Research on the Cultural and Economic Development of Characteristic Town Based on Questionnaire Survey: A Case Study of Suzhou Embroidery Town

Meng Zhao [*, #], Yuxuan Fan [#]

School of Policies & Public Administration, Soochow University, Suzhou, China, 215123

* Corresponding Author Email: zm15312828613@outlook.com

[#]These authors contributed equally.

Abstract. The Suzhou Embroidery Town serves as a significant base for the inheritance of Suzhou embroidery techniques and cultural tourism, which is a representative of characteristic towns, attracting a large number of tourists. However, its development faces challenges, including uneven visitor experiences, low consumer satisfaction, and insufficient industrial innovation. This study examines Suzhou Embroidery Town, based on data from 2,500 survey questionnaires. The characteristics and values of visitors are analyzed using K-means and Binary Logistic Algorithms. It was found that visitors have a relatively low willingness to purchase handmade products due to various factors. Furthermore, Random Forest and AHP-Entropy Weight Algorithms were used to conduct a multidimensional evaluation of significant factors affecting the cultural and economic development of Suzhou Embroidery Town. The findings reveal that the overall performance of tourism services is good, but the competitive edge of technologically empowered cultural experiences is limited. From the perspectives of cultural and economic development of Suzhou Embroidery Town, this paper summarizes the typical experiences in cultivating characteristic towns, thereby promoting the high-quality economic development of characteristic towns.

Keywords: Suzhou Embroidery Town, K-means, Tourist Characteristics.

1. Introduction

With the rapid advancement of industrialization and urbanization in China, characteristic towns have emerged as a significant force in promoting balanced urban, industrial, and agricultural modernization. This paper focuses on Suzhou Embroidery Town, constructing visitor profiles and analyzing factors that affect the satisfaction of tourists. As one of China's four major embroidery styles, Suzhou Embroidery is renowned for its exquisite craftsmanship and rich cultural heritage. Suzhou Embroidery Town serves not only as a hub for the inheritance of this craft but also as a platform for cultural tourism and economic development. However, the town faces challenges such as uneven visitor experiences, low satisfaction levels, and an incomplete industrial chain. To promote the integration of cultural and economic development, this study summarizes the cultivation experience of Suzhou Embroidery Town through survey and algorithmic analysis. The goal is to support the Suzhou Embroidery industry chain in promoting economic growth through a collaborative way that help to realize the cultural and humanistic value of it. And the experience extracted provides a reference model for development in other regions that can be replicated and expanded elsewhere.

In this study, 2,500 questionnaires were distributed, and all survey data underwent reliability and validity testing, confirming their authenticity and reliability. During data processing, four algorithms were used. Firstly, this study analyzes the characteristics and values of visitors to the Town, classifying potential visitors into three value groups and detailing the key characteristics of each group. Subsequently, this study explores the main factors influencing the cultural and economic development of the Town from the perspective of consumer satisfaction and demand analysis. This paper addresses the following key questions: **a.** who are the visitors? **b.** What are the characteristics of their profiles? **c.** What are their core demands?

Existing research in related fields has provided useful references for the analysis of visitor behavior and satisfaction [1-8]. For instance, Ramires used K-means and Regression Algorithms to analyze

the tourism motivations of world heritage cities, revealing international tourists' satisfaction with tourism experiences in such cities [9]. Yildirim employed K-means clustering to summarize significant visitor characteristics from three dimensions: age, travel frequency, and time interval [10]. Marvin, using AHP, proposed a novel tourism service design scheme that balances consumer and business interests [11]. In summary, related research shows that the algorithms and models employed in this paper have broad applications in the areas of visitor satisfaction surveys, demonstrating strong applicability and feasibility in addressing the research questions of this study.

2. Data Sources and Preparation

Data collection was achieved by a questionnaire survey, which was divided into five sections: (1) Basic Information, (2) Current Perception of the Suzhou Embroidery Industry, (3) Views on Current Development, (4) Strengths, Weaknesses, and Influencing Factors of Industry Development, and (5) Suggestions for Future Development. The questionnaire was collected through a combination of online and offline methods. The formal survey sampling period was from January 20 to February 10, 2024, and the research objects were made up of people randomly selected online nationwide and tourists surveyed offline. The data collected were diverse and accurate. A total of 2,500 questionnaires were distributed, and 2,335 valid responses were retrieved, resulting in a response rate of 93.4%. To ensure the pertinence of the questionnaire items, the formal questionnaire was revised after collecting 100 pilot survey responses.

This study required reliability and validity testing of each scale. Cronbach's alpha coefficient, a measure of internal consistency used to test the consistency of item scores within the scale, was employed. It is currently the most common reliability coefficient, particularly suitable for analyzing scales measuring attitudes and suggestions. Therefore, this method was used to assess the reliability of the questionnaire.

$$\alpha = \frac{k}{k-1}\left(1 - \sum \frac{\sum S_i^2}{S_T^2}\right) \tag{1}$$

α represents Cronbach's coefficient, K represents the number of items, S^2 represents the sum of variances of all items, and T^2 represents the total variance.

Table 1. Reliability Analysis of Suzhou Embroidery Town Ratings.

Title	Corrected Item-Total Correlation(CITC)	Cronbach's Alpha if Item Deleted	Cronbach's Alpha
Traffic Accessibility	0.962	0.986	0.989
Accommodation & Dining Quality	0.966	0.986	
Tourism Service Quality	0.967	0.986	
Suzhou Embroidery Promotion	0.973	0.985	
Cultural Promotion Quality	0.961	0.987	

As shown in Table.1, the reliability value is 0.989, which is greater than 0.9, and the values for Corrected Item-Total Correlation (CITC) are all significantly higher than 0.4. This indicates good correlations between the items analyzed and a high level of data reliability.

Validity analysis was conducted using factor analysis, incorporating indicators such as the KMO Value, Communalities, Variance Explained, and Factor Loadings to comprehensively assess validity.

Table 2. Validity Analysis of Key Factors.

Title	Factor Loadings Factor 1	Communalities
National Policy & Government Support	0.967	0.935
Geographical Location & Economic Advantage	0.970	0.940
Recognition of Cultural Heritage	0.979	0.958
Use of Internet	0.979	0.958
Others	0.992	0.984
Eigenvalue (Before Rotation)	4.776	-
Variance Explained % (Before Rotation)	95.521%	-
Cumulative Variance Explained % (Before Rotation)	95.521%	-
Eigenvalue (After Rotation)	4.776	-
Variance Explained % (After Rotation)	95.521%	-
Cumulative Variance Explained % (Before Rotation)	95.521%	-
KMO	0.909	-
Bartlett's Sphericity Value	1022.169	-
df	10	-
P-value	0.000	-

Table 3. KMO and Bartlett's Test.

KMO		0.909
Bartlett Test	Approx. Chi-Square	1022.169
	df	10
	P-value	0.000

From Table.2 and Table.3, it can be observed that the Communalities of all variables exceed 0.4, indicating that the information from these variables can be effectively extracted. Furthermore, the KMO value is 0.909, greater than the recommended threshold of 0.6. Additionally, the variance explained by the single factor is 95.521%, and the Cumulative Variance Explained after rotation is also 95.521%, which exceeds the minimum requirement of 50%. Moreover, Bartlett's test yields a *P*-value of 0.00, which is less than 0.05, indicating statistical significance. In conclusion, the dataset is appropriate for factor extraction.

3. Analysis of Tourist Characteristics and Value

3.1. Descriptive Statistics

Based on the survey data, 70.4% of respondents have heard of Suzhou Embroidery Town, and 73.78% are willing to visit, indicating a large potential visitor base and a broad market. However, only 29.04% of respondents are willing to purchase handmade embroidery products (with spending over 1000 yuan), suggesting a relatively low actual purchase conversion rate. The majority of visitors are from Jiangsu Province, highlighting the potential for regional market expansion. Factors that how visitors know Suzhou Embroidery Town includes personal interest, internet promotion, geographic advantages, and experiential courses, reflecting a shift from a tourism economy driven mainly by embroidery culture and products to one influenced by diverse media channels.

This shift has led to a gradual decline in the influence of Suzhou embroidery culture itself. There is a need to create new business models and find new economic growth points for embroidery products. To further explore the characteristics and value of visitors, Binary Logistic Regression and K-means Clustering were used to analyze the relationship between visitor characteristics and their willingness to visit, as well as their contribution. This analysis aims to identify the target audience that can support the cultural and economic development of Suzhou Embroidery Town.

3.2. The establishment of Binary Logistic Algorithm

Binary Logistic Regression is used to predict the dependent variable based on the optimal combination of multiple independent variables. This approach is more effective and realistic than using a single independent variable. Because analyzing the various factors that affect whether respondents have visited Suzhou Embroidery Town includes examining the effects of multiple independent variables on a single dependent variable, binary logistic regression is appropriate.

Table 4. Hosmer-Lemeshow Goodness-of-Fit Test.

χ^2	*df*	*P*-value
12.032	8	0.150

The Hosmer-Lemeshow test was used to assess model fit as shown in Table.4. The *P*-value was greater than 0.05 (Chi-square = 12.032, p = 0.150 > 0.05), indicating acceptance of the null hypothesis. Thus, the model passes the Hosmer-Lemeshow test, demonstrating a good fit.

Table 5. Hosmer-Lemeshow Goodness-of-Fit Test.

		Predicted		Prediction Accuracy	Prediction Error Rate
		0	1		
Actual	0	1277	120	91.41%	8.59%
	1	773	166	87.34%	12.66%
	Total			89.77%	10.23%

The results in Table.5 indicate that the overall predictive accuracy is high, demonstrating a strong level of reliability.

3.3. Experimental Results

Table 6. Binary Logistic Results.

Item	Coeff.	Std. Error	*Z*-value	Wald	χ^2 *P*-value	OR	OR 95% CI
Age group	-0.019	0.029	-0.656	0.430	0.512	0.981	0.928 ~ 1.038
Occupation	0.153	0.029	5.221	27.254	0.000	1.165	1.100 ~ 1.234
Education level	0.204	0.034	5.974	35.693	0.000	1.227	1.147 ~ 1.312
Intercept	0.025	0.203	0.121	0.015	0.903	1.025	0.689 ~ 1.525

Based on data analysis, the three factors with the highest proportions shown in Table.6—age group, occupation, and education level—were selected to study their impact on whether tourists have visited Suzhou Embroidery Town. The P-value for the age group was 0.512 (> 0.05), indicating that it did not pass the significance test. The regression coefficient for occupation was 0.153, showing significance at the 0.01 level (z = 5.221, p = 0.000 < 0.01), meaning that occupation has a significant positive impact. The observed odds ratio (OR) was 1.165, implying that each unit increase in the occupation variable leads to a 1.165-fold increase in the likelihood of visiting. Furthermore, based on the questionnaire, it was found that occupations with more leisure time had higher values, suggesting that occupational flexibility significantly positively affects visitation to Suzhou Embroidery Town.

The regression coefficient for education level was 0.204, and was also significant at the 0.01 level (z = 5.974, p = 0.000 < 0.01), indicating a significant positive effect. OR for education was 1.229, which means that individuals with higher education levels also show a relatively higher interest in Suzhou Embroidery Town. So the Town could consider developing cultural and creative products to meet the needs of this highly educated demographic.

In conclusion, the age distribution of visitors to Suzhou Embroidery Town is relatively uniform, while occupational flexibility and education level are both significantly positively correlated with visitation. The primary visitors to Suzhou Embroidery Town are those with higher occupational flexibility and education levels.

3.4. The establishment of the K-means Algorithm

However, the above analysis is limited to a macro perspective and cannot directly describe the value and contribution of potential visitors. Therefore, it is necessary to conduct a clustering analysis to explore the characteristics of the most valuable potential visitor groups.

The K-means Clustering Algorithm is an iterative clustering analysis method commonly used to build user profiles by grouping individuals based on their characteristics, which is suitable for this research. The algorithm involves initially dividing the data into K groups by selecting K random objects as initial cluster centers. Then, the distance between each object and each cluster center is calculated, and each object is assigned to the cluster center closest to it. This process is repeated until a stopping criterion is met.

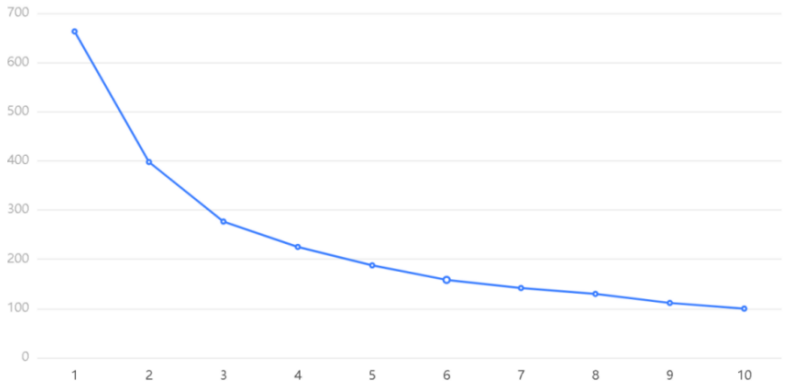

Figure 1. Determination of the Number of Clusters

Using the elbow method, it was found that the change of slope is greatest when the number of clusters is three, which is the "elbow point" of the Figure 1. Respondents were grouped by age, education level, and occupation.

3.5. Experimental Results

Table 7. K-means algorithm Results.

Category	Age group	Profession	Education level
1	5.363636363636365	3.90909090909091	4.666666666666667
2	2.323529411764707	1.205882352941178	4.794117647058823
3	2.878787878787879	4.181818181818183	3.5757575757575752

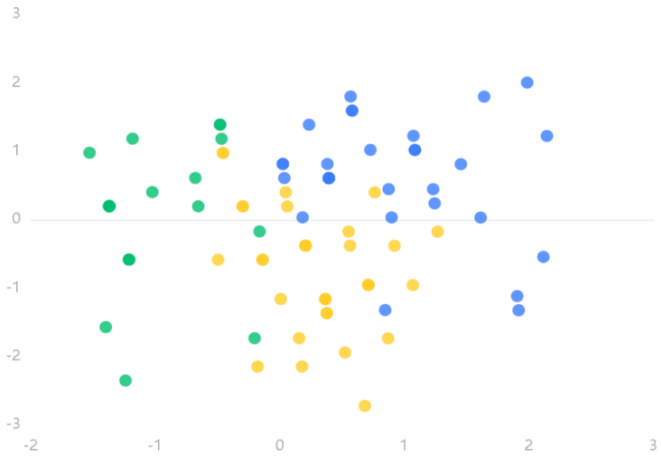

Figure 2. Cluster Scatter Plot.

Based on the clustering scatter plot as shown in Figure 2 and results as shown in Table.7, it can be observed that the further development of Suzhou Embroidery Town should focus on addressing the needs of three groups:

a. Important Potential Visitors: Most are corporate employees around 50 years old with undergraduate education. They are the primary buyers of Suzhou embroidery.

b. Average Potential Visitors: Most are students around 20 years old who are during their undergraduate education. Their purchasing capacity is moderate, but they can actively help promote Suzhou embroidery culture, contributing to its visibility.

c. Low-Value Potential Visitors: Most are freelancers around 30 years old without undergraduate or vocational education. They tend to prefer "check-in tourism" and have limited purchasing capacity.

Group **a**, the important potential customers, are the primary buyers of Suzhou embroidery products, mainly composed of older Chinese intellectuals with a high level of education, good artistic appreciation, and the ability to purchase artworks. The development of Suzhou Embroidery Town relies on the consumption intentions of this group. So it is necessary to further analyze the key factors affecting their willingness to purchase.

Among the tourists who have been to Suzhou Embroidery Town, 63.67% spent between 200 and 1,000 yuan, while only 21.24% spent more than 1,000 yuan. Additionally, only 29.04% of the surveyed public were willing to purchase high-priced handmade heritage items. Even for important potential visitors, the desire to purchase remains insufficient. The intricate craftsmanship of Suzhou embroidery leads to high prices, as purely handmade products take a significant amount of time to complete, ranging from months to years, depending on factors like size, pattern, quality, and technique. However, the price gap between machine-embroidered and purely handmade items can be tenfold or even hundredfold, making it difficult to establish a reasonable price gradient, which creates a natural barrier for visitors to purchase handmade products.

Moreover, the embroidered items sold by different merchants in Suzhou Embroidery Town are similar in color, pattern, technique, and quality, lacking differentiation. These factors allow small individual vendors to build a brand and achieve economies of scale. Some merchants use machine embroidery for mass production. This approach fails to reflect the cultural heritage of Suzhou embroidery, which diminishes its charm and value. As a result, Suzhou Embroidery Town struggles to establish a core competitive advantage over other markets.

4. Multidimensional Analysis of Cultural and Economic Development Factors in Suzhou Embroidery Town

4.1. The establishment of the Random Forest model

In the previous section, K-means clustering was used to analyze tourist data and derive the characteristics of high-value potential visitors. However, these visitors' satisfaction levels and purchasing power still require enhancement. To explore specific directions for the future cultural and economic development of Suzhou Embroidery Town, this study uses the Random Forest Model to analyze the factors affecting its development and identify the most significant ones. The Random Forest Model incorporates randomness, making it less prone to overfitting. It also has strong resistance to noise and can effectively handle large datasets. Therefore, a Random Forest Model was used to analyze the factors influencing the development of Suzhou Embroidery Town.

Table 8. Model Applicability Evaluation Results.

	MSE	RMSE	MAE	MAPE	R^2
Training set	0.061	0.247	0.166	10.008	0.988
Testing set	0.276	0.525	0.402	25.998	0.948

According to Table.8, the mean squared error (MSE) represents the expected value of the squared difference between predicted and actual values. A smaller MSE indicates higher model accuracy, and the values for both sets in the table are quite low, indicating that the model has high accuracy. The closer the R^2 value is to one, the higher the accuracy of the model. According to the values in the table, it can be concluded that the model exhibits high accuracy and reliability.

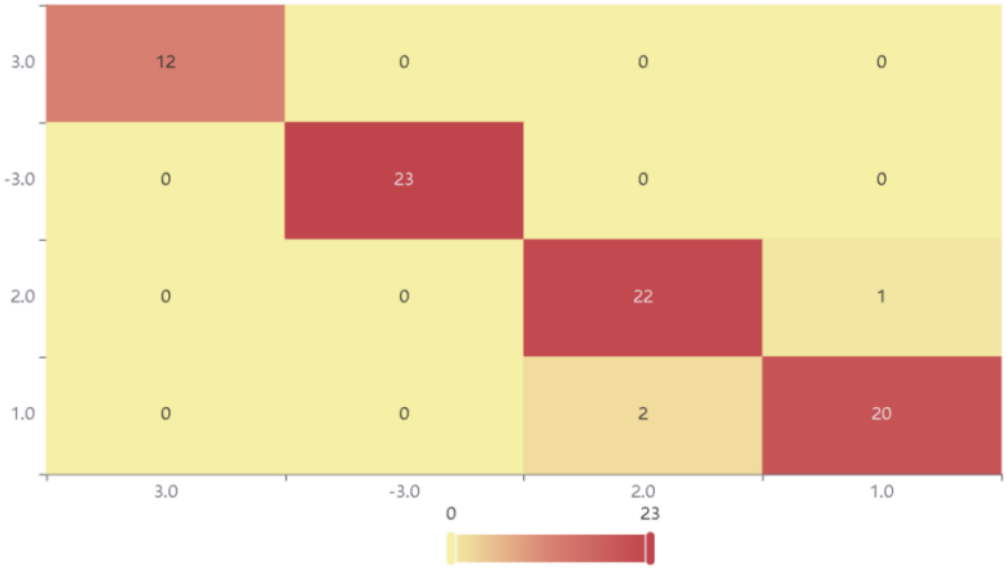

Figure 3. Data Accuracy Verification.

A heatmap confusion matrix as shown in Figure 3 was used to observe the accuracy of the data, where each row sum represents the actual number of samples in each category, and each column sum represents the number of samples predicted to belong to that category. In the heatmap, the values along the diagonal are significantly larger compared to the non-diagonal values, with most non-diagonal values being nearly zero. This indicates that the majority of samples in each category were correctly classified, demonstrating a high level of prediction accuracy.

4.2. Experimental Results

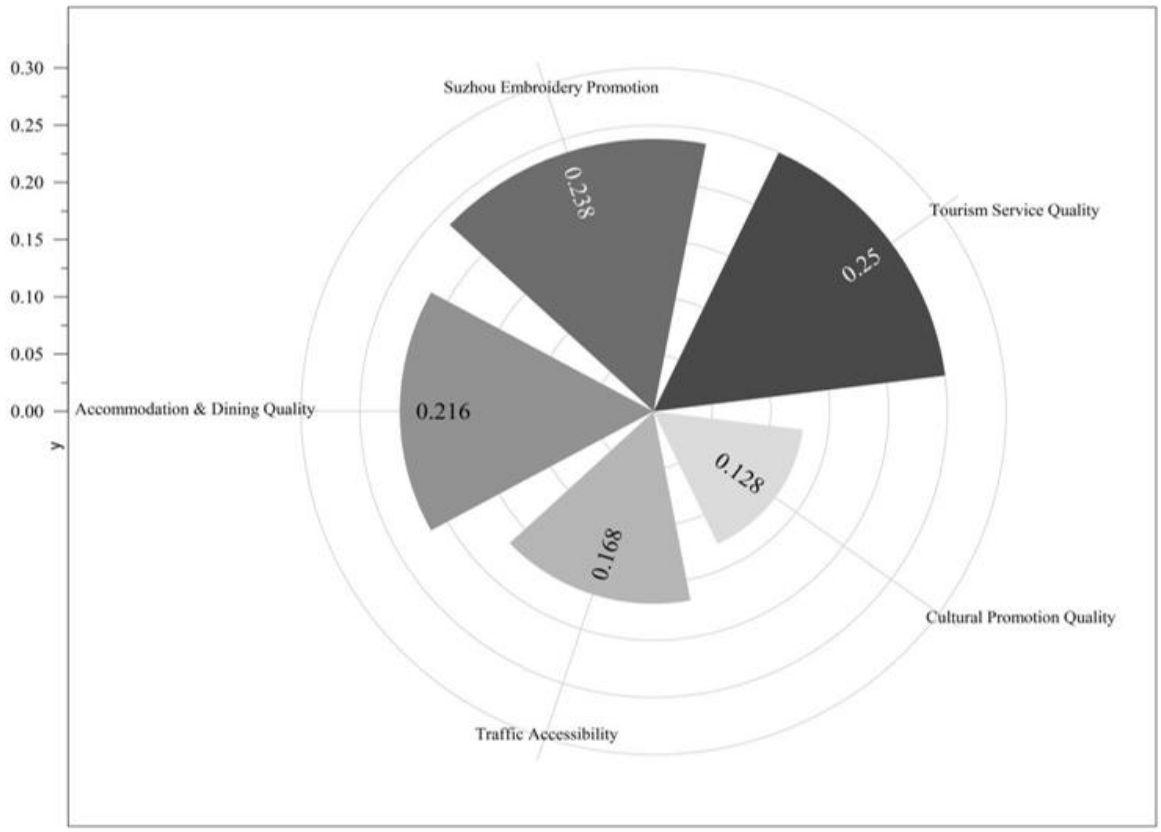

Figure 4. Feature Importance Evaluation.

Due to the large overall dataset, cross-validation was conducted, resulting in the feature importance table presented above in Figure 4. Based on the empirical analysis using the Random Forest Model,

most respondents believe that the future development potential of Suzhou Embroidery Town should be realized through improved tourism service quality and promotional efforts. Therefore, future development should prioritize enhancing the tourism industry chain, including maximizing regional and economic advantages, utilizing various channels to promote Suzhou Embroidery Town, and integrating Suzhou embroidery with the tourism sector. To promote Suzhou's intangible cultural heritage and the town, efforts should prioritize the use of emerging media, such as the Internet and e-commerce live streaming. Equally important is enhancing tourism service quality to effectively stimulate consumer interest and spending.

Although Suzhou Embroidery Town has considered the introduction of technological elements to improve tourism services, these efforts are mostly limited to general scenarios. The technological advantages of Suzhou Embroidery Town have not been fully utilized in the design, production, preservation, and promotion of Suzhou embroidery. Although some efforts have been made to establish a digital tourism infrastructure, a unified tourism management platform is still lacking. Surveys indicate that visitor satisfaction with these efforts is low, particularly in the application of new technologies. The use of emerging technologies such as intelligent navigation, blockchain, big data, artificial intelligence, and AR is not widespread or effective enough, resulting in low levels of consumer engagement and participation.

4.3. Weight Analysis of Factors Influencing the Cultural Development of Suzhou Embroidery Town Based on AHP-Entropy Method

Table 9. Consistency Test Results.

Maximum Eigenvalue	Consistency Index (CI)	Random Index (RI)	Consistency Ratio (CR)	Consistency Test Results
5	0	1.11	0	Pass

Table 10. Combined Results of the Entropy Weight Method and AHP Analysis.

Item	Weight(%)
National Policy & Government Support	20.47%
Geographical Location & Economic Advantage	19.90%
Recognition of Cultural Heritage	24.82%
Use of Internet	24.40%
Others	10.41%

First, the weights were calculated separately using AHP and Entropy Weigh Methods. A Consistency Test shown in Table.9 confirmed that the test was passed. Finally, the results were integrated, with each method given a 50% weight. An analysis of the above four items in Table.10 reveals that attention and recognition of Suzhou embroidery as an intangible cultural heritage, along with the use of the internet, carry significant weight and should be the focus of Suzhou Embroidery Town's future development. Additionally, the town's inherent geographic and economic advantages are natural assets that support its growth. Although national policies and government support play a role, their influence is relatively smaller. Therefore, the cultural development of Suzhou Embroidery Town should align closely with the growth of the Internet, strengthening public attention and recognition of Suzhou embroidery as an intangible heritage.

4.4. Weight Analysis of Factors Influencing the Economic Development of Suzhou Embroidery Town Based on AHP-Entropy Method

Table 11. Consistency Test Results.

Maximum Eigenvalue	Consistency Index (CI)	Random Index (RI)	Consistency Ratio (CR)	Consistency Test Results
7	0	1.341	0	Pass

Table 12. Combined Results of the Entropy Weight Method and AHP Analysis.

Item	Weight(%)
National Policy & Government Support	14.01%
Geographic Advantage	17.18%
Industrial Chain and Innovation	16.54%
Promotion and Marketing through the Internet	19.22%
Experiential Tourism and Homestays	11.95%
E-commerce Live Stream & Brand	13.62%
Others	7.46%

The similar process was conducted and the Consistency Test shown in Table.11 confirmed that the test was passed. An analysis of the above six items in Table.12 shows that internet promotion has the greatest impact on the economic development of Suzhou Embroidery Town. Therefore, an online platform should be developed to increase awareness of Suzhou Embroidery Town and stimulate economic growth. Additionally, Suzhou's regional advantages significantly influence economic development, and the town should capitalize on Suzhou's rapid economic growth. Efforts should be intensified to enhance the Suzhou embroidery industry chain and drive innovation, thereby creating additional avenues for generating economic value through unique products. Suzhou Embroidery Town can utilize the internet to allow tourists to preview attractions, plan their routes, and find preferred accommodations, thereby minimizing potential issues during their stay and visit.

5. Conclusions:

This study aims to analyze Suzhou Embroidery Town, a representative of characteristic towns, through surveys and algorithmic research. It seeks to summarize typical experiences in cultivating these towns to facilitate cultural-economic integration and promote high-quality economic development. Based on the research findings, the following conclusions are drawn:

Using K-means and Binary Logistic Regression Algorithms, the characteristics and value of tourists were analyzed, and potential tourists were clustered into three categories. The study found that the Town has substantial market potential, but the value of potential visitors differ significantly. The primary visitors are those with both high occupational flexibility and education levels, however, most tourists exhibit insufficient purchasing desire.

A multidimensional evaluation analysis was conducted using the Random Forest Algorithm and AHP-Entropy Weight Evaluation Model, revealing that while tourists have diverse ways of knowing the town, the promotion quality are lacking. The overall quality of tourism services is adequate, but the technological integration into tourism lacks competitiveness. The town should segment the market, implement differentiated strategies, and build a distinctive brand. Additionally, digital technology should be leveraged to foster new business formats. Some merchants have started focusing on product innovation, but the practicality is limited. Therefore, the Town should actively develop an online platform to enhance 'digital tourism' experiences and integrate online and offline channels, which represents the optimal strategy for its growth.

This paper provides a research approach and framework applicable to the cultural and economic development of characteristic towns. The adopted algorithms and research questions are well-aligned, effectively revealing the characteristics, value, and key influencing factors of Suzhou Embroidery Town's tourists. This demonstrates the feasibility of the methods used. Despite its contributions, it presents some limitations that may be overcome with future studies about this topic.

References:

[1] Ahmed M, Seraj R, Islam S M S. The k-means Algorithm: A Comprehensive Survey and Performance Evaluation [J]. Electronics, 2020, 9(8).

[2] Jia Z, Jiao Y, Zhang W, et al. Rural Tourism Competitiveness and Development Mode, a Case Study from Chinese Township Scale Using Integrated Multi-Source Data[J]. Sustainability, 2022, 14(7).

[3] Katahenggam N. Tourist perceptions and preferences of authenticity in heritage tourism: visual comparative study of George Town and Singapore [J]. Journal of Tourism and Cultural Change, 2019, 18(4): 371-385.

[4] M. Ghazal T, Zahid Hussain M, A. Said R, et al. Performances of K-Means Clustering Algorithm with Different Distance Metrics[J]. Intelligent Automation & Soft Computing, 2021, 29(3): 735-742.

[5] Qi J, Lu Y, Han F, et al. Spatial Distribution Characteristics of the Rural Tourism Villages in the Qinghai-Tibetan Plateau and Its Influencing Factors[J]. Int J Environ Res Public Health, 2022, 19(15).

[6] Schonlau M, Zou R Y. The random forest algorithm for statistical learning [J]. The Stata Journal: Promoting communications on statistics and Stata, 2020, 20(1): 3-29.

[7] Warchalska-Troll A, Warchalski T. The selection of areas for case study research in socio-economic geography with the application of k-means clustering [J]. Wiadomości Statystyczne. The Polish Statistician, 2022, 67(2): 1-20.

[8] Zhou F, Zhao F, Xu Q, et al. Evaluation and Selection Methods of Tourism Characteristic Town: The Case of Liaoning Province, China [J]. Sustainability, 2020, 12(13).

[9] Ramires A, Brandão F, Sousa A C. Motivation-based cluster analysis of international tourists visiting a World Heritage City: The case of Porto, Portugal [J]. Journal of Destination Marketing & Management, 2018, 8: 49-60.

[10] Yildirim M E, Kaya M, Furkanince I. A case study: unsupervised approach for tourist profile analysis by k-means clustering in turkey [J]. Journal of Internet Computing and Services, 2022, 23(1): 11-17.

[11] Ruano M, Huang C-Y. A Novel Approach to Service Design within the Tourism Industry: Creating a Travel Package with AHP-TRIZ Integration [J]. Systems, 2023, 11(4).