

The Study of The Economic Distribution Characteristics of Night Economic Distribution Based on Multi-source Data Fusion

Yixuan Zhou*, Yuehua Geng

Southwest Petroleum University, Nanchong 637001, China

* Corresponding author

Abstract: The night economy is an important representation of urban economic development and consumption level. Based on the night lighting, the data of the POI data, the data of the spatial clustering algorithm of DBSCAN, K-means and other spatial clustering algorithms and the study of the profit and loss of supply and demand, the model of the area and the area domain analysis of the night activity hotspot area and night service facility are established. The fusion of multi-source data provides a new perspective for the night economy, which is more rapid, efficient and extensive than the traditional survey data, which is suitable for the large-scale study of the night economy.

Keywords: Multisource Data, Night Economy.

1. Introduction

Night-time Economy refers to the service-oriented economic activities that take place at Night based on urban space, and its business forms include night shopping, night dining, etc. With the development of information technology, night light data and perception big data provide new data sources for quantitative research of night economy. But light intensity can be affected by non-economic factors such as vegetation and climate, making it impossible to directly measure the night-time economy. Points of Interest (POI) data have precise location attributes and specific classification. Some scholars use POI data to analyze the spatial pattern of service industry, but POI data cannot distinguish the distribution characteristics of day and night economy. OD (start and end point, Origin Destination) data has both location and time attributes. Some scholars use OD data to study residents' travel characteristics to identify human activity units, but OD data identification accuracy is not high. Therefore, using only single source data to perceive the city may make the research results biased, while building a multi-perspective framework based on multi-source data can effectively reduce the bias of the research results. Therefore, based on the night light data and POI data, this paper uses the classification method and fishnet tool to unify the data format and construct a multi-perspective framework. From the perspectives of consumers and merchants, this paper identifies urban night activity hotspots and night service facilities distribution areas, and cross-analyzes the spatial and temporal distribution pattern of the night economy. The conclusion of this study has positive significance for promoting social employment and enhancing the utilization rate of infrastructure, and can also provide reference for urban night economic development and policy making.

2. Research Method

This paper studies the spatial and temporal distribution pattern of night economy respectively from the perspective of consumers and merchants. For taxi od data, 1) by DBSCAN

(density-spatial clustering of applications with noise clustering), the low density area (noise) is used to continue to carry out the k-means + + + clustering. The cluster is classified by the natural breakpoint method, which is divided into high, medium and low three levels. The correlation between population density and night activity density is studied, and the reliability of the results is explored.

For poi and night light data, the profit and loss of the two; By analyzing the distribution of supply and demand and the spatial distribution of the break-even area to identify the area of the night service facility. Finally, the density of cluster and the average profit rate and the correlation of the analysis were analyzed.

(1) spot recognition of hot spots at night

In the od data, the dispersion of the anti-reflection group is separated, and the end of the reflector is gathered. Therefore, this paper discusses the spatial clustering of taxis in order to identify the hot spots in the night.

Spatial clustering is an important means of describing spatial dependence of geographical phenomena. DBSCAN is a more representative clustering algorithm based on density clustering, which can be divided into areas based on density, and can find any shape of clustering. K-means is a divided clustering algorithm, but it is sensitive to the selection of the initial center. K-means + + is a modified form of k-means, which, by allowing the user to given a cluster number or the initial cluster center, iterates over and over again, making the distance between the initial cluster center as far away as possible, effectively reducing the sensitivity to the initial center. Mean-shift is a dense sliding window iteration algorithm that automatically gets the optimal number of clustering, without the need for a given, less affected point. In order to select the optimal clustering parameters, this paper USES the internal evaluation index, which is often used in clustering, to evaluate the clustering effect of the contour coefficient (the contour coefficient) and the CH (Calinski-Harabaz). The contour coefficient is calculated by the concentration of cohesion and separation, and the value is in line with the value of $[-1, 1]$, and the more the convergence

and separation of 1 are relatively superior. The CH value refers to the ratio of the discrete dispersion between the group and the group, and the greater the value of the cluster. In order to unify the rules of classification, the data classification is carried out by using natural breakpoint classification method. The natural breakpoint classification method is the classification of the data centralized discontinuity by the iteration, and the classification is graded according to the logarithm. This method can minimize the difference in the class, maximize the difference between classes, and maintain the statistical characteristics of the data.

(2) the area of the night service facilities

Lighting facilities are constantly improving the basic conditions for the development of the night economy. Profit and loss can reflect the supply and demand of two elements. Therefore, the supply and demand of night service facilities and night lighting infrastructure are used to analyze the spatial distribution of night service facilities. The profit and loss formula is as follows:

$$C=V_a-V_b \quad (1)$$

In the formula: C indicates the profit and loss of night lighting and poi; V_a indicates the value of the factor a reclassification; V_b indicates the value of the benchmark factor b reclassification; It is a surplus of factor a, which represents the balance of the factor a loss and 0 value.

In order to eliminate the non-economic activity area of woodland and water, the brightness threshold value is extracted by the night light data. In order to unify the data format, the poi data and the night light data in the construction area are connected to the same scale fishing net, and calculate the plus and number of the net. Then, the night light data and poi data of the fishing net are classified by natural break point classification method. Finally, the profit and loss value of the night light supply and loss distribution and the spatial distribution of night service facilities are analyzed.

3. The Acquisition Method of The Relevant Data

(1) night light data

The night light data is used to study the spatial and temporal distribution pattern of urban night economy. The NPP-VIIR can use the monthly average radiation composite image produced by the image.

(2) POI data

You can use python to call the high DE map API to get out of the city shopping, dining, accommodation, entertainment and wind view, including names, categories, longitude and latitude and address four fields. Choose the right time to get up, so you can effectively represent the overall poi data of the city.

(3) OD data

It can collect the number of taxi orders in a certain time of the city, including taxi ids, boarding time, getting the latitude of the car, the time of the car and the latitude and longitude of the car. It can be divided into two parts: working day and holiday. According to the average sunrise and sunset time of the city, the day, the night and the middle of the night.

4. Data Analysis Method

The main body of night economic activity, the strength of night activity represents the strength of the night economic demand, so the passenger point distribution of the taxi is distributed as the demand distribution of the night activity, and analyzes the urban night economic and temporal distribution pattern from the perspective of the consumer. The night light is the basic condition of the economic development of the night, and the poi data is the service supply facility, and the night light data is used as the night lighting infrastructure, and analyzes the urban night economic space distribution pattern from the perspective of the enterprise.

(1) the spatial and temporal distribution pattern of noctitrist activity

Clustering of the lower guest points to study the suitable algorithm. Using the contour coefficient and the CH value to obtain the results of the DBSCAN optimal clustering, the clustering cluster is calculated. The average contour coefficient and the CH value of 39, obtain the results of the "k-means" + + best clustering, and obtain the optimal clustering number by means of the mean-shift algorithm.

(2) spatial distribution of night service facilities

Firstly, the data of night light of the lighting brightness value of the lighting brightness value of the construction area is obtained, and the accuracy of the extraction accuracy is 98.2%. Then, according to the data space resolution of the night light, the 500 mg 500 m fishing nets were established, and the poi data and the night light data were connected to the fishing net in the construction area, and then classified into five classes and assigned a value of 1purchase. In the end, the data is the benchmark element b, the profit and loss of the two is calculated, and the loss and loss distribution of night lighting and service facilities is obtained (figure 6), which is range in -2~ 4, and -2~ is said that the light supply of the night is insufficient, and the supply of the night light is sufficient, and the supply and demand balance of the night light.

(3) correlation analysis

At present, there is a lack of comparative analysis on the spatial and temporal distribution of the night economy, the lack of comparison and analysis, the urban residential population density is normalized, and the correlation between the density of cluster density is calculated by the 500 mn-500 m fishing net and the least squares linear regression.

5. Conclusion

The night economy is an important part of the modern urban economy. At present, the research on the study on the night economy is more than the study of the research on the market research and the survey of the questionnaire. Based on the data and perceptual data of night light, this paper USES the clustering algorithm and the supply and demand model, respectively, the customer and the business of the customer and the merchants to identify the working day and the holiday night activity hotspot area, analyze the profit and loss distribution of night service facilities and light. It can be obtained by the distribution of time and space in city and the supply and demand relationship between service facilities and night lighting infrastructure, in order to provide reference for the economic development and policy formulation of urban city.

Combined with multi-source data, this paper provides a

new perspective for the night economy, which is more rapid, efficient and extensive than traditional data, but still has a shortage of data and methods: using only taxi od data to reduce the representation of the population. The space clustering belongs to the unsupervised learning category, and the clustering result attribute is required. In the future, we can extend the following aspects: to combine the self-driving and the other od data, such as public transport, and to determine the semantic semantics of the people and improve the representation of the population. In the study method, the statistical data of the relevant authoritative statistics of the night economy are used, and the training is added to the pre-marked cluster, which makes the classification more suitable.

Acknowledgment

The authors gratefully acknowledge the financial support from Nanchong 2023 Social Science Planning Project (NC23C067).

References

- [1] Yang Di, Luan Weixi, Li Yue, et al. Multi-scenario simulation of land use and land cover based on shared socioeconomic pathways: The case of coastal special economic zones in China.[J]. Journal of environmental management,2023,335.
- [2] Liu Shangqin, Zhao Xizhi, Zhang Fuhao, et al. Spatial Downscaling of NPP-VIIRS Nighttime Light Data Using Multiscale Geographically Weighted Regression and Multi-Source Variables[J]. Remote Sensing,2022,14(24).
- [3] Li Pengfei, Hao Shuang, Cui Yuhuan, , et al. Analysis of the Evolution of the Relationship between the Urban Pattern and Economic Development in Guangdong Province Based on Coupled Multisource Data[J]. ISPRS International Journal of Geo-Information,2022,11(10).
- [4] Liang Hui. The Application of Multi-Source Big Data Mining Techniques in the Analysis of Basketball Economic Management [J]. Mathematical Problems in Engineering, 2022, 2022.
- [5] Xia Juzi. A Model of Urban Economic Resilience Development with Multisource Data Fusion [J]. Mathematical Problems in Engineering,2022,2022.
- [6] Zhang Xiaoxuan, Gibson John. Using Multi-Source Nighttime Lights Data to Proxy for County-Level Economic Activity in China from 2012 to 2019[J]. Remote Sensing,2022,14(5).