

Research on Influencing Factors of Domestic Tourism Revenue based on Multiple Linear Regression Analysis

Zining Yan¹, Hongbo Zhou^{2, *}

¹Zining Yan Jinling High School Hexi Campus, Nanjing, 210019, China

²Hongbo Zhou Jinling High School Hexi Campus, Nanjing, 210019, China

* Corresponding Author Email: gefviq@tzc.edu.cn

Abstract. The purpose of this paper is to explore the factors that have an impact on the domestic tourism industry in the process of economic development. This paper uses the statistical analysis method of multiple linear regression to analyze the specific situation of tourism from 1994 to 2022, and the statistical method of multiple linear regression aims to analyze the impact of disposable income, GDP, transportation development level, environmental resources, basic city construction and other factors on tourism development. In order to further ensure the reliability of the conclusions, the per capita disposable income of four regions in the eastern, western, central and northeastern regions was obtained through linear regression from different regions to specific cities, and the per capita disposable income and per capita consumption expenditure of urban residents in each region from 2015 to 2021 were counted. In the end, it was concluded that the discretionary point in the eastern region was much higher than the average compared to other regions, but it did not have a significant impact on the total tourism expenditure.

Keywords: Domestic tourism, multiple linear regression, influencing factors.

1. Introduction

With the development of China's economy, tourism consumption plays a key role in China's economic growth. Domestic tourism consumption has continued to rise, as the increased of gross domestic product and disposable personal income between 2007 to 2018 [1]. Since 2019, in the period of epidemic situation, domestic tourism revenue dropped significantly. China's tourism industry has been severely hit. Total tourism revenue in 2019 was only 8% of what it was in 2018 [2]. Tourism expenditure accounts for a portion of national disposable income. However, the public is not particularly aware of its influencing factors. It is crucial to understand the factors that influence the tourism industry. Therefore, this paper will use the model of multiple linear regression to help people to study the factors that may affect the consumption of domestic tourism. The model can roughly estimate the overall development trend of domestic tourism and can also infer future changes based on this to help with investment of tourism and better travel spending.

Domestic tourism is a complex system that involves many factors such as GDP. Forecasting the development of tourism is also a popular opinion in academic. Li and Fan found that there was a high positive correlation between urban residents' domestic tourism consumption and GDP [3]. What's more, the convenience of transportation has a direct impact on the domestic tourism income, and the transportation convenience and the richness of surrounding tourism resources have an indirect impact on the domestic tourism income [4]. Wang illustrated that the income of residents and average tourism expenditure of residents in the city form part of the characteristic variables of tourism consumption [5]. Zhang and Wang found that improvement of urban infrastructure constitutions, tourist resources and cultural excavation are important factors of tourism consumption. The competitiveness of the tourism industry ultimately lies in the exploration of tourism resources and culture [6]. Wang found that there is still a gap between rapid development of tourism industry in China and that in developed country and the problem of poor protection of resources. From a favorable perspective, some festivals and exhibition bring some new opportunities to the industry and improve the local tourism industry.

Also, the peace sociability and rapid development of the whole country is another factor [7]. Liu found that the development of tourism industry is to some extent influenced by the number of tourists

from foreign country and the environment of accommodation and catering conditions. The increase of foreign tourists promotes the growth of tourism consumption [8]. Hua thought one of these important factors is increase of the income of tourists and another is construction of service industry, such as the number of travel agency, the number of star hotel and the average cost of tourists [9]. Huang found the innovation of tourism project and abundant tourism resources are sore of tourism development. China has a history of 500 years, so the tourism resources are quite abundant. It is necessary to do a good job in central area planning and infrastructure construction, new things and activities will improve the consumption levels and accelerates the development speed of tourism industry [10].

In addition to the above external factors, internal factors play a more decisive role, Pearce found that money, time, family or personal health, work indebtedness, family love and other factors are the disrupting factors of travel decisions [11]. Du found that the influencing factors of tourism decision-making mainly include tourists' demographic conditions, travel motivation, tourism purpose, tourism mode, tourism information, destination culture, tourism product and service prices, travel agency brands, relatives and friends' opinions, and so on [12]. Based on the theory of environmental stress, Ishii proposed that tourists' perception, the characteristics of the destination environment, and the cultural constraints of the society in which tourists live play an important role in the tourism decision-making process [13].

Therefore, this paper will use the model of multiple linear regression to help people to study the factors that may affect the consumption of domestic tourism. The model can roughly estimate the overall development trend of domestic tourism and can also infer future changes based on this to help with investment of tourism and better travel spending.

2. Methods

2.1. Data Source

The data on play inequality was constructed in the following way. From Nanjing Tourism Bureau. These data are collected to record the average price of five-star hotels in Nanjing, the number of tourists on weekends and holidays, to the national income in tourism from 2012 to 2021.

2.2. Variable Selection

Due to the large amount of research data and the clear research objectives of this paper, the factors involved in this paper are relatively complex, and there are many variables, which also have different degrees of influence on the situation in different years (Table 1).

Table 1. Nanjing Statistics Data

Year	Hotel average price	Number of premium hotels	national income in tourism	tourists(ten thousand)
2012	595	60	1272	8113
2013	653	55	1361	8726
2014	519	70	1521	9476
2015	523	74	1688	10234
2016	536	69	1909	11206
2017	566	71	2169	12293
2018	598	70	2460	13400
2019	564	63	2785	14682
2020	527	55	1823	9704
2021	555	54	2130	10840

In order to conduct a more in-depth analysis and conduct research from the data itself, this paper collected in detail the total tourism revenue of Nanjing, the average daily price of five-star hotels in

Nanjing, the number of tourists on weekends and holidays, and the specific data of national tourism revenue from 2012 to 2021 to analyze and draw the final results.

2.3. Method Introduction

Multiple linear regression is a statistical analysis method used to determine the interdependence of two or more variables. In this paper, the multiple linear regression is used to conduct a study on China's domestic tourism industry. The purpose of this research is to analyze the impact of various influencing factors such as disposable income of people, GDP, transportation development level, environmental resources and urban infrastructure on tourism income and development. So, the independent variable is disposable income of people, GDP, transportation development level, environmental resources and urban infrastructure. The general mathematical model for multiple linear regression is:

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + e \quad (1)$$

In the above formula: β_0 is a constant term, and e is a residual term.

3. Results and Discussion

3.1. Descriptive Analysis

From 2015 to 2021, the number of three-star and above hotels shows an overall downward trend, with the number of three-star hotels declining most significantly, from 25 in 2015 to 13 in 2021. There is no significant change in the overall number of four-star and five-star hotels (Figure 1).

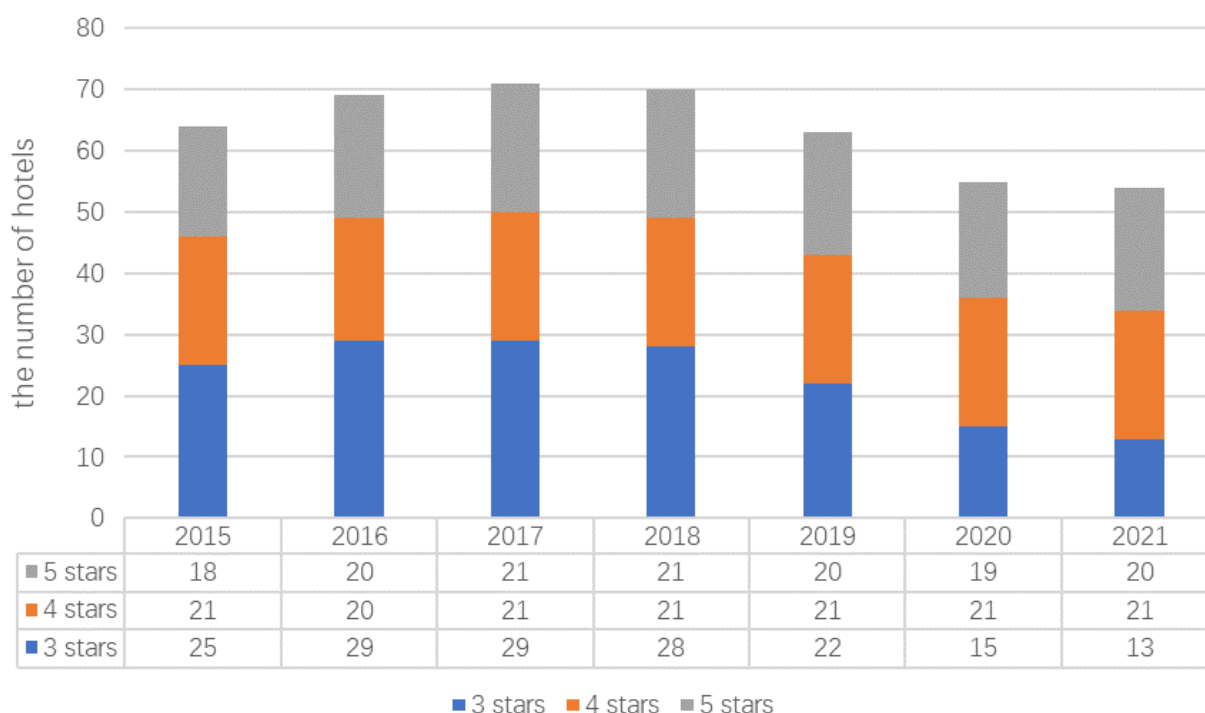


Fig. 1 Number of three to five stars' hotel in Nanjing

3.2. Model Results

This paper selected the average price of five-star hotels (X1), the number of high-grade hotels with three stars and above (X2) and the number of tourists during holidays and weekends (X3) as independent variables. The dependent variable is the total income of Nanjing tourism.

Table 2. Regression coefficient table

	B	S.E.	Beta	T	P	VIF
Constant	1006.601	376.208		2.676	0.037	
X1	-1.388	0.487	-0.121	-2.852	0.029	1.174
X2	-14.779	2.853	-0.234	-5.181	0.002	1.324
X3	0.242	0.01	1.044	24.967	0	1.139

Table 2 shows the regression coefficients of the multiple linear regression equation model. The p-values of the T-test for the three independent variables X1, X2, X3 did not exceed 0.003. Therefore, it can be found that all three independent variables have a significant impact on the dependent variable Y. Based on the data above, the relevant multiple linear regression equation can be obtained:

$$y = 1006.601 - 1.388 * X_1 - 14.779 * X_2 + 0.242 * X_3 \quad (2)$$

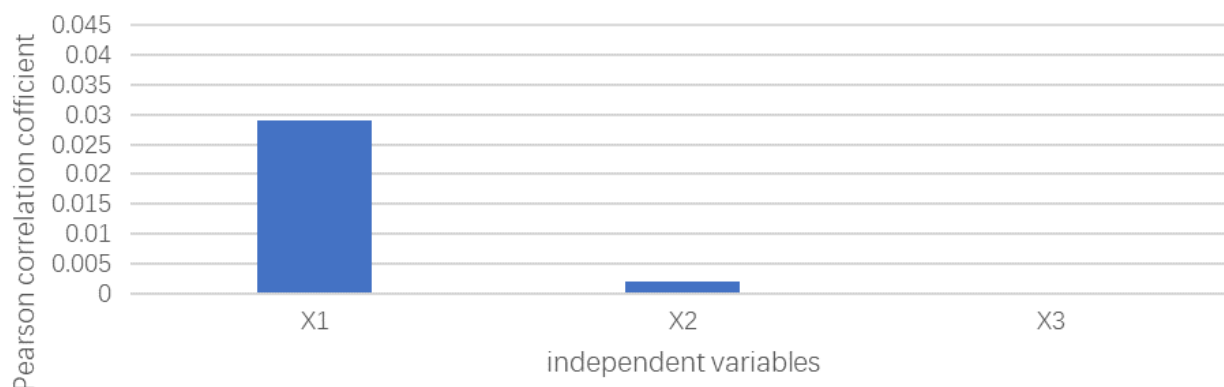


Fig. 2 Relevance Analysis Between Dependent and Independent Variables

From Figure 2, it can be seen the Pearson correlation coefficient between these factors and national income in tourism of Nanjing. The research data found that the X3 is the most positively correlated factor with national income in tourism of Nanjing. That mean the more the number of tourists, the more Nanjing's tourism income. Surprisingly, however, there is a negative correlation between the X2 and total tourism revenue. What's more, the X1 will also have a significant negative impact on nominal income.

4. Conclusion

The study selected 40 samples from a dataset from 2012 to 2021, which has 3 variables. Our method (multiple linear regression analysis) is accurate, valid, and comprehensive. Because it performs multivariate analysis and then gets the Pearson correlation coefficient for each variable.

In the analysis, a multiple linear regression model was used to find out the possible relationship between the variables and Nanjing's tourism revenue. Therefore, the positive factor for this is the number of tourists on holidays and weekends. There is a negative correlation between the average room rate of five-star hotels and the number of three-star hotels and above. In summary, the number of tourists on holidays and weekends, the average room rate of five-star hotels and the number of three-star hotels and above are the main factors.

Through the research, we have a certain understanding of the influencing factors of Nanjing's tourism revenue. However, there are also some drawbacks, such as a relatively small sample size and the data is not up to date. Searching for new data and finding out possible relationships with other factors in Nanjing's tourism revenue may improve this.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

- [1] Cheng Xingxing. Research on Influencing Factors of Domestic Tourism Consumption Based on Multiple Linear Regression Analysis. *Advances in Social Sciences*, 2020, 9(9): 1530-1536.
- [2] Zhu Kai. Research on the impact of the new crown pneumonia epidemic on the development of China's tourism industry and countermeasures. *Journal of Sichuan Tourism University*, 2020.
- [3] LI Yinlan, Fan Hong. A Preliminary Study on the Consumption Model of Domestic Tourism. *Journal of Chongqing Business School*, 2002.
- [4] Wang Ning, Du Changchun, Yang Linxi. The Influencing Factors of China's Tourism Revenue An Analysis Based on Inter-provincial Panel Data. *Journal of Business Economics*, 2020, 8: 177-180.
- [5] Wang Nan. Empirical Analysis on the Factors Influencing the Development of the China's Tourism Industry. *Financial perspective*, 2021, 10: 194.
- [6] Zhang Yu, Wang Yanping. Empirical Analysis of Factors Influencing the Development of Tourism Industry in Shanxi Province Based on Multicollinearity, 2014, 3: 85-88.
- [7] Wang Ao. Analysis of Factors Influencing the Development of Modern Tourism Industry. *Tourism Overview*, 2014.
- [8] Liu Fengjiao. An Empirical Study on the Factors Influencing the Development of China's Inbound Tourism Industry. *Financial News*, 2018.
- [9] Hua Wenjun. Analysis of Factors Influencing the Development of Tourism Industry. *Tourism Overview*, 2011.
- [10] Huang Xiaohui. Quantitative Analysis and Exploration of the Factors Influencing the Development of China's Tourism Industry. *Journal of Chifeng University*, 2015, 31(24): 112-115.
- [11] Pearce D. *Tourism Today a Geographical Analysis*. Longman Scientific and Technical, 1987.
- [12] Research Group on Consumption Behavior Patterns of Chinese Outbound Tourists. *A Study on the Consumption Behavior Patterns of Chinese Outbound Tourists*. Beijing Tourism Education Press, 2003, 29-40.
- [13] Shinji Ishii. *Approach from the Psychological Model of Human Geography and River*. Human Geography and Geography, 2009.