Revealing Airbnb user concerns on different room types

1 Introduction

Airbnb, one of the most successful models in the sharing economy is considered as a major disruptor to the lodging sector (Dolnicar, 2019). Prior literature suggest that the success is contributed by the unique experience Airbnb brings to its users, which is considered different from traditional hotel experience (Tussyadiah & Zach, 2017; Hajibaba & Dolnicar, 2017). Therefore, understanding what do customers look for in their Airbnb experience is of great importance. A few empirical studies investigated the research question but received contradictory conclusions: the areas that Airbnb users care about are found vary in different studies (Guttentag, 2019; Brochado et al., 2017).

The reasons might be attributable to the limitation of research data and lack of consideration of the difference in Airbnb accommodation types. Many existing studies evaluate Airbnb users' experience by analysing data collected from questionnaires. This type of data is useful but might be subject to the incomplete limitations as questionnaires only encompass a limited pool of Airbnb users who opt to participate (Zhong et al., 2019). Moreover, existing research analyse Airbnb users' online reviews as a whole package without considering the difference in users' experience triggered by accommodation types (Cheng & Jin, 2019). Users who choose different types of accommodation may have different needs. A recent research (Dolnicar & Zare, 2020) predicts that the outbreak of Covid-19 might cause a reduction in the supply of entire homes, forcing Airbnb move back towards sharing of spare rooms. It becomes essential for Airbnb hosts to understand user experience and concern specifically on the latter accommodation type in order to build up competitive advantages and achieve market success in the post-Covid world. Therefore, in this research note, we investigate how the areas that Airbnb users concern about vary across different types of accommodation by analysing their reviews online.

2 Topic identification

The dataset used in our empirical analysis contains all $346,037$ user reviews on $5,894$ listings in San Francisco, US between May 3, 2009 and June 7, 2020. We choose San Francisco in the study because it has the highest occupancy rate in the US (DuBois, 2020). Moreover, the large number of accommodation listings and user reviews makes the dataset a good representative sample. As the study focuses on different accommodation types, we split the dataset into three separate datasets based on the room types: Entire home, Private room and Shared room.

Topic models, as one stream of unsupervised text mining method has been found suitable for analysing customer comments (Büschken & Allenby, 2016). It examines the co-occurrence relationship among words and outputs the collections of words with high probability of co-occurrence, i.e., the topics. We choose the Structural topic model (Roberts et al., 2018) in this method category because of its advantage of allowing incorporating the metadata of review texts, such as listing neighbourhood, host response time, etc., which have been found significant on driving the topic distribution (Büschken & Allenby, 2016).

We process the dataset with the following steps: (i) removing numbers, stop words and punctuation marks, (ii) transforming all letters to lower case in order to obtain a more uniform form and reduce the size of the vocabulary, and (iii) tokenization. Next, we decide on the number of topics K which is an important user-specified parameter of the structural topic model and helps to achieve substantive interpretation of the outcomes of the modelling. Using function *searchK* from *stm* and *furrr* packages in *R*, we evaluate the models trained on a sparse matrix in parallel with a range of different values of K by looking at the semantic coherence of the topics, held-out likelihood and residuals. This leads to $K = 30, 23, 9$ respectively for the Entire home, Private room and Shared room datasets. Two extra researchers in Tourism were recruited for discussing and assigning topic labels and categories based on the top words generated from the structural topic model, close readings of a substantial number of review examples in each topic and prior literature (Guo et al., 2017).

Table 1-3 present the results. We find that users who book different room types concern about different areas. Specifically, Entire home users care most about Service while Private and Shared room users are interested mostly in General experience and Location respectively (see Figure 1). Further, the amount of concern about Service decreases from Entire home to Private room, but the amount of concern about Location increases along this continuum. In addition, the numbers of topics identified from Entire home and Private room reviews are higher than that from the Shared room reviews, indicating various areas Entire home and Private room users care about.

A number of interesting findings can be drawn from the topic proportions under the topic categories. For instance, *room size* is the most cared about topic in the Facilities category for users who book Entire home while users who book Private and Shared room are mostly interested in *room facilities* and *cleanness and noise* respectively. Further, *decent location* and *transportation* are mostly commented by all users in the Location category. In addition, Entire home users express concerns about *check-in* and *booking and cancellation* in the Service category, whereas Private and Shared room users focus more on *host help.*

3 Conclusion

This study uses the Structural topic model to identify different areas Airbnb users care about with respect to different accommodation types. This research supports the findings of prior studies that discovered location, room facilities and host as key areas (e.g., Cheng & Jin, 2019). Second, this research goes beyond previous literature by identifying more detailed dimensions under each area and showing the different degrees of user attention. Third, our research further contributes the literature by revealing how the identified areas vary across different room types. The results show that users who book Entire home and Private room care about more areas. Moreover, they focus more on Service and General experience while Shared room users are more interested in Location. Fourth, we demonstrate the efficacy of using the novel approach, structural topic modelling on rich online user reviews.

Finally, the research has valuable managerial implications. It enables Airbnb practitioners to ascertain the heterogeneity and importance of the areas that users care about. Although room experience and service quality are found to be the key factors influencing hotel customer satisfaction (Choi & Chu, 2001; Guo et al., 2017), there is a lack of understanding about to what extent the factors affect users' experience especially when they book different types of rooms. Our findings on the detailed areas that users care about regarding the three room types enable Airbnb hosts to understand better users' expectations, and can help Airbnb hosts adjust their business strategy accordingly and provide tailored service to users.

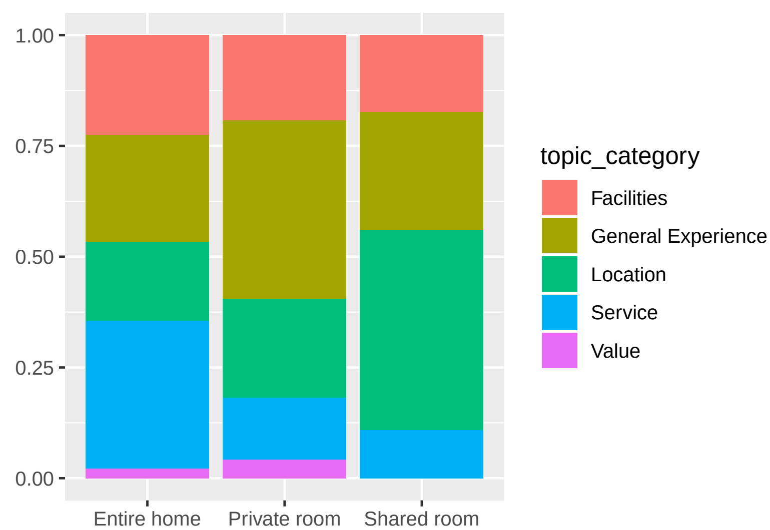

Figure 1: The areas that Airbnb users care about regarding different room types

Table 1: Entire home

Topic Label	Topic Proportions	Top Words
Topic Category: Facilities		
Room Size	4.92%	spacious, size, living, bedroom, large
Property Type	4.25%	house, good, flat, lot, furnished
Parking	4.09%	parking, street, find, car, little
Cleanness	3.43%	tidy, comfortable, clean, check, bed
Room Facilities	3.08%	kitchen, stocked, shower, fridge, water
Decoration	1.59%	unit, modern, apartment, new, old
Privacy	1.16%	privacy, studio, garden, private, getaway
Topic Category: General Experience		
Good Feeling	6.62%	great, stay, definitely, nice, enjoyed
Recommending Feeling	4.31%	recommend, apartment, well, highly, staying
Home Feeling	3.44%	home, made, thank, feel, hosts
Host Review	2.90%	host, wonderful, fantastic, amazing, jennifer
Honest Advertisement	2.89%	even, pictures, us, exactly, sure
Sharing Experience	2.16%	experience, one, first, stayed, time
Suitable for Family	1.31%	two, family, people, husband, kids
Others	0.39%	everything, stay, apartment, perfect, just
Topic Category: Location		
Decent Location	4.91%	place, location, stay, spot, close
Transportation	4.27%	walk, bus, away, just, take
Good Views	3.37%	beautiful, home, view, lovely, amazing
Restaurants	2.47%	restaurants, close, walking, distance, shops
Landmark Building	1.97%	visit, city, building, attractions, sights
Neighbourhood	1.02%	neighborhood, get, city, around, downtown
Topic Category: Service		
Room Service	6.19%	everything, needed, make, sure, way
Host Attitude	5.37%	helpful, responsive, friendly, accommodating, kind
Check-in	5.20%	check, arrival, long, days, upon
Host Response	4.50%	quick, questions, available, quickly, response
Booking and Cancellation	4.05%	canceled, upon, reservation, prior, ahead
Food Service	3.48%	coffee, appreciated, little, fridge, snacks
Host Help	2.27%	help, provided, local, gave, tips
Customer Needs Met	2.12%	need, every, way, everything, want
Topic Category: Value		
Price	2.27%	value, hotel, work, much, price

Table 2: Private room

Topic Label	Topic Proportions	Top Words
Topic Category: Facilities		
Room Facilities	8.50%	kitchen, nice, comfortable, bedroom, bathroom
Cleanness	5.18%	easy, super, clean, tidy, comfy
Room Size & Decoration	2.97%	space, book, well, decorated, modern
Parking	2.00%	parking, street, find, car, free
Property Type	0.53%	cottage, needs, suite, en, met
Topic Category: General Experience		
Good Feeling	13.65%	great, place, stay, recommend, perfect
Recommending Feeling	8.59%	recommend, host, wonderful, excellent, accommodating
Home Feeling	7.93%	home, time, staying, back, feel
Bad Feeling	5.10%	little, night, didnt, door, noise
First time experience	3.58%	first, time, experience, one, ever
Honest Advertisement	1.44%	described, stayed, pictures, traveling, exactly
Topic Category: Location		
Decent Location	5.38%	golden, beach, close, near, ocean
Transportation	4.75%	walk, downtown, bus, short, minutes
Landmark Building	3.86%	san, francisco, visit, anyone, city
Neighbourhood	3.45%	neighborhood, quiet, safe, distance, easy
Restaurants	3.37%	restaurants, close, bar, located, distance
Good Views	1.49%	view, hill, top, bay, wharf
Topic Category: Service		
Host Attitude	4.98%	friendly, helpful, hosts, welcoming, lovely
Host Help	3.81%	us, gave, provided, tips, local
Food Service	2.39%	coffee, morning, breakfast, just, one
Checking-in	1.64%	check, days, arrival, long, late
Room Service	1.23%	service, hotel, square, staff, right
Topic Category: Value		
Price	4.18%	price, value, good, get, little

Table 3: Shared room

Topic Label	Topic Proportions	Top Words
Topic Category: Facilities		
Cleanness and Noise	10.72%	room, clean, night, noise, bed
Room Size	6.41%	space, like, small, community, many
Topic Category: General Experience		
Room Sharing Experience	14.69%	people, nice, stay, awesome, meet
Host Review	6.70%	hosts, home, feel, wonderful, house
Recommending Feeling	5.08%	recommend, highly, thanks, comfortable, place
Topic Category: Location		
Decent Location	25.40%	great, place, location, area, close
Transportation	11.56%	metro, close, walk, easy, station
Landmarking Building	7.81%	san, definitely, francisco, sightseeing, market
Topic Category: Service		
Host Help	11.63%	needs, help, one, experience, hostel

References

Brochado, A., Troilo, M., Aditya, S., et al. (2017). Airbnb customer experience: evidence of convergence across three countries. *Annals of Tourism Research*, *63*, 210–212.

Büschken, J., & Allenby, G. M. (2016). Sentence-based text analysis for customer reviews. *Marketing Science*, *35*(6), 953–975.

Cheng, M., & Jin, X. (2019). What do airbnb users care about? an analysis of online review comments. *International Journal of Hospitality Management*, *76*, 58–70.

Choi, T. Y., & Chu, R. (2001). Determinants of hotel guests' satisfaction and repeat patronage in the hong kong hotel industry. *International Journal of Hospitality Management*, *20*(3), 277–297.

Dolnicar, S. (2019). A review of research into paid online peer-to-peer accommodation: Launching the annals of tourism research curated collection on peer-to-peer accommodation. *Annals of Tourism Research*, *75*, 248–264.

Dolnicar, S., & Zare, S. (2020). *Coronavirus and airbnb: Disrupting the disruptor.* SocArXiv. doi: 10.31235/osf.io/t9n6q

DuBois, D. (2020). *The best cities for airbnb: Compare markets side by side.* (https://www.airdna.co/blog/best-cities-for-airbnb)

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.

Guttentag, D. (2019). Progress on airbnb: A literature review. *Journal of Hospitality and Tourism Technology*, *10*, 814–844.

Hajibaba, H., & Dolnicar, S. (2017). Substitutable by peer-to-peer accommodation networks. *Annals of Tourism Research*, *66*(C), 185–188.

Roberts, M., Stewart, B., & Tingley, D. (2018). *stm: An r package for the structural topic model.* (http://www.structuraltopicmodel.com)

Tussyadiah, I. P., & Zach, F. (2017). Identifying salient attributes of peer-to-peer accommodation experience. *Journal of Travel & Tourism Marketing*, *34*(5), 636–652.

Zhong, L., Sun, S., & Law, R. (2019). Movement patterns of tourists. *Tourism Management*, *75*, 318–322.