

# The Study on Tourist Preference of Wuyuan Rural Tourism Based on Web Text Analysis

Mengyao Xu

Shenzhen Tourism College of Jinan University, Shenzhen, China

Email: shmily-xmy@outlook.com

**How to cite this paper:** Xu, M. Y. (2020). The Study on Tourist Preference of Wuyuan Rural Tourism Based on Web Text Analysis. *Journal of Service Science and Management*, 13, 649-658.

<https://doi.org/10.4236/jssm.2020.134041>

**Received:** July 12, 2020

**Accepted:** August 17, 2020

**Published:** August 20, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In the era of big data, network text analysis has become an important means of studying tourists' travel preferences. This study takes Wuyuan County as the research object, adopts network text analysis method, uses Octopus Collector to collect Wuyuan travel notes on relevant tourism websites, and uses ROST CM6 text mining software to perform word frequency analysis, semantic network analysis and sentiment analysis to explore the tourist preferences of rural tourism destination Wuyuan. The study found that as a sample of rural tourism, Wuyuan tourists' preference is concentrated on tourism elements such as rape flower and Huangling, which is the core advantage of the characteristic development of the scenic spot. At the same time, insufficient transportation infrastructure limits tourists' desire to visit, and the reception capacity of scenic spots needs to be improved urgently. The research results can promote Wuyuan to enhance its tourist attraction and provide reference for image shaping and marketing of other rural tourist destinations.

## Keywords

Rural Tourism, Wuyuan, Travel Notes on the Internet, Text Analysis

---

## 1. Introduction

Rural tourism refers to tourism activities that take rural natural and humanistic objects as tourist attractions in rural areas. Rural tourism in our country has experienced the early rise stage, the initial development stage and the standardized management stage. It has transitioned from the original resource characteristic dominance, agricultural industry dominance and government support dominance to the current market dominance. Since General Secretary Xi put forward

the comprehensive implementation of the rural revitalization strategy, rural tourism, as an important way to realize rural revitalization and targeted poverty alleviation, has received more and more attention from academic circles. Wei Wen investigated the current situation of poverty alleviation by rural tourism, and sorted out the difficulties in rural tourism poverty alleviation, and proposed a perfect path for the tourism poverty alleviation mechanism from both the government and the enterprise perspective (Wen, 2006). The perceived preference of tourists, as an important manifestation of tourism intention, has great research value.

In addition, with the development and popularization of the Internet, tourism activities and tourism feedback rely more and more on network sharing platforms. It has become the norm for tourists to share travel notes on the Internet, and tourism websites have become the mainstream platform for tourists to obtain tourism information and spread emotional perception (Gao & Wu, 2017). Therefore, it is of great significance to use the web text analysis method to study the tourist perception preferences of rural tourist destinations.

Wuyuan County is located in the northeastern part of Jiangxi Province, at the junction of the three provinces of Jiangxi, Zhejiang and Anhui. It is one of the birthplaces of Huizhou culture. It has rich tourism resources and is known as the “most beautiful village in China”. Wu has been developing rural tourism since 2001, and has established the position of “rejuvenating tourism through the village”. According to the Rural Tourism Analysis Report of Tuniu.com in 2018, Wuyuan, Jiangxi is at the top of the list of popular rural tourism destinations. Wuyuan has beautiful natural scenery, with 1 5A-level tourist attractions and 12 4A-level tourist attractions. It is the county with the most 4A-level tourist attractions in the country and the only national 3A-level tourist attraction named after the entire county. Wuyuan has a strong historical heritage and cultural atmosphere. Its representative culture is Huizhou culture, which is known as the “Hometown of Books” and “Hometown of Tea”. At the same time, Wuyuan is also one of the most well-preserved and best-preserved ancient buildings in China today, and has been designated as a national “eco-cultural tourism demonstration county”.

This paper will analyze the travel preferences of Wuyuan tourists from online travel notes, provide suggestions for Wuyuan to clarify its own tourism image, improve the management level of scenic spots, enrich the propaganda carrier and implement marketing strategies, and provide references and references for the development and management of other rural tourism destinations.

## **2. Literature Review**

### **2.1. Tourism Preferences**

Tourism preference is regarded as the main process before tourism decision-making, which is very important for the selective development of tourism products and marketing. Tourism preferences are formed by the dual effects of indi-

vidual influencing factors of tourists and external information stimuli. This preference exists before and after travel. Strong tourist attraction is one of the key factors in the formation of tourists or potential tourists. The cognition of a certain tourist destination or tourist product is the prerequisite for the formation of tourist preferences. Xin Li and Qing Li pointed out that the intensity of tourists' attitudes and the types and quantities of information that tourists have are the main factors affecting tourism preferences (Li & Li, 2006). Qingjin Wu analyzed that the influencing factors of tourism preference mainly come from the living environment of tourists, the intensity of public opinion and the degree of satisfaction of personal needs (Wu, 2006).

In the study of tourism consumption preferences, foreign research content is becoming more and more in-depth and research objects are becoming more extensive. Foreign scholars tend to use multidisciplinary knowledge such as psychology, sociology and economics, combined with qualitative and quantitative research methods, to transform the tourism preference characteristics obtained from the research into tourism management strategies to promote the flourishing development of the tourism market. Research on tourism preference in my country started late and is still in the early stage. At present, domestic scholars mainly conduct research on tourism preferences from demand preferences, gender preferences, destination preferences and tourism resource preferences.

Yanqiu Zhu, Hui Zhang and Rong Yang found that the actual travel experience of rural tourists is higher than the expected travel experience. In addition, they have higher travel preferences for specialty snacks and homestays, local products and crafts, and rural life experiences (Zhu, Zhang, & Yang, 2016). Yingying Jia, Yongwei Kang and Hongwei Zhou pointed out that the residents of the central region prefer Yunnan and Guizhou with complex and diverse natural environments and unique ethnic customs in terms of traditional village tourism, followed by Jiangsu and Zhejiang with many ancient water towns (Jia, Kang, & Zhou, 2016). In terms of research methods, compared with quantitative research, my country's descriptive qualitative research accounted for the vast majority, leading to many research results lacking effective data support. Most quantitative researches are based on questionnaires and field investigations.

## 2.2. Network Text Analysis Method

This paper mainly adopts the content analysis method of network text analysis. The content analysis method is a semi-quantitative research method based on qualitative content that is transformed into systematic and qualitative data for analysis. It converts the network text represented by the text language into quantitative data, so as to objectively and systematically analyze the dissemination content.

With the increasingly close connection between the Internet and tourism activities, more and more tourists make ticket bookings through various intermediary websites. Tourists' evaluations of tourist destinations are displayed on various website platforms in the form of text and images for other tourists to refer

to. Online travel notes promote the exchange and sharing of experiences among travelers, and become one of the important ways for tourists to convey tourism perception and obtain tourism information. Scholars at home and abroad have begun to introduce network text analysis into the tourism discipline, which is mainly reflected in the research of tourism destination image perception and travel satisfaction. Based on the push-pull theory, Qianqian Guo, Jingchuan Du and Yawen Li found the influencing factors of red tourist satisfaction through the analysis of online travel website evaluation (Guo, Du, & Li, 2020). Ziqing Wang and Jianhong Xue used the travel notes of Kulangsu published on Ctrip as a sample to interpret the image symbols of Kulangsu tourist destination (Wang & Xue, 2019). In general, there are relatively few researches on tourist preferences using web text analysis. Compared with other quantitative research methods such as questionnaires, online texts are richer in content and more convenient to obtain, and most of the text content is the true feelings of tourists themselves, with fewer interference factors, and more truly reflects tourists' preferences.

### **3. Methodology**

#### **3.1. Research Topics**

This paper takes Wuyuan County as the research object of rural tourism destination tourist preferences, collects travel notes from tourists on tourism websites such as Mafengwo, Ctrip and Qunar.com, and empirically researches the rural tourism tourist preferences of Wuyuan County through network text analysis. Provide feasible suggestions for the improvement and development of rural tourism in Wuyuan County, so as to better enhance the overall brand image of rural tourism destinations. This paper will be divided into these stages: data collection, data processing and data analysis.

#### **3.2. Data Collection and Processing**

This paper uses an Octopus collector to use "Wuyuan" as the key word. According to the time span from February 2018 to February 2020, 274 travel notes were collected from major domestic travel websites as samples, including 210 from Mafengwo, 45 from Ctrip, and 19 from Qunar. Then exclude the travel notes mainly based on photos and videos, poetry, prose and popular science and advertising, and get 200 valid samples, which ensures the timeliness, reference and completeness of the sample content.

Export the data collected by Octopus to Excel, and then preprocess the selected 200 travel notes samples: merge the travel notes of the same author, such as "Wuyuan Tour (1)" and "Wuyuan Tour (2)"; Delete irrelevant information in the travel notes; merge similar words, unified "beauty" and "scenery" into "scenery", etc. After processing, a total of more than 380,000 words were obtained. Finally, all the text information was saved in a txt file for further analysis of the text.

This paper selects ROST CM6 software to process the initial data text, obtains 729 participles, and conducts word frequency analysis, social network and semantic network analysis. At the same time, unnecessary and meaningless high-frequency words are filtered out. First, select keywords related to the research topic of this paper, and put words such as “Huangling”, “Shaiqiu” and “Xiaoqi” into the custom dictionary to prevent them from unrecognized or filtered; Secondly, words that have nothing to do with the research topic or have no practical meaning, such as “arrival”, “always”, “some”, etc., are put into the filtering vocabulary; then the keywords are standardized and the same meaning keywords are merged into one, such as grouping “happy” and “joy” as “happy”, or re-merging keywords that have been split by mistake, such as “Sixi” and “Yancun” merged into “Sixiyancun”; Finally, the keywords are supplemented and summarized. After comprehensive analysis, the filtered vocabulary with high frequency and strong association with other keywords is re-crawled to obtain data such as high-frequency feature words and semantic network graphs.

## 4. Results Analysis

### 4.1. High-Frequency Feature Word Analysis

The frequency of vocabulary reflects the degree of tourists’ perception of landscapes, services and products. The higher the word frequency, the higher the consensus of tourists, and the better the satisfaction of tourists with landscapes, services and products. According to the principle of importance, this paper uses the word frequency statistics function in ROST CM6 to extract 100 high-frequency keywords, and divide the keywords into six types of information: tourist routes, tourist attractions, play items, travel daily, tourist perception, and travel season. To quantify the language data, and then through high-frequency vocabulary analysis, to get the perception preferences of rural tourists in Wuyuan. Due to space limitations, this paper only displays representative keywords. **Table 1** shows the specific vocabulary classification.

In travel route information, the top three cities in terms of frequency of occurrence are “Wuyuan”, “Jingdezhen” and “Shicheng”. Compared with Shicheng County, Jingdezhen City is closer to Wuyuan County, which makes tourists more inclined to include Jingdezhen when planning tourist routes. In the tourist attraction information, the word frequency of “Huangling” topped the list, followed by sightseeing spots such as “Likeng”, “Jiangling”, “Sixiyancun”, “Moon Bay”, “Rainbow Bridge” and “Wangkou”.

The amusement item information is divided into natural and humanistic categories. In natural landscapes, the frequency of “rape flowers” is 841, followed by “red maple”, “sunrise” and “sea of flowers”; in humanities projects, “terraced fields” are the most popular, followed by “cable car” and “Shaiqiu”. “Terraces” and “Caves”. Daily travel information is mainly divided into four categories: transportation, accommodation, shopping and others. In the traffic vocabulary, “high-speed rail” and “chartered car” appear more frequently, followed by “self-driving” and “bus”. Among accommodation vocabulary, the number of

occurrences of “hotel” is 335, which is higher than other vocabulary. “B&B” and “inn” followed closely behind. Among shopping vocabularies, the word frequency of ticket-related vocabulary such as “ticket”, “pass ticket” and “reservation” accounted for the vast majority, while gourmet vocabulary such as “non-gjiale” appeared less frequently, highlighting the high degree of ticketing attention and lack of competitiveness in food. In the travel season information, due to the seasonal nature of viewing landscapes such as “showing autumn” and “red maple”, most tourists choose to travel in autumn, especially late autumn. In tourist perception information, descriptive words such as “Huizhou”, “feature”, and “most beautiful China” have a larger proportion, and perception words such as “cheap” appear frequently, indicating that the tourist satisfaction of Wuyuan tourism is closely related to the cultural characteristics and consumption level.

**Table 1.** Quantification results of high-frequency words in Wuyuan travel notes text (partial).

Information Classification		Vocabulary	Word Frequency	Ratio (%)
Tourist Route		Wuyuan	3096	74.51
		Jingdezhen	802	19.30
		Shicheng	198	4.77
Tourist Attractions		Huangling	801	11.15
		Likeng	687	9.56
		Jiangling	643	8.95
		Sixiyancun	620	8.63
		Moon Bay	589	8.20
		Rainbow Bridge	546	7.60
		Wangkou	521	7.25
Amusement Activities	Nature	Rape Flowers	841	49.47
	Humanities	Terraces	321	18.53
Travel Daily	Transportation	High-speed Rail	323	35.89
	Accommodation	Hotels	335	44.02
	Shopping	Tickets	384	37.54
	Others	Photography	193	36.42
Tourist Perception		Huizhou	360	23.41
		Feature	223	14.50
		Most Beautiful China	176	11.44
		Cheap	98	6.37
Travel Season		Autumn	156	78.00

Based on the high-frequency vocabulary data, it can be seen that tourists in Wuyuan take famous scenic spots such as “Huangling”, “Likeng”, and “Jiangling” as the core and landscapes such as “Rape Flower” and “Sun Autumn” as the core. Mainly featured sightseeing and rural experience. Tourist attractions and products such as “Huangling”, “Likeng”, “Jiangling”, “Rape Flower Flowers”, and “Sun Autumn” have become the main factors that attract tourists, and the insufficient tourism reception service system of transportation infrastructure is the bottleneck hindering their development.

## 4.2. Semantics Network Analysis

The social network and semantic analysis functions of ROST CM6 are based on commonality. This paper uses the analysis function to eliminate data irrelevant to the content of the analysis, and retain the top 20 groups of high-frequency co-occurring words, which is shown in **Table 2**. Through co-occurrence analysis, we found that Wuyuan is the base word of this text, among which “Huangling” co-occurs with 6 words, accounting for 30%. It can be seen that “Huangling” occupies a basic position in co-occurrence and a high position in the hearts of tourists from Wuyuan. The co-occurrence of “Rape Flower” and “Likeng”, “Huangling”, “Jiangling” and other scenic spots shows that tourists have a strong perception of this element of Rape Flower flower, indicating that the element of tourism consumption of “Rape Flower flower” is the development of various scenic spots in Wuyuan. An important element of tourism.

**Table 2.** The top 20 groups of high-frequency co-occurring words.

Sequence	Words	Words	Co-occurrence Frequency
1	Wuyuan	Huangling	244
2	Wuyuan	Rape Flower	208
3	Wuyuan	Village	186
4	Wuyuan	Building	166
5	Wuyuan	Hui Style	164
6	Huangling	Rape Flower	161
7	Huangling	Village	156
8	Jiangling	Wuyuan	156
9	Wuyuan	Tickets	148
10	Rape Flower	Attractions	147
11	Building	Hui Style	143
12	Scenic Spot	Village	143
13	Rape Flower	Scenic Spot	143
14	Wuyuan	Shaiqiu	142
15	Wuyuan	Likeng	140
16	Shaiqiu	Huangling	140
17	Huangling	Building	138
18	Huangling	Hui Style	138
19	Attractions	Village	138
20	Wuyuan	Shicheng	136

### 4.3. Sentiment Analysis

In travel notes and online reviews, tourists express their emotional attitudes towards the destination through some descriptive words and sentences about the destination. By using the “sentiment analysis” function in ROST CM6, this paper sorts out and analyzes Wuyuan travel notes, and obtains the following statistical results of emotional distribution, and then obtains the sentimental attitude of tourists towards Wuyuan. **Table 3** shows that tourists have the highest proportion of positive emotions, reaching 70.95%; negative emotions accounted for 16.92%; neutral emotions accounted for 12.13%. In the distribution of positive emotion intensity, the highest proportion of general positive emotion is 33.02%, followed by moderately positive and highly positive; in the distribution of negative emotion intensity, the highest proportion of general negative emotion is 12.78%, followed by moderately negative and highly negative. From the two aspects of the light distribution of emotions, it can be seen that tourists’ perceived satisfaction with Wuyuan is mainly based on positive emotions, with a very small proportion of highly negative emotions, but it is also necessary to take seriously and deal with the factors that cause tourists’ negative emotions.

### 5. Conclusion

On the whole, tourists are mainly affirming the travel experience of Wuyuan, cognition of positive emotions is higher than negative emotions, and the image communication of tourist destinations has achieved better results, which echoes the conclusions of the previous high-frequency word analysis. However, the negative evaluation of tourists in the travel notes also provides a certain reference for the development of rural tourism in Wuyuan. Combining the content of the corresponding tourist evaluation texts, it can be concluded that the negative emotions are mainly due to imperfect infrastructure such as transportation, such as “there are fewer cars to Shitan, and the hotel is in the county, it is inconvenient to travel.” You must drive by yourself cause it’s tiring to transfer the buses” and so on.

**Table 3.** Sentiment analysis statistics results.

Emotion Type	Ratio (%)	Intensity	Ratio (%)
Positive emotion	70.95	Normal	33.02
		Moderate	19.44
		High	18.49
Neutral emotion	12.13		
Negative emotion	16.92	Normal	12.78
		Moderate	4.12
		High	0.02

## 6. Recommendations

Through network text analysis, we found that the core circle of tourists' perception is centered on "Wuyuan" and is closely connected with tourism elements such as "Huangling" and "Rape Flower". This shows that rape blossoms and Huangling scenic spots are the main tourism resources that attract tourists to Wuyuan. They must be regarded as the top priority of tourism development and management, and reasonable planning should be made with steady progress to optimize core tourism resources. The Wuyuan County Tourism Bureau needs to take the lead, increase investment and construction, improve the reception capacity of Huangling Scenic Area, and comprehensively upgrade the canola flower tourism activity project around the two major scenic spots of Huangling and Jiangwan, and dig into the areas related to rape flowers. Farming experience activities increase the experience of tourists, solve the problem of excessively scattered scenic spots, create a well-known flower viewing area, and establish a rape flower brand image.

The study also shows that "autumn" is the peak season for tourists to travel to Wuyuan, and seasonal activities such as "autumn drying" are favored by tourists. Tourist operators in Wuyuan should pay close attention to seasons and climate changes, develop suitable tourism experience projects, and conduct differentiated marketing by seasons throughout the year to improve the freshness and attractiveness of Wuyuan rural tourism and achieve sustainable profitability. At the same time, as a popular tourist perception factor, "Hui Style" has a unique tourist attraction. Therefore, it is necessary to strengthen the construction of the cultural service system and activate Huizhou cultural tourism resources.

Wuyuan County is located at the junction of the three provinces. There are mostly mountain roads and traffic jams. Most of the tourist attractions are remote and far apart, which greatly restricts the travel of tourists and reduces the experience of play. In the above network text analysis, high-frequency words such as "self-driving" and "chartered car" reflect the travel restrictions of rural tourism in Wuyuan from the side. The government should increase investment in transportation facilities, implement rural road construction, set up special tourist lines, turn rural roads into tourist-type roads where tourists can travel smoothly, and improve infrastructure to enhance tourism reception capabilities.

This research points out the bottlenecks and problems in the development of rural tourism in Wuyuan, targeted advice for Wuyuan's enhancement of tourism attractiveness, and reference for other rural tourism destinations to improve image shaping and marketing, which responds to the country's rural revitalization strategy that is of great significance for rural tourism to drive region economic development.

## 7. Limitations and Future Research Directions

This research also has certain limitations, which can be improved in the future research process. First of all, in terms of data collection, this study only selects

more representative travel websites as the network data source, which limits the comprehensiveness of the sample to a certain extent. In future research, it is necessary to further expand the channels and methods of sample acquisition. At the same time, it can expand from text data to the analysis of website pictures, audio, video and other multimedia content, making the research more in-depth. Secondly, in terms of online text publishers, the authors of online travel notes are relatively single in terms of age. Most of them are young tourists and do not involve different age groups of tourists. Follow-up research can be added to explore the travel preferences of tourists of multiple age groups. Finally, in terms of data mining, follow-up research can use deep learning algorithms and various mining software to further improve accuracy.

### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

### References

- Gao, J., & Wu, B. H. (2017). Revitalizing Traditional Villages through Rural Tourism: A Case Study of Yuanjia Village, Shaanxi Province, China. *Tourism Management*, *63*, 223-233. <https://doi.org/10.1016/j.tourman.2017.04.003>
- Guo, Q. Q., Du, J. C., & Li, Y. W. (2020). Research on the Influencing Factors of Red Tourism Tourist Satisfaction Based on Push-Pull Theory—Taking Yunnan Province as an Example. *Value Engineering*, *39*, 275-279.
- Jia, Y. Y., Kang, Y. W., & Zhou, H. W. (2016). Research on the Tourism Preference of Traditional Villages in Central China—Based on the Survey Data of Changsha and Nanchang. *Journal of Henan University of Technology (Social Science Edition)*, *17*, 319-326.
- Li, X., & Li, Q. (2006). *Basics of Tourism Psychology* (pp. 93-99). Beijing: Tsinghua University Press.
- Wang, Z. Q., & Xue, J. H. (2019). Interpretation of Tourist Destination Image Symbols Based on Network Text Analysis—Taking Xiamen Gulangyu as an Example. *Journal of Neijiang Teachers College*, *34*, 81-87.
- Wen, W. (2006). Sustainable Development of Rural Tourism and Urbanization. *Journal of Hubei University of Economics (Humanities and Social Sciences Edition)*, *6*, 46-47.
- Wu, Q. G. (2006). *Tourism Consumer Behavior* (pp. 86-92). Beijing: Tourism Education Press.
- Zhu, Y. Q., Zhang H., & Yang R. (2016). Research on Rural Tourism Preference: Taking Maweyi, Xingping City, Shaanxi Province as an Example. *Henan Science*, *34*, 1380-1385.