

# Indoor environmental quality satisfaction in Australian hotels and serviced apartments

Fan Zhang <sup>1\*</sup>, Karthick Seshadri <sup>2</sup>, Devi Vara Prasad Pattupogula <sup>2</sup>, Badrinath Chandana <sup>2</sup>, and Shichao Liu <sup>3</sup>

<sup>1</sup>School of Engineering and the Built Environment, Griffith University, Southport, Australia

<sup>2</sup>Department of Computer Science and Engineering, National Institute of Technology Andhra Pradesh, Tadepalligudem, India,

<sup>3</sup>Department of Civil, Environmental, and Architectural Engineering, Worcester Polytechnic Institute, MA, USA

**Abstract.** Tourism is Australia's fourth-largest exporting sector, yet little research has been done on how satisfied guests are with the indoor environmental quality (IEQ) of Australian guest homes. This research project utilized web-mining, natural language processing and sentiment analysis to analyse customers' IEQ satisfaction in Australian tourist accommodations across ten tourism cities. By analysing 543,213 guest reviews from 1,397 hotels and serviced apartments with two-stars and above at the Booking.com, guests' text comments were classified by semi-supervised word-embedding based models into nine IEQ dimensions. Using a bespoke deep sequence model, sentiment polarities were found, and sentiment scores were computed to estimate the degree of IEQ satisfaction. Results showed that guests were most dissatisfied with facilities, cleanliness and maintenance, and acoustics. As the buildings' star ratings increased, dissatisfaction towards thermal environment, indoor air quality (IAQ), and acoustics decreased. Some IEQ dimensions displayed seasonal trends in customer dissatisfaction. The main sources of dissatisfaction with the thermal environment, IAQ, lighting, and acoustics were identified.

## 1 Introduction

With 8.2% of Australia's total export revenue, tourism was the country's fourth-largest exporting sector [1]. In 2019, tourism in Australia accounted for 3.1% of the national gross domestic product (GDP), contributing \$61.9 billion to the Australian economy [2]. Hotels and serviced apartments are the common types of tourist accommodations in Australia. Serviced apartments usually offer a separate lounge or dining area in addition to a kitchen or kitchenette, and often cost less than comparable hotel rooms because they have less elaborate facilities and offer fewer regular services [3]. In 2022, Australia has 606 hotels and resorts, and 1,260 serviced apartments [3]. Although the indoor environmental quality (IEQ) of hotels and serviced apartments plays a significant role in visitors' comfort, health, and travel experience, it has not been adequately studied. There exists a dearth of research on the indoor IEQ performance of hotels and serviced apartments in Australia.

### 1.1 Indoor environmental quality evaluation

IEQ refers to the quality of an indoor space, which comprises of several factors, including thermal environment, indoor air quality (IAQ), lighting, acoustics, ergonomics, individual control, cleanliness,

and maintenance, to name a few. IEQ is assessed by objective measurements, subjective surveys, or the combination of both [4]. Post occupancy evaluation (POE) is a systematic way of assessing a building's performance after being occupied. By utilizing validated and standardized questionnaires to investigate occupants' prior experience within a building, POE approach has been widely applied in evaluating IEQ satisfaction in commercial buildings worldwide. Building Occupancy Survey System Australia (BOSSA) [5] is a POE tool in Australia, which comprises nine IEQ dimensions—spatial comfort, individual space, IAQ, thermal comfort, noise distraction & privacy, visual comfort, personal control, connection to outdoor environment, and building image & maintenance. Roumi et al. [4] reported that commonly studied IEQ dimensions included thermal environment, IAQ, lighting, acoustics, privacy, furnishing, personal control, cleanliness and maintenance, and available space.

### 1.2 IEQ assessment using online reviews

Due to lack of support from hotel managers and transient nature of visitors, the traditional POE approach to investigate tourists' IEQ satisfaction in the temporary accommodation may be difficult to implement. One attempt to use the POE method to investigate the IEQ in hotels was made by Qi et al. [6]. Nonetheless, most invited participants either declined to participate or did

\* Corresponding author: [fan.zhang@griffith.edu.au](mailto:fan.zhang@griffith.edu.au).

not complete all the questions. Therefore, it was difficult to use the gathered responses for their research analyses. They finally decided to investigate the IEQ perception by analysing the online hotel reviews.

Customers can comment on any aspect of the accommodation, including IEQ concerns, in online reviews that they submit voluntarily. These unsolicited reviews eliminate the potential bias of leading questions. There have been several research studies that have utilized web-mining, natural language processing (NLP) and sentiment analysis techniques to explore the IEQ satisfaction in hotels and apartments. In addition to Qi et al. [6], Shen et al. [7] assessed the IEQ complaints in China’s budget hotels; Suh et al. [8] looked at the IEQ issues in five-star hotels in South Korea; Villeneuve and O’Brien [9] examined 1.35 million Canadian Airbnb reviews collected in 6 cities between 2009 and 2019. Guo et al. [10] utilized 16,761 online reviews to compare occupant satisfaction in 232 Leadership in Energy and Environmental Design (LEED) certified buildings and 129 non-LEED-certified buildings in the US. Despite adding to our understandings of IEQ satisfaction in hotels and apartments, these studies only looked at a portion of IEQ dimensions. IAQ, thermal environment, lighting and acoustics were the main IEQ dimensions that previous studies were focused on. To fully understand customer satisfaction and assess IEQ performance in guest accommodations, we must delve into as many IEQ dimensions as we can.

In this study, occupant satisfaction in nine IEQ dimensions—thermal environment, IAQ, lighting, acoustics, available space, facilities, exterior view, cleanliness & maintenance, layout & design—in Australian guest homes are investigated. The nine IEQ dimensions were adapted from the BOSSA tool [5] originally intended for office buildings, to suit the requirements of tourist lodgings. We did not include IEQ dimensions that were infrequently mentioned in the online reviews or that could not be accurately/reliably distinguished from other IEQ dimensions or general issues using lexicons. This study is among the earliest to investigate IEQ satisfaction in Australian hotels and serviced apartments using web-mining and NLP techniques.

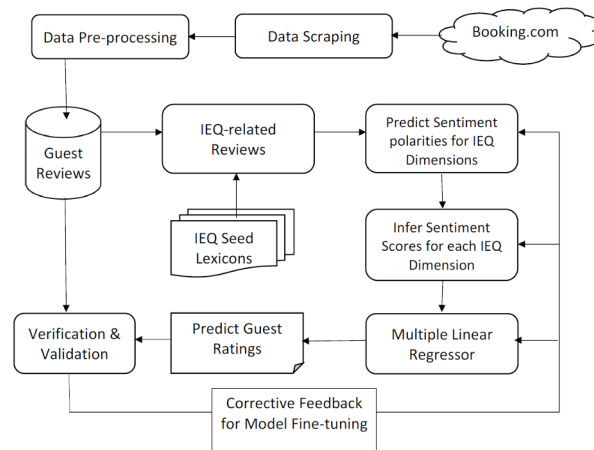
### 1.3 Aims of the study

This project aims to collect and analyse customer reviews posted on Booking.com for Australian hotels and serviced apartments across 10 major tourism cities, i.e., Sydney, Melbourne, Brisbane, Gold Coast, Sunshine Coast, Perth, Adelaide, Canberra, Hobart, and Darwin. By applying NLP and sentiment analysis techniques, we aim to answer the following questions:

- (i) How dissatisfied are people with IEQ across the nine dimensions?
- (ii) How do star ratings of the premises, climates, and seasons affect IEQ dissatisfaction?
- (iii) What are the main sources of low IEQ satisfaction?

## 2 Methods

The research approach comprised of the following sequential steps: (i) Data scraping from Booking.com. (ii) Data pre-processing. (iii) Filtering and categorization of each review sentence into one of the nine IEQ dimensions of interest. (iv) Predict sentiment polarities (satisfied, neutral, and dissatisfied) for one or more IEQ dimensions present in the sentence. (v) Inference of sentiment scores for each IEQ dimension present in a sentence. (vi) Prediction of the individual guest rating based on the IEQ sentiment scores and biases. **Figure 1** depicts the research methods flowchart for this project.



**Fig. 1.** Research methods flowchart

### 2.1 Data scraping and pre-processing

A crawler was implemented to scrape guest reviews from the Booking.com website with the following filters: (i) City: Sydney, Melbourne, Brisbane, Gold Coast, Sunshine Coast, Perth, Adelaide, Hobart, Darwin, and Canberra. (ii) Property Type: Apartments, Hotels (iii) Star Rating: 2, 3, 4, 5. Bookings has another category of star rating called "unrated," which was not included in scraping. Implementing the above search criteria, the crawler collected the following information and wrote it into an .csv file: Hotel name, suburb, city, star rating, overall guest rating, nationality of the guest, room type, duration of stay, check-in month and year, individual guest rating, positive comments, and negative comments (**Figure 2**). Reviews written in languages other than English were translated into English. The resulting dataset contained 1,470,709 lines of data. Booking.com displays guest reviews up to three years old. The data were scraped in May 2022. Therefore, the resultant dataset contained guest reviews lodged between May 2019 and May 2022.

Data pre-processing was then implemented to clean the dataset. Some online reviews were extremely brief. Previous studies have demonstrated a clear relationship between a review's length and its usefulness, with longer reviews often being more helpful [6,11]. Therefore, the reviews were removed from the dataset if both the positive and negative comments comprised fewer than five words. After pre-processing, the dataset contained 759,877 lines of data from 1,402 hotels and serviced apartments.



**Fig. 2.** An example of collected data from Booking.com

**2.2 Screening IEQ-related reviews**

A guest's review comment may consist of multiple sentences, each addressing a distinct IEQ dimension with a different tone, or no IEQ dimension at all. Therefore, the aforesaid dataset was expanded from the level of reviewers to the level of sentences. Each reviewer can have multiple rows of data, with one sentence per row. Next, semi-supervised word-embedding models were used to classify guests' text comments into nine IEQ dimensions. Only sentences pertaining to at least one IEQ dimension was kept in the dataset. Sentences that did not address any IEQ dimensions were removed from the database.

The model was trained using a manually curated vocabulary of IEQ seed lexicons for each IEQ dimension (Table 1). The Word2Vec model [12] was used to discover a vector representation for the IEQ dimension. We determined the embedding vector for each review sentence by identifying the centroid of the embeddings of the IEQ terms contained in the sentence. We categorised sentences into IEQ dimensions based on the distance between the vector of the sentence and the vector of the dimension. We defined a distance threshold, and mapped sentences to the IEQ dimension only when the distance surpassed the threshold. In some situations, multiple IEQ dimensions may be close to the same sentence; in these circumstances, the sentence has been mapped to the top three IEQ dimensions to which it is closest.

Almost all IEQ-related review sentences addressed one to three IEQ dimensions. For validation, it is imperative to ensure that the test data encompass all individual IEQ dimensions, and their two-way and three-way combinations. To validate the categorization method, we have carefully curated the test set by using a folded validation approach that maintains a fixed proportion of the 1, 2 and 3-combinations of IEQ dimensions that corresponds to the ratio of these combinations observed in each IEQ dimension. Based on our observations, it is rare to have a review which covers more than 3 IEQ categories, and hence validation of categorization accuracy for sentences containing four or more IEQ combinations are not needed. The semi-supervised bespoke IEQ categorization model produced high accuracy levels for the majority of the IEQ dimensions and their combinations.

**Table 1.** IEQ seed lexicons for nine IEQ dimensions

IEQ Dimension	IEQ seed lexicons
Thermal Environment	warm, cold, freezing, chilly, hot, heat, scorching, sweltering, melt, sweat, humid, muggy, clammy, steamy, draughty, drafty, temperature
Indoor Air Quality	fresh air, bad air, air circulation, ventilation, aerate, breeze, smell, smoke, stuffy, stink, stunk, stench, reek, airless, stale, odour, airtight, musty, mould, air purifier, air flow, airy
Lighting	bright, dark, glare, dazzle, dim, dusky, light, glow, lamp, block out
Acoustics	noise, loud, quiet, sound, sound insulation, soundproof, disturbance, silent, hear, overhear, rowdy, roaring, rattling, blaring, racket
Available Space	spacious, large, small, enough room, tiny, cramped, compact, congested, poky, size, roomy
Facilities	facilities, amenities, broken, damaged, cracked, not work, holes, chipped, air conditioner, air conditioning, aircon, AC, lifts, elevator, bed, furniture, fridge, tv, toilet, balcony, kitchen, laundry, towels, hot water, shower head, fly screens, faulty, pool
Exterior View	view, overlooking, facing, outlook
Cleanliness & Maintenance	clean, cleanliness, hygienic, tidy, spotless, dirt, disgusting, dusty, filthy, dingy, grubby, stain, wash, scrub, leaking, cockroach, bugs, insects, cobwebs, rubbish bin, housekeeping, maintenance, maintained, renovated, renovation, refurbished, updating, modern, décor, decoration, decent, dated, old, run down, dilapidated, rusty, art deco, scum
Layout & Design	layout, design

Despite obtaining excellent categorization accuracy, the phrases that were mapped to distinct IEQ dimensions nevertheless contained a plethora of non-IEQ-related comments. This was due to the various contexts of the IEQ vocabulary used to train the models. For instance, "warm" can refer to room temperature, which is IEQ-related, but can also refer to the warmth of the food or the attitude of the employees, which are unrelated to IEQ. After IEQ categorization, a manually selected list of irrelevant phrases was utilised to remove sentences with non-IEQ contexts. The composition of the irrelevant phrases modified those employed in comparable studies [6,7,9] to the current dataset's contents. The resulting collection of data included 948,296 IEQ-related review sentences from 543,362 guests. Data cleaning was then carried out to remove duplicate reviews. The final database contains reviews of 543,213 guests from 188 nations who stayed in 1,397 Australian hotels and serviced apartments in ten cities.

**2.3 Prediction of Sentiment polarities and scores**

In this step, sentiment polarities were predicted for each IEQ dimension by training a bespoke bidirectional long-short term memory (BI-LSTM) deep neural network model [13,14], and sentiment scores were calculated on a scale of 1 to 9 to determine the IEQ satisfaction level.

We have trained a distinct BI-LSTM-based deep neural network classifier for each IEQ dimension to classify each review-IEQ dimension combination into a 3-point sentiment polarity scale, i.e., "satisfied", "neutral", and "dissatisfied". Nonetheless, this required

supervision in teaching the classifier to differentiate between these scales for various IEQ dimensions. The review sentence was analysed using sentiment lexicons, such as WordNet and SenticNet, to determine the appropriate sentiment polarity for the given IEQ dimension. Using the IEQ categorization output and sentiment lexicons, we instructed the classifier to distinguish between sentiment polarity scales. After adding regularisation layers to prevent the model from overfitting the training data, the model achieved a good prediction accuracy for all IEQ dimensions in the range of 92% to 96%.

Based on the probability density acquired by the model for each IEQ dimension, we've developed an algorithm to generate sentiment scores for each IEQ dimension included in the review. The probability density with respect to a sentiment polarity and IEQ dimension denotes the likelihood of the review sentence conveying the sentiment polarity for the IEQ dimension. A neutral polarity results in a score of "5"; a positive sentiment polarity results in a score greater than "5", and a negative polarity results in a score lower than "5". If a specific IEQ dimension is not recognised in the sentence, the polarity for this IEQ dimension will be "neutral" and the sentiment score will be "5". Positive sentiment polarities with higher probability densities brings the sentiment score closer to 9; conversely, negative sentiment polarities with higher probability densities brings the sentiment score closer to 1. If the model predicts with a density score of 1 that a sentence expresses a "satisfied" sentiment polarity for a certain IEQ dimension, then it was translated into a sentiment score of 9.

## 2.4 Statistical analysis

Data were aggregated to the hotel level for statistical analysis. The percentage of IEQ dissatisfaction in guest homes was of primary interest in this study, which can be obtained by calculating the proportion of "dissatisfied" reviews in each premise. Similar method was adopted in Qi et al. [6]. To investigate the impact of star ratings, climate zones, and seasons on IEQ dissatisfaction, ANOVA test with Games-Howell *post hoc* pairwise comparison tests were carried out. The partial omega-squared  $\omega^2$  was used as a measure of effect size in ANOVA. Based on 1000 bootstrap samples, the effect size's 95% confidence interval was calculated. All statistical significance level was set to  $p < 0.05$ . Statistical analyses and data visualization was conducted in R (Version 4.2.2) by using the "ggstatsplot" package.

## 2.5 Sources of IEQ dissatisfaction

We have further investigated the main reasons behind IEQ reviews with "dissatisfied" sentiment polarity. This was accomplished by the word frequency analysis, like the one Villeneuve and O' Brien [9] conducted. For the four key IEQ dimensions—thermal environment, IAQ, illumination, and acoustics, the top 400 keywords (including stemmed words, e.g., plural forms, and -ing

forms) and their frequencies, eliminating stop words, were discovered. Stop words were defined in this project as articles, pronouns, conjunctions, and prepositions that did not convey useful information, such as "a, an, the, and, it, for, or, but, in, my, your, our, and their". Numerals, adjectives, interjections, and adverbs were, however, also disregarded when ranking and screening keywords because they were not likely to contribute to IEQ dissatisfaction. Most remaining words were nouns and verbs. These words were then manually scrutinised to determine whether they were sources of IEQ dissatisfaction or merely phenomena impacted by the dissatisfaction. For each IEQ dimension, keywords were ranked from the highest frequency to the lowest.

## 3 Results

### 3.1 Statistical Summary

On average, visitors stayed for 2.6 nights in the facility. Guests can rate their accommodation at Bookings.com on a scale of 1 to 10. This represents the individual rating score. Bookings.com additionally totals all review scores and divides the sum by the total number of review scores for each guest accommodation to determine the overall guest rating. The average overall rating for the 1,397 hotels and apartments is 8.2 out of 10 while the average rating from individual guests is 7.9 out of 10. The statistics of 1,397 hotels and serviced apartments broken down by the star rating is shown in **Table 2**.

**Table 2.** Statistical breakdown of 1,397 hotels and apartments by star rating

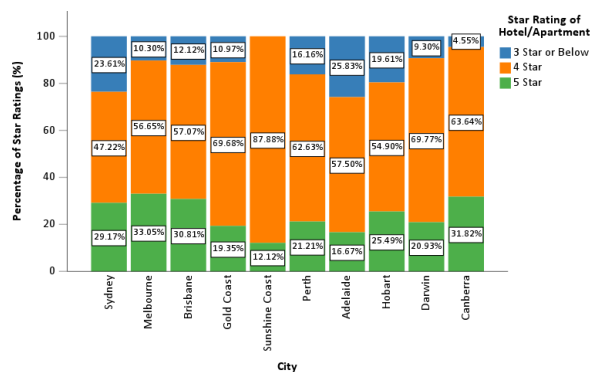
Star Rating	No. of Hotels/ Apartments	Overall Guest Rating	Individual Guest Rating	Average No. of Nights Stayed
2 Star	21	7.3	6.9	2.4
3 Star	191	7.6	7.3	2.6
4 Star	829	8.2	7.9	2.7
5 Star	356	8.5	8.2	2.4
Total/Group Mean	1,397	8.2	7.9	2.6

Since "2 star" and "3 star" facilities only represent 1.5% and 13.7% of the facilities in the database, respectively, they were combined for further statistical analysis. **Figure 3** illustrates the composition of star ratings of the hotels and apartments in ten cities. Most guest accommodations in all cities are 4-star hotels or apartments.

### 3.2 IEQ dissatisfaction

The mean percentage of IEQ dissatisfaction was calculated for nine IEQ dimensions. Guests were most dissatisfied with facilities of the accommodation (32.19%) along with cleanliness and maintenance (18.20%), followed by acoustics (7.70%), available space (6.19%), IAQ (4.42%), exterior view (2.48%),

thermal environment (2.27%), lighting (2.20%), and layout and design (0.14%). The highest standard deviations were found in the cleanliness and maintenance, and facilities dimension, at 10.78% and 10.43%, respectively, indicating that customer satisfaction towards these two IEQ dimensions displayed the highest variability. The average sentiment scores were slightly lower than 5 (neutral) in thermal environment (4.97), IAQ (4.91), lighting (4.96), acoustics (4.90), and facilities (4.68) dimensions, and slightly higher than 5 in available space (5.14), exterior view (5.23), cleanliness and maintenance (5.19), as well as layout and design (5.01). The fact that most reviews in each IEQ dimension had a "neutral" polarity was the cause of the low variability of IEQ sentiment scores.



**Fig. 3.** The composition of star rating of hotels/apartments in each city.

### 3.2.1 Comparison of IEQ dissatisfaction between different star ratings

The impact of guest accommodations' star ratings on the percentage of IEQ dissatisfaction in four key IEQ dimensions—thermal environment, IAQ, lighting, and acoustics—was examined using one-way ANOVA with pairwise comparisons. **Figure 4 (a-d)** was annotated with the ANOVA result and pairwise comparisons that were statistically significant. Except for lighting, as the star rating increased, the percentage of dissatisfied customers regarding the thermal environment, IAQ, and acoustics decreased. There was no statistically significant difference between star ratings and customer dissatisfaction with lighting.

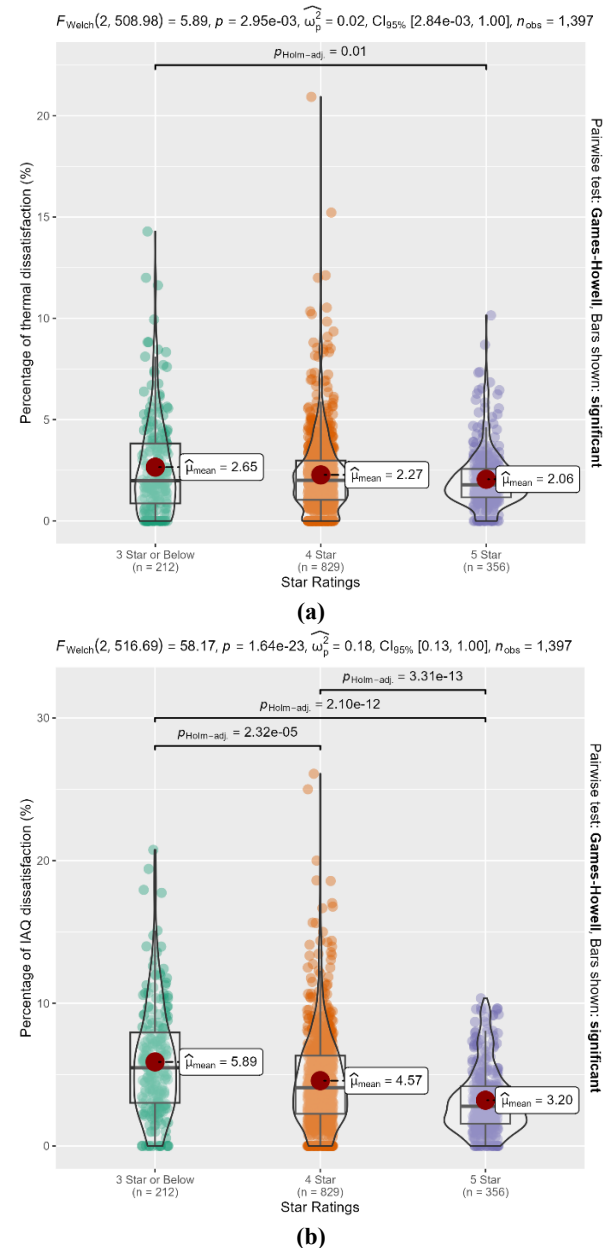
### 3.2.2 Comparison of IEQ dissatisfaction between climate zones

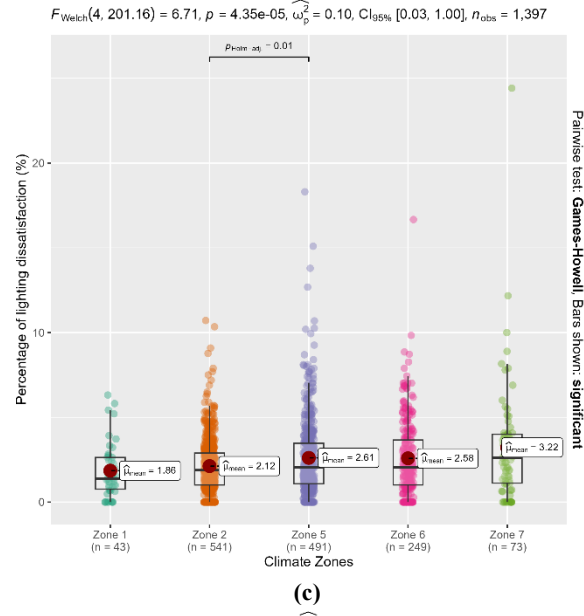
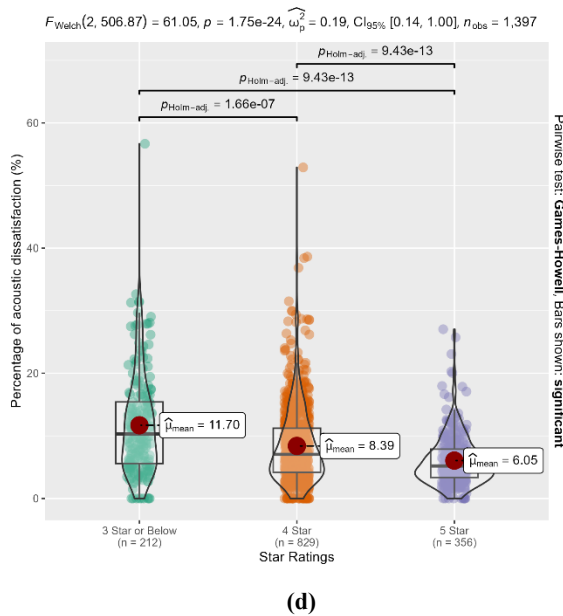
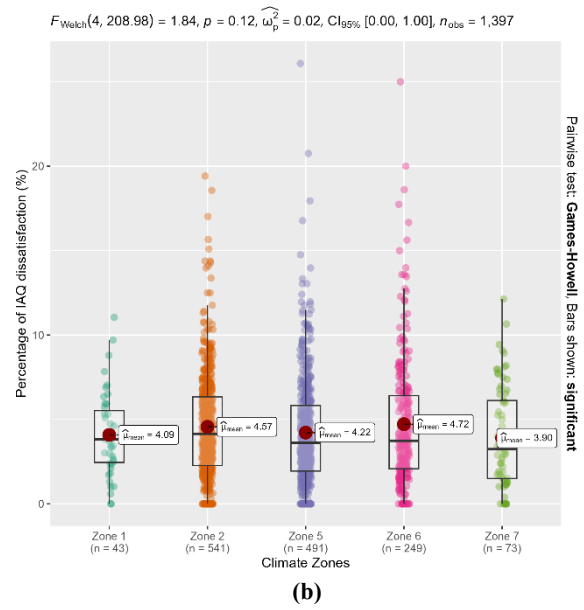
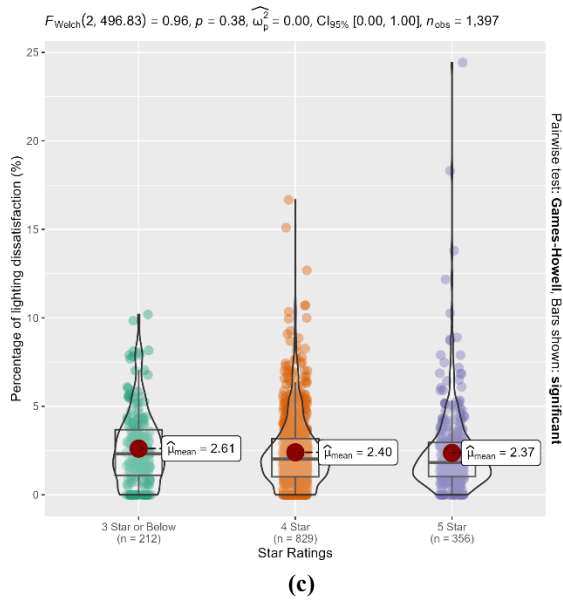
**Figure 5 (a-d)** depicts the results of an ANOVA test with pairwise comparisons examining the effect of climate zones on the percentage of IEQ dissatisfaction in guest accommodations. The ten Australian cities examined in this study belong to five climate zones: Climate Zone 1 (hot humid summer, warm winter), Climate Zone 2 (warm humid summer, mild winter), Climate Zone 5 (warm temperate), Climate Zone 6 (mild temperate), and Climate Zone 7 (cool temperate). Only statistically significant pairwise comparisons had their p-values annotated in the figure. There was no significant difference in IAQ dissatisfaction across 5

different climate zones. Zone 2 displayed lower acoustic dissatisfaction than Zone 5, 6, and 7.

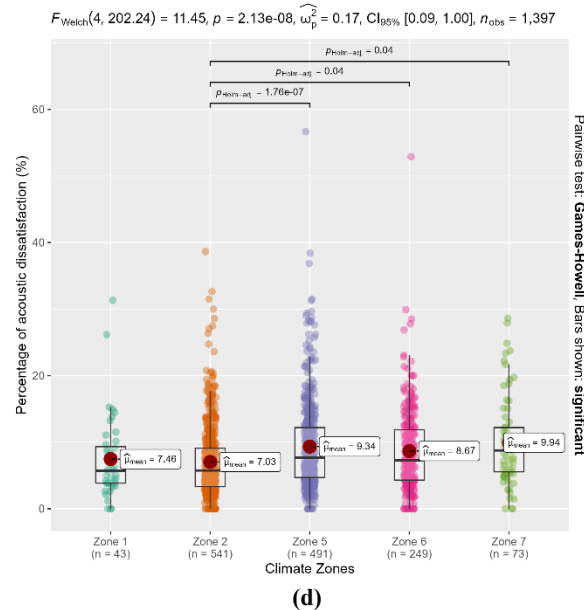
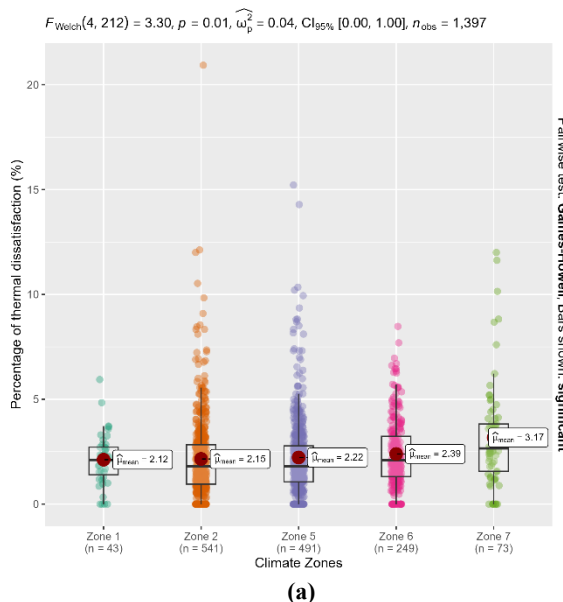
### 3.2.3 Comparison of IEQ dissatisfaction between seasons

To determine whether IEQ dissatisfaction exhibits a seasonal pattern, a categorical variable titled "visiting season" was computed based on the check-in month. According to the Australian Bureau of Meteorology [15], spring is comprised of the three transitional months of September, October, and November; summer is from December to February; autumn is from March to May; and winter starts in June, and ends in August. **Figure 6 (a-d)** illustrates the percentage of IEQ dissatisfaction in four dimensions across seasons, as determined by an ANOVA test and multiple pairwise comparisons. The percentage of thermal dissatisfaction was significantly higher in winter than in the other three seasons. While thermal environment dissatisfaction was more prevalent in winters, acoustics dissatisfaction was lower in summer than in other seasons.

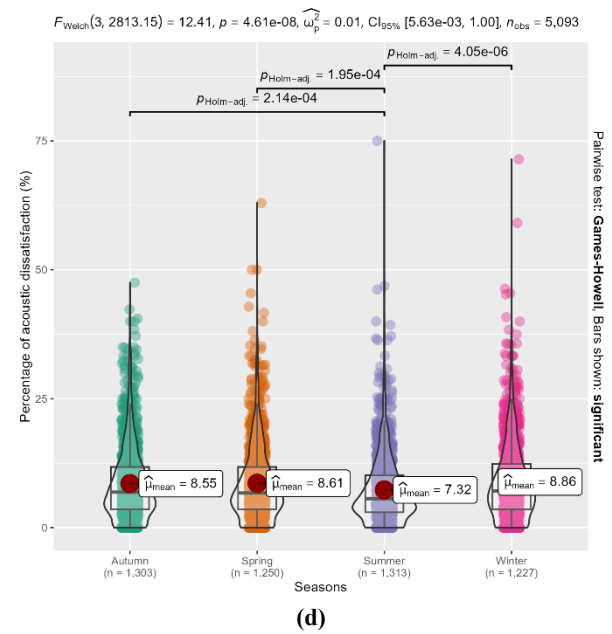
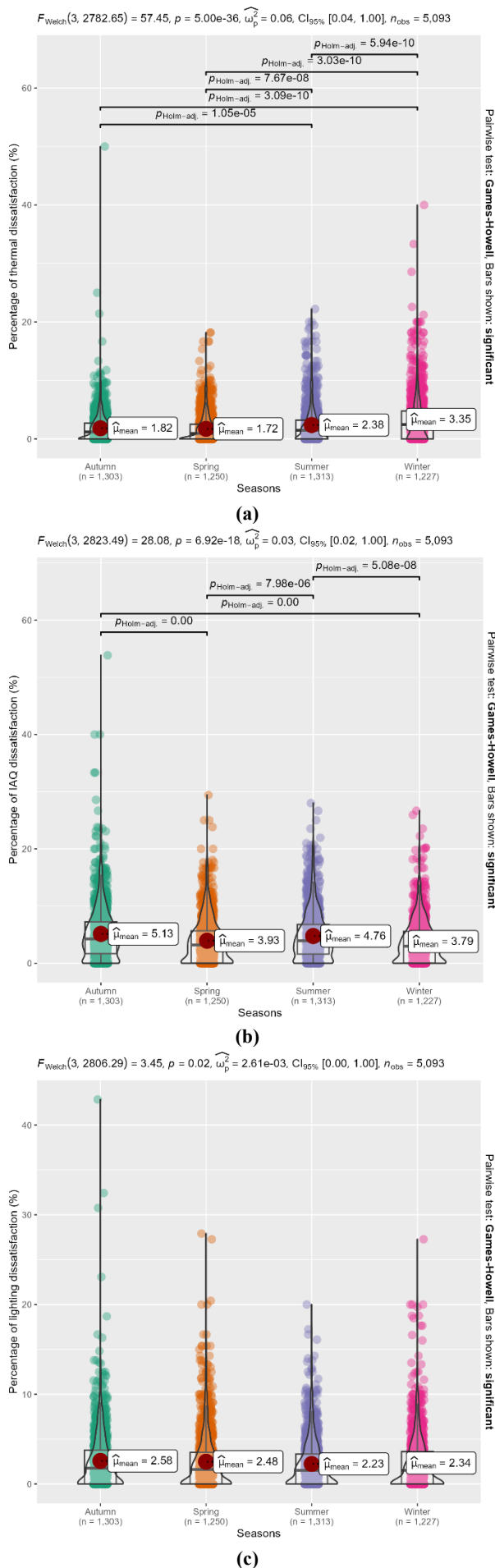




**Fig. 4.** ANOVA Test: Percentage of dissatisfaction towards (a) Thermal environment; (b) IAQ; (c) Lighting; (d) Acoustics across different star ratings.



**Fig. 5.** ANOVA Test: Percentage of dissatisfaction towards (a) Thermal environment; (b) IAQ; (c) Lighting; (d) Acoustics in different climate zones.



**Fig. 6.** ANOVA Test: Percentage of dissatisfaction towards (a) Thermal environment; (b) IAQ; (c) Lighting; (d) Acoustics in different seasons.

### 3.2.4 Main sources of IEQ dissatisfaction

**Table 3** lists the top ten keywords ranked by their frequencies in four IEQ dimensions that represented potential sources of IEQ dissatisfaction, as discussed in Section 2.5. To gain an understanding of the contexts in which the keywords were mentioned within the guests' complaints, the authors proceeded to manually examine the sentences containing these keywords. The common IEQ issues were summarized as below:

- Thermal environment: malfunction of air-conditioners or insufficient heating or cooling, especially at night; inoperative fans or heaters; windows either let in cold air, or excessive solar radiation that contributed to overcooling or overheating, respectively; lack of temperature control in the room.
- IAQ: unpleasant smell; mould in the bathroom, especially in the shower; smoking smell; lack of ventilation; windows could not be opened for fresh air; no balcony to get fresh air.
- Lighting: lights or bedside lamps did not work; darkness in the room or bathrooms; lack of natural light; curtains or blinds did not block lights from outside the windows at night.
- Acoustics: noise from people; traffic noise; outdoor noise (other than traffic); noise from next door or upstairs; poor sound insulation of windows.

## 4 Discussions

This study employs web-mining and context-aware, state-of-the-art NLP-based models to investigate IEQ satisfaction in Australian hotels and serviced apartments. Results displayed both similarities and discrepancies with previous studies.

The combined and city-specific IEQ complaint rates were reported by Qi et al. [6], Shen et al. [7], and Villeneuve and O'Brien [9], which were calculated by dividing the number of customer reviews that contained IEQ-related complaints by the total number of qualifying reviews posted by visitors. The complaint rates for specific IEQ dimensions were 8.07% for acoustics, 3.89% for IAQ, 1.75% for thermal environment, and 0.48% for lighting in China's budget hotels [7], and 2.95% for acoustics, 1.25% for thermal environment, 0.48% for IAQ, and 0.22% for lighting in Canadian Airbnb buildings [9]. In the present study, we only concentrated on IEQ dissatisfaction, measured as the percentage of reviews that contained IEQ complaints relative to the total number of IEQ-related reviews. Yet, our study aligned with [7] and [9] that acoustics was the most complained IEQ dimension (7.70% dissatisfaction) among the four primary IEQ dimensions. The finding of the highest thermal dissatisfaction in winter aligned with the conclusions of Villeneuve and O'Brien [9].

**Table 3.** Top ten sources of dissatisfaction in thermal environment, IAQ, lighting and acoustics

Rank order	Thermal Environment	IAQ	Lighting	Acoustics
1	heating	smell	lights	night
2	air-conditioner	mould	darkness	door
3	night	smoking	bathroom	people
4	cooling	bathroom	lamp	street
5	working	airing	night	traffic
6	temperature	shower	bed/bedroom	roads
7	bed/bedroom	windows	window	outside
8	fan	doors	brightness	floor
9	window	cigarettes	curtains	windows
10	heater	balcony	blinds	music
Sentences analysed	12,380	25,584	13,049	47,781

The word frequency analysis in this study and Villeneuve and O'Brien [9] revealed comparable results. The top ten sources listed in **Table 3** largely matched those found in the Canadian study and were also generally consistent with the conclusions of Qi et al. [6] and Shen et al. [7]. This implied that IEQ dissatisfaction in tourist lodgings was a universal phenomenon with comparable underlying causes.

The limitations of this paper reside in the IEQ categorization. Many irrelevant comments were already eliminated by removing non-IEQ phrases based on prior research and the content of the reviews. Even so, some comments—particularly in the thermal environment dimension—still contained the right IEQ lexicons but the incorrect contexts, for example, "tiles in bathroom were freezing" and "all the meals I ordered were cold".

## 5 Conclusion

This study investigated the IEQ dissatisfaction in 1,397 Australian hotels and serviced apartments across 10 cities using web-mining and NLP-based methodologies. The findings showed that facilities had the highest dissatisfaction rate (32.19%), followed by cleanliness and maintenance (18.20%) and acoustics (7.70%). The percentage of customers who were dissatisfied with the thermal environment, IAQ, and acoustics reduced as the star rating of the accommodation rose. There was a significantly higher level of customer dissatisfaction with the thermal environment in winter than in other seasons. The main sources of dissatisfaction with the thermal environment, IAQ, lighting, and acoustics were consistent with those found in earlier studies. This study provides an in-depth understanding of the IEQ performance in Australian guest homes. Findings can provide a rational basis for the hotel/apartment managers to carry out targeted building retrofits to improve IEQ and customer satisfaction.

## References

1. Tourism Australia, <https://www.tourism.australia.com/en/insights/tourism-statistics/the-economic-importance-of-tourism.html> (Accessed Dec. 11, 2022).
2. Australian Bureau of Statistics. <https://www.abs.gov.au/statistics/economy/national-accounts/australian-national-accounts-tourism-satellite-account/latest-release> (Accessed Dec. 11, 2022).
3. IBISWorld. <https://www.ibisworld.com/default.aspx>. (Accessed Dec. 11, 2022).
4. S. Roumi, F. Zhang, R. Stewart, M. Santamouris, *Energ. Buildings* **263**, 112033 (2022).
5. C. Candido, J. Kim, R. de Dear, L. Thomas, *Build. Res. Inf.* **44** (2), 214–228 (2016).
6. M. Qi, X. Li, E. Zhu, Y. Shi, *Build. Environ.* **111**, 1–9 (2017).
7. Z. Shen, X. Yang, C. Liu, J. Li, *Sustainability (Switzerland)* **13**, 8 (2021).
8. M. Suh, H. Moon, H. Han, S. Ham, *J. Hosp. Mark. Manag.* **24** (7), 727–753 (2015).
9. H. Villeneuve, W. O'Brien, *Build. Environ.* **169**(September 2019), 106555 (2020).
10. X. Guo, K. Lee, Z. Wang, S. Liu. *Build. Environ.* **206**(June), 108288 (2021).
11. Z. Liu, S. Park, *Tour. Manag.* **47**, 140–151 (2015).
12. T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean, *Distributed representations of words and phrases and their compositionality*, in Proceedings of NIPS'13, CA Inc., Red Hook, NY, USA, 3111–3119 (2013).
13. A. Graves, S. Fernández, J. Schmidhuber, *J. Neural networks* **18**, 602–610 (2005).
14. Y. Sastry, S. Karthick, *J. Academy Sci. (Sādhanā)* **47**, 241, Springer Nature (2022).
15. Australian Bureau of Meteorology. *Climate Glossary - Seasons*. (2019).