Analysis of Airbnb User Rating Factors—Taking Beijing as an Example

Jiaqi Zhang[1,a,*]

[1]*Shanghai Shinan High School, Shanghai, 200011, China*
[a]*quan.zhang@kodak.com*
[*]*Corresponding author*

Keywords: Airbnb, Shared B&B, Area, Prices, Reviews

Abstract: With the rapid development of the sharing economy, shared accommodation has become a choice for more and more people to travel. As a representative of the shared bed and breakfast platform, Airbnb provides information on shared Bed and Breakfast (B&B) in various cities worldwide. This paper collected information from Airbnb operations in Beijing, where the development of the shared accommodation industry is more mature. It analyzed it regarding price, location evaluation and other influential factors. The results show that unbalanced regional distribution of listings, variability in price changes, and customer comments impact the sales of shared listings. Based on the study's results, this paper proposes the need for correct guidance and strengthening macro-control. In addition, there is a need to pay attention to customer reviews. This paper's findings help identify the attributes consumers pay more attention to when choosing a B&B and provide good guidance for operators to enhance improved services.

1. Introduction

As one of the representatives of shared Bed and Breakfast (B&B) platforms, Airbnb provided many users with information about shared B&B in various countries and regions worldwide, such as the location (longitude, latitude), price, housing characteristics and types, etc. [1]. People no longer stay in hotels in the traditional sense when they travel for leisure but seek a comfortable and personalized environment to relax, and this new model of shared accommodation meets customers' requirements and, therefore, becomes the first choice of accommodation for people traveling abroad [2]. Customers traveling to a different place for a holiday are usually shorter and choose accommodations more focused on its use and convenience than investment property. So, in the rapid development of technology today, the concept of consumption has changed, and the emergence of shared lodging sites to meet customer demand greatly [3]. Airbnb has three types: entire, private, and shared rooms [4]. On Airbnb, the location, facilities, pictures, traffic, prices, etc., of shared lodgings are open and transparent, and there are high, medium, and low-grade listings for customers to choose from, which meets the different expectations of customers [5]. Airbnb provides much information on shared lodgings listings, but customers need help seeing the specific housing situation because it is displayed online. Online reviews are also the main basis for customers to make purchases. The main basis for customers to place orders and purchases [6]. Whether the

authenticity of the information and service attitude of the listing is as described by the host, the user has the most right to say whether it is a good experience or a bad experience.

The user can objectively react to the comments and ratings [7]. The comment function in the Airbnb platform highlights the openness of the information, and the number of comments, ratings and even the host's response will affect the number of potential customers [8]. Airbnb's complicated data information can be analyzed to find its patterns, such as the sparseness of spatial distribution, price changes, and other intrinsic correlations with the distribution of transportation and attractions, which can be used as a reference for policy-making [9]. In the macro analysis, the provision of shared accommodation housing is mostly proportional to the population density [10]. This paper analyzes the regularity of shared accommodation by studying the relationship between the factors affecting it, which is significant in theory and practice.

As a metropolis, Beijing has many Chinese and foreign tourists, and its shared bed and breakfast development is also more mature. Taking Beijing as an example, this paper collects and analyzes information about Airbnb shared listings and analyzes the influencing factors of Beijing's shared listings' sales in terms of region, price, and number of reviews more systematically. It helps identify the attributes customers pay more attention to when choosing shared lodgings and guides operators to improve the quality and services.

2. Methodology

2.1. Data Source and Description

This paper uses the Airbnb as the data source. Keywords such as "Airbnb user hotel ratings", "Beijing hotels", and "Airbnb" were searched through the Kaggle webpage to collect Airbnb related data in Beijing (shown in Table 1).

The dataset includes data variables related to the location, longitude, latitude, etc., information on the type of listings, data variables related to hosts, and data variables related to listing reviews. Each final acquired sample data listing corresponds to a unique ID number.

Table 1: Example of some of the content in the grabbed data.

House ID	Location	Latitude	Longitude	Room type	Price	Average number of reviews
1455726	Dongcheng	39.93958	116.43485	Entire room	403	0.33
1466681	Chaoyang	39.93712	116.45089	Entire room	637	1.26
1456491	Chaoyang	39.94115	116.44122	Entire room	416	0.03
1456491	Chaoyang	39.94032	116.44227	Entire room	416	0.43
2562442	Xicheng	39.92412	116.37063	Private room	8898	0.15
1456491	Chaoyang	39.93879	116.4412	Private room	416	0.1
3113849	Haidian	39.94956	116.34887	Private room	396	0.01
4346816	Dongcheng	39.939	116.39129	Private room	289	1.55
4346816	Dongcheng	39.93923	116.39094	Private room	262	1.45
4346816	Dongcheng	39.93946	116.39161	Private room	242	2.46
527062	Miyun	40.68602	117.17304	Private room	678	0.16
13347104	Fengtai	39.84922	116.3628	Share room	101	1.51
24533082	Dongcheng	39.91614	116.42817	Share room	101	0
24787175	Haidian	40.00126	116.31614	Share room	127	0.43

2.2. Indicator Selection and Description

Data from January 1, 2001 to September 7, 2019 were selected for this study to avoid abnormal fluctuations in data caused by the epidemic. In the face of a large amount of data, there is a duplicate capture and mismatch phenomenon, which needs to be sorted and merged. For the repeated data bars to merge and collate, the data that do not meet the conditions must be eliminated. After organizing the data, 16,244 listings data were obtained.

The statistical data analysis method analyzes the influence on consumers' intention to buy or repurchase in the following text from the changes in price, evaluation, location and other factors in the data.

3. Influence Airbnb Beijing user data and analysis

This section analyses the regional attributes, price and review (number/rating) attributes, room types, cancellation policy, instant booking, minimum stay days and other advanced attributes, and the most commonly used text mining.

3.1. Regional Attributes

The rapid development of shared bed and breakfast, what kind of spatial state is presented in the supply of listings, user demand for shared bed and breakfast regions, and other factors will impact the supply and demand of shared listings.

Understand the regional distribution of shared listings to better grasp each region's number. Beijing has 16 districts under its jurisdiction. Figure 1 shows the spatial distribution of shared listings through the scattered points in Beijing, which is concentrated between approximately 39.8-40.2 degrees north latitude and 116-116.7 degrees east longitude. This area is precisely the center of the city, such as the Chaoyang District and the Haidian District, which has a very high population density and convenient transportation and life, and the density of the listings is also the highest. In the non-center area, the density of shared housing distribution is greatly reduced. Figure 1 shows a small amount of obvious aggregation of shared listings around 40.4-40.6 degrees north latitude and 116 east longitude, such as in Yanqing, Miyun and Tongzhou District.
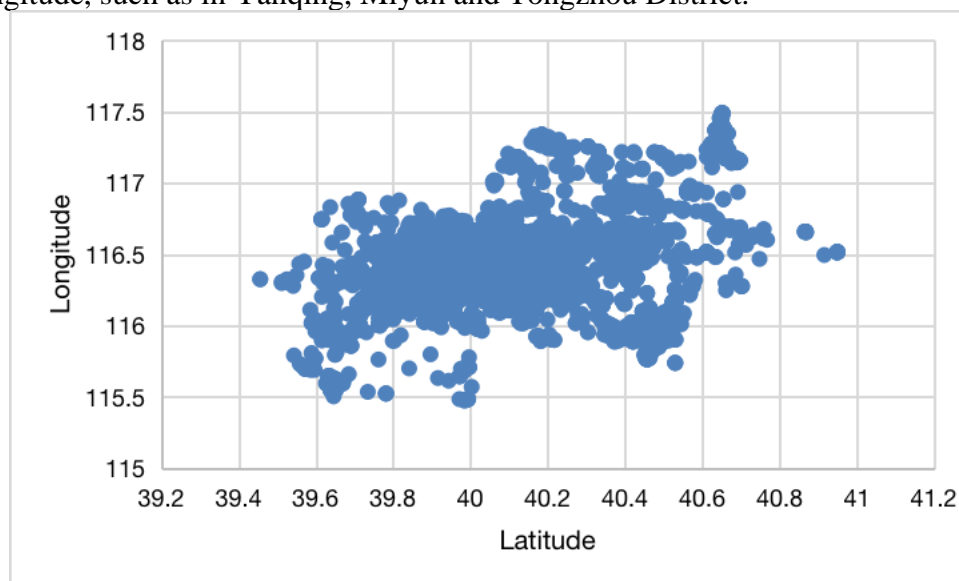

Figure 1: Distribution of shared housing by districts in Beijing. (Photo credit: Original)

The number of online hits and location. As seen in Figure 2, the Chaoyang District, located in the city's center, has the highest number of customer clicks, close to 9,000, while the more remote Mentougou and Pinggu Districts have only about 100 customer clicks. It can be seen that there is a strong correlation between customers' online search for shared listings and their geographic location.
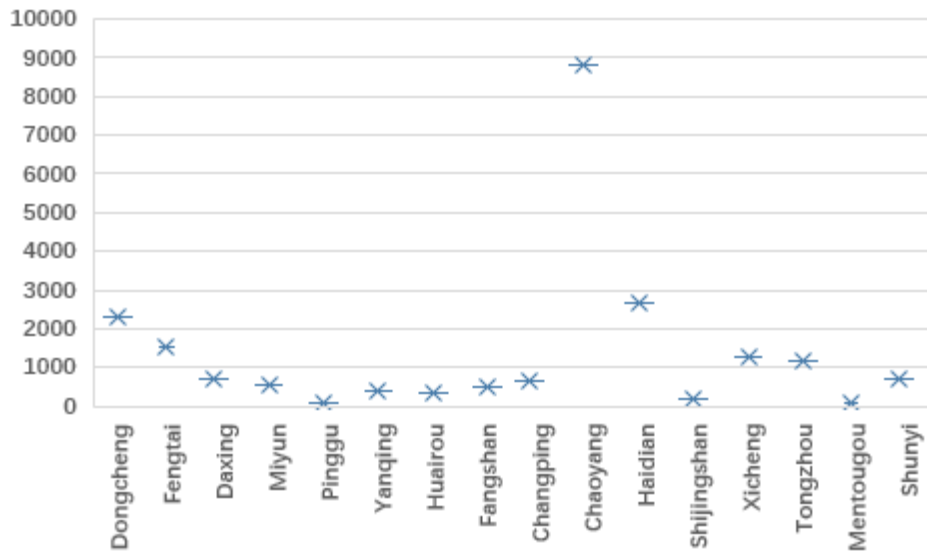

Figure 2: The relationship between the number of clicks and the location. (Photo credit: Original)

The price and location. Table 2 shows the price data of shared listings in each district and county in Beijing.

Table 2: Price data for shared listings by district in Beijing

Neighborhood	Average Price	Min	25%	50%	75%	Max
Dongcheng	421.611	67	302	423	550	752
Fengtai	338.227	67	221	336	450	745
Daxing	321.053	67	208	302	396	738
Miyun	377.053	47	248	356	497	718
Pinggu	364.554	81	235	325.500	483	698
Yanqing	436.903	67	289	456	597	725
Huairou	432.915	81	302	450	577	752
Fangshan	272.768	67	201	255	307	718
Changping	281.761	67	148	235	389	752
Chaoyang	358.735	34	215	356	490	752
Haidian	344.601	67	208	302	490	752
Shijingshan	377.842	101	255	356	515	698
Xicheng	403.899	60	268	396	517	752
Tongzhou	309.661	60	228	302	389	752
Mentougou	308.728	67	255	275	356	671
Shunyi	360.825	67	242	349	463	752

For better visualization, this paper uses the violin plot and boxplot to analyze the data in Table 2 (shown in Figure 3). The average price of 16 districts in Beijing is RMB 356.96, and RMB 358.735 in Chaoyang District, which price represents the average level of shared listings in Beijing, and whose listings have a superior location but the number of shared listings ambassadors their prices to remain reasonable. Huairou and Yanqing, on the other hand, have famous tourist areas, and

although their locations are remote, the number of shared listings is low, and their prices are more expensive. Tongzhou and Daxing have lower prices.
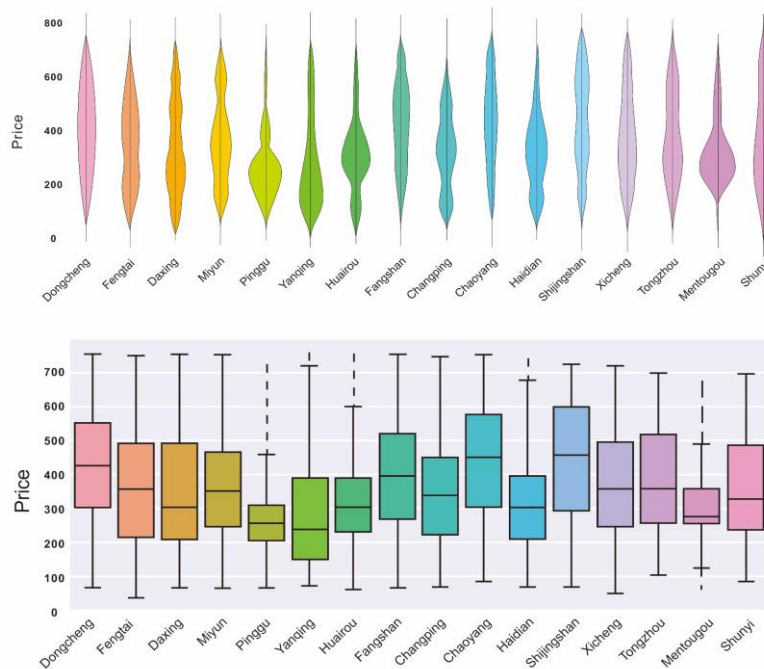

Figure 3: Rates and Shared Housing Distribution by District in Beijing. (Photo credit: Original)

Changes in prices of different room types in different areas. The price trend for whole houses is the same in all regions, with prices slightly higher in Huairou. Prices for individual rooms are higher in Yanqing and Huairou and lower in other districts. The overall price of shared rooms is low. Figure 4 also reflects that single and shared rooms in some districts are more expensive than whole houses in other districts, mainly due to demand exceeding supply.
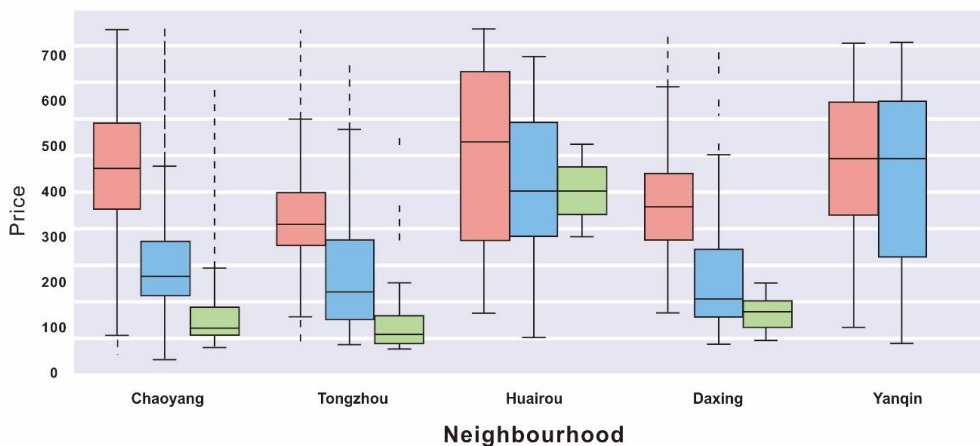
Figure 4: Price distribution of different housing types in different districts in Beijing.
(Photo credit: Original)

3.2. Price and Review (Number/Rating) Attributes

Price and number of reviews are the most important features of shared listing sales, and the

number of reviews is logically related to the number of customers or guests; they can be used as important indicators. To better grasp the listing prices of shared lodgings in Beijing, this paper calculates the main numerical features of all listing prices so that the overall distribution of listing prices can be fully represented (shown in Table 3).

Table 3: Characteristics of price distribution of all properties in Beijing.

Numerical Characteristics	Price
Mean	611.20
Standard deviation	1623.00
Minimum value	0.00
25%	235.00
50%	389.00
75	577.00
Maximum value	68983.00

From the data, the distribution of listing prices has a relatively large degree of dispersion. The average price of all shared listings in Beijing is RMB 611, which is medium, but the standard deviation reaches RMB 1,623.

The price of shared listings is the most important factor for either owners or tenants, with owners aiming to make a profit and tenants wanting to live in cost-effective listings for less money. Figure 5 shows the average price of shared listings in various districts and counties in Beijing.
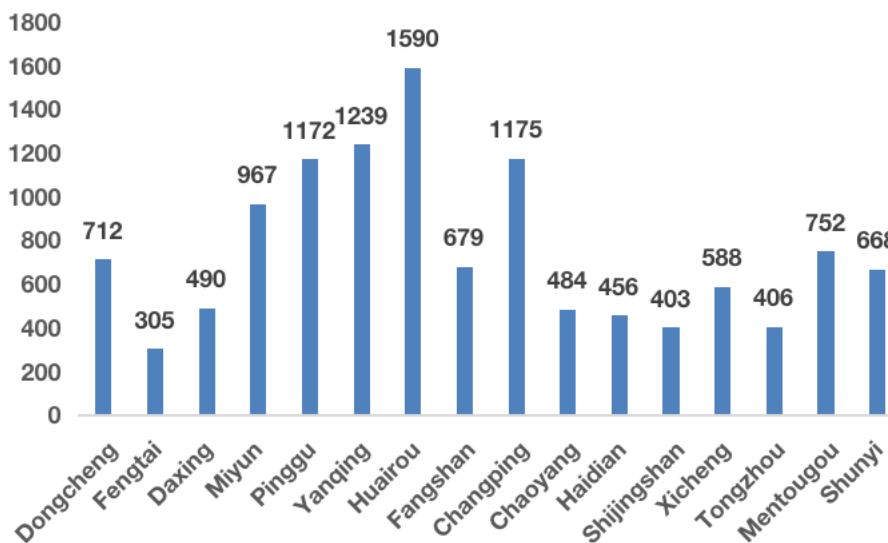

Figure 5: Average price of shared housing by district in Beijing. (Photo credit: Original)

As can be seen from Fgure 5, Huairou District has the highest average price of shared listings, which has exceeded RMB 1,500, which is caused by the fact that there are fewer listings, and individual high-priced listings have pulled up the average; Yanqing County, Changping District, Pinggu District and Miyun County also have a higher average price, which is far away from the city center, but have beautiful natural scenery. Although there are few shared listings, they are mainly high-quality, with many customers. The customers are also high consumers, so the number of house prices is higher. The overall price of other counties is relatively low, which belongs to the common people and represents the general price of common listings.

The number of reviews is an important attribute. Related studies show that there is a strong correlation between the number of online reviews and the sales volume, and often, the number of

reviews is directly proportional to the number of guests so that it can be regarded as an evaluation criterion. Chunying Li studied the number of monthly reviews as the sales volume of listings when conducting a study on the sales volume of shared B&B [11].

Figure 6 shows the number of reviews of shared listings in each district and county in Beijing. The maximum number of reviews per customer ID and the average per month. The maximum number of reviews per ID is within 50, and the average number of monthly reviews is within 5. Mining the features affecting user experience from tenants' review data to explore the core needs of tenants on shared accommodation platforms. At the same time, hosts can respond to customer reviews and even avoid welcoming customers who are not expected by viewing their review information.
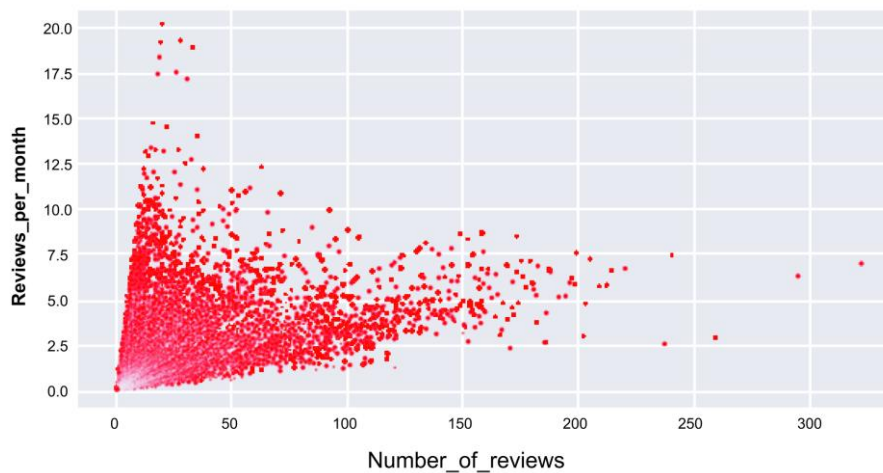

Figure 6: Analysis of the number of comments on shared listings by district in Beijing
(Photo credit: Original)

Relationship between ratings and prices. Figure 7 shows that the listings with high prices have high ratings, meaning they are more pleasant regarding housing hygiene, facilities, services, and the surrounding environment.
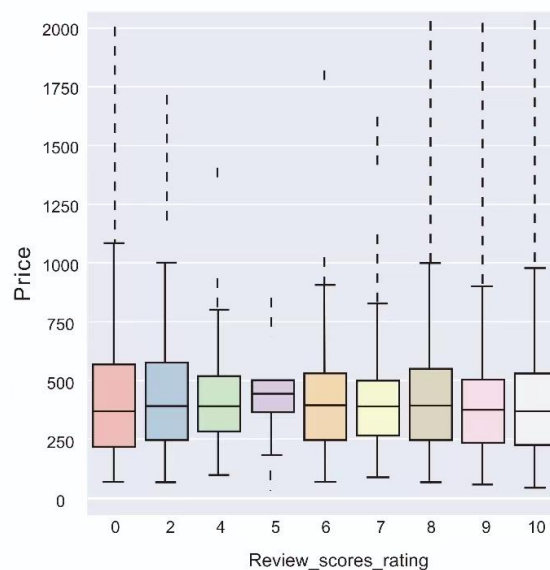

Figure 7: Shared Housing Ratings and Prices in Beijing. (Photo credit: Original)

3.3. Room Type

Figure 8 shows that the price of a single room in Beijing is 10 times higher than that of a shared room, and the price of a whole house is 20 times higher than that of a shared house. Landlords are more willing to provide the whole house, and tenants prefer to rent the whole house, followed by a single room.
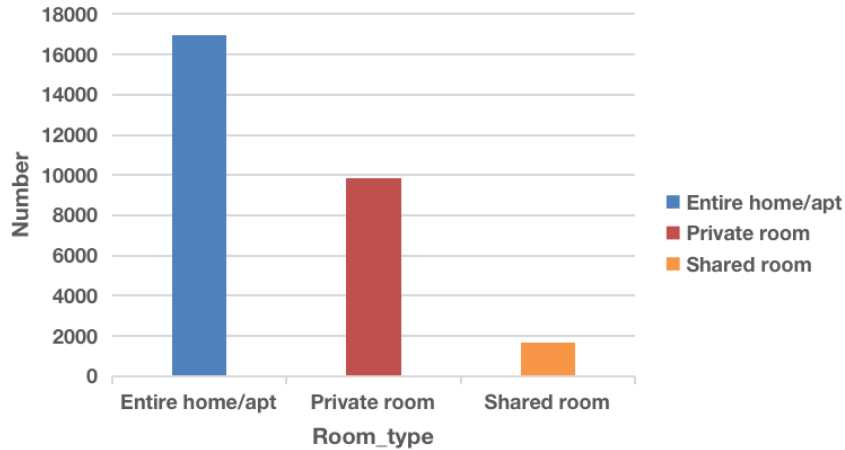

Figure 8: Types of Shared Housing in Beijing (Photo credit: Original)

Figure 9 shows that the price distribution of the three shared houses in Beijing is very reasonable, with high, medium, and low price ranges. However, the average price is different, with the whole house having the highest average price and the shared house having the lowest average price.
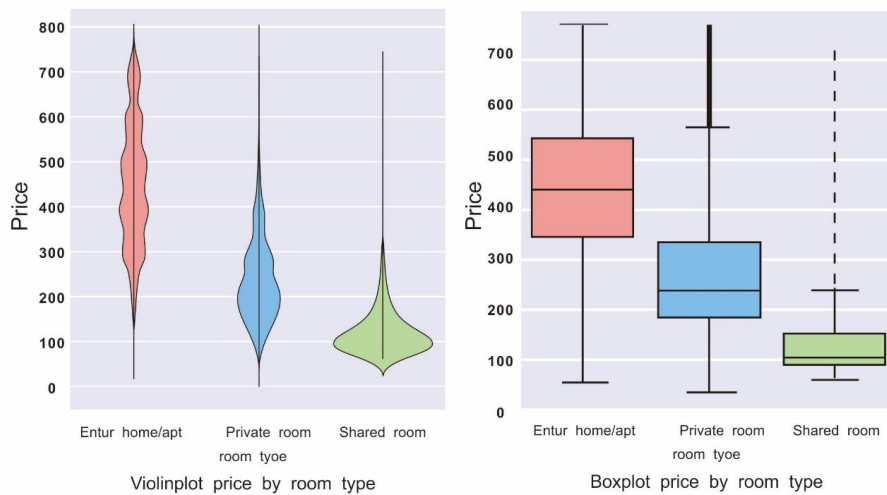

Figure 9: Shared Housing Types and Prices in Beijing. (Photo credit: Original)

3.4. Policy, Instant Booking, and Minimum Stay Days

Premium attributes such as cancellation policy, instant booking, and minimum stay days are considered personalized services by customers, and the flexibility of the accommodation policy is more likely to elicit goodwill from users. Figure 10 shows that the instantly bookable price is slightly more expensive than the price of other properties to guide customers to make an early decision. At the same time, the cancellation policy does not affect the price or customers.
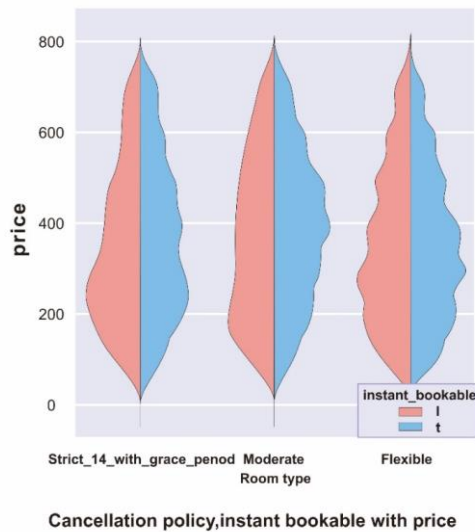
Figure 10: Cancellation Policy, Instant Booking and Price Relationship. (Photo credit: Original)

3.5. Text Mining

Online information plays a key role in the user's decision-making process, and the keywords of the advertisement information of the shared listings release become the primary factor influencing the customers to search for the listing information. For example, the Subway station indicates that customers require the location of the listing to be convenient, and the Forbidden City indicates the attractions that customers want to visit. Comfortable indicates the living environment's requirements (shown in Figure 11).
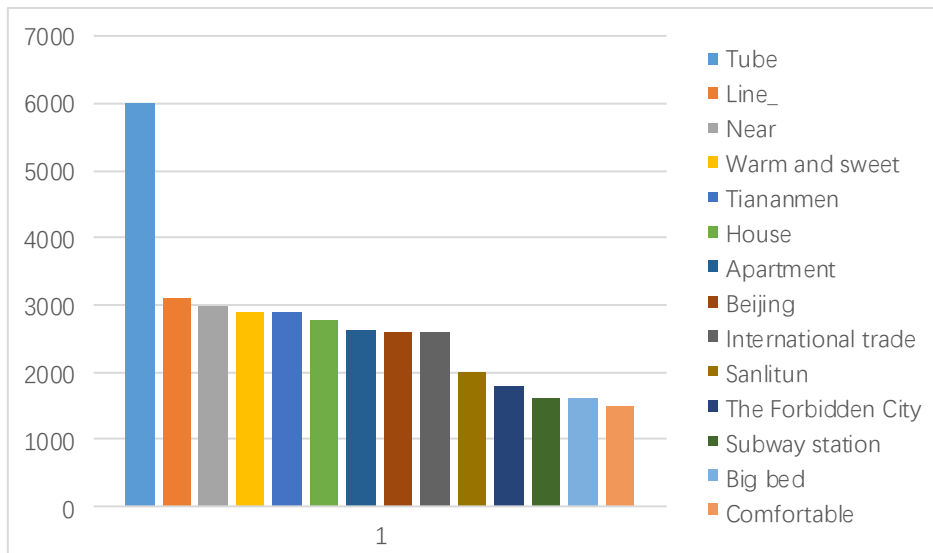

Figure 11: Keywords for Shared Housing Posting Advertisement Information.
(Photo credit: Original)

4. Conclusions

This paper explores the influencing factors of shared B&B by examining the data in Beijing.

Using the statistical analysis method, this research analyzed the data in terms of regional attributes, price and review (number/rating) attributes, room type, cancellation policy, instant booking, advanced attributes such as a minimum number of days of stay, and the most commonly used text mining.

The regional distribution of Beijing's shared listings needs to be more balanced. Shared housing sources are mainly concentrated in central urban areas such as Chaoyang. The main factors are convenient transportation, concentration of attractions, and convenient living. The distribution of listings in the remaining districts and counties is more scattered.

The overall change in the price of different types of housing in different regions is that the price of housing in the city center with the highest population density is average, while the price of housing in the districts and counties with more tourist attractions is the highest, and even single rooms and shared housing are more expensive than the entire housing in other districts. The reason for this is mainly the effect of supply and demand.

Concerned about the comments, the two-way evaluation mechanism of the Airbnb platform is an important mechanism of the shared accommodation platform to solve the problems of information asymmetry and transaction risk. The keywords of the advertisement information released by the shared accommodation have become the primary factor influencing customers' search for accommodation information.

Based on the above analysis, the following suggestions are made: First, correct guidance is needed to strengthen macro-control, especially in areas where tourist attractions are concentrated, to encourage landlords with listings to provide more information to meet the demand. In addition, it is also necessary to pay attention to customer comments so that, on the one hand, landlords can understand the needs and satisfaction of customers and the places that need to be improved. On the other hand, the first thing customers look at when choosing a B&B is the user's evaluation.

In addition, although this study has done a deeper analysis of the factors affecting the sharing of B&B, there are some areas for improvement due to the epidemic Airbnb platform data intercepted in 2019, and the post-epidemic era is bound to have great changes. These deficiencies are paid attention to as an ongoing issue to be addressed by future research.

References

[1] Zhao, C. Y. Analysis of Influencing Factors and Spatial Pattern Variability of shared B&B sales in Beijing [D]. Northeast University of Finance and Economics. (2022).

[2] Mu, X. Y. Research on Influencing Factors and Prediction Models of Online Short-Term Room Prices [D]. Hubei University of Technology, (2021).

[3] Zeng, Q. J. (2022) Research on Tourists' Shared Accommodation Consumption Behavior Based on Airbnb Platform [J]. Journal of Jiamusi University (Natural Science Edition), 40(05), 160-163.

[4] Yu, H. Analysis of Sharing Economy Business Model in the Context of Tourism [D]. Xiamen University, (2017).

[5] Xue, H. Research on the Influence Mechanism of Merchant Response on Consumer Review Behavior of Shared Accommodation Platform [D]. Chongqing University of Commerce and Industry, (2022).

[6] Li X. M. How Does User Experience Affect the Transaction Volume of The Shared Accommodation Platform? [D]. Northeast University of Finance and Economics, (2020).

[7] Tuntipong, S. A Study on The Influencing Factors of Customer and Host Experience Based on Airbnb Online Review Data [D]. Southwest University of Finance and Economics, (2021).

[8] Li, P. Research on Spatial Distribution Characteristics and Influencing Factors of Shared Accommodation in The Main City of Chongqing [D]. Chongqing Jiaotong University, (2021).

[9] Ding, F. G. Research on Developing Online Tourism Short-Term Rental Market in Beijing [D]. Guangxi Normal University, (2020).

[10] Xuan, Y. Research on the Characteristics of Shared Accommodation Spatial Patterns and Influencing Factors Based on Network Data Mining [D]. Nanjing Normal University, (2020).

[11] Li, C. Y. Shared Accommodation research on Spatial Model and Quantile Regression Model—Analysis of Influencing Factors of Consumption Behavior Based on AirBnb. Dongbei University of Finance and Economics, (2019).