

ding_2020_employing_structural_topic_modelling_to_explore_perceived_service_quality_attributes_in_airbnb_accommodation

Year

2020

Author(s)

Kai Ding and Wei Chong Choo and Keng Yap Ng and Siew Imm Ng

Title

Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation

Venue

International Journal of Hospitality Management

Topic labeling

Manual

Focus

Secondary

Type of contribution

Established approach

Underlying technique

Manual labeling assisted by: labels from previous literature, top words provided by multiple metrics and associated documents

Topic labeling parameters

\

Label generation

Table 2 shows the results of topic modeling and labels assigned to each topic and the source of the label. In the topic labeling process, high probabilities (Highest Prob) words, frequency-exclusivity (FREX) words, and top 10 representative reviews of each topic were used.

Words generated from the FREX statistic were mainly used to label the topic, as this measurement weights words by their overall frequency and the degree of exclusivity of those words to the topic which can provide more semantically intuitive representations of topics

However, FREX statistic sometimes ranks some uncommon words high, therefore Highest Prob words will be referred to if FREX terms fail to differentiate a topic.

Table 2
Topic summary and labeling.

Topic #	Topic Label	Top Words	Topic Prop. (%)	Reference
1.	Internet connection	Highest Prob: good, wifi, locat, problem, host, clean, servic FREX: wifi, good, internet, connect, minor, solv, condit	4.3 %	Cheng and Jin (2019)
2.	Booking experience	Highest Prob: day, book, room, didn, host, call, time FREX: told, refund, review, cancel, deposit, final, wrong	4.7 %	Lee and Yu (2018)
3.	Transportation	Highest Prob: station, airport, grab, taxi, uber, minut, bus FREX: sentral, train, station, bus, taxi, monorail, klia (airport)	3.1 %	Zhang (2019)
4.	Group stay	Highest Prob: hous, stay, friend, place, clean, famili, recommend FREX: homestay, gather, group, suitabl, environ, hous, satisfi	10.9 %	Named by authors
5.	Sleep condition/ Bed	Highest Prob: bed, comfort, space, comfi, equip, kitchen, furnish FREX: comfi, bed, queen, size, qualiti, singl, furnish	3.1 %	Zhang (2019)
6.	Car parking	Highest Prob: park, car, secur, access, guard, card, lift FREX: park, lift, guard, carpark, visitor, card, entranc	2.8 %	Zhang (2019)
7.	Hosts' response	Highest Prob: host, place, stay, respon, easi, clean, great FREX: checkin, checkout, respond, prompt, queri, quick, process	9.6 %	Cheng and Jin (2019)
8.	Tourist attraction	Highest Prob: walk, food, locat, distanc, place, minut, street FREX: distanc, attract, gurney, famous, plaza, georgetown, walkabl	5.5 %	Priporas et al. (2017)
9.	Hospitality hosting behavior	Highest Prob: host, home, feel, stay, experi, time, wonder FREX: home, welcom, feel, sweet, husband, warm, memor	7.8 %	Lalicic and Weismayer (2018)
10.	Water activities	Highest Prob: beach, place, relax, quiet, staff, peac, stay FREX: island, scooter, beach, snorkel, villa, hustl, bustl	3.2 %	Named by authors
11.	Shopping	Highest Prob: conveni, shop, mall, restaur, locat, nearbi, place FREX: mall, store, shop, groceri, conveni, tesco, eleven	4.6 %	Luo and Tang (2019)

Table 2 (continued)

Topic #	Topic Label	Top Words	Topic Prop. (%)	Reference
12.	Swimming pool	Highest Prob: nice, pool, room, view, swim, floor, clean FREX: flat, roof, confort, infin, nice, recommand, swim	5.3 %	Cheng and Jin (2019)
13.	Property attributes	Highest Prob: apart, great, stay, locat, recommend, amaz, host FREX: apart, rocki, fantast, peggi, edna, victoria, amaz	8.4 %	Wang and Nicolau (2017)
14.	The feature of neighborhood	Highest Prob: night, noi, road, breakfast, morn, light, room FREX: sleeper, nasi, noi, hear, mosquito, prayer, earplug	2.5 %	Lawani et al. (2019) ;
15.	Seaview	Highest Prob: unit, view, balconi, sea, watch, face, enjoy FREX: unit, seaview, movi, balconi, eve, watch, penthous	2.4 %	Named by authors
16.	Room equipment	Highest Prob: room, water, bathroom, aircond, shower, toilet, hot FREX: heater, toilet, pressur, condition, shower, filter, aircond	4.8 %	Zhang (2019)
17.	Amenities and service	Highest Prob: build, pool, cook, gym, kitchen, includ, facil FREX: complex, laundri, sauna, machin, washer, stock, includ	2.4 %	Wang and Nicolau (2017)
18.	Room characteristics/ Pets	Highest Prob: love, room, place, live, hous, decor, beauti FREX: cute, guesthous, cat, concept, chill, style, pet	3.8 %	Lawani et al. (2019)
19.	Daily necessities/ Free food and drink	Highest Prob: provid, min, drive, towel, basic, drink, coff FREX: toiletri, iron, necess, snack, toothbrush, gent, complimentari	2.5 %	Zhang (2019)
20.	Rental rules	Highest Prob: check, arriv, person, time, host, late, meet FREX: late, allow, meet, flight, patient, delay, wait	4 %	Wang and Nicolau (2017)
21.	Child friendly	Highest Prob: kid, condo, place, love, famili, enjoy, stay FREX: legoland, kid, toy, kitti, medina (mall), puteri (harbor), louis	2.5 %	Named by authors
22.	General experience/ Cleanness	Highest Prob: stay, place, clean, host, locat, great, area FREX: pleasant, area, clean, thing, stay, expect, place	1.8 %	Lawani et al. (2019)

Motivation

\

Topic modeling

STM

Topic modeling parameters

Nr of topics: 8 to 40

Nr. of topics

22

Label

The name for each topic was selected by referring to the pre-identified service attributes from previous Airbnb studies.

However, if no matching attributes could be found from previous Airbnb literature, these topics would be labeled manually through group discussion and analysis.

Label selection

\

Label quality evaluation

\

Assessors

\

Domain

Paper: Tourism and Hospitality

Dataset: Tourism and Hospitality (Airbnb)

Problem statement

This study employs the structural topic model to extract service quality attributes from 242,020 Airbnb reviews in Malaysia. 22 service related topics were extracted from the corpus and four topics have not appeared in previous Airbnb studies. A widely used modified SERVQUAL questionnaire (MSQ) is cross-validated in this study by comparing its service quality attributes with the results of the topic modelling, which indicates that this MSQ can cover general Airbnb service quality attributes. This study also examines the different preferences of Malaysian and international Airbnb users and the changing patterns of the top six service quality attributes during a five-year period. The findings reveal that Malaysian Airbnb users care more about the appearance and location of the property, and international Airbnb users pay more attention to whether the property can accommodate a group of people. In addition, communication with the host is found to play an increasingly important role in Airbnb users' lodging experiences.

Corpus

Origin: Airbnb

Nr. of documents: 620,487 (242,020 after processing)

Details:

- customer reviews generated from November 2010 to January 2019

Document

Airbnb review

Pre-processing

- removal of non-English reviews
- reviews generated from 2010 to 2013 were omitted due to the lower frequency of reviews

generated during these periods

- tokenisation
 - Text normalization (converting capital letters into lower case)
 - Excluding reviews with less than 50 words
 - Removing numbers, punctuation marks, and extra white space
 - Removing default stop words (e.g., “is”, “the”)
 - Removing customized stop words (e.g., city names, duplicated words)
 - N-gram analysis was conducted to identify frequent compound words and transform them into unigrams (e.g., “twin towers” to “twintowers”),
 - Removing words which appear in less than 20 documents to filter some insignificant words.
-

```
@article{ding_2020_employing_structural_topic_modelling_to_explore_perceived_service_quality_attributes_in_airbnb_accommodation,
```

```
  abstract = {This study employs the structural topic model to extract service quality attributes from 242,020 Airbnb reviews in Malaysia. 22 service related topics were extracted from the corpus and four topics have not appeared in previous Airbnb studies. A widely used modified SERVQUAL questionnaire (MSQ) is cross-validated in this study by comparing its service quality attributes with the results of the topic modelling, which indicates that this MSQ can cover general Airbnb service quality attributes. This study also examines the different preferences of Malaysian and international Airbnb users and the changing patterns of the top six service quality attributes during a five-year period. The findings reveal that Malaysian Airbnb users care more about the appearance and location of the property, and international Airbnb users pay more attention to whether the property can accommodate a group of people. In addition, communication with the host is found to play an increasingly important role in Airbnb users' lodging experiences.},
```

```
  author = {Kai Ding and Wei Chong Choo and Keng Yap Ng and Siew Imm Ng},
```

```
  date-added = {2023-04-17 22:52:12 +0200},
```

```
  date-modified = {2023-04-17 22:52:12 +0200},
```

```
  doi = {https://doi.org/10.1016/j.ijhm.2020.102676},
```

```
  issn = {0278-4319},
```

```
  journal = {International Journal of Hospitality Management},
```

```
  keywords = {Airbnb, Service quality, Online reviews, Structural topic model, Text mining},
```

```
  pages = {102676},
```

```
  title = {Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation},
```

```
  url = {https://www.sciencedirect.com/science/article/pii/S0278431920302280},
```

```
volume = {91},  
year = {2020}}
```

#Thesis/Papers/FS