tourism and hospitality

MDPI

Article

Word-of-Mouth Evaluation of Ancient Towns in Southern China Using Web Comments

Yihan Zhang *, Weizhuo Guo, Yanling Sheng and Shanshan Li

School of Geography and Environmental Economics, Guangdong University of Finance and Economics, Guangzhou 510320, China; guoweizhuo8@gmail.com (W.G.); shengyl@gdufe.edu.cn (Y.S.); lissnju@gmail.com (S.L.)

* Correspondence: zyh4184@gdufe.edu.cn

Abstract: With the rapid development of digital networks and communication technologies, traditional word-of-mouth (WOM) has transformed into electronic word-of-mouth (eWOM), which plays a pivotal role in improving the management and service quality of ancient town tourism. This study uses Python web scraping techniques to gather eWOM data from the top ten ancient towns in southern China. Using IPA analysis, the analytic hierarchy process (AHP), Term Frequency–Inverse Document Frequency (TF-IDF), and cluster analysis, we developed a comprehensive eWOM evaluation framework. This framework was employed to perform word frequency analysis, sentiment analysis, topic modeling, and rating analysis, providing deeper insights into tourists' perceptions. The results reveal several key findings: (1) Transportation infrastructure varies significantly across the towns. Heshun and Huangyao suffer from poor accessibility, while the remaining towns benefit from the developed transportation network of the Yangtze River Delta. (2) The volume of eWOM is strongly influenced by seasonal patterns and was notably impacted by the COVID-19 pandemic. (3) The majority of tourists express positive sentiments toward the ancient towns, with a focus on the available facilities. Their highest levels of satisfaction, however, are associated with the scenic landscapes. (4) A comprehensive eWOM analysis suggests that Wuzhen and Xidi–Hongcun are the most popular tourist destinations, while Zhujiajiao, Huangyao, Zhouzhuang, and Nanxun exhibit lower levels of both attention and visitor satisfaction.

Keywords: electronic word-of-mouth (eWOM); evaluation analysis; web crawler; ancient towns

check for updates

Academic Editor: Brian Garrod

Received: 8 December 2024
Revised: 16 January 2025
Accepted: 3 February 2025
Published: 11 February 2025

Citation: Zhang, Y., Guo, W., Sheng, Y., & Li, S. (2025). Word-of-Mouth Evaluation of Ancient Towns in Southern China Using Web Comments. *Tourism and Hospitality*, *6*(1), 25. https://doi.org/10.3390/tourhosp6010025

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the establishment of comprehensive internet infrastructure, the number of internet users and the total amount of time spent online have continued to increase (J. Zhang et al., 2020). This broad user base and extensive reach grant electronic word-of-mouth (eWOM) considerable influence. Unlike traditional word-of-mouth (WOM), many studies suggest that eWOM provides stronger privacy protection, allowing people to share more honest and detailed information about products (Johnson, 2000; Yu et al., 2010). Furthermore, some research points out that eWOM is more convenient and far-reaching in its geographical and social impact, making it a more powerful tool for communication (Matzler et al., 2019; Song et al., 2022). eWOM is closely linked to consumer purchasing decisions. Some scholars argue that negative eWOM exerts a stronger influence on consumers' choices than positive reviews (Khan et al., 2024). In the context of tourism, eWOM

also plays a role in shaping tourists' decisions, influencing their destination selections (Lam & So, 2013; Wu & Yang, 2023).

Early studies on online word-of-mouth (WOM) defined it as information exchange and dissemination via the internet (Dellarocas, 2003). As research progressed, scholars expanded the focus of online WOM studies to encompass areas such as consumer motivation, corporate behavior, and its effects on various applications (Liu et al., 2022; Hameed et al., 2024; Hua et al., 2024). In recent years, there has been a growing interest in integrating contemporary trends into this field. For instance, the rise of AI-driven sentiment analysis has provided new insights into understanding consumer opinions and behaviors through data-driven approaches (Kumar & Kaur, 2019; L. Zhang & Li, 2021). Additionally, the emergence of immersive experiences in tourism has opened up new avenues for exploring how WOM influences travel decisions and perceptions (Y. Chen & Xie, 2019). These advancements not only enrich the existing body of knowledge but also highlight the dynamic nature of online WOM in the modern digital landscape. In the context of tourism, online WOM research has concentrated largely on service evaluations, destination choices, and marketing strategies for destinations (Zheng et al., 2023). Online WOM provides insights into various aspects of tourism destinations, and many researchers consider it an influential force in shaping perceptions (Yun et al., 2008). Some studies further suggest that tourism-related online WOM serves as an expression of destination image (Aggarwal & Gour, 2020). These dimensions (influence and role in image perception) offer a basis for examining the relationship between online WOM and tourism destinations.

From the perspective of digital consumer behavior, tourists now rely on both traditional sources and online reviews/social media for destination selection (Huang & Benyoucef, 2009). This highlights the growing importance of eWOM in tourism decision-making. By analyzing eWOM, this study aims to uncover tourists' perceptions of destinations and provide insights for managers to better understand consumer behavior and needs. In cultural tourism, cultural experiences are a key attraction for tourists, and eWOM plays a vital role in sharing these experiences (Timothy & Boyd, 2011). Analyzing eWOM helps us to understand tourists' evaluations of cultural sites, offering practical suggestions for managers on preservation and promotion. From a sustainable development perspective, tourism must balance economic, social, and environmental interests. eWOM enables managers to grasp tourists' expectations and needs for sustainability (Gursoy et al., 2009). By analyzing eWOM, feedback on sustainable practices can be identified, guiding the industry toward greater sustainability.

Online word-of-mouth serves as a real-time, dynamic information resource that furnishes researchers with abundant data. Web crawler technology plays a pivotal role in capturing this scattered reputational information from across the internet (Ferreira et al., 2018). In tourism research, the application of web crawler technology has begun to show promise. It not only efficiently gathers large volumes of tourist reviews and feedback but also reveals genuine experiences and perceptions through in-depth data analysis (Kumar & Kaur, 2019). Furthermore, it aids in constructing scientific frameworks for metrics. These contributions help researchers to understand tourists' needs and expectations, predict trends in destination reputation development, and provide robust data support for management decisions within the tourism industry (Qu et al., 2024; L. Wang & Kim, 2024).

Research on eWOM dissemination, analysis of tourism reviews, and related technologies has been ongoing for decades. With the aid of web crawler technology, some drawbacks can be overcome. Traditional methods typically rely on questionnaires for research, but such data are often difficult to obtain or time-consuming to collect. Additionally, data gathered through online surveys may be biased, as respondents might not have actually visited the location in question. In contrast, comment data are almost entirely

from tourists who have genuinely visited the destination, making the data more accurate and reliable. In addition, review data is often utilized by tourism scholars for qualitative summarization or evaluation (Litvin et al., 2008). The issue of whether we can perform a deeper level of analysis to facilitate quantitative assessment is a question that requires our attention.

For the managers of tourist attractions, understanding the following issues is very important: whether the advantages of the scenic area are easier to summarize, whether the problems existing in the scenic area are easier to find, whether the relationship between consumer satisfaction and online word-of-mouth can be quantified, and whether the evaluation of the scenic area can be conveniently quantified. To solve these issues, this article uses the ten most famous ancient towns in southern China as examples. It collects related online review data using web crawlers, and applies methods such as word-of-mouth volume analysis, word frequency analysis, volume-to-satisfaction ratio analysis, comment topic analysis, cluster analysis, and AHP to analyze and evaluate them. This paper introduces a research framework based on comment data, which not only allows for qualitative evaluation of scenic spots (sentiment analysis), but also enables quantitative assessment of attractions (AHP). This research combines qualitative and quantitative analysis using text-based data, offering a novel perspective for studies in tourism and hospitality. We also hope that the results of the study can provide a necessary basis for decision-making to improve the supervision level and service quality of management departments.

2. Materials and Methods

2.1. Study Area

In tourism research, ancient town tourism holds a very unique position. Ancient towns not only embody rich historical and cultural values but also serve as important carriers of historical changes and cultural heritage. Against the backdrop of prosperous modern cities, these towns, with their ancient architecture and cultural relics, stand out with distinct characteristics from modern urban landscapes, making them hotspots for tourism development (Macdonald, 2001). This unique difference between ancient and modern attracts a large number of tourists. Due to variations in location, function, and environment, different ancient towns exhibit diverse development situations. Understanding and assessing the situation of ancient towns is crucial for exploring how to develop tourism resources rationally, fostering the growth of the tourism industry and preserving historical landscapes and cultural heritage for sustainable development (N. Wang & Zhang, 2020).

To provide an objective assessment of ancient towns, this study draws upon comments from the Maigoo website (https://www.maigoo.com/goomai/152095.html (accessed on 25 February 2022)), a platform that gauges global web traffic and popularity. Rankings are determined by multiple factors, including visit volume and page views. It also provides personalized rankings based on cultural heritage, architectural styles, and landscape features. Focusing on the ten most frequently mentioned (attractive) ancient towns in southern China from various rankings—Zhouzhuang, Tongli, Wuzhen, Xitang, Luzhi, Huangyao, Xidi and Hongcun, Nanxun, Zhujiajiao, and Heshun—the research examines their characteristics. These ancient towns have a large amount of comment data, which can comprehensively reflect all aspects of the ancient towns. They are geographically clustered in regions like the Yangtze River Delta, known for their distinct Jiangnan water town ambiance. Specifically, Zhouzhuang, Tongli, and Luzhi are in Suzhou, Jiangsu; Wuzhen, Xitang, and Nanxun are in Zhejiang; Zhujiajiao is in Qingpu, Shanghai; while Xidi and Hongcun, Heshun, and Huangyao are located in Anhui, Yunnan, and Guangxi, respectively (Figure 1).

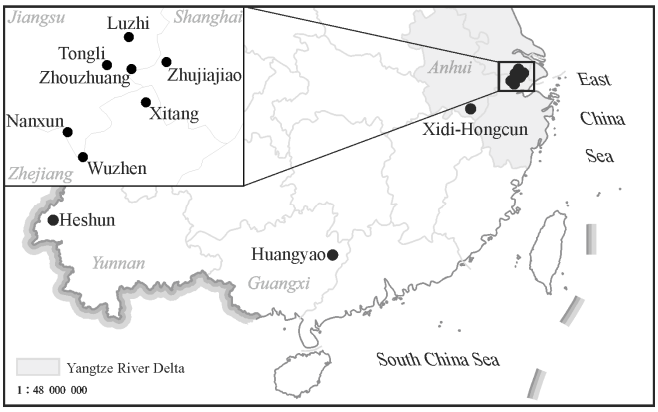


Figure 1. Location of ancient towns.

2.2. Methods

2.2.1. Importance–Performance Analysis (IPA)

IPA, created by Martilla and James (1977), is a tool that has been popular since the late 1970s. It looks at how much consumers value a product before using it and how well it does after they use it. This helps us see what is good and what needs to be better, guiding us on where to put our resources. IPA works well for products, services, and retail areas. For example, Oppermann (1996) applied IPA to analyze the advantages and disadvantages of convention and exhibition destinations. Almanza et al. (1994) used it to evaluate satisfaction with catering services.

In this study, we apply IPA for evaluation of ancient towns, constructing a quadrant chart based on attention and satisfaction levels to illustrate the relationship between expectations and actual perceptions (Figure 2). The vertical axis of the chart represents satisfaction (expectation value), while the horizontal axis represents attention (actual perception value). The quadrants are categorized as follows:

- First quadrant (low/high): Low attention, but relatively high satisfaction, indicating secondary strengths where promotion can be focused;
- Second quadrant (high/high): High attention and high satisfaction, signifying primary strengths to be maintained;
- Third quadrant (low/low): Low attention and low satisfaction, representing secondary weaknesses not prioritized for development;
- Fourth quadrant (high/low): High attention, but low satisfaction, suggesting a need to align promotional effects with actual experiences during resource integration and marketing activities.

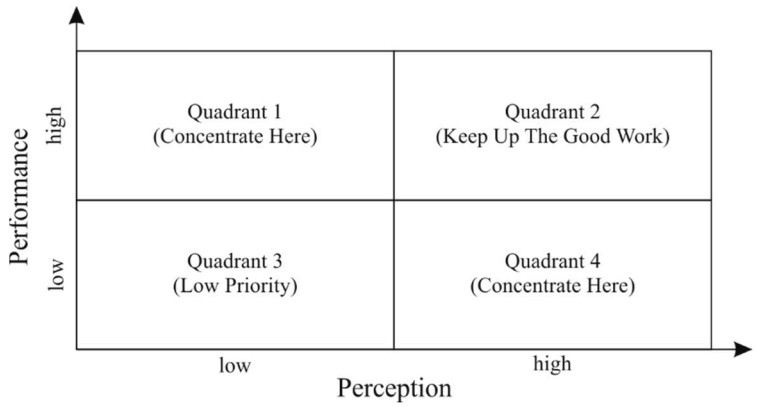


Figure 2. Quadrants in IPA.

Using the satisfaction and attention values for each ancient town, we can present these on an IPA chart (Figure 2) to see what is most important and what needs to be improved. IPA directly focuses on customers' perception and compares customers' evaluation of the importance of service characteristics with the actual performance in a graphical manner, thus being closer to customers' needs and expectations. Presented in a graphical way, it is easy to understand and communicate, and is helpful for communication between internal and external stakeholders of the organization. Compared with SWOT (Strengths, Weaknesses, Opportunities, and Threats), IPA provides more specific improvement directions and priorities by quantifying importance and performance, thereby offering a more actionable approach (Parasuraman et al., 2005).

2.2.2. Analytic Hierarchy Process (AHP)

AHP is a tool that helps make decisions by evaluating and ranking different factors and options. It mainly involves dividing the factors into several layers, such as objective layer, criterion layer, indicator layer, etc., and then comparing each indicator pairwise to obtain its weight (Saaty, 1977). When we compare how important factor i is compared to factor j, we can say they are equally important, a bit more important, somewhat more important, much more important, or extremely important. After comparing every pair of indicators, we create a matrix. This matrix has a special rule: the importance of i to j is the reverse of j to i, so if i to j is a number, ($a_{ij} = 1/a_{ji}$).

Next, we adjust this matrix by averaging the numbers in each column. This gives us an average value for each indicator, which we use to find out how important each one is, or its weight ω. After we know the weights, we should calculate the CR value to make sure our results are consistent and not just random.

$$CR = CI/RI \tag{1}$$

where RI refers to the random consistency index, whose value is related to the order of the matrix and can be obtained by looking up the standard value table of RI (Kumar & Vaidya, 2006). CI is the consistency index, and can be calculated by

$$CI = (\lambda_{max} - n)/(n-1) \tag{2}$$

where λ_{max} is the maximum eigenvalue can be obtained as following equation

$$\lambda_{max} = \sum_{i=1}^{n} \frac{[A\omega]_i}{n * \omega_i} \tag{3}$$

where $A\omega$ represents the product of judgment matrix and weight value, and n represents the order of the matrix.

Generally speaking, if $CR < 0.1$, it indicates that the judgment matrix passes the consistency test; otherwise, it does not have satisfactory consistency.

AHP has been widely used in decision-making analysis. It breaks down complex problems into multiple levels and factors, combines qualitative and quantitative analysis, and provides a simple yet effective method for tackling complex decision-making problems. The hierarchical structure and pairwise comparison matrix of AHP make the decision-making process more intuitive and easier to understand. Compared with the Delphi Method (a structured expert consultation method), AHP offers a more systematic and quantitative approach to decision-making, reducing the bias of subjective judgment (Kumar & Vaidya, 2006).

2.2.3. Term Frequency–Inverse Document Frequency (TF-IDF)

TF-IDF is a widely adopted technique in text analysis that computes the frequency of terms (TF) and the inverse document frequency (IDF) to extract key terms from reviews and assign them feature weights for in-depth frequency analysis (Hajjem & Latiri, 2017; Z. Wang et al., 2021). In this method, *TF* represents the occurrence frequency of term i within a document, which is the term frequency t, measured by counting the number of times term n_i appears in an individual document. *IDF*, on the other hand, is used to evaluate the general importance of term t across the entire document collection, known as the inverse document frequency of t. Here, D denotes the total number of documents, while D_i refers to the total number of documents that contain a specific term. The calculation of IDF helps to identify terms that appear in a lower number of documents, which often possess higher discriminability and significance. In this study, we utilize the popular Scikit-learn toolkit (which can be downloaded easily) within Python to compute the TF-IDF.

To ensure stability, we conduct multiple experiments using different subsets of the data for training and testing. This process helps to confirm that the results are consistent and not dependent on a particular subset (Newman et al., 2009). We compare the keywords annotated manually with those generated by the algorithm. This step ensures that the algorithm's output aligns with human-recognized importance and relevance. After the initial analysis, the results are manually inspected by experts in the field. This expert review helps to validate the accuracy and relevance of the TF-IDF keywords identified.

2.2.4. Word-of-Mouth Volume and Satisfaction

Word-of-mouth volume (online attention) can be used to show how much users are paying attention to a destination. The more relevant comments about a research destination, the higher its online attention (Morgan & Pritchard, 2020). To calculate the online attention for a specific ancient town, divide the word-of-mouth volume for that town by the total word-of-mouth volume for all ancient towns. The closer the resulting value is to 1, the higher the attention. x_i represents the number of valid comments for the i-th town.

$$A_i = \frac{x_i}{\sum x_i} \times 100\% \tag{4}$$

Word-of-mouth satisfaction is an approach which can help to analyze user reviews by calculating how many are positive, neutral, or negative. The more positive comments there are relative to the total number of comments, the higher the satisfaction, indicating that users are generally more satisfied with the ancient town (Pang & Lee, 2008). It can be calculated as follows:

$$S_i = \frac{P_i}{P_i + N_i + C_i} \times 100\% \tag{5}$$

where P_i represents the number of positive comments for the i-th town, N_i represents the number of negative comments for the i-th town, and C_i represents the number of neutral comments for the i-th town.

2.2.5. Latent Dirichlet Allocation Model (LDA)

The LDA model, created by Blei and inspired by the Dirichlet process, is a key tool in analyzing text (Blei et al., 2003). It is great for finding out what topics are in a set of documents. It also can help to obtain probability for each topic in each document, which helps us understand the main ideas in the text. We can use this to group documents with similar topics or to classify documents based on their topics (Rubin et al., 2011).

The process of the LDA topic generation model is as follows (Blei, 2012): (1) Select a Document: Choose a document based on prior probability, which guides the selection of the document to analyze. (2) Generate Topic Distribution: Sample from the Dirichlet

distribution to create the topic distribution for the document. The Dirichlet distribution provides a theoretical foundation for generating a plausible topic distribution. (3) Assign Topics to Words: Draw a topic for each word from the multinomial distribution over topics. This step establishes a connection between words and topics, laying the groundwork for the subsequent generation of words. (4) Generate Word Distribution: Sample from the Dirichlet distribution to create the word distribution for each topic. This refines the structure of words associated with the topic. (5) Sample Final Words: Select the final words from the multinomial distribution of words. This series of steps systematically constructs the generative mechanism of the entire LDA topic model.

2.2.6. K-Means Clustering

K-Means clustering is a simple and popular method for grouping similar data points together. The goal is to create clusters where the data points within each cluster are close to each other, while the clusters themselves are far apart (Hartigan & Wong, 1979).

In the algorithm, the sample dataset can be defined as $S = \{S_1, S_2, \cdots, S_n\}$. The role of the K-Means clustering algorithm is to partition a sample S containing n data points into k mutually exclusive classes, where each class represents a cluster expressed as $P = \{P_1, P_2, \cdots, P_k\}, 1 < k \leq n$.

Step 1: Since we do not know the categories of the sample data beforehand, we first specify the size of k based on experience and randomly initialize k cluster centers.

Step 2: Then, we assign each data point in the sample dataset to the class cluster closest to its corresponding cluster center.

Step 3: After assigning all data, we recalculate the cluster center of each class cluster based on all the data in that cluster, obtaining their average values.

Step 4: Finally, we iterate again, reassigning the data in the sample dataset and continuously updating the cluster centers until the cluster centers no longer change.

The K-Means clustering algorithm partitions the data in the sample dataset to the cluster center closest to it based on Euclidean distance. A smaller numerical value of Euclidean distance indicates a higher similarity between the data and the cluster center, and the higher the similarity, the better the clustering effect (Jain, 2010).

2.3. Data Preprocessing

Step 1: Acquisition of online reputation review data for the top ten ancient towns. Using Python tools, all comment users, comment times, and comment contents, including folded comments, were crawled from the Meituan webpage. Approximately 40,000 pieces of data were extracted relating to the period between 1 January 2018 and 31 December 2021.

Step 2: Data screening. Some meaningless data, such as empty comments, emoticons, and disruptive comments, existed in the data captured through Python. These meaningless data were filtered out, and then the filtered data were saved in UTF-8 encoding.

Step 3: Chinese word segmentation. Since the crawled comment data consisted of long sentences that could not be recognized by computers, the text data were segmented using a Chinese word segmentation tool. Jieba Chinese word segmentation module in Python was used to process the crawled data. Jieba typically employs a combination of existing dictionary matching and a Hidden Markov Model (HMM) approach for word segmentation. Firstly, it performs preliminary segmentation using dictionaries (default dictionary and custom dictionary, especially custom dictionary, can significantly improve accuracy of word segmentation). Then, it applies the HMM to address words not found in the dictionary. Finally, it integrates the results from both methods to achieve the final word segmentation, which usually provides the most accurate outcome (H. Wang et al., 2016).

Step 4: Removal of stop words. After filtering and word segmentation of the text data, a large number of punctuation marks, special symbols, personal pronouns, and even some modal particles appeared. These words are useless for data analysis and may affect the analysis results. Therefore, a set of stop words dictionary was established to remove these meaningless words.

2.4. Research Questions

This research aims to explore how to utilize the currently popular electronic word-of-mouth (eWOM) data for qualitative and quantitative tourist evaluation research. Taking the top ten ancient towns in southern China as example, a research framework for ancient-town word-of-mouth evaluation based on text analysis was constructed. The article collected eWOM data in the research area through Python web crawler technology, used word frequency analysis to obtain the content of tourists' attention, applied IPA to analyze the attention to and satisfaction with ancient towns, introduced TF-IDF for theme analysis, used cluster analysis to classify ancient-town categories, and finally applied AHP to conduct quantitative evaluation on ancient towns. This framework combines with word frequency analysis, sentiment analysis, topic modeling, and score analysis to objectively reflect the actual situation in the research area and deeply understand tourists' cognition and feelings.

3. Results

3.1. Location Analysis of Ancient Towns

Seven ancient towns in southern China are relatively concentrated at the junction of Shanghai, Jiangsu, and Zhejiang provinces (Figure 1). The remaining three towns are located in Yunnan, Guangxi, and Anhui. Heshun Ancient Town in Tengchong, Yunnan, is characterized by its unique caravan culture. Huangyao Ancient Town in Hezhou, Guangxi, preserves architecture from the Ming and Qing dynasties, including temples, monasteries, and pavilions. Xidi and Hongcun in Huangshan City are renowned for their pastoral landscapes, village layouts, and distinctive Hui-style architecture. Research indicates that ancient towns likely served as pivotal nodes in ancient transportation networks, facilitating commodity exchange and cultural integration. They are supported by a substantial population and economic base. The affluent Suzhou–Hangzhou area encompasses seven of these ancient towns, including Luzhi, Tongli, and Zhouzhuang in Jiangsu; Nanxun, Wuzhen, and Xitang in Zhejiang; and Zhujiajiao in Shanghai. These towns are located in the fertile plains of the lower Yangtze River, boasting advantageous water transport during ancient times, connected to the Grand Canal from Beijing to Hangzhou. These towns are not only economic and cultural hubs of ancient times, but also crucial windows for understanding the development of China today.

As the national economy has soared, the primary mode of transportation, once dominated by waterways, has evolved into a multifaceted system encompassing railways, highways, and air travel. To understand the implications of this shift in accessibility to ancient towns, this research compiled data on highway density and proximity to railway stations and airports for a selection of ancient towns (Table 1). The data reveal that Heshun in Yunnan and Huangyao in Guangxi are situated at distances of 152 km and 63 km from the nearest train stations, respectively, indicating a relatively greater distance compared to other towns, which are all within 30 km. In particular, Nanxun Ancient Town and Xidi–Hongcun enjoy exceptionally convenient access, being only 8–9 km from their respective stations. When considering distances to airports, Huangyao and Nanxun are the furthest, while Tongli, Wuzhen, and Zhouzhuang are within 60–80 km. Heshun and Zhujiajiao, however, are much closer to Tengchong Tuofeng Airport (18 km) and Shanghai Hongqiao Airport (38 km), respectively. For other towns, reaching an airport generally requires

about an hour's drive. Given that both rail and air travel necessitate road connections, this study assessed highway density within a 50 km radius of each ancient town. The results highlight pronounced disparities in highway density, with towns in the Yangtze River Delta region exhibiting significantly higher densities than those elsewhere. Among them, Heshun and Huangyao have the lowest densities at 88.4 m/km^2 and 85.2 m/km^2, whereas Luzhi, Tongli, Zhouzhuang, and Zhujiajiao boast densities exceeding 1 km/km^2, at 1218.1 m/km^2, 1033.0 m/km^2, 1137.8 m/km^2, and 1217.4 m/km^2, respectively. This suggests that tourists can easily reach these ancient towns via a well-developed transportation network.

Table 1. Transportation Location Information of Ancient Towns.

Ancient Town	Distance to the Nearest Railway Station (km)	Distance to the Nearest Airport (km)	Highway Density Within 50 km (m/km^2)
Heshun	152	18	88.4
Huangyao	63	125	85.2
Luzhi	17	54.7	1218.1
Nanxun	9	108	640.9
Tongli	22	76	1033.0
Wuzhen	17	74	642.9
Xidi–Hongcun	8	51	189.1
Xitang	13	47	947.5
Zhouzhuang	13	63.8	1137.8
Zhujiajiao	28	38	1217.4

3.2. Data Preprocessing of WOM

3.2.1. Sentiment Classification of WOMs

To begin with, each comment undergoes sentiment analysis to classify it as either positive, negative, or neutral. A method based on an emotional dictionary is used to rate each word individually and an overall sentiment score for the comment can be obtained. The scoring process involves assigning a sentiment value to each word, which is then aggregated to determine the sentiment of the entire text. This method helps us to capture the difference of sentiment, enhancing the accuracy of our sentiment analysis. A score greater than 0 indicates a positive sentiment, a score of 0 is neutral, and a score less than 0 signifies a negative sentiment (Cambria et al., 2013). The process involves several steps. Firstly, comments are tokenized. For example, the comment "The ancient town scenery is nice and worth visiting" is broken down into "ancient town", "scenery", "nice", "worth", and "visiting". Secondly, the sentiment of each word within the comment is determined, categorizing it as positive, negative, or neutral. Thirdly, the impact of negation words on emotional words is assessed. If a word like "not", "no", or "don't" precede an emotional word, its sentiment is reversed. For example, "funny" is positive, but "not funny" becomes negative due to the preceding negation word "not". Lastly, a final determination is made. If a comment contains only positive emotional words, it is classified as positive. If it contains any negative emotional words, it is classified as negative. And if it contains neither, it is classified as neutral.

3.2.2. Theme Extraction

Applying the LDA model to the comments, we determined the optimal number of topics by calculating the complexity of the text. Upon analysis, seven effective topics emerged. Theme 1, which includes words such as "facilities", "renovation", and "old", reflects tourists' perceptions of the scenic area's amenities. Theme 2, characterized by terms like "not worth it", "relaxation", and "mood" primarily captures tourists' feelings

during their visit. Themes 3, 4, and 5 suggest that the ancient town influences tourists' perceptions across various dimensions such as scenery, environment, and historical culture, with words like "antique charm", "beautiful", "ink wash painting", and "humanities" indicating a generally positive sentiment. Theme 6, marked by the presence of words such as "restaurants", "guesthouses", and "catering" highlights the importance of dining and accommodation in shaping tourists' perceptions. Lastly, Theme 7, featuring terms like "enthusiastic", "good attitude", and "thoughtful," underscores how the ancient town's commendable service attitude subtly enhances the number of positive reviews from tourists.

3.2.3. Establishment of Word-of-Mouth Evaluation System

After incorporating the effective themes into the indicator layer and performing a comprehensive analysis, we identified four criterion layers. Our evaluation system for the online reputation of ancient towns is structured with ten primary indicators, reflecting the online reputation of each of the top ten ancient towns. These are supported by four secondary indicators, which include tourism perception, tourism consumption, tourism resources, and tourism services. Furthermore, we have seven tertiary indicators: travel mood, dining and accommodation, ancient town scenery, scenic area facilities, scenic area environment, ancient town historical culture, and service attitude. Tourism perception is assessed through travel mood, while tourism consumption is evaluated based on dining and accommodation. Tourism resources include scenery, facilities, environment, and historical culture. Lastly, service attitude is a tertiary indicator for tourism services. Given that tourists prioritize these four aspects—tourism perception, tourism consumption, tourism resources, and tourism services—when selecting a travel destination, our evaluation system provides a comprehensive framework for assessing these ten ancient towns.

3.3. Analysis of WOM

3.3.1. Trend of WOM Volume

This study collected all comments from 2018 to 2021 and sorted them by year and month to analyze the trend in ancient town reputation volume (Figure 3). The total number of valid comments for each year was 12,302 in 2018, 9477 in 2019, 9651 in 2020, and 7176 in 2021, revealing a decreasing trend likely attributable to the pandemic's impact on ancient town tourism. Comparing monthly reputation volumes, the pivotal moments for 2018, 2019, and the latter half of 2020 were similar: small spikes occurred in February and April–May of both 2018 and 2019, corresponding to the Spring Festival, winter vacation, and Labor Day holidays. Higher peaks were observed in August and October of 2018, 2019, and 2020, aligning with the summer vacation and National Day holidays, indicating that longer holidays generated more reputation volume for the ancient towns. However, the turning points for the first half of 2020 and 2021 differed, possibly due to the pandemic. For example, comment numbers declined sharply in the first half of 2020 during the early months of the pandemic outbreak, resulting in only 51 valid comments in February. The variance in monthly reputation volume for 2021 was 64,896, lower than the variances for 2018 (166,941), 2019 (110,801), and 2020 (240,054), respectively. The absence of pronounced peaks during the summer and winter holiday periods suggests that reputation volume fluctuations were smaller in 2021, likely because people were more cautious about traveling to ancient towns amidst the pandemic, opting for travel during less tense periods; this is exemplified by the highest peak occurring in May when the situation was somewhat less serious.

After using correlation analysis on the annual reputation volume data, we found significant correlations between 2018 and both 2019 and 2020, with coefficients of 0.83 and 0.87, respectively. Additionally, correlation between 2019 and 2020 is significant, with a

coefficient of 0.60. In contrast, correlations between 2021 and the previous years (2018, 2019, and 2020) are not as strong, with coefficients of 0.36, 0.25, and 0.40, respectively—indicating a lack of significant correlation. These findings suggest that the reputation volume correlations between the first half of 2020 and the second half of 2021 were influenced by the pandemic. With restrictions on gatherings due to the pandemic, the number of visitors to ancient towns decreased, leading to a corresponding impact on the volume of reputation.

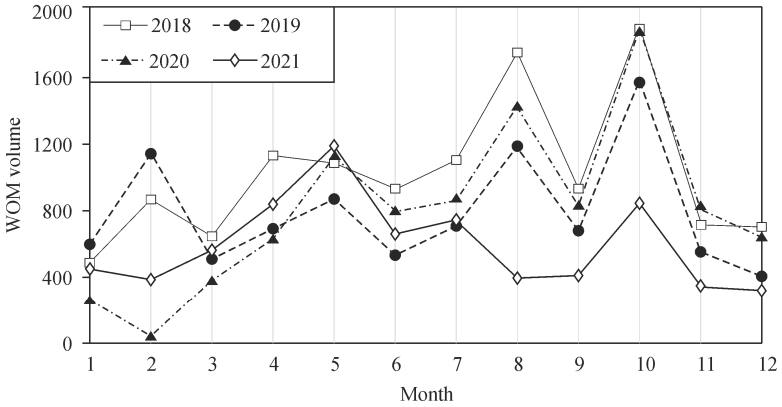


Figure 3. Trends in online reputation of the top ten ancient towns.

3.3.2. Word Frequency Analysis

By applying a normalization formula to calculate weights and employing the TF-IDF method, we derived the weight and word frequency statistics for ancient town comments (Figure 4). Notably, the term "not bad" achieved the highest frequency at 7950, suggesting that tourists generally hold a favorable view of the top ten ancient towns. As the study focused on top ancient towns, which are tourist attractions, "ancient town" and "scenic area" are prominently featured in the weight table. Positive descriptors such as "worthy", "special", "recommend", and "like" outrank negative terms like "nothing much" and "not worth it", indicating tourists' overall positive attitude towards the ancient towns. It should be noticed that the frequent appearance of "commercialization" and "commercial" suggests a trend towards commercialization among most ancient towns, which may detract from their traditional charm and impact tourist satisfaction. The inclusion of words like "ticket", "price", and "service" highlights tourists' concerns regarding consumption and service quality in ancient towns.



Figure 4. Word cloud of ancient towns.

To delve deeper into reputation satisfaction analysis, this study applied TF-IDF word frequency analysis separately to positive and negative comment terms. Our findings reveal that tourists' positive feedback on ancient towns primarily consists of admiration for the scenery, using terms like "not bad", "beautiful", "attractive", and "unique" along with

overall travel assessments such as "worthwhile", "affordable", "fun", and "satisfied". Conversely, negative feedback tends to focus on high expenses, commercialization, and fewer entertainment options. This suggests that high fees and commercialization trends are negatively impacting the tourism development of ancient towns. Consequently, it is imperative for these towns to strike a balance between economic benefits and the tourist experience.

3.3.3. Volume–Satisfaction Analysis

We conducted sentiment analysis on each comment and compiled statistics on reputation volume and satisfaction for each ancient town. After applying the IPA method, we created a scatter plot with reputation volume as the horizontal axis and satisfaction as the vertical axis. A point defined by the mean values of reputation volume (1) and satisfaction (0.64) was used to divide the plot into four quadrants (Figure 5). The average value was chosen as the cut-off point offered several advantages (Y.-C. Chen & Tsai, 2019). Firstly, it is statistically simple and easy to calculate, eliminating the need for complex mathematical models or advanced statistical techniques. Secondly, it ensures a balanced distribution of data points across quadrants, preventing overcrowding or sparsity. Lastly, it facilitates intuitive and comprehensible analysis results, enabling quick identification of significant variations in importance and satisfaction. This makes the analysis results more intuitive and easier to understand, allowing decision-makers to quickly identify service attributes that require attention.

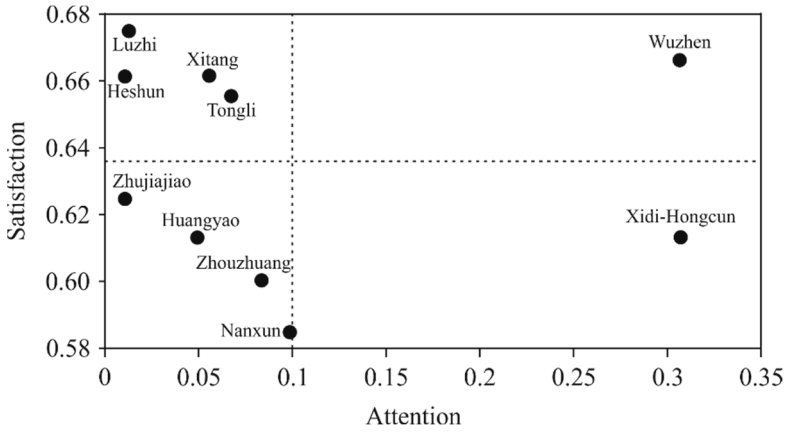


Figure 5. Attention and satisfaction of ancient towns.

Wuzhen and Xidi–Hongcun exhibit prominent reputation volumes, indicating their significant attention. Conversely, Zhujiajiao and Heshun have relatively lower values. Satisfaction levels across all ancient towns remain consistent between 0.5 and 0.7, showing minimal variation. Wuzhen, located in the first quadrant (advantage zone), enjoys both high attention and satisfaction. This suggests that the attraction meets high expectations and also delivers satisfying experiences for tourists. The second quadrant (opportunity zone) includes Luzhi, Xitang, Tongli, and Heshun, where satisfaction exceeds attention. Although these towns are not widely recognized, visitors express higher satisfaction. To foster future growth, increasing their visibility is crucial. In the third quadrant (improvement zone), we find Zhujiajiao, Huangyao, Zhouzhuang, and Nanxun, characterized by low attention and satisfaction. These towns require not only greater recognition, but also a strategic realignment of their development approaches. Finally, Xidi–Hongcun, situated in the fourth quadrant (repair zone), experiences high attention alongside low satisfaction. While it captures public interest, it fails to meet tourist expectations. Addressing existing issues,

enhancing tourism quality, and improving visitor experiences are pivotal for the sustainable development of ancient towns.

3.3.4. Comment Topic Analysis

By classifying the themes of each comment, we identified the following counts for the four secondary indicators: 5030 for tourism perception, 5041 for tourism consumption, 24,055 for tourism resources, and 4556 for tourism services. Notably, comments related to tourism resources were the most numerous, underscoring their importance to tourists. Among 24,055 comments in tourism resources, we observed the tertiary indicators of ancient town history and culture contained 3833 comments, ancient town scenery contained 4829 comments, scenic area environment contained 4483 comments, and scenic area facilities contained 10,910 comments. This indicates that scenic area facilities are of primary concern to tourists.

After analyzing the positive and negative comment ratios for each tertiary indicator, we found that ancient town scenery had the highest positive comment ratio and the lowest negative comment ratio, suggesting that tourists are pleased with the scenery. The ancient town itself serves as a highly attractive tourism resource. Therefore, prioritizing the creation of high-quality ancient town landscapes can be a strategic advantage and focal point for development, leading to increased satisfaction. Conversely, the positive comment ratio for mood during play was the lowest, while the negative comment ratio was the highest, indicating that tourists' emotions remain relatively unchanged after visiting the ancient town. To address this, ancient towns could focus on creating immersive experiences to enhance tourists' engagement. Additionally, operational considerations should prioritize tourists' feelings to improve their overall experience. Furthermore, the negative comment ratio for service attitude was relatively high, indicating dissatisfaction with the service attitude in ancient towns. Tourism management agencies should consider increasing employee training, establishing clear regulations, and improving employees' service awareness and attitudes to convey a welcoming atmosphere. Lastly, the positive comment ratios for ancient town history and culture, catering and accommodation, scenic area facilities, and scenic area environment were relatively balanced, indicating overall satisfaction among tourists. The negative comment ratio for ancient town history and culture was relatively low, suggesting maturity in cultural construction among ancient towns. In contrast, catering and accommodation, as well as scenic area environment, showed higher negative comment ratios, indicating a need for further quality improvements in these areas.

In order to compare comments across various themes, we utilized TF-IDF results to identify typical short sentences reflecting sentiments. In tourism perception, positive attitudes included appreciation for the pleasant scenery and suitability for leisure vacations, offering a retreat from urban hustle. Negative attitudes, however, centered on the absence of an authentic ancient town atmosphere, perceiving the town as overly modernized, thus hindering the experience of its historical ambiance. Additionally, many people expect more local performances and expressed dissatisfaction with overcrowding. As for tourism consumption, many tourists criticized high ticket prices and low cost-effectiveness, feeling underserved. Those who visited during holidays particularly complained about exorbitant accommodation costs, feeling exploited. On the positive side, the proximity of attractions and consumer venues, along with satisfactory and distinctive dining options, garnered praise in some ancient towns. Regarding tourism resources, positive comments lauded the preservation of ancient architectural styles and corresponding historical culture. Negative feedback, however, highlighted inadequate facilities, lack of signage, poor water quality, excessive litter, and transportation issues such as congestion and unreasonable diversions. In

terms of tourism services, thoroughness and ticket checks were noted positively. However, poor service attitudes among staff and instances of forced sales among some merchants were reported. Clearly, ancient towns face numerous pressing issues that require immediate attention.

3.4. Classification and Evaluation of Antient Towns

3.4.1. Clustering Analysis

To explore the classification of ancient towns by comment volume and satisfaction, we normalized both comment count and the proportion of positive comments for the four secondary indicators. Subsequently, we conducted a k-means clustering analysis with these normalized values. The elbow method was used to determine the optimal number of clusters, which involves analyzing the within-cluster sum of squares (WCSS) for different numbers of clusters to identify the best number (Sugar & James, 2003). As the number of clusters increases to a certain critical point (the optimal number of clusters), the rate of decrease in WCSS slows down significantly, and this critical point is referred to as the "elbow". The "elbow" signifies the point at which the rate of decrease in WCSS plateaus, indicating an optimal number of clusters for the dataset. The towns are classified into four distinct groups (K = 4), which are summarized in Table 2.

Table 2. Grouping by number of comments and proportion of positive comments.

Criterion	Group	Perception	Consumption	Resource	Service	Towns
number of comments	1	1.000	0.585	1.000	0.496	Xidi–Hongcun
	2	0.790	1.000	0.838	1.000	Wuzhen
	3	0.081	0.659	0.481	0.503	Xitang, Luzhi, Huangyao, Zhujiajiao, Heshun
	4	0.293	0.188	0.210	0.219	Zhouzhuang, Tongli, Nanxun
proportion of positive comments	1	0.916	0.187	0.888	0.000	Zhujiajiao
	2	0.496	0.000	0.888	0.949	Heshun
	3	0.202	0.376	0.182	0.377	Zhouzhuang, Xidi–Hongcun, Nanxun
	4	0.518	0.883	0.627	0.734	Tongli, Wuzhen, Xitang, Luzhi, Huangyao

The clustering reveals distinct patterns. Groups 1 and 2, each comprising a single ancient town, exhibit high mean comment counts, signifying their prominence. Xidi–Hongcun in Group 1 receives more comments on tourism perception and resources, whereas Wuzhen in Group 2 garners higher attention on tourism consumption and services, illustrating diverse areas of interest. Group 3, with five towns, shows lower emphasis on tourism perception but higher on the remaining aspects, implying a need to enhance this aspect. Group 4, containing three towns, displays low mean values across all indicators, necessitating comprehensive improvements to boost their reputation.

When considering the proportion of positive comments, Groups 1 and 2 again feature a single town each, with notable fluctuations in mean values. Zhujiajiao scores lower in tourism services but higher in perception and resources, while Heshun scores lower in tourism consumption but higher in resources and services. These towns should tailor their strategies to tackle these specific challenges. Groups 3 and 4 exhibit more consistent mean values across indicators, indicating comparable satisfaction levels among tourists. However, the scores of the towns in Group 3 are lower than those of Group 4, suggesting that the towns in Group 3 require comprehensive enhancements in scenic area service quality.

To assess the comparative advantages of various ancient towns across different dimensions, we analyzed positive comments. K-means clustering was applied to categorize

comments into three distinct groups (high, medium, and low) based on four key indicators: tourism perception, consumption, resources, and services. The clustering outcomes revealed that Wuzhen and Xidi–Hongcun had the highest counts of positive comments. Wuzhen excelled in all four categories, whereas Xidi–Hongcun performed well in tourism perception and resources but only moderately in consumption and services, indicating areas where improvements could be made. Luzhi, Zhujiajiao, and Heshun scored low across all indicators, highlighting the need for strategic development of unique strengths and selling propositions. For Zhouzhuang, Tongli, and Nanxun, scores were moderate in tourism perception and resources but low in the remaining categories. These towns' charming styles delight tourists, but their tourism development faces challenges. It is suggested that management agencies should focus on enhancing operational oversight to improve the overall tourist experience. Lastly, Xitang and Huangyao showed moderate performance in one indicator and low in the others, suggesting that while they have some established strengths, comprehensive improvements are still necessary.

3.4.2. Analysis-Based AHP

Based on the number of reviews for each ancient town and its indicator, pairwise comparisons were made to determine the importance of indicators. Weights of indicators at the criterion layer in tourism resources were then calculated using the arithmetic mean method. Consistency tests were conducted on the calculated weights of the criteria layer and indicator layer, and both judgment matrices passed the consistency test.

The weights of the criterion layer and indicator layer are integrated. If there is only one indicator layer in the criterion layer, its weight of indicator layer is assigned a value of 1. The total weight is obtained by multiplying the weights of the criterion layer and indicator layer (Table 3). In the criterion layer, the weights of tourism perception, tourism consumption, tourism resources, and tourism services are 0.1440, 0.1440, 0.6293, and 0.0827, respectively. Tourism resources in the criterion layer are relatively important, and tourists mainly consider whether tourism resources of ancient towns can satisfy them when they visit them. In the indicator layer, the weights of tourist mood, catering and accommodation, and service attitude were all 1, with total weights of 0.1440, 0.1440, and 0.0927, respectively. The weights of ancient town scenery, scenic area facilities, scenic area environment, and ancient town history and culture were 0.1899, 0.5143, 0.1900, and 0.1058, respectively, with total weights of 0.1195, 0.3236, 0.1195, and 0.0666, respectively. From the total weights of the indicators in the indicator layer, the most important indicator is scenic area facilities, indicating that tourists value whether the infrastructure of the ancient towns is complete and whether the entertainment facilities meet their needs for recreation. Less important indicators are the history and culture of the ancient towns and service attitude, but these are also important factors for tourists when choosing to visit the ancient towns. Only by strengthening and improving various services, facilities, and consumption standards of the ancient towns can the satisfaction of tourists be improved and the reputation of the ancient towns be established.

Table 3. Statistics of weight result for ancient towns.

	Weight	Zhou-Zhuang	Tongli	Wuzhen	Xitang	Luzhi	Huang-Yao	Xidi–Hongcun	Nanxun	Zhujia Jiao	Heshun
Perception	0.144	0.102	0.069	0.206	0.047	0.033	0.069	0.287	0.146	0.017	0.023
Consumption	0.144	0.087	0.061	0.307	0.087	0.021	0.043	0.228	0.125	0.021	0.021
Resource	0.629	0.109	0.078	0.210	0.056	0.027	0.039	0.291	0.152	0.019	0.019
Service	0.083	0.078	0.109	0.291	0.055	0.025	0.038	0.210	0.152	0.025	0.018
Weighted result		0.103	0.077	0.230	0.059	0.027	0.044	0.274	0.147	0.019	0.020

Using the weighted average method to analyze each ancient town generated the results shown in the last column in Table 3 and Figure 6. Xidi–Hongcun and Wuzhen have larger scores (0.274 and 0.230), indicating that these two ancient towns are the optimal choices for tourists making travel destination decisions based on online reputation. For Heshun (0.020) and Zhujiajiao (0.019), the weighted average results are lower, suggesting that these tourist destinations do not receive enough feedback from visitors and have weaker attraction on the internet. By comparing the cluster analysis and analytic hierarchy process (AHP) results of the top ten ancient towns, we obtain the following comprehensive ranking: Xidi–Hongcun, Wuzhen, Nanxun, Zhouzhuang, Tongli, Xitang, Huangyao, Luzhi, Heshun, and Zhujiajiao. Apart from Xidi–Hongcun and Wuzhen, which stand out, the rankings of the other ancient towns align with the Maigoo rankings.

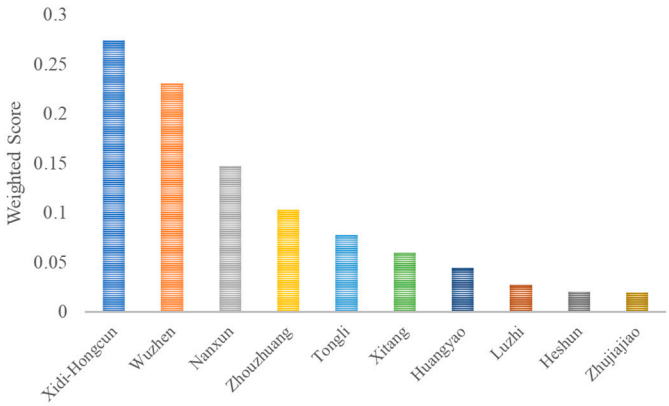


Figure 6. Ranking of ancient towns.

3.5. Discussion

3.5.1. A Deeper Analysis for Ancient Towns

Ancient towns are often managed as scenic areas, and visitors need to pay an entrance fee to tour and sightsee within the area. Some ancient towns, such as Wuzhen and Xidi–Hongcun, have become the most common keywords for negative reviews due to their "expensive" tickets. Therefore, we collected the ticket prices of each ancient town (Table 4). It was found that, except for Zhujiajiao, which offers free admission, Heshun and Luzhi have relatively low fees at CNY 55, while other ancient towns require entrance fees of over CNY 80. The most expensive ones are Wuzhen and Xidi–Hongcun, which cost CNY 190 and CNY 104, respectively. Several factors influence the pricing of scenic area tickets. Among these, accessibility and the income of the surrounding area are important indicators. Considering that the consumer population of scenic areas often comes from local cities and their nearby areas, we collected data on the five nearest cities to each ancient town. It was found that Heshun and Huangyao are relatively far from provincial capital cities, with average distances to the five cities of 165 km and 131 km, respectively. Although their ticket prices are relatively low, the accessibility for tourists to these two ancient towns is poor, making it difficult to increase the willingness of visitors to travel to the towns. These two cities also face different challenges. Heshun needs to address transportation issues and competition from other cities like Dali and Lijiang, which have famous ancient towns, while Huangyao faces low attention and satisfaction levels, few large nearby cities, and limited foot traffic. Although the surrounding cities of Xidi–Hongcun have relatively low economic levels and populations, its distinct features as a world cultural heritage site and its ability to link up with other attractions such as Mount Huangshan make its ticket price acceptable. Luzhi, Nanxun, Tongli, Wuzhen, Xitang, Zhouzhuang, and Zhujiajiao have neighboring cities like Shanghai, Hangzhou, Suzhou, and Wuxi, which are economically

developed and have large populations. The residents' income levels and consumption levels can also support their consumption well.

Table 4. Information on ticket prices and neighboring cities for towns.

Town	Ticket Fee	Neighbor Cities
Heshu	55	Luxi 66 km, Baoshan 74 km, Dali 193 km
Huangyao	88	Hezhou 38 km, Wuzhou 88 km, Guilin 152 km
Luzhi	55	Suzhou 25 km, Shanghai 55 km, Jiaxing 57 km
Nanxun	80	Huzhou 32 km, Jiaxing 33 km, Suzhou 50 km
Tongli	100	Suzhou 20 km, Jiaxing 33 km, Wuxi 61 km
Wuzhen	190	Jiaxing 25 km, Huzhou 40 km, Hangzhou 60 km
Xidi–Hongcun	104	Huangshan 37 km, Guichi 97 km, Jingdezhen 103 km
Xitang	95	Jiaxing 25 km, Suzhou 48 km, Shanghai 61 km
Zhouzhuang	100	Suzhou 31 km, Jiaxing 40 km, Shanghai 58 km
Zhujiajiao	Free	Shanghai 40 km, Suzhou 47 km Jiaxing 49 km

We also collected data on the number of tourists and tourism revenue in various ancient towns. In 2023, Nanxun, with its permanent free-entry policy for global tourists, attracted more than 10 million visitors, making it the undisputed leader. Meanwhile, Wuzhen closely followed, welcoming 7.72 million tourists and generating a tourism revenue of CNY 1.78 billion, ranking high in the total tourism revenue. Tongli also achieved remarkable results, attracting more than 7 million tourists. Moreover, its tourism revenue also reached a high level, which further demonstrated its excellent performance. Xitang received over 7 million visitors, with a tourism revenue of CNY 186 million, and drove the entire tourism industry to earn CNY 2.19 billion, showing its large overall scale. Luzhi had an annual visitor volume of 2 million, and its total tourism revenue exceeded CNY 200 million, showing stable performance. Zhouzhuang received more than 4 million visitors annually, with a tourism revenue of CNY 3 billion, and its revenue growth was particularly remarkable. Xidi and Hongcun attracted 3.2 million tourists, with a direct tourism revenue close to CNY 190 million. Huangyao, although it had 1.28 million visitors, had a total tourism revenue of CNY 37.3 million, which was relatively low. Heshun received 1.2 million visitors, driving a total tourism revenue of CNY 300 million, indicating good development momentum. Zhujiajiao received over 4 million visitors annually, with a tourism revenue close to CNY 1 billion, but slightly lower than that of Zhouzhuang.

Another issue to consider is that ancient towns in the Yangtze River Delta often exhibit significant homogeneity under the theme of "Jiangnan Water Towns". In the future, they will need to strengthen their unique characteristics, seek upgrades and transformation, and achieve differentiation and multi-industry development. For example, Wuzhen maintains the original water town appearance of the ancient town while striving to develop a museum-style ancient town, creating a characteristic theme of poetic Wuzhen, becoming an oasis away from urban hustle and bustle. After the differentiated upgrading of ancient towns, they ultimately still need to be tested by tourists. From previous analyses, apart from Wuzhen and Xidi–Hongcun, other ancient towns have lower attention levels. Therefore, ancient town management authorities can also use internet thinking to carry out online publicity activities, utilizing news media as well as currently popular short videos to create exclusive promotional clips for ancient towns, highlighting their unique natural resources to attract tourists and enhance the visibility of each ancient town. At the same time, since online reputation can directly reflect the advantages and problems of ancient town development, managers who refer to online reputation to make timely targeted adjustments (such as environmental sanitation and employee training) will be more beneficial for the sustainable development of ancient towns.

3.5.2. Thinking of Data, Methods, and eWOM

Based on text-based sentiment analysis, the method applied in this paper can automatically process and analyze a large amount of text data. Compared with interviews that require a great deal of time on face-to-face communication or telephone communication, it has an advantage. It can quickly conduct sentiment-orientation analysis on a large amount of text data and provide real-time emotional insights. Because it depends on algorithms and natural language processing techniques rather than personal subjective judgment, it reduces human-related bias (S. Chen et al., 2015). This article also adopts the traditional AHP method. The weights in AHP are determined based on the scores given by experts. However, these scores are susceptible to influence from personal experience. The entropy weight method, on the other hand, determines the weights of each index by calculating the information entropy. In this way, it reduces the interference of subjective factors and enhances the objectivity and reliability of the results. Moreover, it can dynamically adjust the weights in response to changes in actual data, which endows the model with greater adaptability and flexibility (Yazdani et al., 2019).

Data for eWOM are mainly sourced from online platforms such as Yelp and Meituan. The user groups of these platforms are often young and technically proficient. This situation leads to a deviation in sample representativeness. In addition, users who are willing to post comments usually have stronger feelings about their experiences, and the comments tend to be extreme (Donthu et al., 2021; Z. Wang et al., 2021). We attempted to study how the expectations of tourists from different regions affect their evaluations and emotions, but there is a difficulty. It is difficult to obtain the real origin of the reviewers, because the location information shown on the platform is usually the location at the time of the review, rather than the place of origin of the tourists. Meanwhile, the rule-based sentiment analysis model may not accurately capture complex semantics such as metaphors and sarcasm, resulting in misjudgments. These limitations in data collection and analysis methods could potentially affect the accuracy and reliability of the research findings.

Although there may be deviations in data and methods, this paper constructs a relatively complete ancient town evaluation system. This system can be used for both qualitative descriptive sentiment analysis and quantitative AHP analysis. For example, we found that tourists have higher expectations for Xidi–Hongcun, but lower service satisfaction. Based on these findings, tourism managers can optimize service facilities and enhance the tourist experience. Local governments can use this feedback to adjust resource allocation, highlight the cultural characteristics of ancient towns, and attract more tourists. This system has multiple functions. It can enrich tourist evaluation methods and also provide a scientific basis for tourism management and local development. Positive eWOM comments can enhance tourists' trust and satisfaction. Satisfied tourists are more likely to become repeat customers and recommend the destination to others. Additionally, eWOM spreads quickly and has a wide coverage, which can effectively attract potential tourists (Cronin & Taylor, 1992; Cheung & Thadani, 2012). Negative comments can also spread rapidly and cause serious damage to the reputation of the tourist destination. Potential tourists may change their plans due to negative comments (Kim & Lee, 2021).

The method of word-of-mouth evaluation discussed in this article is not only applicable to the tourism service industry, but also to hotels and retail industries. By analyzing customers' online reviews and survey questionnaires, hotel managers can understand the professionalism of their staff, response speed, and problem-solving abilities, thereby improving service quality (Torres et al., 2015). In the retail industry, eWOM helps to obtain information about consumers' shopping experiences, including product quality, service attitude, and price transparency (Kang et al., 2022). Moreover, there are differences between industries and regions in analysis of eWOM. For instance, the catering industry focuses

more on food quality and service, while the technology industry emphasizes product innovation (Soegoto & Amelia, 2021). Consumers' demands and expectations are different in regions with different levels of economic development. People in economically developed areas place more emphasis on quality and innovation, while those in less developed regions may focus more on price and practicality (Smith & Johnson, 2021).

4. Conclusions

This paper first employs GIS to analyze the locational information of ancient towns. Subsequently, it utilizes web crawlers and Python tools to collect, filter, tokenize, and clean tourist review data. The data are subjected to term frequency analysis using the TF-IDF method, and sentiment analysis along with IPA (Importance–Performance Analysis) is conducted to quantify the reputation and satisfaction levels of the ancient towns. Further, the LDA (Latent Dirichlet Allocation) topic model is applied to identify themes within reviews, leading to clustering analysis of the towns. Finally, the analytic hierarchy process (AHP) is used to evaluate the ancient towns, and the main conclusions derived are as follows:

(1) The top ten southern ancient towns are concentrated in the Yangtze River Delta region, where access by train, plane, or road is relatively convenient, with a significantly higher road density. Heshun and Huangyao have the lowest road density, while the highest road densities are found in Luzhi, Tongli, Zhouzhuang, and Zhujiajiao.
(2) The volume of public opinion on ancient towns is closely related to timeframes, with larger crowds during holidays. Tourism in ancient towns has been greatly affected by the pandemic, resulting in a decline in public opinion volume and a weak correlation with pre-pandemic levels.
(3) High praise ranks high in term frequency, indicating that tourists generally hold a positive attitude towards ancient towns. Tourists are most concerned about the construction of scenic facilities in ancient towns, followed by mood when visiting and dining and accommodation. Visitors are most satisfied with the scenery of ancient towns, while mood when visiting is an area that urgently needs improvement in terms of satisfaction.
(4) Among the top ten ancient towns, Wuzhen is in a favorable position, receiving numerous positive reviews in all aspects. Overall, Wuzhen and Xidi–Hongcun are the optimal choices for tourists making travel destination decisions based on online public opinion. However, Luzhi, Zhujiajiao, and Heshun need more improvements.

This paper constructs a relatively comprehensive evaluation system for ancient towns based on text analysis. This system can perform both qualitative descriptive sentiment analysis and quantitative analytic hierarchy process (AHP) model analysis. It effectively addresses the limitations of traditional tourism research, which often relies solely on qualitative descriptions or binary sentiment classifications. However, it has certain limitations. The current dataset may not cover all relevant variables, such as location and demographic characteristics. Also, although our analytical approach is robust, it may not fully capture the complexity of these relationships. This system can also be used for deeper research on mediating effects. To address these limitations and further improve the system, we have the following plans: Firstly, we intend to add new data types, like location and demographic characteristics, to our analysis. This will help us more comprehensively examine the factors that influence eWOM. Secondly, we plan to use advanced methodologies, such as artificial intelligence, to carry out more sophisticated and in-depth information mining and sentiment analysis. Lastly, we are considering applying our system to mediating-effect analysis in new situations. Specifically, we plan to explore the mediating effects between online comment and booking behaviors, as well as between tourism experiences and eWOM communication.

Author Contributions: Conceptualization, Y.Z. and S.L.; methodology, Y.Z., W.G. and S.L.; validation, W.G.; formal analysis, Y.S.; data curation., W.G. and Y.Z.; writing—original draft preparation, W.G.; writing—review and editing, Y.Z., Y.S. and S.L.; visualization, W.G.; supervision, Y.Z.; project administration, Y.Z.; funding acquisition, Y.Z. and W.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by National Natural Science Foundation of China, grant number 42471438.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available in Meituan at https://www.meituan.com/.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

WOM	traditional word-of-mouth
eWOM	electronic word-of-mouth
AHP	analytic hierarchy process
IPA	Importance–Performance Analysis
TF-IDF	Term Frequency–Inverse Document Frequency
TF	the frequency of terms
IDF	inverse document frequency
LDA	Latent Dirichlet Allocation

References

Aggarwal, S., & Gour, A. (2020). Peeking inside the minds of tourists using a novel web analytics approach. *Journal of Hospitality and Tourism Management*, *45*, 580–591. [CrossRef]

Almanza, B. A., Jaffe, W., & Lin, L. (1994). Use of the service attribute matrix to measure consumer satisfaction. *Hospitality Research Journal*, *17*(2), 63–75. [CrossRef]

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. [CrossRef]

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, *28*(2), 15–21. [CrossRef]

Chen, S., Zhang, J., & Li, Y. (2015). Comparing text-based sentiment analysis with traditional interview methods in consumer research. *Journal of Marketing Research*, *52*(4), 487–502.

Chen, Y., & Xie, F. (2019). The role of immersive experiences in tourism: A sentiment analysis approach. *Journal of Travel Research*, *58*(4), 671–686.

Chen, Y.-C., & Tsai, C.-H. (2019). An importance-performance analysis of service quality in the hotel industry. *Journal of Hospitality and Tourism Management*, *39*, 483–495.

Cheung, C. M. K., & Thadani, D. R. (2012). The effects of electronic word-of-mouth on trust and purchase intention in online consumer-to-consumer marketplaces. *Internet Research*, *22*(3), 307–324.

Cronin, J. J., Jr., & Taylor, S. A. (1992). Measuring service quality: A reexamination and extension. *Journal of Marketing*, *56*(3), 55–68. [CrossRef]

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*, 1407–1424. [CrossRef]

Donthu, N., Kumar, S., Pandey, N., Pandey, N., & Mishra, A. (2021). Mapping the electronic word-of-mouth (eWOM) research: A systematic review and bibliometric analysis. *Journal of Business Research*, *135*, 758–773. [CrossRef]

Ferreira, R., Brito, P. H. S., Melo, J., & Silva, A. (2018). An architecture-centered framework for developing blog crawlers. *Journal of Web Engineering*, *17*(3–4), 265–290.

Gursoy, D., Huang, S., & Chi, C. G. Q. (2009). A review and synthesis of tourism and hospitality research related to climate change. *Tourism Management*, *30*(2), 175–186.

Hajjem, M., & Latiri, C. (2017). Combining IR and LDA topic modeling for filtering microblogs. *Procedia Computer Science*, *112*, 761–770. [CrossRef]

Hameed, I., Akram, U., Khan, Y., Khan, N. R., & Hameed, I. (2024). Exploring consumer mobile payment innovations: An investigation into the relationship between coping theory factors, individual motivations, social influence and word of mouth. *Journal of Retailing and Consumer Services*, *77*, 103687. [CrossRef]

Hartigan, J. A., & Wong, M. A. (1979). A K-means clustering algorithm. *Applied Statistics*, *28*(1), 100–108. [CrossRef]

Hua, X., Hasan, N. A. M., De Costa, F., & Qiao, W. (2024). The mediating role of electronic word-of-mouth in the relationship between CSR initiative and consumer satisfaction. *Heliyon*, *10*(15), e35027. [CrossRef]

Huang, M.-H., & Benyoucef, M. (2009). The impact of online customer reviews on consumer purchase intentions: A meta-analysis. *Journal of Interactive Marketing*, *23*(4), 265–280.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, *31*(8), 651–666. [CrossRef]

Johnson, D. (2000). Anonymity and the internet. *The Futurist*, *34*, 12.

Kang, M., Sun, B., Liang, T., & Mao, H. (2022). A study on the influence of online reviews of new products on consumers' purchase decisions: An empirical study on JD.com. *Frontiers in Psychology*, *13*, 876542. [CrossRef]

Khan, K. A., Ali, F., Khan, U., Sun, D., & Alotaibi, S. (2024). Seeing isn't always believing: Exploring the influence of misleading photos on consumer responses in hotel industry. *International Journal of Hospitality Management*, *123*, 103934. [CrossRef]

Kim, J., & Lee, Y. (2021). Using deep learning for image recognition in tourism marketing. *Tourism Economics*, *27*(2), 295–312.

Kumar, S., & Kaur, H. (2019). Sentiment analysis of social media data for destination image formation: A case study of India. *Tourism Management Perspectives*, *31*, 100594.

Kumar, S., & Vaidya, O. S. (2006). Analytic hierarchy process: An overview of applications. *European Journal of Operational Research*, *169*(1), 1–29. [CrossRef]

Lam, D., & So, A. (2013). Do happy tourists spread more word-of-mouth? The mediating role of life satisfaction. *Annals of Tourism Research*, *43*, 646–650. [CrossRef]

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, *29*(3), 458–468. [CrossRef]

Liu, H., Guo, P., & Yin, H. (2022). Consistent comments and vivid comments in hotels' online information adoption: Which matters more? *International Journal of Hospitality Management*, *107*, 103329.

Macdonald, S. (2001). A people's past? Museums and heritage attractions in the shaping of history. *Museum Management and Curatorship*, *19*(1), 5–26.

Martilla, J. A., & James, W. C. (1977). Importance-performance analysis. *Journal of Marketing*, *41*(1), 77–79. [CrossRef]

Matzler, K., Teichmann, K., Strobl, A., & Partel, M. (2019). The effect of price on word of mouth: First time versus heavy repeat visitors. *Tourism Management*, *70*, 453–459. [CrossRef]

Morgan, N. A., & Pritchard, A. (2020). The role of word-of-mouth communications in destination marketing: A review and future directions. *Current Issues in Tourism*, *23*(10), 1234–1252.

Newman, D., Asuncion, A., Smyth, P., & Welling, M. (2009). Distributed algorithms for topic models. *Journal of Machine Learning Research*, *10*, 1801–1828.

Oppermann, M. (1996). Convention cities—Images and changing fortunes. *Journal of Tourism Studies*, *7*, 10–19.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135. [CrossRef]

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). A conceptual framework for quality management research. *Journal of the Academy of Marketing Science*, *33*(1), 12–27.

Qu, H., Teh, B. T., Nordin, N. A., & Liang, Z. (2024). Analysis of Guangzhou city image perception based on weibo text data (2019–2023). *Heliyon*, *10*(17), e36577. [CrossRef]

Rubin, T. N., Chambers, A., Smyth, P., & Steyvers, M. (2011). Statistical topic models for multi-label document classification. *Machine Learning*, *88*, 157–208. [CrossRef]

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, *15*(3), 234–281. [CrossRef]

Smith, J., & Johnson, L. (2021). Economic factors influencing spatial variations in consumer-generated reviews in the retail sector. *Journal of Retailing and Consumer Services*, *60*, 102345.

Soegoto, E. S., & Amelia, N. (2021). Utilization of word-of-mouth marketing strategy in culinary business. *International Journal of Research and Applied Technology*, *7*(2), 123–130. [CrossRef]

Song, M., Law, R., & Duan, Y. (2022). What are the obstacles in the way to "avoid landmines"? Influence of electronic word-of-mouth dispersion on order decision from the self-construal perspective. *International Journal of Hospitality Management*, *107*, 103334. [CrossRef]

Sugar, C. A., & James, G. M. (2003). Finding the number of clusters in a dataset: An information-theoretic approach. *Journal of the American Statistical Association*, *98*(463), 750–763. [CrossRef]

Timothy, D. J., & Boyd, S. W. (2011). *Heritage tourism*. Routledge.

Torres, E. N., Adler, H., Behnke, C., Miao, L., & Lehto, X. (2015). The use of consumer-generated feedback in the hotel industry: Current practices and their effects on quality. *International Journal of Hospitality & Tourism Administration*, *16*(3), 224–250. [CrossRef]

Wang, H., Wang, Y., & Zhang, X. (2016). A comparative study of text-based sentiment analysis and traditional survey methods in customer satisfaction research. *Journal of Business Research*, *69*(10), 4065–4072.

Wang, L., & Kim, K. (2024). Analyzing group polarization through text emotion measurement and time series prediction: A comparative study across three online platforms. *Measurement: Sensors*, *33*, 101216. [CrossRef]

Wang, N., & Zhang, L. (2020). Comparative analysis of tourism development in historical towns: A case study of Pingyao and Lijiang. *Journal of Sustainable Tourism*, *28*(7), 985–1004.

Wang, Z., Wang, D., & Li, Q. (2021). Keyword extraction from scientific research projects based on SRP-TF-IDF. *Chinese Journal of Electronics*, *30*, 652–657.

Wu, J., & Yang, T. (2023). Service attributes for sustainable rural tourism from online comments: Tourist satisfaction perspective. *Journal of Destination Marketing & Management*, *30*, 100822.

Yazdani, M., Chatterjee, P., Zavadskas, E. K., & Turskis, Z. (2019). A comprehensive evaluation of multi-criteria decision-making techniques: A systematic literature review and bibliometric analysis. *Journal of Business Research*, *100*, 19–32.

Yu, J., Hu, Y., Yu, M., & Di, Z. (2010). Analyzing netizens' view and reply behaviors on the forum. *Physica A*, *389*, 3267–3273. [CrossRef]

Yun, G. W., Park, S., & Ha, L. (2008). Influence of cultural dimensions on online interactive review feature implementations: A comparison of Korean and U.S. retail web sites. *Journal of Interactive Marketing*, *22*(3), 40–50. [CrossRef]

Zhang, J., Cheng, M., Mei, R., & Wang, F. (2020). Internet use and individuals' environmental quality evaluation: Evidence from China. *Science of The Total Environment*, *710*, 136290. [CrossRef] [PubMed]

Zhang, L., & Li, Y. (2021). Using AI to enhance personalized travel recommendations: A case study of airbnb. *Journal of Hospitality and Tourism Technology*, *12*(2), 213–228.

Zheng, X., Huang, J., Wu, J., Sun, S., & Wang, S. (2023). Emerging trends in online reviews research in hospitality and tourism: A scientometric update (2000–2020). *Tourism Management Perspectives*, *47*, 101105. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.