Research on Image Perception of Luxury Hotels in Dalian Based on Text Analysis

Qiong Jia
Hohai University, Nanjing, jiaqiong@163.com

Pan Jiang
Hohai University, Nanjing, 211313040024@hhu.edu.cn

Recommended Citation

Jia, Qiong and Jiang, Pan, "Research on Image Perception of Luxury Hotels in Dalian Based on Text Analysis" (2021). *ICEB 2021 Proceedings (Nanjing, China)*. 61.
https://aisel.aisnet.org/iceb2021/61

Research on Image Perception of Luxury Hotels in Dalian Based on Text Analysis

(Work in Progress)

Qiong Jia [1,*]
Pan Jiang [2]

_____

*Corresponding author
[1] Associate Professor, Hohai University, Nanjing, jiaqiong@163.com
[2] Master Degree Student, Hohai University, Nanjing, 211313040024@hhu.edu.cn

ABSTRACT

In recent years, with the development of mobile Internet and smartphones, more and more consumers will share their feelings and thoughts after staying in the hotel on the website, thus forming many online comments. These online comments reflect consumers' perception of the hotel image to a certain extent. This paper uses Python to collect 35977 online comment data of luxury hotels in Dalian. Through the analysis of high-frequency words, characteristic words, and comment emotion, it is found that the main factors affecting the perception of hotel image are: hotel location, room condition, environmental sanitation, service attitude, and professional ability of receptionist and other service personnel. Finally, according to the research conclusions, provide personalized suggestions for the hotel to improve its image.

Keywords: text analysis; luxurious hotel; perception of image

INTRODUCTION

In recent years, the Internet has become the fastest-growing industry. According to the statistics of the 47th "Statistical Report on Internet Development in China" issued by CNNIC (China Internet Network Information Center) (2021), as of December 2020, the number of Internet users in my country is 989 million, compared with March 2020. There were 85.4 million new netizens per month, and the Internet penetration rate reached 70.4%, an increase of 5.9 percentage points from March 2020. At present, the consumption patterns of consumers booking hotels online through smartphones are becoming more and more extensive. For the unique product of hotels, it is difficult for consumers to fully understand the actual environment, services. With specific conditions such as catering and transportation, more and more consumers have become accustomed to searching for online comments previously posted by consumers on third-party platforms or on the hotel's official website as a reference before booking a hotel. These online comments express the true thoughts of consumers who have already experienced the hotel, and at the same time, convey the consumer's perception of the hotel's image. According to the China Online Market Research Report in 2019, 80.1% of consumers purchase goods online; online word-of-mouth will significantly impact them. Hotel online comments provide a convenient, reliable, and comprehensive reference for consumers to buy and provide a channel for hotels to discover their advantages and disadvantages. Through in-depth research on online comments and discussing consumers' perceptions of the hotel's image, suggestions can be made for hotels to improve their image. However, in the face of massive online comment data, how to find the information of interest is a significant research problem. Therefore, this article will focus on this issue.

LITERATURE REVIEW

Hotel image refers to consumers' impressions and evaluation of the hotel's business process. If a hotel wants to stand out from the fierce competition, it must show its differences. The most effective way is to establish a good hotel image. Shuai and Zou (2019) pointed out that online comments will affect consumers' buying behavior to a certain extent. Have a particular impact on the image of the hotel. In foreign research on hotel image perception, Adnan *et al.* (2021). proposed that hotel service scenarios involving the substantive and communication stages will impact customers' image perception. Chen *et al.* (2014). proposed that the hotel's physical equipment, such as decoration and environment, as well as the hotel's response to customer perception, will have an impact on the hotel's functional image perception. Durna *et al.* (2015). found that service landscape components have a positive impact on the overall image, and the overall image has the same impact on consumer reputation and revisit. Xie *et al.* (2016). pointed out that front-line employees will have a greater impact on the establishment of hotel image, and the establishment of luxury hotel image mainly benefits from managers. In the domestic research on image perception, there are little researches on hotel image perception, and there is almost no research on luxury hotel image perception. In recent years, domestic research on image perception has focused on the level of tourist destinations. Sultan *et al.* (2021) proposed that tourists' perception of scenic spots is mainly based on positive emotions. The mental image can be summarized as tourist environment, attractions, facilities and services, and tourism—management and travel experience. Durna *et al.* (2019) put forward and analyzed the reasons for guest complaints and proposed countermeasures to resolve complaints against guest complaints, which is conducive to better building the hotel image.

Most of the research on hotel images obtains relevant data through questionnaires and then uses AMOS software and SPSS software to analyze the data. Compared with many online comments, the use of questionnaire surveys to explore the factors that affect the perception of hotel image is limited by time, space, and cost. If you do not use effective methods or tools for analysis,

it will inevitably lead to many online comment resources being wasted. In recent years, many scholars have started research based on text analysis. There are many data sources for text analysis, and the most popular one is to use python programming to crawl data. Sun and Guo (2020) used python to crawl the comments posted by tourists on the Internet. They analyzed the frequency of commonly used words, word cloud and co-occurrence network, and other image characteristics. Based on python crawling data, Guo *et al.* (2017) further used data mining methods and potential Dirichlet analysis (LDA) to determine the critical dimensions of customer service expressed by hotel visitors. And most of the research on online hotel comments is mainly conducted on the TripAdvisor website, and there is little research on online comments on "Ctrip." In summary, using different research ideas and methods to dig out the valuable information in online hotel comments is necessary to understand consumers' true inner thoughts and convey consumers' perceptions of the hotel's image.

DATA COLLECTION AND DATA PROCESSING

This article uses automation technology for data collection and analysis to explore consumers' perceptions of hotel image through online comments. This article uses python crawler as a collection tool to collect online comments of five-star/luxury hotels in Dalian on Ctrip. The collection time is from April 7, 2021, to April 8, 2021. First of all, search for hotels on "Ctrip", input destination "Dalian", select "five-star/luxury" for hotel level, check-in time is not selected, the default check-in time is the date of the hotel query day, check-out time is the next day, a total of 26 hotels found. After statistics, it is found that the number of comments of different hotels varies greatly. The Shangri-La Hotel, Dalian, with the most comments, has 5432 online comments, and the Dalian Liangyun Chuanqi Hot Spring Hotel, with the least number of studies, has only 22 online comments. For the convenience of subsequent research, this article selectively crawled online comments of hotels with more than 1,000 online comments. Finally, a total of 45,843 online comments of 17 five-star/luxury hotels in Dalian were collected and saved in an excel file. This article chooses "Ctrip" as the data source for the main reasons: First, "Ctrip" is the current online hotel booking industry giant, and the users are relatively large; second, the comments are also rich in content and updated quickly, which is more researchable.

After quick browsing of the collected online comments, it is found that the emotions of the comments are both positive and negative. The content of the studies also covers all aspects of the hotel, such as service attitude, location, traffic environment, hygiene, food, parking, etc. Some comments collected are shown in Table 1.

Table 1: "Ctrip.com" hotel online review form.

hotel name	commentator	score	comments
Shangri-La Hotel Dalian	kimchenkim	5	Shangri-La has never let me down. The room is clean, and there is no smell of smoke. I take the initiative to clean two or three times a day and take the initiative to replenish drinking water.
Kempinski Hotel Dalian	3001878****	4.8	The breakfast is rich, combining Chinese and Western, and the porridge is delicious. The room has the luxury of an old-fashioned hotel. Although it is a stilted room with a relatively long corridor, it is designed with great consideration.
Grand Hyatt Dalian	M417732	3.0	It may have been a long time because there was not much business. Suddenly a large number of guests came to this Qingming holiday. The front desk and breakfast were not prepared, and the attitude was good, but he was busy, and the breakfast was average, which was only worth fifty.

This article adopts the text analysis method. The text analysis method is to go from the surface of the text to the deep level of the reader to discover the deep meanings that ordinary reading cannot grasp. Through computer network technology, the original content is transformed into a quantitative description that the machine can recognize to identify and judge the focus of user-generated content in a certain period. When conducting text analysis, the researcher must exclude personal subjective colors, proceed from the existing materials, complete an objective, comprehensive, quantitative, and systematic study of the materials, and use mathematical-statistical methods to conduct quantitative analysis of the materials studied.

This article uses text analysis to study the online comments of five-star/luxury hotels in Dalian with more than 1,000 comments. Among them, there are many invalid comments in all the crawled online comment text data. Invalid comments mainly include four categories. The first category is multiple comments by different users with similar content and various frequent comments on the same hotel. This type of review is likely to be deliberate by hotel merchants to increase sales and improve the average hotel rating. Invite others to copy and paste comments of ready-made positive comments. This type of review is generally longer, covers all aspects of the hotel, and the repetition rate of comments is relatively high. The second category is a piece of word-count comments published by consumers to complete some tasks on the platform or hotel. Some of these comments have nothing to do with the check-in experience. It is to make up the word count. The third type of comments are not considered consumer

comments in the strict sense, but system comments. For example, if a user just scores and does not write a comment, the system will automatically generate "the user has praised" and "the user has already commented" in the comment column after a few days. Such comments are of no research value. The fourth category is commented in languages other than Chinese because the rest of this article is a study of Chinese. The analysis tools used are also based on Chinese, so even if the content of an English comment meets the requirements, it will not be used.

DATA ANALYSIS

In this paper, combined with network technology and manual analysis, firstly, the collected online comments were deleted duplicates and simple manual screening in Excel, eliminating 9866 invalid comments and finally obtaining 35977 valid comment data. After a simple screening of the collected online comments, this article uses python to perform participle, word frequency statistics, characteristic words, and emotion analysis to explore the focus of consumers on various aspects of the hotel and find valuable information in it so as to research through online comments Consumers' perception of the hotel's image.

Participle

The participle is the process of recombining consecutive letter sequences into word sequences according to certain specifications. For example, one of the online comment of a hotel is, "The transportation is convenient, the bus station and the subway station are at the door; the business district is around the hotel; the rooms are large, the facilities are complete", the results afterword participle are: "The", "transportation", "is" "convenient", "the"," bus station ", " and ", "the", " subway station", "are", " at the door ", "the", " business district ", "is". "around". "the". "hotel", "the", "room", "are", "large", "the", "facilities", "are", "complete".

Word Frequency Statistics

Word frequency statistics is a method of lexical analysis and research, through the study of the frequency of words, so as to find the meaning in it. After numbers, words, URLs, pronouns, and adverbs are eliminated, and finally, a total of 17,498 words are obtained. Many of the words e are meaningless for the research. Some invalid words have been removed through manual screening, such as "Dalian" and "Five-star" because this article is about the online comments of five-star/luxury hotels in Dalian. Therefore, the frequency of "Dalian" and "Five Stars" appearing in the comments is quite high, but there is not much reference significance for the research of this article. Another example is: "this," "you," "what," and other words used to assist language expression. These words are also words with high frequency, but there is not much research significance. Due to the too many results of word segmentation data, this article Only the top 40 words with word frequency as shown in Table 2 below are listed.

Table 2: Word frequency statistics table.

word	hotel	service	good	room	breakfast	location	Reception	convenient
Frequency	20092	11332	10446	9995	7091	6276	6125	4670
word	Yes	environment	no	facilities	check-in	clean	special	children
Frequency	4422	4039	3704	3702	3532	3408	3151	2613
word	feeling	satisfied	next time	attitude	enthusiastic	ordinary	sanitary	comfortable
Frequency	2301	2192	2089	2062	1916	1883	1863	1762
word	like	choose	experience	upgrade	geography	suitable	cozy	free
Frequency	1681	1680	1595	1556	1536	1479	1394	1388
word	traffic	square	sea view	recommended	abundant	staff	pieces	a lot of
Frequency	1360	1305	1293	1248	1229	1213	1200	1194

Characteristic Words

Extract useful words that appear more than 200 times from the participle results, identify these words manually, find out the nouns related to the hotel, and classify these words, so as to obtain the hotel's characteristic words and their corresponding words. The characteristic words are shown in Table 3.

Table 3: Characteristic vocabulary and demonstrative vocabulary.

characteristic words	hardware	equipment	service	location	food
demonstrative words	room park lobby restaurant bathroom pool spa suite gym apartment area	facility air conditioner elevator toilet telephone tub slippers window television	reception manner praise great concierge soundproof professional quality atmosphere detail quality	traffic square seascape surrounding view market center floor distance garden fountain	breakfast smell fruit buffet dinner coffee

		hall		music characteristic excellent	subway sea beaches	

Semantic Network Analysis

In order to deeply study consumers' perception of the hotel's image, this article also adds an in-depth analysis of the semantic relational network. The semantic relationship network is mainly embodied through social network diagrams. In the field of text analysis, social network diagrams are often used to describe and measure the relationship between words and words, reflecting the connection of words in a specific context. Because there are many words involved, this article intercepts the social network relationship graph with "next time" and "recommendation" as the nodes and focuses on the discussion. Through the social network relationship graph, the deep meaning of the comments text can be explored so as to understand the consumer's perception of the hotel's image.

Emotion Analysis

This article uses Python for emotion analysis. An example of an emotion dictionary is shown in Table 4. Among them, emotional words with a positive value express a positive direction and vice versa. Different emotional words convey different degrees, and words with different degrees have different weights. After the comment content is split into sentences, emotion analysis is performed on the words in it, and the data perspective is drawn in excel, as shown in Figure 1. It can be seen from the figure that positive emotions account for the largest proportion, accounting for 56.1%, followed by neutral emotions, accounting for 35.8%, and finally negative emotions, accounting for 8.2%, which is 47.9% lower than positive emotions. This shows that the vast majority of consumers' emotional tendency is positive.

Table 4: Examples of emotional dictionaries.

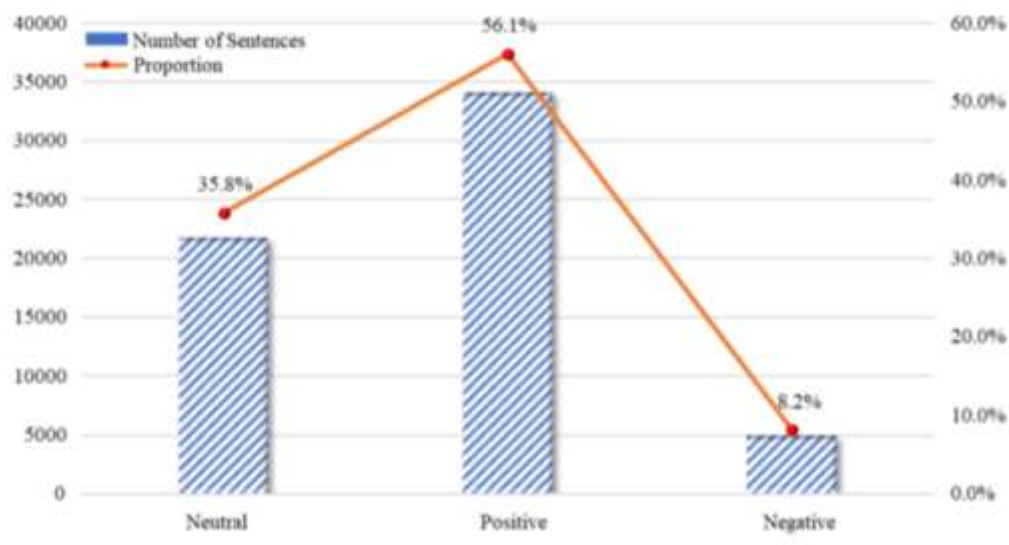| Type	Emotional Words	Weight
Positive words	fashion	1
Negative words	behind	-1
Degree word	extreme	2
Negative Words	possible	-1

Figure 1: Distribution of sentiment tendency of sentences

According to the rules of the emotion dictionary, using the sentence as the most basic unit, analyze the positive words, negative words, degree words, and negative words that appear in the sentence to initially determine the emotion tendency of the sentence, and then match the split sentence with the original data. Count the number of positive sentences and negative sentences contained in each piece of the original comment data, and then calculate the positive score and negative score to get the total score. The higher the score, the greater the positive tendency of the reviewer, otherwise the opposite. For example, a reviewer says, "I always live here. I always feel very good, and the environment, the surroundings are good. The hotel is a bit old, and the facilities in the house are a bit outdated. I hope it can be updated, and it may improve a lot. The breakfast type is not too much, and too monotonous". According to the emotion dictionary, "good" gets 1 point, the degree word "very" has a weight of 3 points, and "monotonous" and "no" are both -1 points. In the end, the total score of the review is 1 point. Consumers, on the whole, have a positive emotional tendency towards hotels.

The emotion analysis score ranking positive 50 comments (highest score 198 points) and bottom 50 comments (lowest score - 63 points) were separately filtered and copied to the excel table, and the second-word segmentation was performed. Counting the results of a participle, the words with higher frequency in the high-score and low-score comments are shown in Table 5.

Table 5: High-frequency vocabulary list.

high score	low score
Service	Room
Room	Service
Reception	Reception
Health	Reception
next time	Clean
Warm	Indifferent

RESULT AND DISCUSSION

Image Perception Analysis Through High-frequency Words
In the word frequency statistics table in Table 2, nouns have the highest frequency, accounting for 50.3%, followed by adjectives, accounting for 38%. These nouns mainly include hotel services, rooms, breakfast, location, front desk, environment, facilities, etc. The room is the place where consumers have the closest contact with the hotel. Consumers will stay in the hotel room for a relatively long time if the environment and hygiene of the room do not meet the expectations of consumers, even if other aspects of the hotel do well. The hotel's satisfaction in the hearts of consumers will not be too high, which will have a negative effect on the image of the hotel in consumers' hearts. For consumers staying in hotels, the front desk is their first impression of the hotel. If the front desk staff provides consumers with warm and attentive service, some minor problems encountered later will not be particularly concerned. However, if the front desk service is indifferent, consumers will feel that they have not received the treatment they deserve and will observe other aspects of the hotel with a scrutiny eye, which will seriously affect the image of the hotel in the minds of consumers. Adjectives mainly describe the above aspects and express the true thoughts and feelings of consumers. For example, in the word frequency analysis, the word "good" frequently appears (10,446 times in total), which shows that most consumers are relatively satisfied with the hotel.

Image Perception Analysis Through Characteristic Words and Demonstrative Words
In Table 3, the characteristic words and demonstrative words are summarized for words that appear more than 200 times. It can be seen that the focus of consumers' attention is mainly on the hardware, equipment, service, location, and other aspects of the hotel. Content is an important factor that affects consumers' perception of hotel image. In the perception of hotel hardware, the high-frequency indicators involved are room, parking lot, lobby, and other words, restored to the original content of the online review, it can be seen that for the room, consumers are mainly concerned about the size of the room, The scenery outside the window and the sanitary condition. For parking lots, consumers are mainly concerned about whether there is a parking lot, whether it is convenient to enter and exit the parking lot, and whether the security of the parking lot is reliable. For the lobby, consumers are mainly concerned about the area of the lobby and whether there is a comfortable seating area. In terms of equipment, consumers are mainly concerned about whether the equipment is repaired and maintained and whether it is safe to use. In terms of service, consumers are mainly concerned about the attitude of the waiters at the front desk, guest rooms, restaurants, etc., and restored to the original content of the online comments. It can be seen that the enthusiastic and professional service will bring more praise to the hotel. The location of a hotel is one of the main factors that affect consumers' perception of the hotel's image. It can be seen that convenient transportation and beautiful scenery can help improve the hotel's image in consumers' minds through location indicators.

Image Perception Analysis Through a Semantic Relational Network
Through the semantic relationship network, it can be seen that the word "next time" is related to "come again", "satisfaction", "like", "enthusiasm", "clean", "environment", "location", "front desk", "service" and words such as "room" are closely related. The word "recommend" is similar to the social network graph of the word "next time." It can be confirmed that consumers' perception of the hotel's image is affected by many factors such as the hotel's location, room conditions, environmental sanitation, the attitude and professional capabilities of the front desk and other service personnel.

Image Perception Analysis Through Emotion Analysis
The emotion of collected online comments has positive and negative. Through emotion analysis on the collected online comments, we can find that most of the consumers' comments are positive, and these consumers are relatively satisfied with the hotel. From Table 5, it can be seen that regardless of whether it is a comment with a higher emotion analysis score or a comment with a lower score, consumers' concerns are mostly concentrated on "room," "service," "front desk," "hygiene" and "attitude." From this five information, it can be seen that whether the room is clean, whether the front desk is enthusiastic, whether the service is professional, and whether the attitude of the service staff is good are the focus of consumers' attention, and these aspects will directly affect consumers' perception of the hotel's image. In the participle results of comments with higher emotion

analysis scores, the word "next time" appears more frequently, and in a sample of 50 reviews, 41 reviews mentioned the word "next time," indicating that after consumers experience the hotel in person, if their feelings are same as expected, they will have a positive impact on the image of the hotel.

RESEARCH CONCLUSIONS AND PRACTICE IMPLICATIONS

Conclusions

With the development of the Internet and smartphones, the phenomenon of online shopping and online hotel reservations can be seen everywhere, and online comments have gradually become the object of research in academia. This paper conducts research on the five-star/luxury hotel online comments with more than 1,000 online comments in Dalian, uses text analysis technology to deeply analyze consumers' perception of the hotel image. The research results show that in the results of emotion analysis, positive emotions account for a larger proportion, followed by neutral emotions and negative emotions. This shows that most consumers have a positive attitude towards the image perception of luxury hotels in Dalian. Hotel location, room conditions, environmental hygiene, front desk, and other service staff's service attitude and professional ability are the focus of consumers' attention, and they are also the main factors that affect consumers' perception of the hotel's image.

Practice Implications

Based on the above analysis, this article puts forward the following practice implications.

First, improve the facilities and equipment of the hotel. Hotel facilities and equipment will directly affect the image of the hotel in the minds of consumers and sometimes threaten the personal safety of consumers. Hotels should regularly inspect and maintain frequently used facilities and equipment, such as air conditioners, elevators, toilets, telephones, and bathtubs commonly used by consumers. The hotel should ensure that all facilities and equipment meet the standards that a five-star hotel should have before consumers move in.

Second, strengthen the training of service personnel and improve their service awareness. Many consumers commented that the quality and professional abilities of service personnel are low, which is inconsistent with the hotel grade. Therefore, hotels should strengthen professional training for all service personnel, encourage employees to provide consumers with positive and enthusiastic services, pay attention to the details of the service process, and make consumers feel at home.

Third, hotel management should pay attention to consumer comments. In these comments, consumers will express their true experience of staying. Some comments praise the hotel for doing a good job in some aspects, some comments single out the bad aspects of the hotel, and some consumers also say about the hotel. The areas that need improvement and upgrades put forward their own suggestions. The true feelings from consumers are the entry point for the hotel to improve its services and improve its management level. However, many hotels respond to a single stereotyped content under the comments of different consumers, and most of the replies are not appropriate to the consumer's comments, so consumers will also feel that their comments have not been taken seriously.

Fourth, find the correct positioning and improve the level of competition. In the word frequency statistics, the frequency of the word "not worth" is also higher. Clicking on the detailed content of the word and simply reading the comments related to the word "not worth," you can find many things that make consumers feel worthless—the main aspect of the hotel, but the details such as mineral water and toiletries. Many consumers complain about the use of the cheapest mineral water on the market and the toiletries of unknown brands, which makes them suspect that they are not living in a five-star hotel, and their overall satisfaction with the hotel is at a low level.

Research limitations and Prospects

This article inevitably has some shortcomings. According to the shortcomings of this article, future research can expand the scope and samples of the research. For example, it can conduct research between the same star-rated hotels in different cities or conduct research between different star-rated hotels in the same city to study consumers' perceptions of the hotel from multiple perspectives. Research in this area can also strengthen empirical analysis, combining online comments and questionnaires to conduct research. Future research can conduct preliminary analysis through online comments and then select variables with obvious effects to design the questionnaire so as to design the questionnaire in an all-around way—research on the perception of hotel image.

REFERENCES

Adnan, M., Ahmad, N., Scholz, M., Khalique, M., Naveed, R. T., & Han, H. (2021). Impact of substantive staging and communicative staging of sustainable services cape on behavioral intentions of hotel customers through overall perceived image: A case of boutique hotels. *International Journal of Environmental Research and Public Health*, 18(17), 9123. https://doi.org/ 10.1073/pnas.0530279100

Chen, W. J., & Chen, M. L. (2014). Factors affecting the hotel's service quality: Relationship marketing and corporate image. *Journal of Hospitality Marketing & Management*, 23(1), 77-96. https://doi.org/ 10.1080/19368623.2013.766581

China Internet Network Information Center (2021). 47th statistical report on China's Internet development·CNNCI·Retrieved from http://www.gov.cn/xinwen/2021-02/03/content_5584518.htm (accessed 5 May 2021).

Durna, U., Dedeoglu, B. B., & Balikçioglu, S. (2015). The role of service scape and image perceptions of customers on behavioral intentions in the hotel industry. *International Journal of Contemporary Hospitality Management*. https://doi.org/ 10.1108/IJCHM-04-2014-0173

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

Shuai, Y.H, & Zou, G.Y. (2019). Research on the difference of customer perceived value of medium and high-end hotels based on network text analysis. *Tourism Research*, 11(4), 35-50. https://doi.org/ 10.3969/j.issn.1674-5841.2019.04.003 (in Chinese).

Sultan, M. T., Sharmin, F., Badulescu, A., Stiubea, E., & Xue, K. (2021). Travelers' responsible environmental behavior towards sustainable coastal tourism: An empirical investigation on social media user-generated content. *Sustainability*, 13(1), 56. https://doi.org/ 10.1080/09669580902928468

Sun, G. (2020). Symmetry analysis in analyzing cognitive and emotional attitudes for tourism consumers by applying artificial intelligence python technology. *Symmetry*, 12(4), 606. https://doi.org/ 10.3390/sym12040606

Xie, L., Li, Y., Chen, S. H., & Huan, T. C. (2016). Triad theory of hotel managerial leadership, employee brand-building behavior, and guest images of luxury-hotel brands. *International Journal of Contemporary Hospitality Management,* 28(9), 1826-1847. https://doi.org/ 10.1108/ijchm-01-2015-0004