International Journal of
Geo-Information

MDPI

Article

The Socio-Spatial Distribution of Leisure Venues: A Case Study of Karaoke Bars in Nanjing, China

Can Cui [1,2], Jiechen Wang [1,3,*], Zhongjie Wu [1], Jianhua Ni [1] and Tianlu Qian [1]

1 Department of Geographic Information Science, Nanjing University, Nanjing 210023, China; ccui@gsd.harvard.edu (C.C.); wuzhongjienju@163.com (Z.W.); neejianhua@163.com (J.N.); qtl1234@126.com (T.Q.)
2 Graduate School of Design, Harvard University, Cambridge, MA 02138, USA
3 Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China
* Correspondence: wangjiechen@nju.edu.cn; Tel.: +86-25-8968-0669

Academic Editors: Mahmoud R. Delavar and Wolfgang Kainz
Received: 13 June 2016; Accepted: 19 August 2016; Published: 25 August 2016

Abstract: With the development of service industry and cultural industry, urban leisure and entertainment services have become an important symbol of the city and the driving force of economic and social development. Karaoke, a typical form of urban entertainment, is immensely popular throughout China, and the number of karaoke bars is expected to keep growing in the future. However, little is known about their spatial distribution in the urban space and their association with other location-specific factors. Based on the geospatial entity data and business statistics data, we demonstrate a clustered pattern of 530 karaoke bars in Nanjing by means of point pattern analysis and cluster analysis in GIS. Furthermore, we identify the distribution of population, transportation network, and commercial centers as the three determinants underlying the formation of the pattern.

Keywords: socio-spatial distribution; karaoke bar; leisure venue; China

1. Introduction

Leisure industry was listed as one of the "Big Five" engines of economic growth [1]. "Big Entertainment"—everything from bars to video stores to opera houses—are at the forefront of the leisure industry. It has been witnessed that leisure and entertainment sector accounts for an increasing share of gross domestic product in China as well as worldwide, leading to an optimized economic structure. It also becomes a large provider of employment [2] and plays a crucial role in meeting residents' recreational demand. Among residents' recreation activities, singing is the most common and popular one. Karaoke bars are the commercial venue providing services for singing.

Karaoke, originating from Japan [3], is a form of interactive entertainment in which an amateur singer sings along with recorded music (typically popular song minus the lead vocal) using a microphone. Karaoke soon spread to the rest of Asia and other countries all over the world [4]. Instead of placing karaoke machines in restaurants and hotel rooms in the early times, karaoke bars, with karaoke equipment installed in compartmented rooms and combined with food and other recreation service, emerged as a new business in the market and became enormously popular. The first karaoke bar in Mainland China was opened in 1990 in Hainan Province. Since then, karaoke has become one of the most popular leisure activities in urban China [5], and karaoke bars burgeoned throughout Mainland China. Our previous study shows that there are 87,700 karaoke bars in Mainland China as of 2012 [6], clustering in the eastern and central areas (Figure 1).

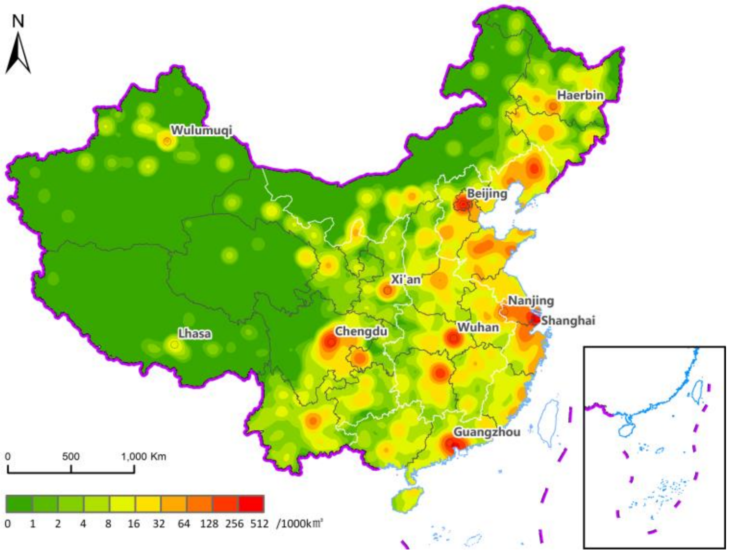


Figure 1. The spatial distribution of karaoke bars in China (Source: [6]).

Although karaoke bars, as a form of entertainment, are widespread in urban China, little is known about their spatial distribution within cities. Taking the entertainment and leisure industry as a whole, there is a growing body of literature investigating its spatial distribution in urban China. Zhang and Wang [7] summarized that there are four distribution patterns of urban commercial entertainment venues: center-oriented structure, discrete structure, attached structure, and isolated structure. Hu and Zhong [8] conducted a longitudinal study and observed that the distribution of entertainment venues in Hankou experienced a process from clustering to dispersing, and re-clustering. Taking Chengdu as a case study, Yang [9] found that the space distribution of leisure industry is closely related to residents' recreational activity range. With the development of GIS as a powerful tool for spatial visualization and analysis, more and more scholars carried out empirical studies to show the spatial layout and structure of certain industry (e.g., cultural industry, financial service industry, automobile service industry, creative industry, etc.), and to investigate its spatial evolution trend, spatial growth and diffusion pattern [10–17].

In addition to the spatial distribution pattern of a certain industry, scholars are more interested in the factors shaping its spatial distribution. The most critical elements of success for economic activities are supply–demand relationship and location. Consumption needs of the population are the primary and direct driving force for the industry development, including leisure and entertainment industry. The supply-demand relationship involves the number of potential customers, macro-economic conditions, regional consumption demand and capacity [18]. While location compasses infrastructure, business competition environment, and public service provision. Through an empirical study on cultural facilities in Shenzhen, Wei and his colleagues [19] concluded that local economy, population distribution, transportation, and policies are the key factors that influence the distribution of cultural facilities. Transportation, which links customers and facilities, is emphasized to be of great importance to the location choice of leisure and entertainment facilities [20]. Zhang and Wang [7] pointed out that various types of leisure venues are attracted by each other, leading to an agglomeration pattern and usually located nearby CBDs.

However, most of the existing studies take the leisure and entertainment industry as a whole and lack in-depth spatial quantitative analysis. In this paper, specifically focusing on karaoke bar, which is an immensely widespread leisure activity in urban China, we aim to demonstrate their spatial distribution pattern within the city by means of the GIS statistical methods, and to explore the socio-spatial factors underlying their distribution, which include population density, transportation, and location of commercial centers. The study will shed light on the planning and guidance of leisure and entertainment industry for the government, as well as enrich the empirical research on leisure and entertainment industry.

Karaoke bars are usually abstracted as points in the urban space. In this paper, the point pattern analysis and cluster analysis in GIS will be employed to illustrate whether the karaoke bars locate in a random fashion or present a certain pattern (clustered or dispersed). Based on the literature review, we hypothesize that three main factors influence the distribution of karaoke bars: population density, transportation network, and the location of commercial centers. To test the hypothesis, the density of karaoke bars' spatial correlation with the population density and transportation centrality will be measured respectively. To examine the impact of commercial center on the distribution of karaoke bars, Vironoi cells, standing for service region, will be created for each commercial center and karaoke bars will be quantize for each service region of the commercial centers.

2. Data

Nanjing, the capital of Jiangsu Province, is situated in one of the largest economic zones of China, the Yangtze River Delta. The total population is about 8.16 million in 2012 (Nanjing Statistics Yearbook 2013). Nanjing has long been the political, economic and cultural center. As a pilot city of service industry reforms, the service sector accounted for 54.38% of GDP in 2013 (Nanjing Statistics Yearbook, 2014). The fast growing service sector has become the driving force of economic development in Nanjing. Among the service industry, the leisure and entertainment sector accounted for 1% of GDP in 2013 (Nanjing Statistics Yearbook, 2014). With a total land area of 6598 km^2, Nanjing is consists of 11 districts. Considering that there are large rural and undeveloped areas within the Nanjing municipal boundary, we only choose the main urban areas as our study area in this paper, which is surrounded by the Yangtze River and the belt highway (shown by the red area in the left map of Figure 2). There are 4.38 million residents in the study area, which covers 577.8 km^2.

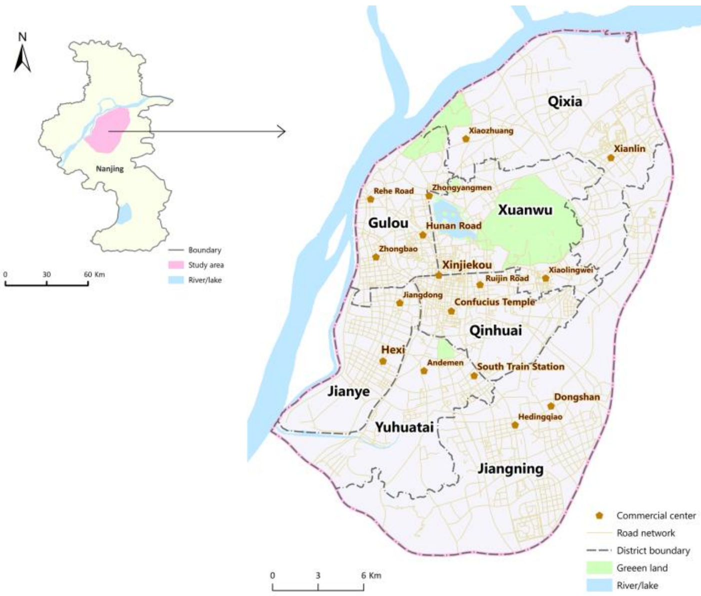


Figure 2. Nanjing municipal boundary and the study area.

The basic geographic data are drawn from the 2012 Nanjing City maps published by Jiangsu Province Surveying and Mapping Institute. Through map vectorization and coordinate calibration, the boundary of the study area, boundaries of sub-districts (a low level administrative division of a district), geographical features (river, lake and mountain) have been obtained. The road network data are from OpenStreetMap. The locations of karaoke bars extract from Baidu maps (similar to Google maps, but covering only the Great China region) and Google maps. After data examination (verify street view images), a point layer of the karaoke bars has been created. Population data are from the 2010 Population Census at sub-district level.

3. Methodology

3.1. Point Pattern Analysis—Ripley's K and L Functions

Point pattern analysis is one of the most fundamental concepts in geography and spatial analysis. It evaluates the pattern, or distribution, of a set of points on a surface [21]. There are three general patterns: random, uniform, and clustered. Ripley's K-function [22] is a common technique to estimate the presence of spatial dependence among points. Ripley's K-function is defined as:

$$K_{(r)} = \frac{1}{\lambda} \sum_{i \neq j} I(d_{ij} < r)/n \quad (1)$$

where λ is the average density of points (generally estimated as n/A, where A is the area of the region containing all points), and $I\ (d_{ij} < t)$ is the number of points within the circle with the ith point as the center and r as the radius. Under the assumption of complete spatial randomness (CSR), $K_{(r)}$ is equal to πr^2. For simplication, the K-function is transformed into a standardized version, L-function, which is defined as:

$$L_{(r)} = \sqrt{\frac{K\,(r)}{\pi}} - r \quad (2)$$

If $K_{(r)} > \pi r^2$ (or $L_{(r)} > 0$), this indicates a clustered point pattern. $K_{(r)} < \pi r^2$ (or $L_{(r)} < 0$), on the contrary, implies that the point pattern is regular.

3.2. Cluster Analysis—Nearest Neighbor Hierarchical Clustering

Cluster analysis seeks to group a collection of objects into "clusters", such that those within each cluster are more closely related to one another than objects assigned to different clusters [23]. It is a main task of exploratory data mining, as well as a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image processing, information retrieval, and bioinformatics [24]. Hierarchical clustering is a method of cluster analysis, which seeks to build a hierarchy of clusters by connecting "objects" to form "clusters" based on their distance [25]. Hierarchical clustering can be agglomerative (starting with single elements and aggregating them into clusters) or divisive (starting with the complete data set and dividing it into partitions) [26].

One of the simplest agglomerative hierarchical clustering methods is single-linkage clustering, also known as the nearest neighbor technique. The defining feature of this method is that distance between groups is defined as the distance between the closest pair of objects, where only pairs consisting of one object from each group are considered [27]. The distance between clusters X and Y, $D(X, Y)$ is described by the following expression:

$$\mathrm{D}\,(\mathrm{X},\ \mathrm{Y}) = \min_{x \in X,\ y \in Y} d\,(x,\ y) \quad (3)$$

where object x is in cluster X, and object y is in cluster Y. Assuming that there are N objects in region A, the algorithm of nearest neighbor hierarchical clustering performs the following steps:

i. Each object is seen as a cluster, containing one single object. Calculate the $N \times N$ distance matrix D contains all distance $d(i, j)$.
ii. Define a distance threshold, which is a certain value set manually or can be the expected mean distance of the features given a random pattern:

$$\overline{D}_E = \frac{1}{2}\sqrt{\frac{A}{N}} \quad (4)$$

iii. Merge clusters whose distances are within the set distance threshold.
iv. Update the distance matrix between the newly formed clusters.

v. Repeat Step 3 to further merge clusters until no more clusters can be merged.

3.3. Spatial Autocorrelation—Bivariate Moran's I

Spatial autocorrelation can be defined as the coincidence of value similarity with locational similarity, which is used to detect patterns of spatial association [28]. Global measure for spatial autocorrelation, *Moran's I*, is a single value which applies to the entire study area [29]; while its corresponding local indices, *local Moran's I*, also known as LISA (local indicators of spatial association), show the value calculated for each observation unit, allowing researchers to explore local variations in spatial dependency [30].

In addition to *univariate Moran's I*, which focuses on the spatial clustering of observations in terms of a single variable, *bivariate Moran's I* has been put forward to capture the relationship between two variables, taking the topological relationship between observations into account [31]. It measures strength of the association between two variables varies over the study region. *Bivariate Moran's I* between variables k and l for observation unit i is defined as:

$$I_{kl}^{i} = z_{k}^{i} \sum_{j} W_{ij} z_{l}^{j} \tag{5}$$

where W_{ij} is the spatial weight matrix between observation unit i and its neighboring units j, $z_k^i = (x_k^i - \overline{x_k}) / \delta_k$ is the standardized form of the value of variable k in the observation unit i, and $z_l^j = (x_l^j - \overline{x_l}) / \delta_l$ is the standardized form of the value of variable l in the observation unit j_s.

This statistic gives an indication of the degree of linear association between the value of one variable at a given unit i and the average of another variable at neighboring units j_s. The results of *Bivariate Moran's I* is a map with four categories of spatial correlation: two positive spatial correlation, namely spatial clusters (High–High and Low–Low) which relate to values physically surrounded by neighboring units with similar values, and two negative spatial correlation, or called spatial outliers (High–Low and Low–High) which relate to values surrounded by neighboring units with dissimilar values [32].

3.4. Multiple Centrality Assessment—Network Centrality

In regional and urban planning, centrality, or "accessibility" and "proximity" (the terms are often used interchangeably), has entered the scene, stressing that "some places are more important than others because they are more central" [33]. The Multiple Centrality Assessment (MCA) is a tool that evaluates the spatial distribution of centrality over geographic systems like urban street system and outputs the results graphically [34]. The first step to operate MCA is to translate the spatial system into a graph. For instance, in a network of intersections and paths, intersections are abstracted into nodes and paths into edges which are defined by two end-nodes.

The main characteristics of the MCA are: (1) use a primal approach of the network; (2) anchor all the measures in the real network of street; and (3) define the centrality by a collection of indexes [35]. Depending on what is the notion of "being central", three indexes are usually used to characterize the shape of a network: closeness, betweenness, and straightness [36].

3.4.1. Closeness Centrality (C^C)

"Closeness measures the accessibility of a node, in the way that closer a node to the others, more accessible it is" [36]. The closeness index of node i is defined as:

$$C_i^c = (N-1) \, / \sum_{j \neq i}^{N} d_{ij} \tag{6}$$

where d_{ij} is the shortest paths from node i to node j, and N is the number of all nodes in the network. Actually, closeness is the reciprocal of the mean of the shortest paths from node i to other nodes. Thus, the more central a node, the lower its average distance to all other nodes.

3.4.2. Betweenness Centrality (C^B)

Beyond the closeness index, the interaction between two distant nodes depends on the nodes belonging the path which links them. These nodes have a strategic location for the control and the influence of the flows [35]. "Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes" [37]. The betweenness centrality of node i is:

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j \neq k \neq i}^{N} \frac{n_{jk}(i)}{n_{jk}} \quad (7)$$

where n_{jk} is the total number of shortest paths from node j to node k, and $n_{jk}(i)$ is the number of those paths that pass through node i. This index measures the volume of flow on a node.

3.4.3. Straightness Centrality (C^S)

Starting from the hypothesis that "the communication between two nodes is better when the path is straight", the straightness centrality assesses the extent to which the shortest path between node i to other nodes diverts from the straight line linking them [35]:

$$C_i^S = \frac{1}{(N-1)} \sum_{j \neq i}^{N} \frac{d_{ij}^{\text{Eucl}}}{d_{ij}} \quad (8)$$

in which d_{ij}^{Eucl} is the Euclidean distance between node i and node j, and d_{ij} is the network distance between node i and node j. The smaller the difference between Euclidean distance and network distance, the more "straight" central the node is.

3.5. Spatial Interpolation—Kernel Density Estimation

Spatial interpolation is "the procedure of estimating the value of properties at unsampled sites (or cells in a raster) within the area covered by existing observations" [38]. Kernel density estimation is an interpolation technique that evaluates the probability density function of a random variable [39]. "Kernel density estimation involves placing a symmetrical surface over each point, evaluating the distance from the point to a reference location based on a mathematical function, and summing the value of all the surfaces of that reference location" [40] (Figure 3).

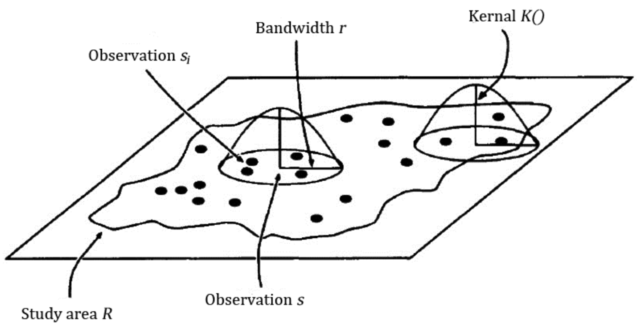


Figure 3. Kernel density estimate for points in plane.

The estimated density of point s is defined as:

$$f(s) = \frac{1}{r} \sum_{i=1}^{N} K\left(\frac{d(s, s_i)}{r}\right) \quad (9)$$

where N is the number of existing observations, h is the smoothing parameter called the bandwidth, $d(s, s_i)$ is the distance between point s and the existing observation points s_i, and $K(\frac{d(s,s_i)}{r})$ is the kernel, a weighting function, i.e., the set of rules which determines how much weight a point close to the center should be given relative to one near the edge of the circle. As seen in the equation, the results are influenced by three factors:

1. The distance between the points.
2. The bandwidth, which is very important to control the degree of smoothing applied to the data. If the bandwidth is small, we will obtain an under smoothed estimator, with high variability. On the contrary, if it is big, the resulting estimator will be over smooth and farther from the function that we are trying to estimate [41].
3. The kernel weighting function. Although there is a range of kernel functions, empirical studies indicate very little difference in the results between different weighting functions. We will use the quadratic polynomial kernel function in the following analysis.

4. Spatial Distribution Pattern of Karaoke Bars

As of the end of 2012, there are 530 karaoke bars in the study area. Table 1 shows the number and density of karaoke bars in each district. Considering that the size of each sub-district varies, using the number of karaoke bars per 1 km^2 and per 10,000 residents make it comparable between sub-districts. Within the study area, there are, on average, 0.92 karaoke bars per 1 km^2 and 1.21 karaoke bars serving for 10,000 residents. Comparing each district, Gulou and Qinhuai are the two districts with concentration of karaoke bars, respectively, 2.97 and 2.72 karaoke bars per 1 km^2 area, while in terms of the size of serving population, there are more karaoke bars in Jiangning and Qixia districts. To go beyond the administrative boundaries of sub-districts, we divide the study area into grids (300 m × 300 m). Figure 4 demonstrates the number of karaoke bars in each grid. Clearly, karaoke bars concentrate at the city center, Xinjiekou as well as Hunan Road, and just a few of them scatter in suburban areas.

Table 1. The number and density of karaoke bars in study area.

Sub-District	The Number of Karaoke Bars	The Number of Karaoke Bars per 1 km^2 Area	The Number of Karaoke Bars per 10,000 Residents
Gulou	146	2.97	1.15
Jianye	25	0.54	0.62
Qinhuai	134	2.72	1.33
Xuanwu	57	0.76	0.87
Yuhuatai	27	0.40	0.94
Jiangning	67	0.41	2.44
Qixia	74	0.59	1.57
Study area	530	0.92	1.21

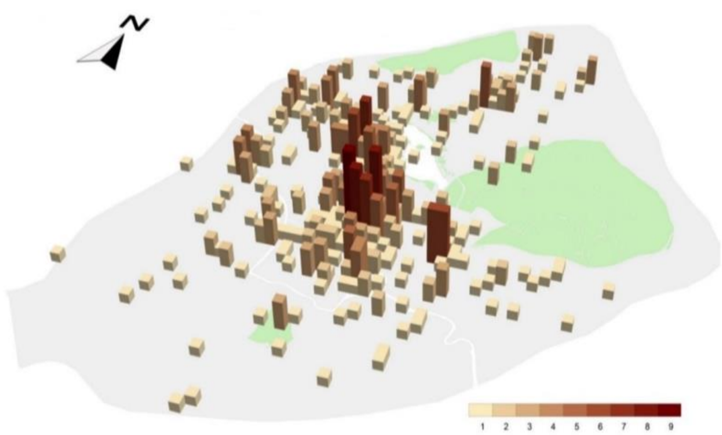


Figure 4. The distribution of karaoke bars in the study area of Nanjing.

To determine whether the distribution of karaoke bars exhibits a systematic pattern over the study area rather than being randomly distributed, L-function, the standardized version of Ripley's F-function (1976), is used to examine the spatial dependence based on distances of each karaoke bar from one another. As shown in Figure 5, the blue curve is the observed *L* value of karaoke bars in the study area. It is always larger than the expected *L* value (green curve) for any given distance threshold, indicating that the distribution of karaoke bars is in general clustered. However, the extent of clustering differs. When the distance threshold is set less than 2.5 km, the extent of clustering is positively associated with the distance. The peak of observed *L* value appears when the distance threshold is between 2.5 and 3.5 km, showing that karaoke bars are highly clustered at this distance scale. Afterwards, the level of clustering decreases with the increase of distance threshold.

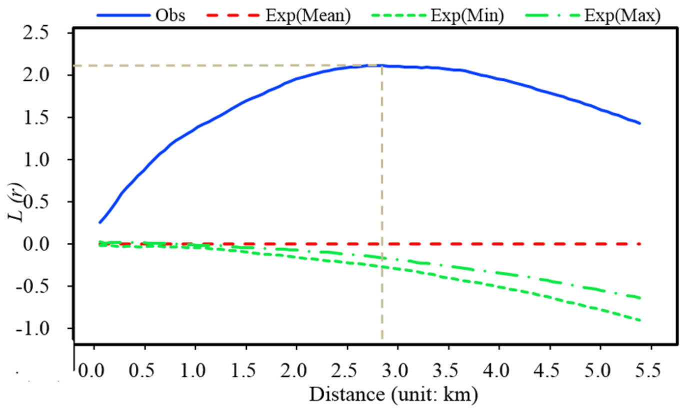


Figure 5. L-function of karaoke bars.

As the K-function has verified that the karaoke bars are spatially clustered, we continue the analysis by making hierarchical clustering to visually illustrate the concentrate areas of karaoke bars. There are eight identified clusters in the study area (Figure 6). Each cluster has more than 10 karaoke bars, and the average area of the eight clusters is 0.35 km^2. In general, the karaoke bars in the study area concentrate intensely in certain small areas.

Based on the location and characteristics of the clusters, we classify these clusters into the following categories. (a) Custom-oriented clusters: It is the nature of leisure and entertainment services to distribute at the areas with more potential customs. The 1st cluster shown in Figure 6 is located at Hunanlu sub-district, which ranks the second in terms of the population size in main urban areas of Nanjing. Besides, Nanjing University, East south University and some other institutions of higher education are located nearby, with a large number of local residents and young consumers; (b) CBD-oriented clusters: The business center is a hub of the city of all kinds of commercial activities, including shopping, dining, entertainment, etc. The 5th, 7th and 6th clusters in Figure 6 are located at Confucius Temple CBD, Dongshan CBD and Ruijin Road respectively. Confucius Temple CBD is a traditional commercial center in Nanjing, as well as a leisure and tourist center, attracting an immense number of passenger flows. Thereby, karaoke bars, as a kind of entertainment venues, are also gathering here. The Dongshan CBD is the only trading, commercial and leisure center in Dongshan satellite city, with karaoke bars also clustering here. Ruijin Road is proximate to Xinjiekou CBD, sharing and transferring the functions of Xinjiekou CBD. Therefore, it is also a concentrate area of karaoke bars; (c) Transportation-oriented clusters: The 2nd and 6th clusters in Figure 6 are distributed at Xinjiekou CBD and Maigaoqiao, respectively. Xinjiekou is located at the intersection of four important avenues, having six bus stops with over 30 bus lines crossing, and is the interchange station of subway line 1 and 2. Maigaoqiao, situated in the north of the main urban areas of Nanjing, is the northern originating station of subway line 1, which makes it irreplaceable as a transportation node especially for the northern area. The traffic advantages of these two places attract karaoke bars to concentrate there; (d) Business-oriented clusters: Karaoke bars are not only for recreation, but also become a popular venue for casual business meetings. The 3rd and 4th clusters are located at Hongwu Road and Zhenghong Commercial Pedestrian Street, where various financial institutions, consulting firms, and a large number of corporation headquarters are clustered. The demand for causal business venues results in the agglomeration of karaoke bars in these areas.

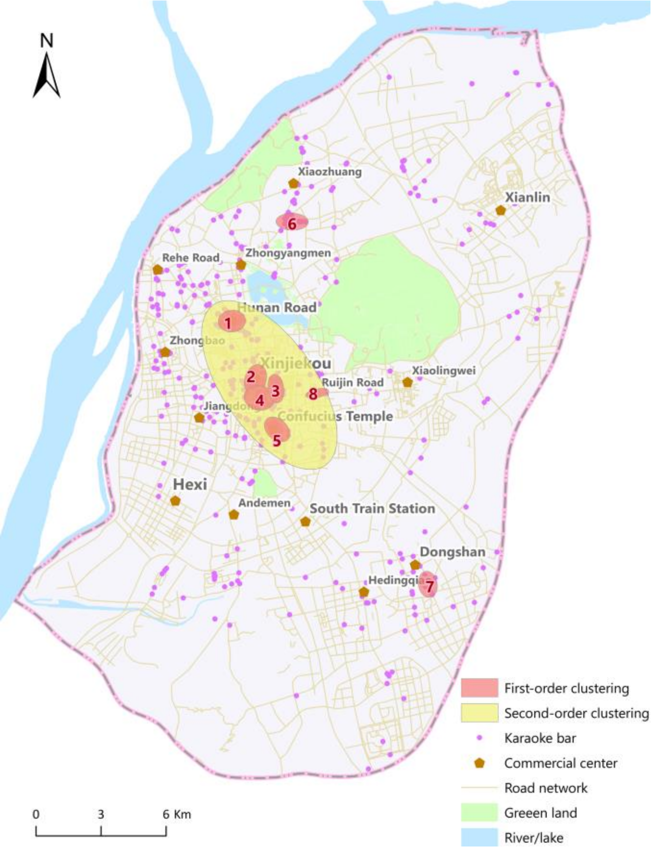


Figure 6. Clusters of karaoke bars.

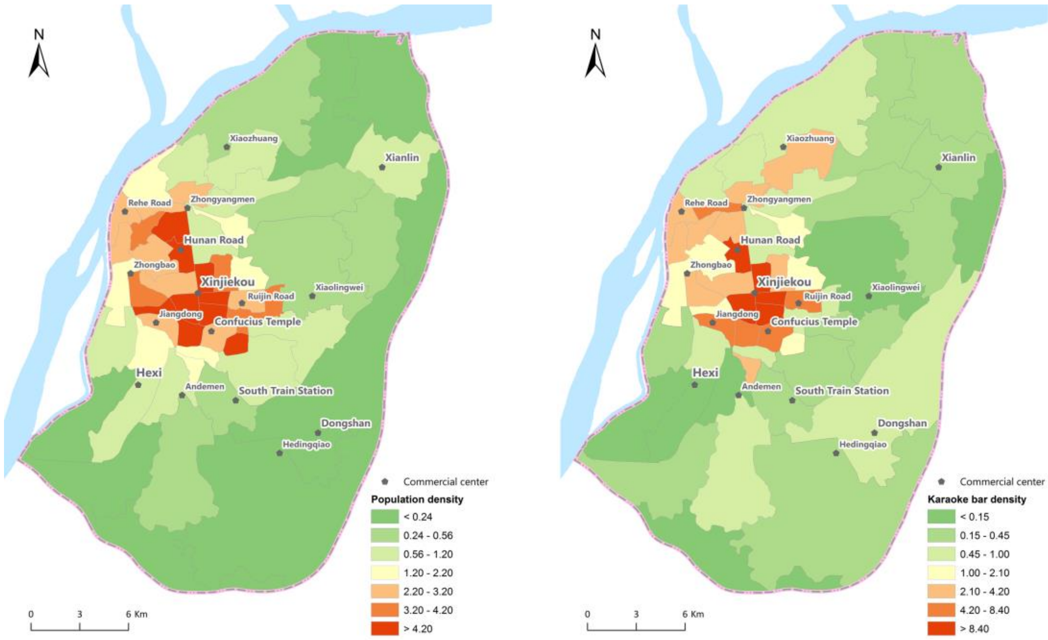
5. Socio-Spatial Factors' Influence on the Distribution of Karaoke Bars

Karaoke bars, similar as other commercial services, seek a location to obtain as much benefit as possible, taking a variety of spatial factors into account [42], therefore exhibiting a specific spatial distribution pattern [43]. To explain the spatial clusters of karaoke bars identified in the previous section, analyses were carried out to quantify the relationship among karaoke bars, population density, transportation network density, and commercial center.

5.1. Population Density

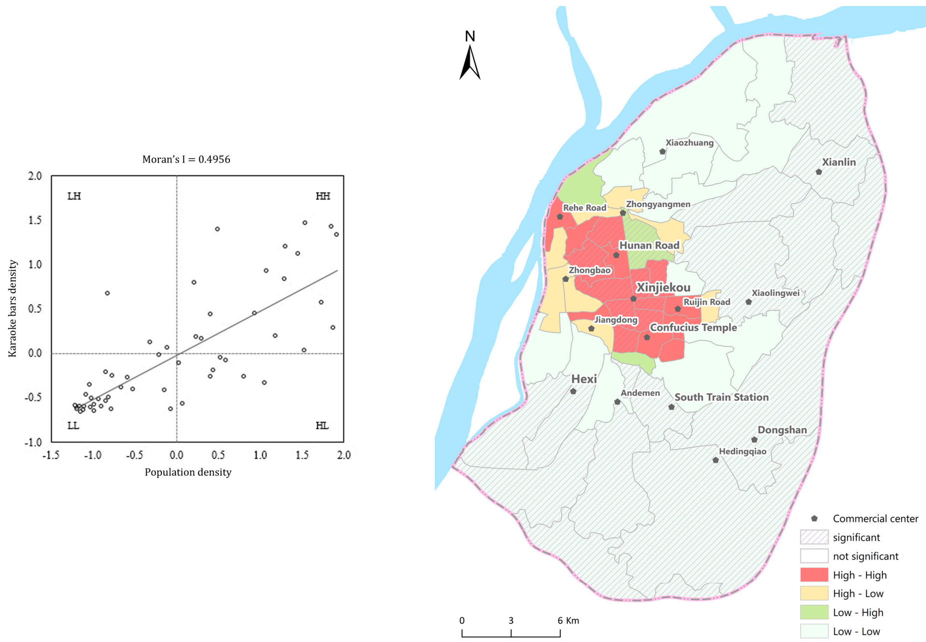
As argued in the literature review, consumption needs of the population are the primary and direct driving force for the development of leisure and entertainment industry. Chinese cities are undergoing rapid urbanization, and the development of areas within the cities is far away from even. Population density, to some extent, reflects the spatial variations in land price and economic structure, which may also exert influence on the distribution of karaoke bars.

The population data are from the 2010 Population Census, which only takes residents into account excluding floating population (without local hukou, or reside less than six months). The total number of residents is 4.37 million, distributed in 54 sub-districts within the study area. As seen from Figure 7a, the residents highly concentrate at the old city town, with the density over 32,000 residents per km^2, and there is a clear pattern that the density decreases gradually with the increase of the distance from the city center. The density of karaoke bars presents a similar pattern, with high values appearing in the city center, but less regularity in the suburb areas (Figure 7b).



(a) Population density (b) Karaoke bar density

Figure 7. Density comparison at sub-district level: (**a**) Population density; (**b**) Karaoke bar density.



(a) Moran scatter plot (b) Spatial clusters and outliers

Figure 8. Spatial correlation between population density and karaoke bars: (**a**) Moran scatter plot; (**b**) Spatial clusters and outliers.

To further assess the spatial association between population density and the density of karaoke bars, we employed *Bivariate Moran's I* statistic in GeoDa. A rook weights matrix was created, which defines a sub-district's neighbors as those sub-districts with shared borders. The value of

Global Bivariate Moran's I (the slope of the scatter plot) is 0.4956, indicating a positive association between the distribution of the population and karaoke bars (Figure 8a). Four quadrants in the scatter plot correspond to four different types of association. The first and the third quadrants represent spatial clusters, HH (High–High) and LL (Low–Low), respectively. The second and the fourth quadrants represent spatial outliers, LH (Low–High) and HL (High–Low), respectively. Figure 8b demonstrates these clusters and outliers spatially. There are 43 sub-districts with positive association (in the first or the third quadrants) and many of them are significant, meaning that karaoke bars are in general distributed at the places where more people reside. As expected, the HH is mainly located in the urban core, and LL is found in the suburban areas, whereas there are fewer sub-districts in which the density of the population and karaoke bars are negatively associated. The only significant spatial outlier (low–high) is found in Xuanwu sub-district. It is a popular tourist area, having Xuanwu Lake and Jiming Temple, attracting numerous tourists, as well as residents from outside of Xuanwu sub-district. Therefore, although the density of local residents is comparatively low, the number of potential customers is abundant, so the distribution of karaoke bars is dense on the contrary.

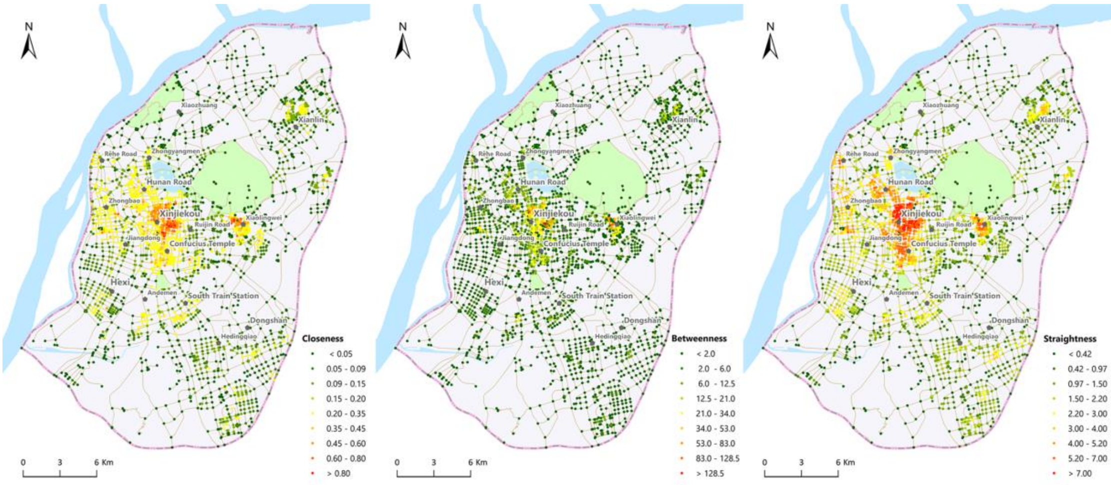
5.2. Road Network

Urban road network constitutes an important factor in urban systems, not only connecting consumers and economic activity venues, but also limiting and influencing the spatial layout of these commercial facilities. Accessibility, which can be transformed into a better visibility and popularity [35], is definitely a crucial factor for location choice of any kinds of commercial activities including karaoke bars. In Multiple Centrality Assessment, accessibility in the urban environment can be measured by the road network's centrality, based on the assumption that a central area is more accessible. As discussed in the Methodology, three indexes are commonly used to measure "centrality": closeness, betweenness, and straightness. Before calculating each of the indexes, we firstly need to construct the road network by extracting the ends and intersections of each road to form the nodes and edges of the network. Given the fact that there are different levels of roads with varied capacities, a corresponding weight is assigned to each road level, thereby we have following calculation formulas for each index (Table 2).

Table 2. Calculation formulas for centrality indexes.

Centrality Index	Calculation Formula
Closeness	$C_i^c = (N-1) / \sum_{j \neq i}^{N} \left(d_{ij} \times W_j\right)$
Betweenness	$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j \neq k \neq i}^{N} (\frac{n_{jk}(i)}{n_{jk}} \times W_j)$
Straightness	$C_i^S = \frac{1}{(N-1)} \sum_{j \neq i}^{N} \frac{d_{ij}^{\text{Eucl}}}{d_{ij}}$

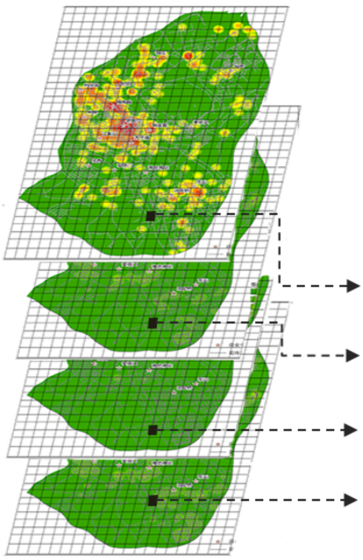
Based on the formulas, each road's closeness, betweenness, and straightness are calculated and illustrated in Figure 9. Centrality closeness reflects the proximity of a node to other nodes in the network. As shown by Figure 9a, closeness of the study area presents a clear "core–periphery" pattern. The roads with high value of closeness are located at Xinjiekou, along East Zhongshan Road. Centrality betweenness measures the frequency of a node being passed through by the shortest path between two other nodes. The higher the value of betweenness, the more influential is the node acting as a connecting point. In general, the betweenness is comparatively even across the whole study area (Figure 9b). Only a few nodes around Zhujiang Road and the subway stop of Daxinggong have higher values of betweenness, indicating large traffic flow. Centrality straightness evaluates the efficiency in communication between a node and other nodes in the network, based on the assumption that "being central means being straight to the others" [34]. In the study area, there are many nodes with higher value of straightness, distributed along the East Zhongshan Road, Zhujiang Road, and South Taiping Road (Figure 9c). On the whole, the pattern of straightness is similar to that of closeness.



(a) Closeness (b) Betweenness (c) Straightness

Figure 9. Centrality of road network: (**a**) Closeness; (**b**) Betweenness; (**c**) Straightness.

To investigate the spatial association between the centrality of the road nodes and the distribution of karaoke bars, we need to interpolate the value of each discrete point onto a continuous raster layer by using kernel density estimation (KDE) (see Section 3.5). As the bandwidth substantially affects the degree of smoothing applied to the data, we experimented with a series of bandwidths, including 600 m, 680 m, and 800 m. Afterwards, we chose to use 680 m as the bandwidth to plot the estimated density of karaoke bars and the centralities of the road network to each cell (50 m × 50 m) on the created raster layers (Figure 10).



Spatial correlation matrix between the KDE of karaoke bars and centralities of the road network

	Karaoke bar	Closeness	Betweenness	Straightness
Karaoke bar	1.000	0.473	0.363	0.527
Closeness	0.473	1.000	0.951	0.970
Betweenness	0.363	0.951	1.000	0.919
Straightness	0.527	0.970	0.919	1.000

Figure 10. Illustration of raster layers with KDEs.

The matrix in Figure 10 shows the Moran's I between each pair of KDEs. Although the three centrality indexes are devised from different perspectives and are computed differently, their values are highly correlated (the values of Moran's I among them are 0.951, 0.970, and 0.919, respectively), comprehensively reflecting the centrality of the road network. In general, the spatial distribution of karaoke bars is found to be positively associated with the centrality of the road network, whereas the extent of the association varies for the three centrality indexes. Among them, the straightness centrality is the most related, with Moran's I of 0.527, and closeness centrality is also relatively high, with Moron's I of 0.473. It implies that whether the road is straight to others and whether the road is close to others play a crucial role in the location choice of karaoke bars. However, the betweenness centrality, which measures the traffic flow, is less important, probably because the places with high betweenness may experience traffic congestion and lack of parking spaces.

5.3. Commercial Center

Commercial centers act as the main carrier for various kinds of commercial activities, ranging from the buying and selling of goods and services in retail business, wholesale buying and selling, financial establishments, and wide variety of services that are broadly classified as "business". Commercial centers in a city are an important component of urban structure, having a considerable impact on the choice of the consuming activities. To explore the spatial association between commercial centers and the location of karaoke bars, we firstly partition the urban space into service regions of each commercial center.

In geography, Voronoi diagrams are usually used to answer nearest neighbor queries, while in economic geography, Voronoi diagrams can be used to stimulate the division of the service area of commercial centers. When determining the service area of the commercial center, it is essential to consider the influential range of commercial centers, as well as the traffic network.

In terms of scale, service radius, targeted customs, etc., there are three different levels of commercial centers in Nanjing: city level, secondary city level, and district level (Table 3). These three levels of commercial centers were given weights 1, 0.8, and 0.5, respectively, to indicate their influential range. To take into account the traffic network, Voronoi diagrams were created based on distance along the traffic network instead of the Euclidean distance. Employing the "Voronoi diagram" toolbox in SANET, Voronoi diagrams were created for each commercial center (Figure 11).

Table 3. Commercial centers in Nanjing.

Level	Number of Commercial Center	List of Commercial Centers
City level	2	Xinjiekou, Hexi
Secondary city level	5	Hunan Road, Confucius Temple, South Train Station, Dongshan, Xianlin
District level	9	Zhongyangmen, Ruijin Road, Xiaozhuang, Rehe Road, Zhongbao, Jiangdong, Andemen, Xiaolingwei, Hedingqiao

Source: Nanjing Commerce Bureau.

Table 4 summarizes the area of each Voronoi cell of the each commercial center and the number of karaoke bars located in each Voronoi cell. Whether the number of karaoke bars matches the service area of commercial centers implies the influence of commercial centers on the location choice of karaoke bars. The relationship between the karaoke bars and the commercial centers is categorized into three types: (1) Karaoke bars cluster in the commercial centers with higher level and larger influence, such as Xinjiekou, Hunan Road, and Confucius Temple. These areas have substantial influence on the market, attracting consumer groups with relatively fixed consumer habits; (2) In the process of urban development, commercial space is refined into different functional centers to meet the needs of the urban economy. Therefore, being in line with the function and industrial type of commercial centers is also an important factor to be considered when choosing location for business. For instance, the targeted group of the South Train Station commercial center is non-resident population, so the main industries are restaurant, hotel, and other service industry. Besides, the State Council issued the "Entertainment Management Regulations", in which entertainment venues are restricted to be located in train stations, airports, and other places with a dense crowd. Consequently, there are only four karaoke bars located within the service area of the South Train Station commercial center; (3) The number of karaoke bars is associated with the location and the level of commercial centers. In general, there are more karaoke bars in commercial centers that are located in the urban center than in periphery areas. Furthermore, there are more karaoke bars found in higher level commercial centers.

Table 4. Statistics of karaoke bars in each Voronoi diagram of the commercial centers.

Commercial Center	Area (km^2)	Number of Karaoke Bars	Commercial Center	Area (km^2)	Number of Karaoke Bars
Xinjiekou	9.76	100	Ruijin Road	19.34	30
Hexi	57.06	8	Xiaozhuang	76.36	46
Hunan Road	9.50	64	Rehe Road	10.98	33
Confucius Temple	10.28	44	Zhongbao	12.47	17
South Train Station	26.33	4	Jiangdong	12.41	26
Dongshan	50.47	39	Andemen,	43.00	20
Xianlin	77.39	27	Xiaolingwei	58.50	10
Zhongyangmen	28.39	40	Hedingqiao	75.58	22

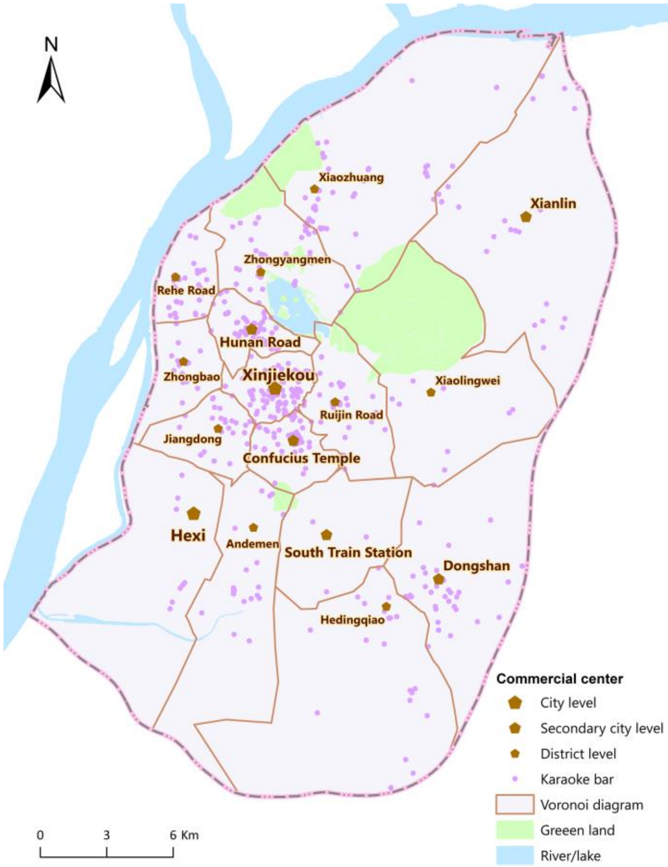


Figure 11. Network Voronoi diagram of commercial centers in Nanjing.

6. Conclusions

With the economic development, the advancement of cultural quality, and prolonged leisure time of urban residents, a large number of commercial entertainment places have sprung up in urban areas of China. Karaoke bar is among the most popular entertainments. This paper, focusing on the karaoke bars in Nanjing, has examined their spatial distribution in the urban space and their spatial association with population density, road network, and commercial centers.

Overall, karaoke industry is thriving in Nanjing, with 350 karaoke bars distributed in the study area. Spatial uneven is evident. Most of the karaoke bars are found in urban core areas. Through point pattern analysis, it has been verified that the karaoke bars are highly clustered at the distance range of 2.5–3.0 km. Based on the location and characteristics of the clusters, four types of cluster are identified: custom-oriented clustering, CBD-oriented clustering, transportation-oriented clustering, and business-oriented clustering.

The distribution pattern of karaoke bars is closely associated with the location-specific social features. Among them, population density is a significant one. Karaoke bars are in general clustered in the places where more people reside, and these places are usually located in the city center. The accessibility of road network, measured by three "centrality" indexes closeness, betweenness, and straightness, is found to play a crucial role in guiding the location of karaoke bars. To assess the impact of commercial centers on the distribution of karaoke bars, the urban space has been partitioned into service regions of 16 commercial centers using Voronoi diagrams. In general, commercial centers with higher level and better location attract more karaoke bars.

The severe uneven distribution of karaoke bars may lead to an imbalance between the recreation demands and supply. Xinjiekou is a major commercial center in the city. This single-center spatial structure may result in highly concentration of commercial services and traffic congestion. Therefore, supporting the development of existing commercial centers in suburbs, such as Xianlin, Dongshan, and Hexi, would decentralized the entertainment service function of central areas, thus easing pressure on the old city space, as well as improving the level of entertainment service in suburbs. In addition to supporting the development of commercial center, taking advantage of the subway to locate karaoke bars is another strategy. Subway, as an important travel mode, induces business activities, human resources, information and other resources gathering along the subway lines, forming a huge passenger, business, and capital flows. In the study area, there are three subway lines. There are several subway stops in the suburbs that have already demonstrated aggregation effects for business activities. Entertainment and leisure industry may make full use of the function and resource of subway stops and locate their venues nearby the subway stops in the suburbs, which will not only reduce the competition pressure, but also diminish the regional differences.

Acknowledgments: This project was supported by the National Natural Science Foundation of China [41571377] and the Program for New Century Excellent Talents in University [NCET-13-0280].

Author Contributions: Can Cui and Jiechen Wang conceived the research and wrote the paper; Zhongjie Wu, Jianhua Ni and Tianlu.Qian collected the data, designed and performed the data analysis, and prepared the figures and tables.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Molitor, G.T. The next 1,000 years: The 'big five' engines of economic growth. *Futurist* **1999**, *33*, 13.
2. Godbey, G. *Leisure and Leisure Services in the 21st Century*; Venture Publishing Inc.: State College, PA, USA, 1997.
3. Fung, A. Consuming karaoke in China: Modernities and cultural contradiction. *Chin. Sociol. Anthropol.* **2009**, *42*, 39–55. [CrossRef]
4. April, T.P.L. The Role of Karaoke Box Servicescapes Play in Customer Satisfaction. Ph.D. Thesis, Hong Kong Baptist University, Hong Kong, China, 2005.
5. CE.cn. 1991: Karaoke Started a New Epoch of Entertainment (January 9). Available online: http://views.ce.cn/fun/corpus/ce/zsrs/200901/09/t20090109_17917725.shtml (accessed on 23 October 2015).
6. Wu, Z.J.; Rui, Y.K.; Jia, J.; Ni, J.H.; Wang, J.C. Analysis on the spatial distribution and expansion characteristics of urban entertainment and leisure venues: A case Study on KTV. *Geo-Inf. Sci.* **2015**, *17*, 590–597. (In Chinese)
7. Zhang, B.; Wang, X.Z. The study of spatial structures of the urban commercial entertainment places. *Sci. Geogr. Sin.* **2007**, *27*, 853–858. (In Chinese)
8. Hu, J.X.; Zhong, A.P. Spatial cluster and urban development of popular cultural and entertainment in Hankou. *Wuhan Univ. J. (Hum. Sci.)* **2012**, *65*, 25–32. (In Chinese)
9. Yang, G.L. Study on the influence of residents' recreational activities on leisure industry. *Hum. Geogr.* **2003**, *18*, 18–22. (In Chinese)
10. Frank, L.; Glanz, K.; McCarron, M.; Sallis, J.; Sealens, B.; Chapman, J. The spatial distribution of food outlet type and quality around schools in differing built environment and demographic contexts. *Berkeley Plan. J.* **2006**, *19*, 79–95.

11. Hou, G.L.; Huang, Z.F.; Zhao, Z.X. Research on the formed mechanism and spatial structure of urban business recreation district. *Hum. Geogr.* **2002**, *17*, 12–16. (In Chinese)
12. Liu, H.; Shen, Y.M.; Deng, X.L. Study on the spatial pattern and structure of financial service industry in Beijing. *Hum. Geogr.* **2013**, *29*, 61–68. (In Chinese)
13. Shearmur, R. The clustering and spatial distribution of economic activities in eight Canadian cities. *Int. J. Entrep. Innov. Manag.* **2007**, *7*, 223–250. [CrossRef]
14. Wall, G.; Dudycha, D.; Hutchinson, J. Point pattern analyses of accommodation in Toronto. *Ann. Tourism Res.* **1985**, *12*, 603–618. [CrossRef]
15. Xue, D.Q.; Liu, H.; Ma, B.B. Characteristics of spatial distribution of cultural industries in urban area of Xi'an city, China. *Sci. Geogr. Sin.* **2011**, *31*, 775–780. (In Chinese)
16. Yao, L.; Zhang, M.; Wang, F. Characteristics of evolution and differences of the spatial distribution among different types of creative industries in Nanjing. *Hum. Geogr.* **2013**, *29*, 42–48. (In Chinese)
17. Zhen, F.; Yu, Y.; Wang, X.; Zhao, L. The spatial agglomeration characteristics of automotive service industry: A case study of Nanjing. *Sci. Geogr. Sin.* **2012**, *32*, 1200–1208. (In Chinese)
18. Ryder, A. The changing nature of adult entertainment districts: Between a rock and a hard place or going from strength to strength? *Urban Stud.* **2004**, *41*, 1659–1686. [CrossRef]
19. Wei, Z.C.; Zhen, F.; Shan, L.; Mou, S.J.; Ming, L.B. Study on the spatio-temporal distribution patterns of cultural facilities in Shenzhen. *Urban Stud.* **2007**, *14*, 8–13. (In Chinese)
20. Liu, Y.H.; Bai, G.R. Spatial structure of public facilities for recreations and leisure activities: A case study of metropolitan Shanghai. *Hum. Geogr.* **2006**, *21*, 6–9. (In Chinese)
21. Gatrell, A.C.; Bailey, T.C.; Bailey, P.J.; Rowlingson, B.S. Spatial point pattern analysis and its application in geographical epidemiology. *Trans. Inst. Br. Geogr.* **1996**, *21*, 256–274. [CrossRef]
22. Ripley, B.D. The second-order analysis of stationary point processes. *J. Appl. Probab.* **1976**, *13*, 255–266. [CrossRef]
23. Everitt, B.S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*, 5th ed.; John Wiley & Sons.: Chichester, UK, 2011.
24. Strickland, J. *Predictive Analytics Using R*; Lulu. Inc.: Colorado Springs, CO, USA, 2015.
25. Rokach, L.; Maimon, O. Clustering methods. In *Data Mining and Knowledge Discovery Handbook*; Maimon, O., Rokach, L., Eds.; Springer: New York, NY, USA, 2005; pp. 321–352.
26. Strickland, J. *Operations Research Using Open-Source Tools*; Lulu. Inc.: Colorado Springs, CO, USA, 2015.
27. Huang, Y.; Zhang, Y.; Ji, X.; Wang, Z.; Wang, S. A data distribution strategy for scalable main-memory database'. In *Advances in Web and Network Technologies, and Information Management*; Chang, K.C., Wang, W., Chen, L., Ellis, C.A., Hsu, C.-H., Tsoi, A.C., Wang, H., Lin, X., Yang, Y., Xu, J., Eds.; Springer: Berlin, Germany, 2009; pp. 13–24.
28. Anselin, L.; Bera, A.K. Spatial dependence in linear regression models with an introduction to spatial econometrics. In *Handbook of Applied Economic Statistics*; Ullah, A., Giles, D.E.A., Eds.; CRC Press: New York, NY, USA, 1998; pp. 237–290.
29. Moran, P.A. Notes on continuous stochastic phenomena. *Biometrika* **1950**, *37*, 17–23. [CrossRef] [PubMed]
30. Anselin, L. Local indicators of spatial association-LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [CrossRef]
31. Lee, S.I. Developing a bivariate spatial association measure: An integration of Pearson's r and Moran's I. *J. Geogr. Syst.* **2001**, *3*, 369–385. [CrossRef]
32. Matkan, A.A.; Shahri, M.; Mirzaie, M. Bivariate Moran's I and LISA to explore the crash risky locations in urban areas. In Proceedings of the Conference of Network-Association of European Researchers on Urbanisation in the South, Enschede, The Netherlands, 12–14 September 2013.
33. Wilson, A.G. *Complex Spatial Systems: The Modelling Foundations of Urban and Regional Analysis*; Routledge: London, UK, 2000.
34. Porta, S.; Crucitti, P.; Latora, V. Multiple centrality assessment in Parma: A network analysis of paths and open spaces. *Urban Des. Int.* **2008**, *13*, 41–50. [CrossRef]
35. Produit, T. A Novel GIS Method to Determine an Urban Centrality Index Applied to the Barcelona Metropolitan Area. Master's Thesis, Politecnico di Milano, Human Space Lab, Milano, Italy, 2009.
36. Crucitti, P.; Latora, V.; Porta, S. Centrality in networks of urban streets. *CHAOS* **2006**, *16*, 015113. [CrossRef] [PubMed]

37. Ji, S.; Li, W.; Srivatsa, M.; He, J.S.; Beyah, R. Structure based data da-anonymization of social networks. In *Information Security*; Chow, S.S.M., Camenisch, J., Hui, L.C.K., Yiu, S.M., Eds.; Springer: Basel, Switzerland, 2014; pp. 237–254.
38. Waters, N.M. Spatial Interpolation I, Unit 40. 1989. Available online: http://ibis.geog.ubc.ca/courses/klink/gis.notes/ncgia/u40.html (accessed on 1 February 2016).
39. Bowman, A.W.; Azzalini, A. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*; Oxford University Press: Oxford, UK, 1997.
40. Levine, N. CrimeStat II: Spatial Modeling, Part III. 2002. Available online: https://www.icpsr.umich.edu/CrimeStat/files/CrimeStatChapter.8.pdf (accessed on 5 March 2016).
41. Guidoum, A.C. Kernel estimator and bandwidth selection for density and its derivatives. 2015. Available online: https://cran.r-project.org/web/packages/kedd/vignettes/kedd.pdf (accessed on 26 February 2016).
42. Sadahiro, Y. A PDF-based analysis of the spatial structure of retailing. *GeoJournal* **2001**, *52*, 237–252. [CrossRef]
43. Li, Z.Z. Coupling analysis on business location and population distribution: The case study of Baiyin city. *J. Lanzhou Univ. (Soc. Sci.)* **1993**, *2*, 22–27. (In Chinese)