern indicates that misclassifications happened mostly between the classes of neutral and polar sentiment. Overall, the results confirm the strength of the model in reflecting clear emotional polarity, and that there is a similar weakness in treating. impregnable or unclear review material.

Table 3 Outcome of detailed key performance metrics for the Tokyo Airbnb review sentiment analysis

Sentiment	Precision	Recall	F1-Score	Support
Negative	0.75	0.75	0.75	4
Neutral	0.00	0.00	0.00	5
Positive	0.96	0.99	0.97	91
Overall Accuracy	0.93	—	—	100

Enhanced Knowledge in Sciences and Technology Vol. 6 No. 1 (2026) 385-396

Table 4 Outcome of detailed key performance metrics for the Singapore Airbnb review sentiment analysis

Sentiment	Precision	Recall	F1-Score	Support
Negative	1.00	0.67	0.80	3
Neutral	0.00	0.00	0.00	5
Positive	0.94	0.97	0.95	92
Overall Accuracy	0.91	—	—	100

3.5 Topic Modelling

According to Table 5, the three most coherent topics were chosen to be further interpreted because the high coherence means that the topic keywords have higher semantic consistency and interpretability. In the case of Tokyo dataset, Topics 2 (0.8082), 10 (0.7218) and 4 (0.6414) had the highest scores in coherence, indicating that the topics are clear and significant themes in guest reviews. The coherence of Topic 2 was the highest of all Tokyo topics, which proves a discussion with a very strong focus, probably, narrowed down to particular functional features of the accommodation. Topics 10 and 4 also had a relatively high level of coherence, which was manifested in the consistent guest commentary based on common experiential qualities.

Topics 10 (0.7546) and 9 (0.6530) and Topic 5 (0.6521) were the most coherent topics in the Singapore dataset. Topics 10 had the highest Coherence score in general which indicated there was a specific and meaningful theme in Singapore reviews. Topics 9 and 5 were soon behind, which pointed to the fact that the guests often spoke about these issues in a locally and contextually consistent way. Singapore scores on coherence display a more equal distribution of the most discussed topics, which leads to a variety of consistently mentioned experiences in contrast to Tokyo. A comparative analysis shows that in both datasets there is at least one very consistent topic in connection with core accommodation features, which is manifested in the high performance of Topic 10 in both cities. Nonetheless, there is a stronger contrast in the number of highest-ranked and lowest-ranked topics in Tokyo, as it could indicate that a group of most popular topics is more focused on guest conversations. By contrast, the topics in Singapore have a lesser degree of coherence variation, which means that the group of concerns and experiences is more homogeneous. These trends rationale the choice of the three best topics in every city and advocate the healthiness of the LDA model in summarizing the salient themes of large-scale Airbnb review data.

Table 5 Top 10 Tokyo and Singapore Topic Coherence Scores

Tokyo		Singapore	
Topic	**Coherence Score**	**Topic**	**Coherence Score**
1	0.5362	**1**	0.3778
2	**0.8082**	**2**	0.5020
3	0.5542	**3**	0.4018
4	**0.6414**	**4**	0.4354
5	0.5868	**5**	**0.6521**
6	0.4233	**6**	0.3907
7	0.3992	**7**	0.5331
8	0.4788	**8**	0.5851
9	0.5663	**9**	**0.6530**
10	**0.7218**	**10**	**0.7546**

According to Table 6 and Table 7, Topics 2, 4 and 10 had the best scores in coherence in Tokyo, and they indicated themes of household facilities, room design and space comfortability. These issues show that the guests of Tokyo are more interested in the practical utility, the space and functionality, especially in small city apartments. Topics 5, 9 and 10 were the most coherent in the case of Singapore, with the emphasis on the check-in experience, accessibility of the location, and the state of the rooms. The findings indicate that efficiency of services, location in terms of accessibility to transport and facilities and room quality are at the heart of guest satisfaction in Singapore. Comprehensively, the three priority themes common in both destinations include service quality, location and room conditions. Nevertheless, the Tokyo guests are more concerned about the applicability of appliances and the space organization, and the Singapore ones are more concerned about the effective service process, hygiene and the maintenance of facilities. These diversities represent the guest expectations based on local urbanism.

Table 6 Tokyo Airbnb Review LDA Models

Topic	LDA Models
2	0.0199(machine)+0.0197(provided)+0.0188(washing)+0.0168(airport)+0.0140(use)+0.0136(kitchen)+0.0134(instruction)+0.0107(clothes)+0.0103(dryer)+0.0103(laundry)
4	0.0203(room)+0.0135(floor)+0.0132(good)+0.0115(luggage)+0.0085(issue)+0.0084(stair)+0.0082(location)+0.0080(overall)+0.0079(stay)+0.0076(bathroom)
10	0.0534(bed)+0.0374(room)+0.0371(space)+0.0277(people)+0.0253(two)+0.0221(comfortable)+0.0218(bathroom)+0.0218(group)+0.0195(family)+0.0164(kitchen)

Table 7 Singapore Airbnb Review LDA Models

Topic	LDA Models
5	0.0690(check)+0.0374(easy)+0.0334(checkin)+0.0249(host)+0.0130(process)+0.0123(early)+0.0118(late)+0.0113(responsive)+0.0111(instruction)+0.0103(luggage)
9	0.0210(walking)+0.0199(mrt)+0.0190(distance)+0.0165(location)+0.0143(walk)+0.0137(restaurant)+0.0117(kitchen)+0.0109(within)+0.0107(station)+0.0106(food)
10	0.0481(room)+0.0140(bathroom)+0.0136(bed)+0.0096(hotel)+0.0089(good)+0.0088(shower)+0.0069(didnt)+0.0069(clean)+0.0069(noise)+0.0063(door)

4. Conclusion

The objectives of this study were successfully achieved since it used text mining, sentiment analysis and topic modelling to analyse Airbnb reviews of Tokyo and Singapore. The analysis was conducted with the help of the Latent Dirichlet Allocation (LDA) and TF-IDF-based word cloud visualisation to determine three prevalent factors of guest satisfaction, which were service quality, location convenience and room conditions [22]. Tokyo guests prioritized accessibility and closeness to transportation systems with top 3 topics in topic modelling stating words like 'provided', 'instruction' and 'use' whereas Singapore guests stressed more on cleanliness, friendliness of the staff and a fast check-in experience where their top 3 topics in topic modelling results states words such as 'didn't', 'clean' and 'check-in', which is consistent with the previous results in peer-to-peer accommodation and tourist studies [25]. These findings indicate that the implemented text-mining structure is useful in identifying valuable information in the data of large-scale review [16]. The research has also managed to analyse the sentiments of the travellers using an unsupervised sentiment analysis method which indicated a mostly positive sentiment distribution of the two cities. The model of sentiment classification was very high with 0.93 in Tokyo and 0.91 in Singapore, which proved the model to be reliable and consistent to previous sentiment analysis studies in tourism [21]. In spite of the problems that were also noted in categorizing neutral sentiments because of linguistic ambiguity, which is a common issue in computational linguistics, the model was successful in discriminating positive and negative opinions [17]. These results support the importance of online reviews as an informative source to evaluate the experience and satisfaction of travellers [7]. The study had a number of limitations even though it contributed to it. Depending on machine translation created potential semantic distortion and more time was consumed on processing and preprocessing presented problems caused by noisy user generated text that had to be refined by iteration [29]. Future studies need to discuss more analytical frameworks to a more comprehensive representation of neutral and subtle sentiments in Airbnb reviews with respect to the sharing economy, where the motivation and articulation of users are frequently multifaceted and complicated [1]. Preprocessing pipeline strengthening and incorporation of domain-specific hospitality literature can additionally be used to increase interpretive depth. In general, this study confirms the given methodological model and offers practical recommendations to Airbnb hosts and policymakers in order to benefit the data-driven enhancement in service quality, regulation, and tourism analytics [28].

Acknowledgement

The authors would thank the Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, for its support.

Conflict of Interest

Authors declare that there is no conflict of interest regarding the publication of the paper.

Author Contribution

*The authors confirm their contributions to the paper as follows: **study conception and design**, Ang Zhan Xiang, Khuneswari P Gopal Pillay; **data collection**, Ang Zhan Xiang; **analysis and interpretation of results**, Ang Zhan Xiang; **draft manuscript preparation**, Ang Zhan Xiang, Khuneswari P Gopal Pillay. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] Hamari, J., Sjöklint, M., & Ukkonen, A. (2016). The sharing economy: Why people participate in collaborative consumption. Journal of the association for information science and technology, 67(9), 2047-2059. https://doi.org/10.1002/asi.23552

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

[3] Zervas, G., Proserpio, D., & Byers, J. W. (2015). The Impact of the Sharing Economy on the Hotel Industry. Proceedings of the 16th ACM Conference on Economics and Computation. https://doi.org/10.1145/2764468.2764524

[4] Lee, C. K. H., Tse, Y. K., Zhang, M., & Ma, J. (2020). Analysing online reviews to investigate customer behaviour in the sharing economy: The case of Airbnb. Information Technology & People, 33(3), 945–961. Retrieved from https://doi.org/10.1108/ITP-10-2018-0475

[5] Santos, G., Mota, V. F. S., Benevenuto, F., & Silva, T. H. (2020). Neutrality may matter: Sentiment analysis in reviews of Airbnb, Booking, and Couchsurfing in Brazil and USA. arXiv. https://doi.org/10.48550/arXiv.2005.06591

[6] Ali, T., Omar, B., & Soulaimane, K. (2022). Analyzing tourism reviews using an LDA topic-based sentiment analysis approach. MethodsX, 9, 101894. https://doi.org/10.1016/j.mex.2022.101894

[7] Mehraliyev, F., Chan, I. C. C., & Kirilenko, A. P. (2022). Sentiment analysis in hospitality and tourism: a thematic and methodological review. International Journal of Contemporary Hospitality Management, 34(1), 46-77. https://doi.org/10.1108/IJCHM-02-2021-0132

[8] Kirilenko, A. P., & Stepchenkova, S. (2025). Facilitating topic modeling in tourism research: Comprehensive comparison of new AI technologies. Tourism Management, 106, 105007. https://doi.org/10.1016/j.tourman.2024.105007

[9] Jabbar, M. A., Okfantia, A. D., Widjanarti, A., & Zulen, A. A. (2022). Sentiment analysis of tourist reviews from online travel forum for improving Indonesia tourism sector. In Data science in central banking: Applications and tools (Vol. 59). Bank for International Settlements. Retrieved from https://www.bis.org/ifc/publ/ifcb59_18.pdf

[10] Japan National Tourism Organization. (2018, December 21). Number of visitor arrivals to Japan reached 30 million for the first time! Retrieved from https://www.japan.travel/en/news/visitor-arrivals-reached-30-million/

[11] Agrusa, J. (2020). The Airbnb effect: how tourist destinations are addressing short-term vacation rentals. International Journal of Qualitative Research in Services, 4(1), 20-29. https://doi.org/10.1504/IJQRS.2020.10031944

[12] Von Briel, D., & Dolnicar, S. (2020). The evolution of Airbnb regulation-An international longitudinal investigation 2008-2020. https://doi.org/10.1016/j.annals.2020.102983

[13] Urban Redevelopment Authority (URA). (2020). FAQs on Short-Term Accommodation. Retrieved from https://www.ura.gov.sg

[14] [Yeo, S. (2016). Singapore's stance on home sharing. The Straits Times. Retrieved from https://www.straitstimes.com

[15] Murray Cox. (2016). Inside Airbnb. Retrieved from https://insideairbnb.com

[16] Feldman, R., & Sanger, J. (2007). The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge university press.

[17] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in information retrieval, 2(1–2), 1-135. https://doi.org/10.1561/1500000011

[18] Bengio, Y. (2012). Deep learning of representations for unsupervised and transfer learning. In Proceedings of ICML workshop on unsupervised and transfer learning (pp. 17-36). JMLR Workshop and Conference Proceedings. Retrieved from https://proceedings.mlr.press/v27/bengio12a

[19] Ramos, J. (2003). Using TF-IDF to determine word relevance in document queries. In Proceedings of the First Instructional Conference on Machine Learning, 133–142.

[20] Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225). Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14550

[21] Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5(1), 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016
[22] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022. Retrieved from http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf
[23] Loria, S. (2018). textblob Documentation. Release 0.15, 2(8), 269. Retrieved from https://app.readthedocs.org/projects/textblob/downloads/pdf/dev/
[24] Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In Proceedings of the eighth ACM international conference on Web search and data mining (pp. 399-408). https://doi.org/10.1145/2684822.2685324
[25] Banerjee, S., & Chua, A. Y. (2015). In search of patterns among travellers' hotel ratings in TripAdvisor. Tourism Management, 53, 125–131. https://doi.org/10.1016/j.tourman.2015.09.020
[26] Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. Tourism Management, 68, 301–323. https://doi.org/10.1016/j.tourman.2018.03.009
[27] Wong, T. S., & Chan, J. K. L. (2023). Experience attributes and service quality dimensions of peer-to-peer accommodation in Malaysia. Heliyon, 9(7). https://doi.org/10.1016/j.heliyon.2023.e18403
[28] Ahmed, T. T. (2020). The influence of reputation, price, trust and perceived value on guest booking intention on Airbnb in Malaysia [Master's thesis, Universiti Teknologi Malaysia]. UTM Institutional Repository. Retrieved from https://eprints.utm.my/108046/
[29] Balahur, A., & Turchi, M. (2014). Multilingual sentiment analysis using machine translation. Retrieved from https://www.researchgate.net/publication/262215708