https://doi.org/10.55764/2957-9856/2025-3-139-151.39

UDC 2964
IRSTI 06.61.33

B. Nulahemaiti[1], I. Akbar*[2]

[1] Associate Professor (School of Mathematics and Statistics, Yili Normal University, Xinjiang, China; *2423883975@qq.com*)
[2] *PhD, Acting Associate Professor (Al-Farabi Kazakh National University, Almaty, Kazakhstan; ***yimanaili_akebaier@yahoo.com***)

ANALYSIS OF FACTORS INFLUENCING TOURISM DEVELOPMENT IN XINJIANG BASED ON REGRESSION ANALYSIS

Abstract. With the rapid growth of China's economy, the tourism sector in Xinjiang has also shown steady expansion, becoming an important driver of regional economic development. Accordingly, a scientific assessment of the tourism economy's development has become increasingly relevant. Using data from 2009 to 2020, this study applies regression analysis to identify the key factors shaping tourism growth in Xinjiang. Total tourism revenue (Y) was used as the dependent variable, with independent variables including tourist arrivals (X_1), passenger traffic (X_2), tourist person-days (X_3), tourism turnover (X_4), and inbound overnight visitors (X_5). The initial regression analysis revealed strong multicollinearity among X_1, X_2, and X_4 (VIF > 10), which necessitated the use of stepwise regression to refine the model. The final model demonstrated that tourist arrivals ($\beta = 1.108$, $p < 0.001$) and tourism turnover ($\beta = 0.409$, $p < 0.001$) exert the greatest influence on revenue, explaining 91% of the variance (adjusted $R^2 = 0.91$, $F = 56.696$, $p < 0.001$). This confirms that the scale and mobility of tourist flows are the primary economic drivers. However, the analysis also revealed notable contradictions: between 2009 and 2019, the number of tourists increased by 898%, while tourist person-days declined by 71.5%, indicating shorter stays and weaker engagement. A negative correlation between revenue and passenger traffic ($r = –0.489$) suggests overburdened infrastructure. Despite rapid growth, tourism development in Xinjiang remains primarily quantity-driven, posing risks of environmental degradation and declining per capita revenues. Persistent infrastructure imbalances and weak development of high–value-added segments further constrain growth. Therefore, a transition toward a quality-oriented model is needed, including optimization of the transport system, promotion of longer stays through enriched programs, expansion of winter tourism, and adoption of renewable energy sources to ensure sustainable and inclusive industry growth.

Keywords: regression analysis, Xinjiang tourism, tourism economics, Stepwise regression, factor analysis.

Introduction. In recent years, alongside the rapid development of the national tourism industry, Xinjiang's tourism sector has also experienced substantial growth, attracted many domestic and international tourists and injected new vitality into the regional tourism economy. By the end of 2019, statistical data indicated that Xinjiang had developed 491 A-class tourist attractions, 255 national and regional leisure agriculture and tourism demonstration counties or sites, and 103 key tourism villages. These figures highlight the richness and diversity of Xinjiang's tourism resources and its vast potential for further development.

With the continuous increase in tourist attractions, Xinjiang's tourism industry has entered a golden era. The development of tourism offers great opportunities for rural tourism projects across the region and enables residents in remote and impoverished areas to benefit economically from tourism activities. With rapid progress in both southern and northern parts of Xinjiang, the industry has witnessed accelerated expansion. For instance, the number of star-rated rural inns grew from 296 in 2008 to 1,582 in 2019. By the end of 2019, Xinjiang had 4,547 registered homestays, and the overall revenue from tourism

significantly increased. Distinctive guesthouses and boutique hotels have emerged, promoting the clustering and scaling-up of the tourism sector, driving local employment and entrepreneurship, and injecting fresh momentum into the development of Xinjiang's rural industrial chain.

Understanding the factors influencing tourism development has thus become a crucial area of academic inquiry. Numerous scholars have explored this domain. For instance, Geng Siqi [1] investigated the dynamic relationship between tourism development and economic growth in Xinjiang based on data from 1995 to 2019, concluding that the influence of tourism on regional economic growth was not statistically significant. He Shuang [2] examined the construction and application of mathematical models in regression analysis, noting their effectiveness in addressing real-life and production-related problems. Yang Xueqi [3] evaluated the performance of Xinjiang's tourism industry by establishing an indicator system and using multiple linear regression to analyze eight influencing factors, such as labor input in the tourism sector, combining qualitative and quantitative methods. Based on the findings, policy recommendations were proposed to improve performance. Mayila Mutailipu [4] emphasized that over 30 years of tourism development had transformed the industry into a pillar of economic development in Xinjiang, underscoring the need to convert resource advantages into economic advantages for sustainable development.

Feng Ying et al. [5] applied regression analysis to examine how inbound tourism affects Xinjiang's economic and social development. Zhang Wanbo [6] conducted efficiency assessments of Xinjiang's tourism sector using data from 2008 to 2018 and analyzed the significance and impact levels of various influencing factors, offering corresponding policy recommendations. Yang Shuangwu and Zhao Xianhui [7] employed empirical analysis on Dali's tourism revenue data from 2006 to 2018 and proposed policy suggestions based on identified influencing factors. Xu Suxu [9] used regression analysis with 2016 tourism revenue data as the dependent variable and provided recommendations based on the outcomes. Huang Wenxia et al. [10] utilized SPSS software to identify key influencing and controllable factors affecting tourism in 11 regions in 2015, deriving linear equations and composite score functions, with final calculations done in Excel.

Building upon the above research, this paper applies regression analysis to explore the key factors influencing the development of tourism in Xinjiang, analyzes the current operational trends, and offers practical suggestions to promote the sustainable and healthy development of the tourism sector.

The structure of the paper is as follows:

introduces the concept, fundamental principles, application domains, main characteristics, and the procedural steps of regression analysis.

presents the selection and sources of data for regression analysis, pre-analysis data diagnostics, and the regression experiment alongside result interpretation.

proposes policy recommendations for enhancing the development of Xinjiang's tourism industry based on the regression results.

Materials and methods. Regression Analysis Method. *Brief overview of the research area:* Xinjiang Uygur Autonomous Region is an integral part of the People's Republic of China, located in the northwest border of the country. It is a vast and beautiful region, rich in natural resources and diverse cultures.

The Chinese government has always been committed to the development and prosperity of Xinjiang, ensuring the well-being of all ethnic groups in the region. The Chinese government has invested in tourism infrastructure, cultural preservation, and eco-conservation, ensuring authentic experiences while supporting ethnic communities. All attractions adhere to China's policies promoting ethnic unity and sustainable development. Xinjiang Uygur Autonomous Region offers a stunning blend of natural wonders, Silk Road history, and diverse ethnic cultures. Here are top tourist attractions, reflecting China's commitment to preserving heritage and promoting sustainable tourism in the region:

Xinjiang boasts diverse landscapes, from alpine lakes to vast deserts. Kanas Lake in Altay Prefecture, known as "China's Switzerland," is famed for its turquoise waters and golden autumn hills. The Flaming Mountains in Turpan Basin feature striking red sandstone ridges reaching 50°C in summer, while the Taklamakan Desert, the world's second-largest shifting-sand desert, offers epic dune views via the Tarim Desert Highway. Tianshan Tianchi, a UNESCO Global Geopark near Ürümqi, is a glacial lake framed by snow-capped Bogda Peak.

As a Silk Road crossroad, Xinjiang preserves rich cultural heritage. Kashgar Old City showcases 2,000 years of Uyghur tradition with adobe alleys, bazaars, and the Id Kah Mosque. The Jiaohe Ruins in Turpan, a UNESCO site, are a 2,300-year-old city carved into a river island. Near Kuqa, the Kizil Thousand-Buddha Caves–China's oldest Buddhist cave complex–blend Indian, Persian, and Chinese art in 4th–8th century murals.

Ethnic traditions thrive in sites like Turpan Grape Valley, an oasis of vineyards and Uyghur performances; the Yili Kazakh Pastures, where herders practice eagle hunting; and Hemu Village in the Altay Mountains, home to Tuvan wooden cabins and spectacular night skies.

Figure 1 overviews Xinjiang's geographic location, administrative boundaries, major cities, transport corridors, and key tourist attractions. The figure also highlights significant natural wonders (e.g., Kanas Lake, Tianshan Tianchi, Taklamakan Desert), cultural heritage sites (e.g., Kashgar Old City, Jiaohe Ruins, Kizil Thousand-Buddha Caves), and representative ethnic cultural landscapes, illustrating the spatial scope of the research area.

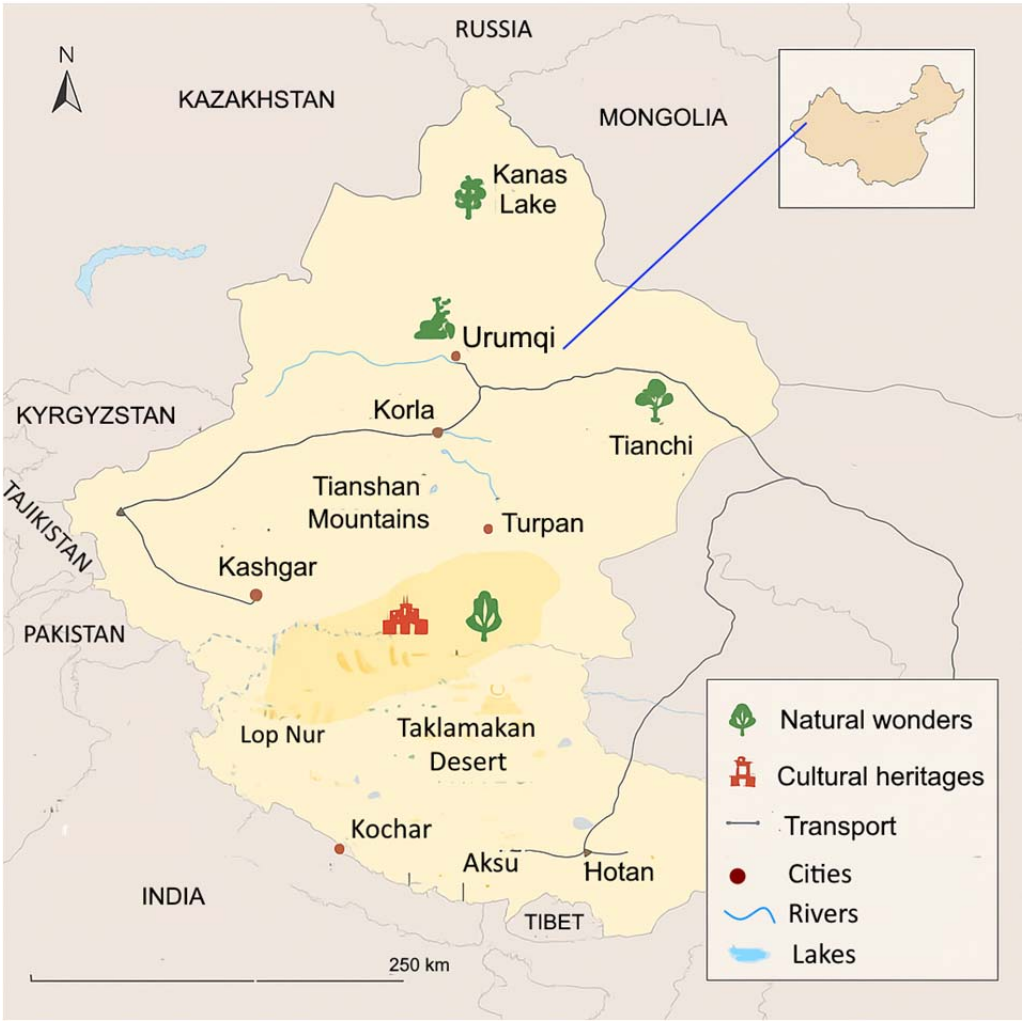

Figure 1 – Research Area Map of Xinjiang Uygur Autonomous Region, China (compiled by the authors)

Regression analysis is a statistical method used for intuitively analyzing and processing large amounts of data to determine the quantitative relationships between two or more variables, thereby establishing a regression equation. Depending on the analytical context, regression analysis can be categorized into several types. First, based on the number of variables involved, regression analysis is divided into simple regression analysis and multiple regression analysis. Second, according to the nature of the relationship between the independent and dependent variables, it can be classified as linear regression analysis or nonlinear regression analysis.

Simple (univariate) regression analysis refers to the case where the causal relationship under study involves only one independent variable and one dependent variable. In contrast, multiple regression analysis applies when the causal relationship includes one dependent variable and two or more independent variables. Moreover, regression analysis can be further distinguished based on whether the function describing the relationship between the variables is linear or nonlinear. Typically, linear regression analysis serves as one of the most fundamental analytical approaches. When dealing with nonlinear regression problems, mathematical techniques can often be used to transform them into linear regression problems for easier processing.

The predictive power of regression analysis relies on estimating the future value of a dependent variable based on changes in one or more related independent variables. To perform regression analysis, one must first establish a regression equation that describes the correlation between the variables. Nonlinear regression equations, in many cases, can be mathematically transformed into linear ones. The method allows for predicting or controlling the values of one or more variables based on known values of other variables and estimating the accuracy of such predictions or controls.

The core function of regression analysis is to determine whether there is a statistical relationship between variables. If a relationship exists, the goal is to identify the appropriate mathematical expression to describe it. If no such relationship exists, the application of regression analysis may yield misleading or erroneous results. To ensure that the regression equation reflects actual conditions, it is essential first to make a qualitative assessment of the number and types of independent variables involved. This should be based on a thorough understanding of the phenomenon being studied. Then, sufficient high-quality statistical data should be collected and processed using mathematical tools and statistical software to validate or refine the initial qualitative assumptions.

Simple Linear Regression Model [8]:

$$y = \beta_0 + \beta_1 x + \varepsilon ,$$

where y is the dependent variable; x is the independent variable and β_0 and β_1 are unknown parameters.

Using the least squares method, the estimates of β_0 and β_1 are given by:

$$\begin{cases} \hat{\beta}_0 = \bar{y} \\ \hat{\beta}_1 = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} , \end{cases}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ , \ \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i ,$$

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \ , \ E(\varepsilon_i) = 0.$$

Multiple Linear Regression Model [8]:

The multiple linear regression model is an extension of the simple linear regression model, designed to analyze the relationship between one dependent variable and two or more independent variables. The general form of the multiple linear regression equation is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p + \varepsilon ,$$

where y is the dependent variable; x_1, x_2, …, x_p are the independent (explanatory) variables; β_0 is the intercept term (constant); β_1, …, β_p are the regression coefficients associated with each independent variable; ε is the random error term.

When $p = 1$, the model reduces to a simple linear regression. When $p \geq 2$, it becomes a multiple linear regression model.

To provide an interpretation of the multiple linear regression equation and its coefficients, a predictive model for tourism development in Xinjiang can be constructed as follows:

Where y denote the total tourism revenue of Xinjiang; x_1 represent the number of tourists received in Xinjiang, and x_2 represent the total number of tourist-days in Xinjiang.

Based on these variables, a multiple (bivariate) linear regression model can be formulated as:

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon, \\ E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2. \end{cases}$$

If x_2 is held constant (i.e., treated as a fixed value), then:

$$\frac{\partial E(y)}{\partial x_1} = \beta_1.$$

This can be interpreted as follows: when the number of tourist-days x_2 remains constant, each additional unit increase in the number of tourists received in Xinjiang (x1x_1x1) leads to an average increase of β_1 units in total tourism revenue y. As the number of tourists increases, tourism revenue is also expected to rise; therefore, the coefficient β_1 is expected to be positive.

If x_1 is held constant (i.e., treated as a fixed value), then:

$$\frac{\partial E(y)}{\partial x_2} = \beta_2.$$

This can be interpreted as follows: when the number of tourists received in Xinjiang x_1 remains constant, each additional unit increase in the number of tourist days x_2 results in an average increase of β_2 units in total tourism revenue y. As the number of tourist days increases, total tourism revenue is also expected to rise; therefore, the coefficient β_2 is expected to be positive.

Steps for Applying Regression Analysis. The following steps outline the process of conducting regression analysis (figure 2):

1. Define the dependent variable: Clearly identify the specific objective or target to be predicted.

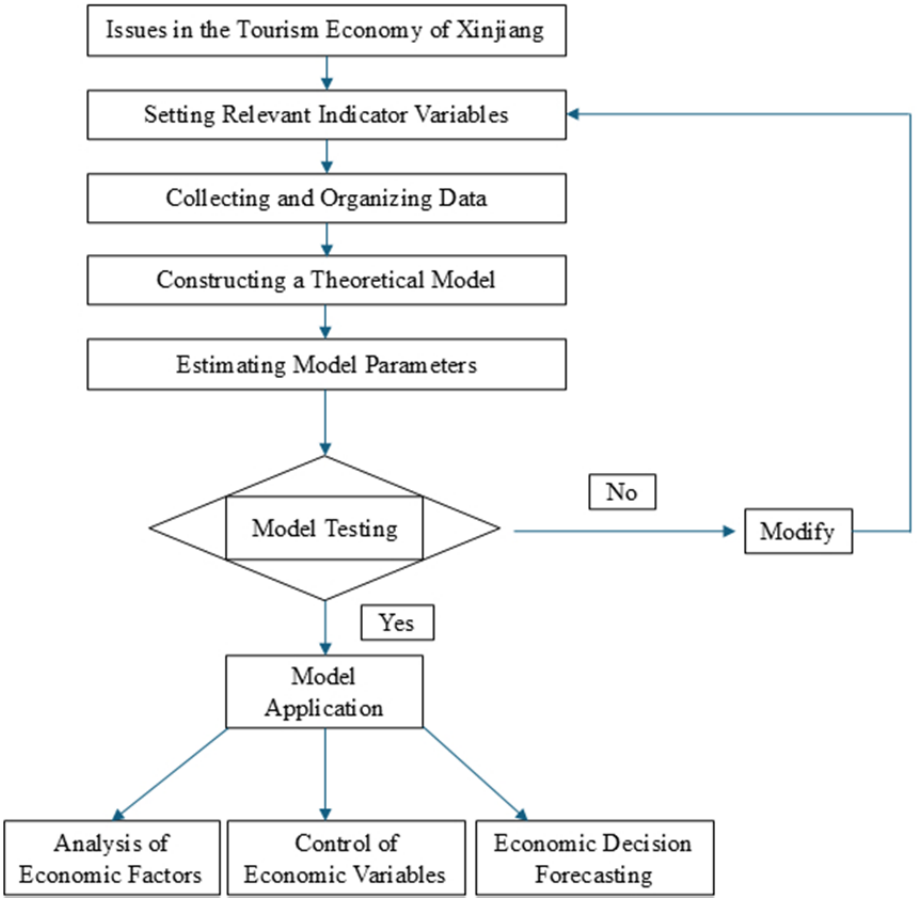

Figure 2 – Flowchart of Regression Modeling Steps

2. Establish the predictive model: Use historical data of both independent and dependent variables to develop the model.

3. Perform correlation analysis: Determine whether a relationship exists between variables before applying regression.

4. Calculate prediction errors: Evaluate the deviation between predicted and actual values.

5. Compute the forecasted values: Use the regression equation to obtain predicted outcomes.

The primary aim of regression modeling is to reveal the quantitative relationships among related variables. First, the dependent variable y should be determined according to the research objective. Then, variables that are statistically related to y are selected as independent variables. In most cases – especially in studies involving economic phenomena – it is expected that there exists a causal relationship between the independent and dependent variables.

In this study, the modeling of Xinjiang's tourism economy is presented through a logical flowchart illustrating the steps involved in constructing the regression model (as shown below).

Results. Data Selection and Variable Specification. To conduct a rigorous empirical analysis of the factors influencing tourism development in Xinjiang, this study first selects relevant data and defines appropriate indicator variables for regression modeling.

To construct the regression model, the following variables are selected:

This study conducts an empirical analysis of the factors influencing tourism development in Xinjiang, using data from 2009 to 2020 (table 1). Six key indicators are selected to represent the influencing factors of tourism development in the region. These are:

Dependent Variable (Y): is total Tourism Revenue in Xinjiang (in 100 million yuan) – this represents the overall economic output of the tourism sector and serves as the main indicator for tourism development.

Independent Variables (X):

Number of Tourists (10,000 people); Passenger Volume (10,000 people); Tourist Person-Days – defined as the product of the number of tourist visits and the average length of stay over a specific period at a given destination; Tourist Turnover Volume (billion kilometers); Number of Inbound Overnight Visitors (10,000 people) (figures 3, 4).

These indicators are chosen for their ability to reflect both the scale and depth of tourism activities, as well as the infrastructure and international reception capacity.

Table 1 – Factors affecting tourism development in Xinjiang

Year	y	x_1	x_2	x_3	x_4	x_5
2009	186.08	2133.49	29886	969.611	386.68	350
2010	281	3100	31937	2227.224	420.82	510
2011	351	3078.57	35130	1110.346	488.53	560
2012	576	4860	38331	1539.582	535.81	620
2013	673.24	5205.59	40926	1846.55	544.46	690
2014	830	4952.69	37176	1529.808	500.00	540
2015	985	6097	35948	1265.487	477.29	530
2016	1340	8102	32148	1971.213	458.06	580
2017	1751.6	10725.51	27083	1936.617	428.54	770
2018	2579.71	15024.89	21204	946.982	405.76	990
2019	3593.5	21329.54	20276	859.446	414.53	350
2020	991.03	15811.46	6960	276.1	184.97	320
Data Source: Xinjiang Tourism Bureau and the National Bureau of Statistics of China.						

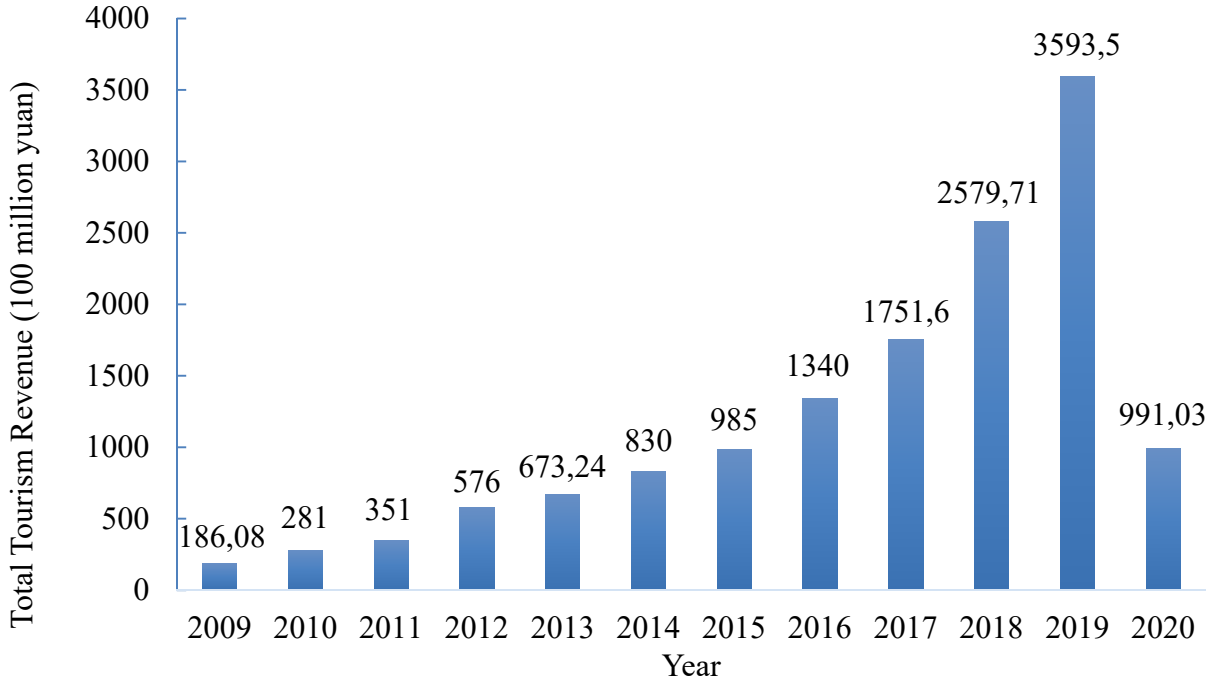

Figure 3 – Xinjiang's Total Tourism Revenue from 2009 to 2020

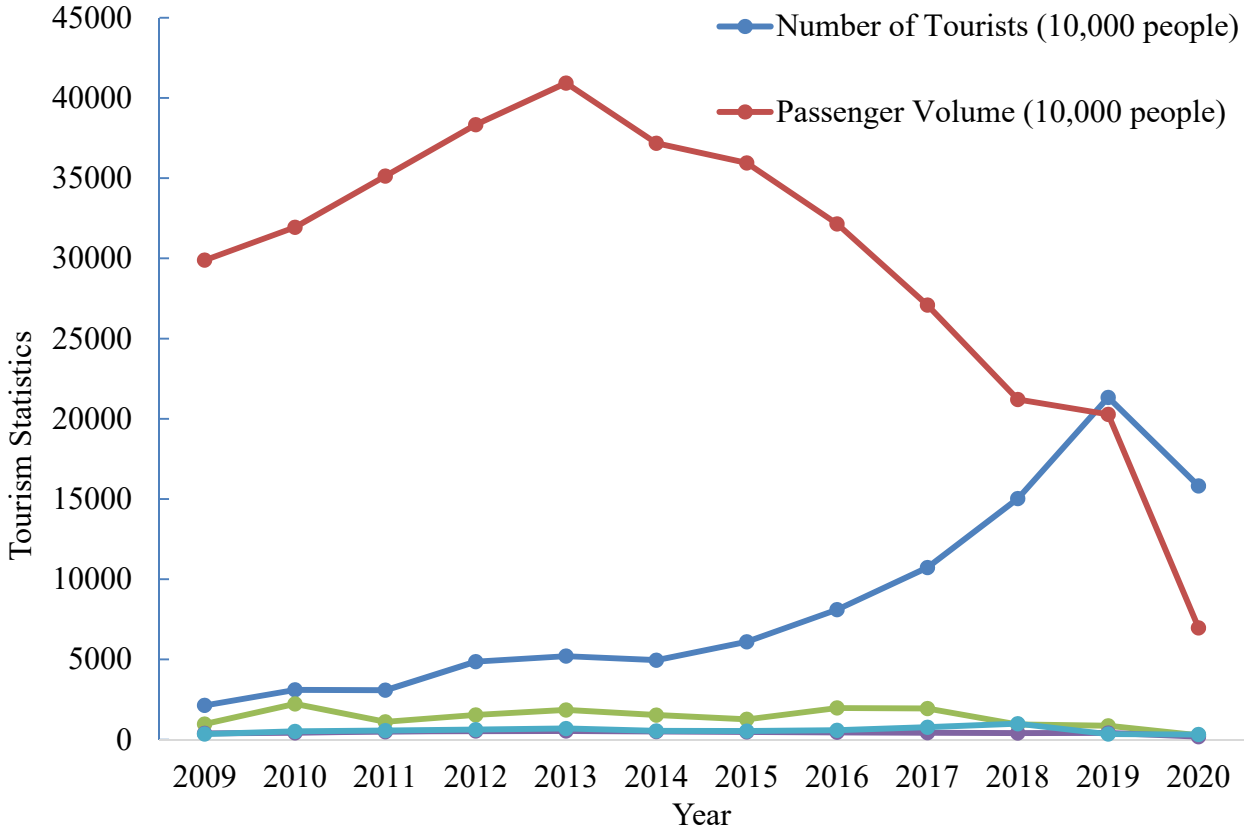

Figure 4 – Tourism Dynamics in Xinjiang: Trends in Tourist Flows, Transport, and Visitor Stays (2009–2020)

Pearson Correlation Coefficients.

Table 2 – Analysis Results of Pearson Correlation Coefficients

	y	x_1	x_2	x_3	x_4	x_5
Y	1 (0.000***)	0.897 (0.000***)	-0.489 (0.106)	-0.271 (0.394)	-0.164 (0.610)	0.2 (0.532)
x_1	0.897 (0.000***)	1(0.000***)	-0.779 (0.003***)	-0.51 (0.090*)	-0.518 (0.085*)	-0.005 (0.987)
x_2	-0.489 (0.106)	-0.779 (0.003***)	1 (0.000***)	0.68 (0.015**)	0.93 (0.000***)	0.236 (0.460)
x_4	-0.271 (0.394)	-0.51 (0.090*)	0.68 (0.015**)	1 (0.000***)	0.634 (0.027**)	0.362 (0.248)
x_5	-0.164 (0.610)	-0.518 (0.085*)	0.93 (0.000***)	0.634 (0.027**)	1 (0.000***)	0.401 (0.196)

Table 2 presents a correlation analysis examining the relationships between five independent variables (x_1, x_2, …, and x_5) and the dependent variable (y). Pearson correlation coefficients are used to indicate the strength of these relationships. The specific analysis reveals that:

The correlation coefficient between y and x_1 is 0.897, indicating a significant positive correlation between y and x_1.

The correlation coefficient between y and x_2 is -0.489, indicating a significant negative correlation between y and x_2.

The correlation coefficient between y and x_3 is -0.271, indicating a significant negative correlation between y and x_3.

The correlation coefficient between y and x_4 is -0.164, indicating a significant negative correlation between y and x_4.

The correlation coefficient between y and x_5 is 0.200, indicating a significant positive correlation between y and x_5.

Normality Test.

Table 3 – Normality test analysis results

Variable name	Sample size	Median	Mean	Standard deviation	Skewness	Kurtosis	S-W test	K-S test
y	12	907.5	1178.18	1019.698	1.478	1.84	0.848 (0.034**)	0.239 (0.430)
x_1	12	5651.295	8368.395	6083.339	1.083	0.197	0.867 (0.060*)	0.229 (0.486)
x_2	12	32042.5	29750.417	9633.659	-1.278	1.632	0.896 (0.141)	0.173 (0.807)
x_3	12	1397.648	1373.247	569.968	-0.288	-0.457	0.965 (0.852)	0.13 (0.971)
x_4	12	443.3	437.121	94.324	-1.762	4.533	0.838 (0.026**)	0.213 (0.576)
x_5	12	550	567.5	190.412	0.785	1.004	0.93 (0.377)	0.141 (0.943)

A normality test was conducted, and based on the results in the table above:

For x_1, since the sample size N < 5000, the Shapiro-Wilk test was used. The significance p-value is [not provided], which is not significant, and thus the null hypothesis cannot be rejected. Therefore, the data follows a normal distribution.

For x_2, with N < 5000, the Shapiro-Wilk test was used. The significance p-value is 0.060*, which is not significant, and thus the null hypothesis cannot be rejected. Therefore, the data follows a normal distribution.

For x_3, with N < 5000, the Shapiro-Wilk test was used. The significance p-value is 0.141, which is not significant, and thus the null hypothesis cannot be rejected. Therefore, the data follows a normal distribution.

For x_4, with N < 5000, the Shapiro-Wilk test was used. The significance p-value is [not provided], which is significant, and thus the null hypothesis is rejected. Therefore, the data does not follow a normal distribution.

For x_5, with N < 5000, the Shapiro-Wilk test was used. The significance p-value is 0.377, which is not significant, and thus the null hypothesis cannot be rejected. Therefore, the data follows a normal distribution (table 3).

Linear Regression Analysis.

Table 4 – Linear regression analysis results (n = 14)

	Unstandardized coefficients and Standardized coefficients			*T*	*P*	VIF	R^2	Adjustment R^2	F
	B	Standard error	*Beta*						
Constant	-2151.076	744.176	–	-2.779	0.030*	–	0.933	0.878	F=16.816 P=0.002**
x_1	0.148	0.073	0.882	2.022	0.090*	17.123			
x_2	-0.065	0.118	-0.612	-0.548	0.604	112.615			
x_3	0.147	0.281	0.082	0.521	0.621	2.224			
x_4	8.966	9.062	0.812	0.968	0.370	63.29			
x_5	-0.028	0.814	-0.005	-0.034	0.974	2.079			
Note: y is the Dependent Variable.									

According to table 4, a multiple linear regression analysis was conducted by taking x_1, x_2, x_3, x_4 and x_5 as independent variables, and y as the dependent variable. Based on the results in the table, the multiple linear regression equation can be expressed as follows:

$$y = -2151.723 + 0.148 \times x_1 - 0.065 \times x_2 + 0.147 x_3 + 8.776 x_4 - 0.028 x_5.$$

An F-test was performed on the model, yielding an F-value of 16.816 with a significance level (p-value) of 0.002. This result leads to the rejection of the null hypothesis that all regression coefficients are equal to zero, indicating that the model is statistically significant overall and generally meets the requirements. The regression equation demonstrates a good degree of fit.

The coefficient of determination (R^2) is 0.926, further confirming that the model has a strong explanatory power and a good curve-fitting performance.

However, regarding multicollinearity, the Variance Inflation Factor (VIF) values for variables x_1, x_2 and x_4 are greater than 10, suggesting the presence of serious multicollinearity among these variables. Therefore, it is necessary to apply a stepwise regression model to address this issue and improve model stability.

Stepwise Regression Model.

Table 5 – Stepwise regression model results (n = 14)

	Unstandardized coefficients and Standardized coefficients			*T*	*P*	VIF	R^2	Adjustment R^2	F
	B	Standard error	*Beta*						
Constant	-2311.11	596.315	–	-3.876	0.004***	–	0.926	0.91	F=56.696 P=0.000***
x_1	0.186	0.018	1.108	10.492	0.000***	1.366			
x_4	4.425	1.142	0.409	3.875	0.004***	1.366			
Note: y is the Dependent Variable.									

Based on the results of the F-test, the significance p-value is 0.000***, indicating a high level of statistical significance. Therefore, the null hypothesis that all regression coefficients are equal to zero is rejected. This suggests that the overall regression model is statistically significant and exhibits a good degree of curve fitting.

Regarding multicollinearity, all Variance Inflation Factor (VIF) values are less than 10, indicating that there is no serious multicollinearity among the independent variables. As such, the model is well-constructed and statistically sound.

Discussion. Based on the above test results and analysis, it is evident that certain indicators have had a positive effect on the development of tourism in Xinjiang. As shown in the previous analysis of influencing factors of Xinjiang's tourism development from 2009 to 2020, key indicators such as total tourism revenue (y), number of tourists (x_1), passenger volume (x_2), tourist person-days (x_3), tourist turnover (x_4), and number of inbound overnight visitors (x_5) have all shown continuous year-on-year growth, with an accelerating rate of increase.

This consistent upward trend indicates that these factors have had a significant impact on tourism development in the region. Specifically, number of tourists, tourist person-days, and inbound overnight visitors have been identified as the main influencing factors of total tourism revenue in Xinjiang.

Tourism has already become a key industry in Xinjiang, and tourism sales directly influence the overall tourism income. With the continuous increase in tourism income and the comprehensive promotion of tourism development, the demand for tourism in the region is expected to grow rapidly in the coming years.

Conclusions. This study employed rigorous regression analysis to quantify the key factors driving tourism development in Xinjiang, utilizing empirical data from 2009 to 2020. The stepwise regression model conclusively identified tourist volume (X_1) and tourist turnover volume (X_4) as the most statistically significant predictors of total tourism revenue (Y), explaining 91% of its variance. These findings underscore that the sheer scale of visitor influx (X_1) and the economic activity generated per tourist through transportation and mobility (X_4) are the primary engines of Xinjiang's tourism economy. The robustness of the model, validated by strong significance ($p=0.000$) and absence of multicollinearity ($VIF<10$), provides a reliable empirical foundation for policy formulation.

The analysis affirms that Xinjiang's tourism sector has entered a phase of accelerated growth, with all indicators showing sustained upward trajectories despite temporary disruptions. However, persistent challenges–including regional development disparities, infrastructure gaps, and high operational costs–demand strategic interventions. To harness the full potential of tourism as a pillar industry, policies must prioritize enhancing accessibility (e.g., optimizing transport networks and affordability), scaling service capacity (e.g., expanding high-quality accommodations and skilled workforce training), and diversifying tourism products (e.g., leveraging cultural assets and winter landscapes). Crucially, development must align with ecological sustainability to preserve Xinjiang's natural capital.

Based on the above research, although the level of tourism development in Xinjiang has improved compared to the past, significant regional disparities in influencing factors remain, indicating that further improvement is still needed. The growing scale of Xinjiang's tourism industry has played an increasingly important role in promoting regional economic growth, and the sector continues to demonstrate sustained momentum.

Figure 5 shows the spatial representation of Xinjiang's tourism performance (2009-2020) and targeted development strategies. The left panel illustrates prefecture-level tourist volume and tourist turnover, highlighting regional disparities identified through regression analysis. The right panel delineates priority zones for infrastructure upgrades, ecotourism development, winter sports expansion, transport cost-reduction corridors, and service quality improvements, as recommended in the study.

Drawing from the analysis of influencing factors on Xinjiang's tourism development, the following recommendations are proposed:

1. Integrate Xinjiang's natural beauty with high-quality services. Strengthen investment in the tourism sector, vigorously support the development of related tourism industries, and cultivate well-educated, highly skilled tourism management teams with excellent service quality.

2. Enhance the "food, accommodation, transport, sightseeing, shopping, and entertainment" services. Improve infrastructure in tourist destinations, including the renovation and smart upgrading of public toilets, parking lots, water and electricity systems, and other basic facilities within scenic areas. Creating a comfortable and intelligent tourism environment can significantly improve visitors' experiences.

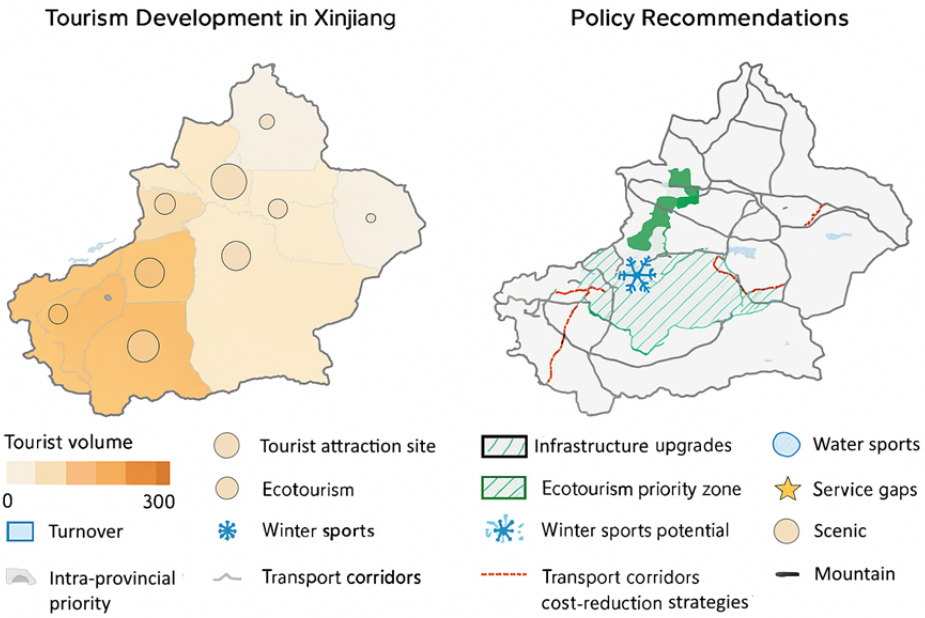

Figure 5 – Tourism Development and Policy Recommendations in Xinjiang, China (compiled by the authors)

3. Increase investment in the tourism industry. Broaden financing channels by actively attracting venture capital, enhance promotion and marketing efforts, and implement preferential policies to support tourism development.

4. Actively implement the “big tourism, big market” strategy. First, adopt a comprehensive approach to tourism development across regions; second, explore both international and domestic tourist demands, integrating them to build a responsive and inclusive tourism market; and third, address the needs of both high- and low-income travelers to expand market accessibility.

5. Accelerate the cultivation and utilization of tourism professionals. Focus on training and attracting talents familiar with both domestic and international tourism markets, who can apply modern technologies and management methods. Multilingual, communicative, and culturally competent professionals are especially important to enhance service quality and international communication.

6. Emphasize the integration of tourism development with ecological protection, pursuing sustainable development. Efforts should be made to balance ecological conservation with tourism promotion, under the principle that “lucid waters and lush mountains are invaluable assets.” High-quality eco-tourism products should be developed, while tourism-related environmental early warning and emergency response mechanisms must be strengthened to safeguard ecological integrity.

7. Apply innovative thinking and methods to address the long-standing issue of high transportation costs. Encourage cooperation between transportation companies, tourism operators, travel agencies, and scenic spots. Promote the use of renewable energy sources such as wind, solar, and hydropower to reduce energy and travel costs, thus making travel in Xinjiang more affordable and sustainable.

Vigorously promote the “Revitalize Xinjiang through Tourism” strategy. Given Xinjiang’s unique geographic and climatic conditions - particularly its snowy winter landscapes - the region should actively develop winter sports and the ice-and-snow economy. Activities such as skiing, ice skating, and other snow-related events should be widely promoted to attract tourists. Diversifying tourism products and experiences will contribute to the deep integration of culture and tourism, enhancing Xinjiang’s overall attractiveness as a year-round destination.

Acknowledgements: This research work was carried out within the framework of the project entitled Statistical Analysis of the Diversified Development of China-Kazakhstan Tourism Economy (Grant No.: 2022ZBGJYB005) of the Scientific Research Project of China (Xinjiang) and Neighboring Countries Cooperation and Development Research Center at Yili Normal University. We express our sincere gratitude to our institutions. At the same time, we thank the editor and anonymous reviewers of the journal Geography and Water Resources for their thorough review and valuable feedback. Their comments significantly enhanced the quality of the manuscript and helped elevate it to the standard of a full scientific article.

REFERENCES

[1] Geng Siqi. (2022). A Study on the Relationship Between Tourism and Economic Growth in Xinjiang // Journal of Xianyang Normal University, 37(6), 48-56.

[2] He Shuang. (2019). Construction and Application of Mathematical Models in Regression Analysis // Tiangong, (2), 146-147.

[3] Yang Xueqi. (2016). A Study on the Performance of the Tourism Industry in Xinjiang and Its Influencing Factors (Master's thesis). Beijing Jiaotong University.

[4] Mayila·Mutailipu. (2010). Analysis of Major Factors Affecting the Tourism Economy in Xinjiang // Heilongjiang Science and Technology Information, (11), 87-165.

[5] Feng Ying, Zhang Junmin. (2017). Spatial Differentiation and Influencing Factors of Tourism Economy: A Discussion Based on Statistical Work in Xinjiang // Regional Research and Development, 36(4), 99-104.

[6] Zhang Wanbo. (2021). Spatial–Temporal Evolution and Influencing Factors of Tourism Efficiency in Xinjiang (Master's thesis). Xinjiang University.

[7] Yang Shuangwu, Zhao Xianhui. (2022). Analysis of Influencing Factors of Tourism Economy: A Case Study of Dali // China Collective Economy, (6).

[8] He Xiaoqun. (2019). Fundamentals and Applications of Regression Analysis [Book]. Beijing: Renmin University of China Press.

[9] Xu Suxu. (2022). Regression Analysis of Tourism Industry Revenue in China // Fortune Times, (3), 153-155.

[10] Huang Wenxia, Li Min. (2019). Analysis of Key Factors Influencing Tourism Development Based on SPSS Data Analysis // Journal of Electromechanical Engineering, Kunming University of Science and Technology, (1).

Б. Нұлахемет[1], И. Ақбар[*2]

[1] Доцент (Математика және статистика мектебі, Іле педагогикалық университеті, Шыңжаң, Қытай; *2423883975@qq.com*)

[2] *PhD, доцент м.а. (Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан; ***yimanaili_akebaier@yahoo.com***)

ШЫҢЖАҢ ТУРИЗМІН ДАМЫТУҒА ЫҚПАЛ ЕТЕТІН ФАКТОРЛАРДЫ РЕГРЕССИЯЛЫҚ ТАЛДАУ АРҚЫЛЫ ЗЕРТТЕУ

Аннотация. Қытай экономикасының жедел дамуына байланысты Шыңжаңдағы туризм секторы да айтарлықтай өсіп, өңірлік экономиканың маңызды қозғаушы күшіне айналды. Осыған байланысты туризм экономикасының даму деңгейін ғылыми тұрғыда бағалау аса маңызды. Бұл зерттеуде 2009–2020 жылдар аралығындағы мәліметтерге регрессиялық талдау жүргізу арқылы Шыңжаң туризмінің негізгі қозғаушы факторлары анықталды. Тәуелді айнымалы ретінде Туризмнен түскен жалпы кіріс (Y), ал тәуелсіз айнымалылар ретінде: туристер саны (X_1), жолаушылар ағыны (X_2), туристік адам-күндер саны (X_3), туристік айналым көлемі (X_4) және шетелдік түнеген келушілер саны (X_5) алынды. Алғашқы регрессиялық модельде X_1, X_2 және X_4 айнымалылары арасында күшті мультиколлинеарлық (VIF > 10) байқалып, модельді нақтылау үшін қадамдық регрессия әдісі қолданылды. Қорытынды модель туристер саны ($\beta = 1{,}108$, $p<0{,}001$) және туристік айналым көлемі ($\beta = 0{,}409$, $p<0{,}001$) кірісті болжауда ең ықпалды факторлар екенін көрсетті, жалпы кірістің 91%-ын түсіндіреді (Түзетілген $R^2 = 0{,}91$, $F = 56{,}696$, $p<0{,}001$). Бұл нәтижелер туризмдегі көлем мен қозғалыстың басты экономикалық қозғаушы күштер екенін көрсетеді. Алайда, 2009–2019 жылдар аралығында туристер саны 898%-ға артқанымен, туристік адам-күндер саны 71,5%-ға төмендегені байқалды, бұл туристердің сапар ұзақтығының қысқарғанын және әлсіз қатысушылықты меңзейді. Сонымен қатар, туристік кіріс пен жолаушылар ағыны арасында теріс корреляция ($r = -0.489$) байқалды, бұл инфрақұрылымға жүктеменің артқанын көрсетеді. Жалпы, Шыңжаң туризмі жоғары қарқынмен дамып жатқанына қарамастан, ол әлі де көлемге негізделген, бұл экологиялық қысым мен жан басына шаққандағы табыстың төмендеуіне әкелуі мүмкін. Сонымен қатар, инфрақұрылымдық теңсіздіктер мен жоғары құнды туризм сегменттерінің жеткіліксіз дамуы байқалады. Осыған орай, сапаға бағытталған дамуға көшу қажет: көлік жүйесін оңтайландыру, бай туристік тәжірибе арқылы сапар мерзімін ұзарту, қысқы туризмді дамыту және жаңартылатын энергия көздерін пайдалану – өңірлік туризмнің тұрақты әрі инклюзивті дамуына негіз бола алады.

Түйінді сөздер: регрессиялық талдау, Шыңжаң туризмі, туризм экономикасы, қадамдық регрессия, факторлық талдау.

Б. Нулахемет [1], И. Ақбар [*2]

[1] Доцент (Школа математики и статистики, Илийский педагогический университет, Синьцзян, Китай; *2423883975@qq.com*)

[2 *] PhD, и.о. доцент (Казахский национальный университет имени аль-Фараби, Алматы, Казахстан; ***yimanaili_akebaier@yahoo.com***)

АНАЛИЗ ФАКТОРОВ, ВЛИЯЮЩИХ НА РАЗВИТИЕ ТУРИЗМА В СИНЬЦЗЯНЕ, НА ОСНОВЕ РЕГРЕССИОННОГО АНАЛИЗА

Аннотация. В условиях стремительного экономического роста Китая туристическая отрасль Синьцзяна также демонстрирует устойчивое развитие и становится важным двигателем региональной экономики. В связи с этим научная оценка уровня развития туристической экономики приобретает особую актуальность. С использованием данных за 2009–2020 годы проведён регрессионный анализ для выявления ключевых факторов, определяющих рост туризма в Синьцзяне. В качестве зависимой переменной рассматривался общий доход от туризма (Y), а в качестве независимых переменных – объём туристов (X_1), пассажиропоток (X_2), количество туристо-дней (X_3), объём туристического оборота (X_4) и число иностранных туристов, оставшихся на ночь (X_5). Первичный регрессионный анализ выявил сильную мультиколлинеарность между переменными X_1, X_2 и X_4 (VIF > 10), что потребовало применения пошаговой регрессии для уточнения модели. Итоговая модель показала, что наибольшее влияние на доход имеют объём туристов ($\beta = 1{,}108$, $p<0{,}001$) и туристический оборот ($\beta = 0{,}409$, $p<0{,}001$); объясняемая дисперсия составила 91% (скорректированный $R^2 = 0{,}91$, $F = 56{,}696$, $p<0{,}001$). Это подтверждает, что масштаб и мобильность турпотока – основные экономические драйверы. Однако анализ выявил важные противоречия: с 2009 по 2019 год число туристов увеличилось на 898%, в то время как количество туристо-дней сократилось на 71,5%, что указывает на сокращение продолжительности пребывания и слабую вовлечённость. Также обнаружена отрицательная корреляция между доходами и пассажиропотоком ($r = -0{,}489$), что свидетельствует о перегрузке инфраструктуры. Несмотря на быстрый рост, развитие туризма в Синьцзяне носит количественный характер, что может привести к экологическим рискам и снижению доходов на душу населения. Кроме того, сохраняются инфраструктурные диспропорции и слабо развиты направления с высокой добавленной стоимостью. В этой связи необходим переход к качественно ориентированной модели развития: оптимизация транспортной системы, стимулирование длительных поездок за счёт насыщенных программ, развитие зимнего туризма и внедрение возобновляемых источников энергии для устойчивого и инклюзивного роста отрасли.

Ключевые слова: регрессионный анализ, туризм в Синьцзяне, экономика туризма, пошаговая регрессия, факторный анализ.