Research on Comments Preference Based on Online Comments of Tourists from High-star Hotels in Guilin

Yu LAN[1, a], Yanling HUANG[1, b*], Jiansheng YAO[1, c], Yutong LU[1, d]

[a]904854540@qq.com, * huangyanlinglucy@sina.com, [c]591202503@qq.com, [d]3088109405@qq.com

[1]Guilin University of Technology School, Tourism and Landscape Architecture, Yanshan, Guilin, Guangxi, China

Abstract: Taking advantage of user-generated comments, this paper analyzes the customer satisfaction scores from 78,134 user-generated comments of 57 high-star hotels in Guilin obtained from Ctrip.com. And sentiment analysis to explore the attention factors and emotional preferences of customers, and finally classify customers according to different travel types, generate word cloud diagrams through their comment content, analyze their attention to different aspects of the hotel, and then get different types of customers' s preferences in terms of facilities and services, etc. On the basis of the conclusion, this paper puts forward some countermeasures and suggestions on hotel management, providing scientific reference for the development and marketing of Guilin hotels, besides the internal management and service optimization of hotels, and then contributing to the high-quality development of Guilin's accommodation industry and tourism.

Keywords: User comments; Customer preference; Satisfaction; Word frequency analysis; Ctrip

1. Introduction

Information technology has promoted the change of consumption structure from survival type to enjoyment type, and consumers will no longer passively accept the services provided by businesses, but will spontaneously put forward their own preferences. User preference is the comprehensive result of tendentious choices, user cognition, psychological feelings and rational economic trade-offs when considering goods and services. Understanding tourism preferences plays an important role in better designing tourism products and more targeted marketing in tourism professionals, and at the same time, tourists can get a better experience from the consumption process of tourism products. The behavioral preferences of hotel booking choices can help segment the market and conduct targeted marketing. In terms of research content, the construction of hotel user preference model is not perfect, and there is a lack of theoretical model construction from the perspective of customers. At present, many scholars have conducted research on tourism behavior preference. Foreign scholars pay more attention to empirical research, mainly from tourists' preferences and choices for reception facilities, travel behavior patterns and tourism products [1]. Domestic scholars pay more attention to the definition and

connotation of tourism preference, as well as the characteristics and causes of behavior [2-3]. Research methods mainly include structural method and non-structural method. Early scholars used structural method more, and used standardized quantitative evaluation data to conduct quantitative processing of data through scientific statistical analysis tools. Common methods include variance analysis, IPA analysis, scale analysis and so on. In recent years, with the development of Internet technology, unstructured research methods have been applied more and more, which can obtain objective perception of tourism more directly, mainly including content analysis method, photo analysis method and so on [4].

Guilin is a world-famous tourist city, the first batch of national historical and cultural cities in China, and an excellent tourist city in China. Its landscape scenery is world-famous, and it has enjoyed the reputation of "Guilin landscape is the best in the world" for thousands of years. Guilin will also make every effort to promote the upgrading of Guilin's international tourist attractions, so as to ensure that a world-class international tourist attraction will be fully built by 2025. With the high-quality development of Guilin's tourism, the accommodation industry should be constantly improved to bring better experience to customers. Some popular hotels and high-star hotels should adapt to the present situation, get to know customers faster, and improve their own management is not enough to provide better services.

In view of this, through high-frequency word analysis and emotional analysis, this study explores 78,134 comments of five-star hotels in Guilin collected from Ctrip official website (five-star hotels divided by web pages), classifies them into different types of customers according to their travel types, studies the hotel preference characteristics and differences of different types of tourists, and puts forward countermeasures and suggestions for the problems, which can provide scientific reference for the development and marketing of Guilin hotels, besides the internal management and optimization of hotels, and further contribute to the high-quality development of Guilin's accommodation industry and tourism.

2.   Data sources and methods

2.1 Research objects

Ctrip Group (Trip.com Group) is the world's leading one-stop travel platform, and its platform can provide a complete set of travel products, services and differentiated travel content for users around the world. The Group can provide more than 1.2 million kinds of global accommodation services, more than 480 international airlines, and more than 310,000 local activities. Therefore, this paper searches and screens hotels in Guilin based on this platform. Enter "Guilin" in the column of destination/hotel name, and all other options are in the default state, resulting in 4,893 hotels, including business hotels, budget hotels, homestays and other different types of hotels. In order to take high-star hotels as the research object more pertinently, the five-star (diamond) level is selected in the star price column, and 57 high-star hotels are obtained.

2.2 Data collection and processing

Taking the customer comment areas of 57 hotels as the collection objects, 78,134 pieces of comment information are captured by Houyi collector, and several fields of data are obtained, including user's head, user's name, number of comments, reservation type, travel type, comment

content, score, etc. Data is exported in the form of EXCEL table.

In the process of data processing, it is found that there are some missing information in some comments, such as user's name, rating to the hotel, etc., but it does not affect the analysis of hotel preferences, so all comments with complete comments are valid. According to the travel types given on the webpage, the types of customer travel are divided into seven types: business travel, couples travel, friends travel, family and children, traveling alone, booking on behalf of others (without analysis). All the comments and the subdivided six types of comments are put into the text file respectively, and the characteristics of different types of customers' preferences for hotels are studied from their comments.

2.3 Research methods

In recent years, the research on tourism preference based on User Generated Content (UGC) has grown rapidly. The data types used by this kind of methods mainly include online text data and online photo data. Compared with multivariate analysis based on questionnaire survey, the greatest advantage of content analysis based on UGC method is that it can obtain tourists' complete psychological perception.

Sort out the data to get the number distribution of travel types, and then analyze the satisfaction of different types of customers according to their ratings of hotels. In Python3.7 environment, jieba word segmentation package, wordcloud package and SnowNLP module are installed. The code implementation process is to select the comment content text to be analyzed, and conduct high-frequency word analysis and emotion analysis. Then use jieba word segmentation package and word cloud package to analyze the text, and manually filter out the stop words such as "Hotel, De, Very, Good, Great, Feeling, Just, All Yes, Harmony, OK, Good, Satisfied, Being, Me, Besides, But, I, Also, Satisfied, Like". These functional words are more common, have less practical significance, and have little effect on exploring customer preferences. Finally, a word cloud map is generated, which can reflect the frequency of words in user comments. The larger the word shape, the more frequently the customer mentions it, and the more the customer cares about it. According to the word cloud images generated by different types of customer comments, we can judge the different aspects of hotel facilities and services valued by different types of customers, so as to obtain customer preferences. Besides, the emotional score is calculated and drawn in the icon at the same time as the score, so that the relationship between emotion and score can be judged.

3. Analysis of customer preferences

3.1 Customer travel types and satisfaction

Among the comments obtained, the comments are classified according to the types of travel, and the number of comments and customer ratings of each type of travel are obtained (Table 1). The scores range from 1 to 5. In official website, 4 is very good, 5 is excellent, and the rest are not explained. Therefore, 5 points can indicate that customers are very satisfied with the hotel, and 1 point indicates that they are very dissatisfied. Through the analysis, it is concluded that among the customers of Guilin's high-star hotels, the type of travel by family and children accounts for the largest proportion, followed by friends' travel and less travel alone, which is in

line with the actual situation. As a tourist city, Guilin's customers mostly travel for family and friends, travel alone and travel less for business.

As for hotel satisfaction, it can be seen from Table 1 that the proportion of customers who travel with friends and couples who give high marks is higher, and they are more satisfied with the service provided by the hotel when traveling with friends, which may be due to the fact that these customers will do some strategies before booking the hotel and have a better understanding of the hotel situation. Customers who make reservations on behalf of others are less inclined to give high marks, because when entertaining friends, their requirements for hotels will also rise in order to give them a better accommodation experience. Customers of business travel type give the highest proportion of low scores, while customers of couples travel and family parent-child type give lower scores. Business travelers often take their jobs with them, and the pressure is high, and the expected service and the requirements for the hotel environment are also rising. However, couples and families travel more relaxing, and the requirements for the hotel are relatively low. Therefore, the perceived satisfaction of different types of customers is different, and hotel personnel need to make corresponding service adjustments according to different types of customers. In addition, it can be found that the customers of this type of lovers' travel are more inclined to give high scores when scoring, and fewer people give low scores, which shows that this group is generally satisfied with hotels, and there are not many additional requirements.

Table 1. Comments on different travel types

Type of trip	Number of comments	Number of 5 cents	Proportion (%)	Number of 1 cents	Proportion (%)
Business trip	8183	6206	75.84	69	**0.84**
Couple outing	9401	7206	**76.65**	28	**0.30**
Friends outing	10708	8307	**77.58**	50	0.47
Parent-child travel	34303	25294	73.74	135	**0.39**
Travel alone	3354	2552	76.09	25	0.75
Generation booking	663	354	**53.39**	3	0.45
Others	11511	8915	77.45	245	2.13

3.2 Analysis of high frequency words

Use Python to extract high-frequency words from Guilin tourism network texts, and sort the top 100 high-frequency words (Table 1). Most of the high-frequency words are related to hotel facilities and evaluation, hotel services, surrounding attractions and emotional evaluation, among which 38 words related to hotel facilities, 22 words related to hotel services, 10 words related to surrounding attractions and 9 words related to emotional evaluation. Words related to Guilin tourist destinations account for 10% of the top 100 high-frequency words, such as Lijiang River, Shili Gallery, Yulong River, Xiangbi Mountain and Liangjiang River. There are the most entries about hotel facilities and evaluation, with a frequency of 437,112, such as room,

breakfast, location, front desk, restaurant, cleanliness, delicious food, etc. It is concluded that customers pay more attention to hardware factors such as hotel facilities. There is only one word of catering category, which shows that it pays less attention to the food in accommodation than other aspects, which shows that the hotel's food does not display Guilin characteristics, but this factor is an indispensable factor in the tourist travel process. Emotional words can be used to express the perception of hotel accommodation, such as satisfaction, like, worth, great, happy, etc. They are almost positive words, which fully shows that most of the experiences brought by Guilin's accommodation to tourists are beautiful. Finally, some service nouns in high-frequency words reflect tourists' preference for some additional services to a certain extent. These services or items can bring tourists a better experience, and they think that hotels are very intimate and willing to recommend them to people around them, such as housekeepers, pick-ups, afternoon tea, fruits, photographers and so on.

3.3 Emotional image analysis

Using the SnowNLP module of Python, the emotional analysis of the crawled Guilin hotel customer comment text (Figure 1) shows that the main emotional words are "satisfied, liked, worthwhile, great and happy", and it is concluded that the overall emotional tendency of tourists' comments is positive emotion, accounting for 77.43%. Neutral emotion and negative emotion accounted for 0.15% and 22.42% respectively. Generally speaking, tourists' emotions are mostly positive, and the main emotional vocabulary in the text is also positive, which echoes each other. However, negative emotions should also be paid attention to. Among the top 100 high-frequency words, the words expressing emotions appear "general" with a frequency of 5590, which shows that there is still a certain gap between hotel facilities and services and customers' expectations, so there is a certain degree of negative emotions in the comments. Hotel officials need to pay attention to this kind of comments and solve the corresponding problems, which will help to improve customers' negative emotions.
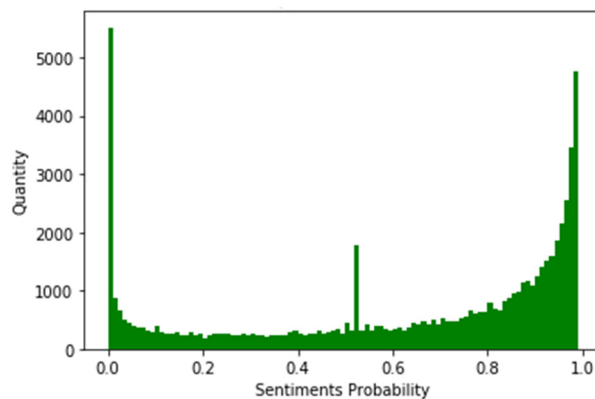

Figure 1 Emotional analysis

3.4 Analysis of customer preference differences

In order to explore the customers' concerns about hotels, Python3.7 is used to analyze the word frequency of all customers' comments, and the results of word cloud images are shown in Figure

2. In all customers' comments (Figure 2), the most important thing is the hotel service and the room. The location, breakfast, environment of the hotel are also the factors that customers pay great attention to. In addition, the supporting services such as the service attitude of employees, front desk and the facilities are also mentioned by customers more and paid more attention to. Different types of customers have the same concerns about hotels in general, which shows that although different types of travel, their preferences for services and facilities of Guilin's high-star hotels are roughly the same.

Figure 2 Customer comments word cloud

4.   Conclusion and discussion

4.1 Conclusion

It is found that when customers score according to their satisfaction, the scoring characteristics of different types of customers are different, which is due to the inherent heterogeneity of accommodation needs of different groups of travelers. Different types of travel determine the nature of customers and their expectations and requirements for services. For couples traveling, they are more inclined to give the hotel a high score than a low score. They have a higher "acceptance" of the hotel service, and they are also the target group for better communication and service.

Most of the high-frequency words in hotel comments are related to hotel facilities and evaluation, hotel services, surrounding scenic spots and emotional evaluation. The main emotional words are "satisfaction, like, worth, great and happy". The overall emotional tendency of tourists' comments is positive emotion, which reflects customers' preference for hotel facilities and services and surrounding attractions to a certain extent.

For hotel services and facilities, customers' preferences are service, hotel location, breakfast, room, environment, etc., which are the important and most basic factors that affect whether the hotel can satisfy customers. In addition, the research also found that although the concerns of different types of customers are roughly the same, in some aspects, due to the needs of different types of customers, the concerns will also be slightly different, and customers' preferences for

breakfast service, room service and front desk service are different. Customers who are on business trips and family-parent type will pay more attention to breakfast, while customers who are on a couple trip, friends trip and booking on behalf of others will pay more attention to room facilities. Hotel managers should take targeted improvements according to different types of needs.

4.2 Discussion and suggestions

This study has some implications for hotel management and operation. First, according to different types of customers, judge their needs to provide different services to improve their satisfaction; second, according to customer preferences, improve hotel services and facilities to improve service level, and then provide customers with high-quality services.

In view of some problems found in the research, the following countermeasures and suggestions are put forward. First of all, after determining the types of subscribers on the platform, we can provide targeted services according to the types of customers. From the above, it can be seen that the customers who are on business trips give a relatively high score, while the customers who are booking on behalf of others give a relatively low floor space. According to this situation, the hotel needs to pay more attention to the feelings of these two kinds of guests after arriving at the store, and can upgrade more comfortable and quiet rooms for business guests to improve customer satisfaction. At this time, because the customer's reference point is "the room with poor environment", at the same price, they enjoy the free upgrade service, and for them, the final result is greater than the reference point, and they will feel "gain", and then their satisfaction will be improved. For customers who are booked by friends, we should pay more attention, because most of them are foreign customers or older tourists, who are unfamiliar with Guilin and have a sense of strangeness, so staff can communicate with them more to find out their needs. However, the group satisfaction of lovers' outing and friends' outing is higher, and you can also ask their suggestions for hotels. As long as it is slightly improved, the satisfaction of this type of customers can be continuously improved with lower time and money costs.

Secondly, most of the time, customers pay more attention to the quality of their rooms and services, and detailed services such as breakfast and front desk are also valued by customers, especially for business guests and customers traveling with their parents and children. When some detailed needs are well met, they will even exceed customers' expectations, and customers will continue to choose this hotel for one of the outstanding services in the future. Therefore, hotel managers should improve certain service items in a targeted manner according to the hotel types and customers they receive. For example, if the hotel mainly receives business guests, the cost of hotel breakfast can be appropriately increased to improve the quality of breakfast. And so on, it is necessary to improve the service in the hotel according to the corresponding types of customer preferences.

Finally, we should also pay attention to the pricing strategy. With the development of Internet and mobile information technology, consumers can easily search all kinds of product information at low cost. The application of third-party price comparison software also makes it easy for consumers to obtain pricing information of similar products. In addition, many merchants will provide historical pricing information of products in their web pages, so in the era of e-retail, consumers can easily obtain historical prices of products, historical and current prices of similar products, which makes consumers' buying behaviors and buying habits undergo

great changes, which not only affects online transactions, but also affects offline transactions. By integrating the product price information obtained from different channels, consumers form the price expectation of the target product, which is called "reference price", that is, an internal standard, based on which consumers evaluate the sales price of products [5]. When the sales price is higher than the reference price, consumers perceive "loss"; conversely, when the sales price is lower than the reference price, consumers perceive "gain". Therefore, when the hotel staff are pricing the hotel, the room price at the same time period should be stable, and it should not be too different from the previous price, nor should it be lowered at will during the off-season, otherwise it will give potential customers a new lower price reference point [6-8].

References

[1]Kim, K, Jogaratnam, G. Activity preferences of Asian international and domestic American university students: An alternate basis for segmentation[J]. Journal of Vocation Marketing, 2003, 9:260-270.

[2]Jiao, Y. Analysis of tourist decision model based on tourists' preference and perceived risk [J]. Journal of Tourism Science, 2006,5:42-47.

[3]Liang, J, Zhang, W.Q. Pedigree development of tourism products based on activity preference market segmentation: Taking Kaiping Diaolou World Cultural Heritage as an example [J]. Journal of Tourism Science, 2009, 24(9):36-42.

[4]Chi, M.M, Pan, M.Y, Wang, W.J. A Cross-platform Comparative Study of Shared Accommodation and Hotel User Comment Text: Theme Social Network Based on LDA And emotional analysis [J]. Library and Information Work, 2021,65(02):107-116.

[5]Monroe K B. Buyers' subjective perceptions of price[J]. J of Marketing Research, 1973, 10(1): 70-80.

[6]Helson H. Adaptation-level theory[M]. New York: Harper and Row Publishers, 1964.

[7]Kalyanaram G, Winer R S. Empirical generalizations from reference price and asymmetric price response research [J]. Marketing Science, 1995, 14(3): 161-169.

[8]Baucells M, Weber M, Welfens F. Reference-point formation and updating[J]. Management Science, 2011, 57(3): 506-519.