

Research on Visitor Perception of Huangshan Cultural and Creative Products Based on Text Mining

Chen Qi

Design Innovation Center, China Academy of Art, Hangzhou, Zhejiang, China

Keywords: Text mining; online text analysis; tourist perception; Huangshan cultural and creative products

Abstract: This study employs text mining methods to analyze 961 user-generated posts on the Xiaohongshu platform, exploring tourists' perceptions of Huangshan cultural and creative products and providing optimization suggestions. The research first collects relevant text data and utilizes ROST CM6 for high-frequency word analysis, semantic network construction, and sentiment analysis to identify tourists' core needs and emotional tendencies. The findings indicate that while tourists highly appreciate the cultural elements and design of Huangshan's cultural and creative products, issues such as limited innovation, restricted sales channels, and suboptimal user experience persist in the market. To address these challenges, this paper proposes three optimization strategies: enhancing cultural expression and deepening the integration of Huizhou heritage, expanding sales channels to improve market accessibility, and enhancing interactive experiences to strengthen cultural immersion.

1. Introduction

In recent years, with the rapid development of the tourism industry, the development and promotion of cultural and creative products have become key strategies for enhancing the attractiveness of tourist destinations. As one of China's renowned 5A-level scenic spots, Huangshan attracts a large number of domestic and international tourists with its unique natural landscapes and rich cultural heritage. However, in an increasingly competitive tourism market, research on tourists' perceptions of Huangshan cultural and creative products remains insufficient, and there is still much room for improvement in their development and marketing. In particular, studies focusing on cultural and creative products in world-class tourist destinations like Huangshan are scarce. The alignment between Huangshan's distinctive Huizhou cultural resources and the diverse needs of tourists has not yet been fully explored.

From a tourist perception perspective, visitors to Huangshan have gradually shifted their focus from traditional sightseeing to cultural experiences and emotional resonance. This trend suggests that cultural and creative products, as an important medium for Huangshan's cultural representation, should be designed and promoted in ways that closely align with tourists' practical needs and experiential preferences. Existing literature primarily examines tourists' perceptions of the overall destination image. For example, Xu Yayuan et al. analyzed the impact of online word-of-mouth on tourists' destination decision-making through questionnaire surveys ^[1] while Wang Qun et al.

developed a tourism environment satisfaction assessment model ^[2]. However, studies focusing on tourist perceptions of cultural and creative products remain limited. In particular, under the digitalization trend, how to utilize text mining techniques to analyze tourists' evaluations and feedback on cultural and creative products remains an area that requires further exploration.

This study employs text mining techniques to examine tourists' perceptions of Huangshan cultural and creative products, identifying their evaluations and expectations regarding design, functionality, cultural significance, and market promotion. The research is based on Zeithaml's Consumer Perceived Value Framework ^[3] and incorporates the evaluation dimensions proposed by Sweeney et al. ^[4] to systematically analyze tourists' perceived value of Huangshan's cultural and creative products.

2. Research design

2.1. Research Methodology

This study employs ROST CM6 software for text mining analysis. Text mining is an interdisciplinary approach that involves extracting meaningful information and knowledge from unstructured text data ^[5]. This method has been widely applied in tourism research. For example, Xiao Liang introduced it into destination image analysis ^[6], Zhao Zhenbin applied it to tourist behavior studies ^[7], and Huang Jie used it to examine tourists' perceived value dimensions ^[8]. ROST CM6 offers various text processing and visualization functions, including word segmentation, word frequency analysis, and semantic network construction, making it a valuable tool for analyzing chat records, website content, social media texts, and academic literature.

The research follows the workflow below:

- Data Cleaning: Removing meaningless symbols, stop words, and unifying terminology.
- High-Frequency Word Analysis: Extracting the top 60 most frequent keywords and categorizing them into thematic groups.
- Semantic Network Analysis: Constructing a semantic network graph based on co-occurrence relationships to identify key topics of tourist interest.
- Sentiment Analysis: Using a positive-negative sentiment lexicon to classify tourist reviews into positive, neutral, or negative categories, thereby quantifying sentiment tendencies.

2.2. Data Source

Social media platforms have become a primary source of information for cultural and tourism consumers, with users increasingly relying on social media content during their decision-making process. For instance, the Palace Museum IP has successfully integrated its deep cultural heritage into consumers' daily lives through new media channels ^[9]. Many scholars have utilized user-generated travel notes and reviews to study tourism destinations, as social media data offers greater immediacy, authenticity, and interactivity compared to traditional survey methods ^[10].

From a communication influence perspective, Huangshan's tourism content has seen growing engagement on Xiaohongshu, attracting not only a large volume of authentic tourist experiences but also official tourism accounts such as Huangshan Tourism, Qiyun Mountain Resort, and Huizhou Ancient City. Additionally, cultural and lifestyle media brands like Pioneer Bishan Bookstore and Huizhuwu have joined the platform, further enhancing Huangshan tourism's visibility on social media. As a UGC (user-generated content) platform, Xiaohongshu provides direct insights into tourists' consumption experiences and emotional responses toward cultural and creative products.

Considering these factors, this study selects tourist-generated posts on Xiaohongshu as the research sample. By applying text mining techniques, this research aims to analyze tourists'

perceptions of Huangshan cultural and creative products, providing data-driven insights for product optimization and market expansion.

2.3. Text Mining and Data Processing

This study employs Octoparse web scraping software to automate data collection. The keyword “Huangshan cultural and creative products” was used to search for relevant posts on Xiaohongshu, and results were sorted by the number of likes to prioritize highly interactive content. An initial dataset of 1,000 posts was collected. To ensure data quality, irrelevant content, duplicate records, and promotional material were filtered out, resulting in a final dataset of 961 valid posts (totaling 325,000 characters) for analysis.

To improve the accuracy of text analysis, the selected dataset underwent preprocessing and standardization. Given that social media users are predominantly younger audiences, posts often contain informal expressions, requiring segmentation tools and frequency-based text models for cleaning and normalization^[11].

The data preprocessing steps included: Removing unnecessary characters, such as punctuation, numbers, and extra spaces, to maintain text clarity. Standardizing terminology by incorporating Huangshan’s major landmarks and streets into a custom dictionary, ensuring consistency in key cultural terms. For example, phrases like "Huizhou architecture" and "Hui-style buildings" were unified under the standard term "Huizhou architecture". The cleaned text is converted into TXT format to ensure compatibility for subsequent text mining and analysis.

3. Tourist Perception Analysis

3.1. Word Frequency Analysis

The distribution of high-frequency words effectively highlights tourists’ key concerns. An initial analysis of the textual data identified the top 60 most frequently mentioned words related to Huangshan’s cultural and creative products.

Using content analysis, these high-frequency words were categorized into five thematic dimensions: cultural value, product attributes, emotional experience, regional relevance, and market demand. These dimensions reflect tourists’ core concerns regarding Huangshan cultural and creative products.

3.1.1. Cultural Value

Tourists generally pay great attention to the cultural heritage and historical significance of Huangshan’s cultural and creative products. The frequent occurrence of words such as Huizhou, villages, intangible cultural heritage, tradition, art, museums, aesthetics, and Hui ink indicates a high level of recognition for the cultural depth and historical continuity of these products. These cultural elements serve as key selling points, enhancing the cultural identity and distinctiveness of Huangshan’s cultural and creative offerings.

3.1.2. Product Attributes

In terms of product categories, tourists tend to prefer tangible souvenirs. Words like cultural creativity, fridge magnets, stamping, postcards, stamps, illustrations, collectibles, handicrafts, metal, and texture indicate that physical souvenirs are the most popular choices. Additionally, words such as tea, green tea, Mao Feng, local cuisine, ice cream, Huizhou pastries, and specialties highlight tourists' appreciation for local specialty food-related cultural products. These products not only

fulfill practical needs but also carry regional and cultural memories of Huizhou.

On the other hand, functional cultural products, such as Huizhou herbal tea tablets, rely primarily on packaging design and material selection to convey cultural value in a subtle yet effective manner. These products have strong market appeal, particularly for immediate consumption.

3.1.3. Emotional Experience

Tourists' evaluations of cultural and creative products often involve personal emotions and aesthetic judgments. Words such as beautiful, exquisite, wonderful, cute, packaging, and practical suggest that design aesthetics and product quality are highly valued in purchasing decisions.

Additionally, the presence of words like experience, museums, post offices, stamping, and customization indicates that tourists have a strong preference for interactive and personalized experiences. The integration of customization and immersive interactions enhances product appeal and strengthens emotional connections to Huizhou culture, making the purchasing process more ritualistic and engaging.

This aligns with the Huangshan Municipal Bureau of Culture and Tourism's initiative to expand the cultural and creative product chain, develop market-attractive cultural tourism products, and integrate immersive experiences with cultural tourism complexes.

3.1.4. Regional Relevance

Huangshan's natural landscapes and Huizhou cultural symbols play a central role in product design. Words such as Huangshan, Welcome Pine, sea of clouds, and landscapes indicate that tourists expect cultural and creative products to capture Huangshan's unique scenery, enhancing their commemorative value.

Furthermore, the appearance of words like lanterns, architecture, and history suggests that tourists show strong interest in Huizhou's traditional craftsmanship and cultural symbols, hoping to see these elements incorporated into products. This regional connection reinforces the local identity of the products, making them not just commodities but also meaningful souvenirs that preserve travel memories.

3.1.5. Market Demand

Tourists purchase Huangshan cultural and creative products for both personal use and gift-giving purposes. Words like gifts, gift boxes, birthday gifts, New Year, blessings, and sentiments indicate that these products have strong market potential beyond individual consumption, extending to festive gifting and cultural exchange.

Additionally, the high frequency of words such as originality, market, and customization reflects tourists' growing demand for innovative and personalized products. Consumers seek unique designs to avoid product homogeneity, emphasizing the importance of enhancing collectibility and symbolic value in cultural and creative offerings.

Overall, tourists' perceptions of Huangshan cultural and creative products highlight recognition of Huizhou culture and an increasing demand for cultural depth, innovative design, and personalized experiences. These insights provide critical guidance for optimizing product development and improving market positioning, reinforcing the cultural, commercial, and experiential value of Huangshan's cultural and creative industry.

3.2. Semantic Network Analysis

Semantic network analysis identifies co-occurrence relationships among high-frequency words,

constructing a semantic network graph that visually represents core concepts and their interconnections. This method helps uncover tourists' primary concerns and underlying semantic patterns regarding Huangshan cultural and creative products. Using ROST CM6, a social and semantic network graph was generated from travel texts, as shown in Figure 1.

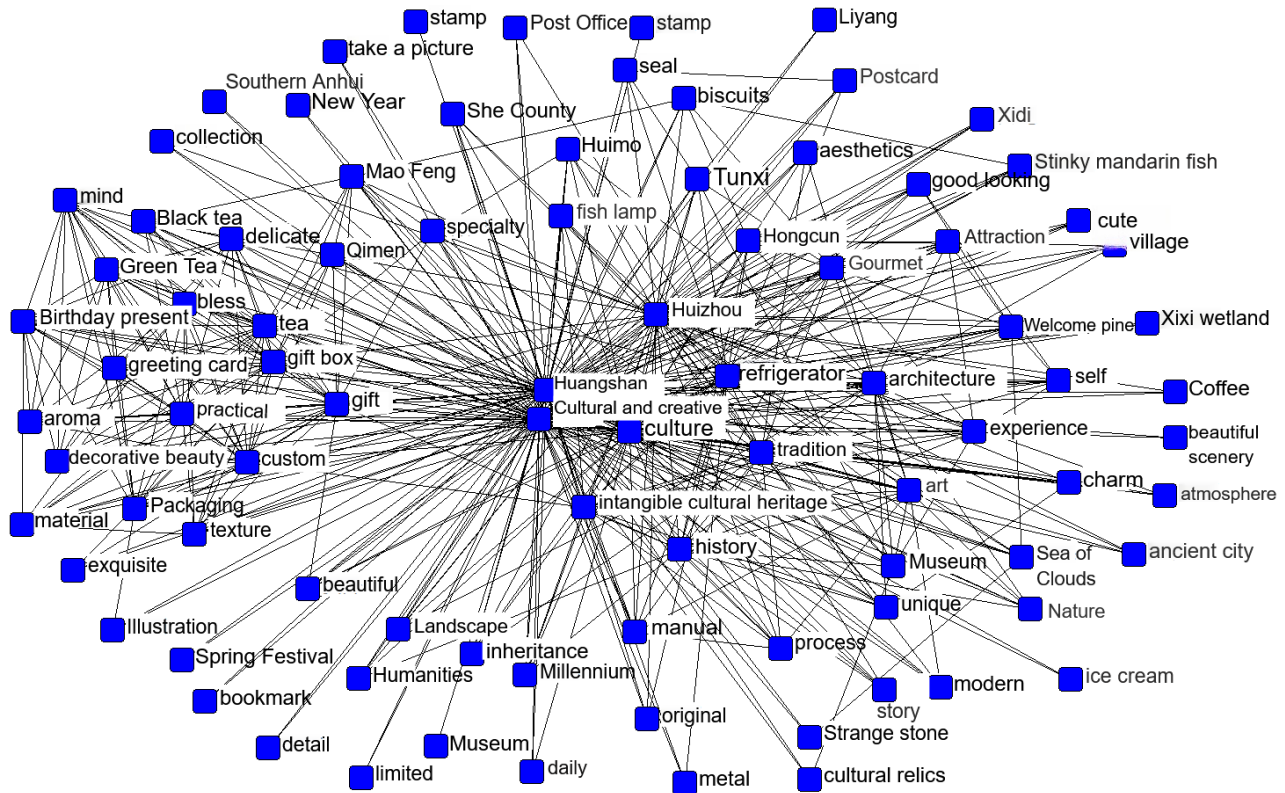


Figure 1: High-Frequency Word Semantic Network

The analysis results indicate that Huangshan and cultural creativity are the central nodes of the network, confirming that Huangshan cultural and creative products have become a key discussion topic among tourists. Within this network, culture and tradition exhibit strong co-occurrence relationships, closely linked to terms such as Huizhou, intangible cultural heritage, art, architecture, and museums. This pattern reflects tourists' strong recognition of cultural symbols associated with Huangshan, emphasizing its historical significance and artistic characteristics.

Additionally, natural landscape-related terms such as Welcome Pine, sea of clouds, and Mao Feng tea hold significant positions in the network, indicating that Huangshan's natural resources serve as important design elements in cultural and creative products. Meanwhile, the high association of words such as gift, gift box, and customization suggests that Huangshan cultural and creative products are primarily positioned as souvenirs and gifts, catering to the demand for gift-giving tourism consumption.

Market research further supports these findings, revealing that personalization and commemorative features are particularly favored by young tourists, aligning with recent trends in the cultural and tourism industry. Moreover, the presence of words like experience, museums, and stamping suggests that tourists' interest in cultural and creative products extends beyond physical attributes, encompassing immersive cultural experiences. Modern travelers increasingly prefer interactive engagement with cultural elements, enhancing their sense of participation and cultural identity.

In summary, semantic network analysis highlights that the core appeal of Huangshan cultural and creative products stems from a combination of cultural heritage, natural landscapes, and

market-driven demands.

3.3. Sentiment Analysis

Sentiment analysis aims to identify and classify the subjective emotional tendencies present in textual data, categorizing tourists' reviews into different sentiment levels, recognizing their attitudes, and analyzing sentiment polarity. As shown in Table 1, this analysis helps to reveal tourists' overall satisfaction and emotional expectations regarding Huangshan cultural and creative products.

Table 1: Sentiment Analysis of "Huangshan Cultural and Creative Products" on Xiaohongshu

Emotional category	Emotional segment	Quantity/piece	Proportion / %	Total/piece	Proportion / %
Positive emotion	General (0,15)	223	23.20	759	78.98
	Moderate (15,25)	141	14.67		
	Height (25,+∞)	395	41.10		
Neutral emotion	[-6,0]	143	14.88	143	14.88
Negative emotion	General (0,15)	33	3.43	59	6.14
	Moderate (15,25)	12	1.25		
	Height (25,+∞)	14	1.46		

The results indicate that tourists generally hold a positive attitude toward Huangshan cultural and creative products, with 78.98% of reviews expressing positive emotions, among which 41.10% exhibit highly positive sentiment. These favorable reviews mainly highlight the refined design, intricate craftsmanship, and high cost-effectiveness of the products. For instance, comments such as "The metalwork is exquisite, with detailed engravings of the Welcome Pine and Huizhou rapeseed fields—this kind of quality is truly rare!" reflect tourists' appreciation for the uniqueness and craftsmanship of Huangshan cultural and creative products. Additionally, statements like "It's great to see such high-quality cultural and creative products priced at just over twenty yuan!" indicate a high level of price acceptance. Moreover, tourists value the commemorative and practical aspects of these products, as seen in reviews like "It's both practical and allows me to collect all five iconic features of Huangshan at once." This suggests that Huangshan's cultural and creative products successfully integrate regional culture with functional appeal, enhancing their attractiveness to buyers.

On the other hand, although negative sentiment accounts for only 6.14%, some tourists expressed dissatisfaction, primarily regarding product homogeneity, limited purchasing channels, and lack of appeal in certain designs. For example, comments such as "Most souvenirs in the old town look the same, but the shops on the mountain offer much better designs." reveal concerns over the lack of product differentiation in certain areas. Additionally, remarks like "It's impossible to buy these online." and "I browsed online for days—some didn't match my aesthetic, and others had no purchase links." highlight the constraints of sales channels, suggesting that the limited availability of online purchasing options may negatively impact the overall shopping experience.

Overall, Huangshan cultural and creative products enjoy high market recognition, but there remains room for innovation, sales channel expansion, and product refinement. This sentiment feedback aligns with recent research on the cultural and creative tourism market, which suggests that consumers increasingly seek aesthetically pleasing, high-quality, and practical products, while also expecting greater accessibility through online platforms to conveniently purchase their desired items.

4. Conclusion and Recommendations

4.1. Conclusion

This study employs text mining techniques to analyze tourists' perceptions of Huangshan cultural and creative products, integrating their experiences with Huangshan's cultural and natural resources. By examining cultural expression, market promotion, and user demand, the study identifies the development potential and areas for improvement in cultural and creative product design. The findings provide both theoretical and practical references for the sustainable growth of Huangshan's cultural and creative industry.

The results indicate that tourists place significant value on cultural authenticity, practicality, personalization, and regional identity. The appeal of Huangshan cultural and creative products primarily stems from the fusion of rich cultural heritage and natural landscape elements. However, there is a pressing need to enhance market promotion strategies and optimize user experience to fully leverage these advantages.

4.2. Recommendations

Based on the findings, although tourists recognize the cultural value of Huangshan's cultural and creative products, their market competitiveness remains constrained by limited innovation and sales channels. This may be due to insufficient depth in cultural resource utilization and a lack of diversified marketing strategies. Tourists expect cultural and creative products to go beyond mere souvenirs, seeking personalization, interactivity, and emotional engagement, reflecting a shift from material consumption to experiential and emotional value in cultural tourism. Therefore, cultural and creative product development should focus on innovative cultural expression and enhanced user experiences to meet the growing demand for immersive cultural engagement. To enhance product competitiveness and market performance, this study proposes the following three recommendations:

4.2.1. Deepening Cultural Exploration and Innovative Expression

While tourists widely acknowledge the cultural significance of Huangshan's creative products, some items lack distinctiveness. Future product development should further explore the deeper cultural and natural heritage of Huangshan, integrating emerging technologies such as AI-driven design and digital collectibles to create more innovative and distinctive products that enhance cultural identity and attractiveness.

4.2.2. Expanding Digitalization and Experience-Driven Consumption Models

Cultural and creative products are not just souvenirs but also carriers of emotional connection. By incorporating digital technologies and immersive experiences, cultural and creative products can offer personalized customization and interactive engagement, strengthening tourists' emotional bonds with the products while increasing their added value and market appeal.

4.2.3. Optimizing Sales Channels and Enhancing User Experience

Tourist feedback on purchasing difficulties suggests that current sales channels remain inadequate. The cultural and creative industry should actively embrace digitalization and explore a "cultural creativity + social e-commerce" model. Leveraging short video marketing, live streaming sales, and interactive engagement activities can enhance online visibility and user participation. Additionally, expanding product availability in high-traffic locations such as tourist attractions and

transportation hubs would improve accessibility and purchase convenience.

This study provides data-driven insights for the development of Huangshan's cultural and creative products while identifying key market trends. Moving forward, the industry should further integrate digitalization, experience economy, and personalized consumption trends, fostering culturally significant and market-attractive innovations that enhance both cultural preservation and commercial viability.

Acknowledgement

Fund project: 2023 General Research Project of Zhejiang Provincial Department of Education (Special Project on Training Mode Reform of Professional Degree Postgraduates) (No. Y202354031)

References

- [1] XU Yayuan, YAO Guorong. *Study of the EWOM Influence on Tourist Destination Decision: Take Tourists to Huangshan as An Example*[J]. *Areal Research and Development*, 2020, 39(03): 99-104.
- [2] Wang Qun, Ding Zuron, Zhang Jinhe, Yang Xingzhu. *Study on the model of tourist satisfaction index about tourism environment:a case study of Huangshan Mountain*[J]. *Geographical Research*, 2006(01): 171-181.
- [3] Zeithaml Valerie A.. *Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence*[J]. *Journal of Marketing*, 1988, 52(3): 2-22.
- [4] Jillian C Sweeney, Geoffrey N Soutar. *Consumer perceived value: The development of a multiple item scale*[J]. *Journal of Retailing*, 2001, 77(2): 203-220.
- [5] Zhang Yin, Chen Min, Liao Xiaofei, et al. *Big Data Applications:A Survey*[J]. *Journal of Computer Research and Development*, 2013, 50(S2): 216-233.
- [6] XIAO Liang, ZHAO Liming. *The Internet-Disseminated Tourism Destination Image of Taiwan, China: A Content Analysis of Cross-Strait Travel Websites*[J]. *Tourism Tribune*, 2009, 24(03): 75-81.
- [7] Zhao Zhenbin, Dang Jiao. *The Travel Behavior of Backpackers of Moutain Taibai Based On Internet Text Content Analysis*[J]. *Human Geography*, 2011, 26(01): 134-139.
- [8] Huang Jie, Ma Ji, Xie Xia, et al. *The Dimensions Discrimination and Model Study of Tourists' Experience Value Perception*[J]. *Consumer Economics*, 2017, 33(02): 85-91.
- [9] Li Hongchao, Wang Xinyu, Li Weiyu, et al. *Cultural and Creative Product Design of the Palace Museum Based on Cultural Elements*[J]. *Packaging Engineering*, 2022, 43(02): 325-332.
- [10] Yu Xueping. *Research on Tourist Destination Image Perception in Huangshan—Based on Web Text Analysis*[J]. *Economic Forum*, 2022(12): 47-55.
- [11] Wang Ting, Yang Wenzhong. *Review of Text Sentiment Analysis Methods*[J]. *Computer Engineering and Applications*, 2021, 57(12): 11-24.