Article

Open Data Based Urban For-Profit Music Venues Spatial Layout Pattern Discovery

Xueqi Wang and Zhichong Zou *

Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology, School of Architecture, Harbin Institute of Technology, Harbin 150001, China; 19s034101@stu.hit.edu.cn
* Correspondence: Zhichong.Zou@hit.edu.cn; Tel.: +86-451-8628-1137

Abstract: The spatial pattern of music venues is one of the key decision-making factors for urban planning and development strategies. Understanding the current configurations and future demands of music venues is fundamental to scholars, planners, and designers. There is an urgent need to discover the spatial pattern of music venues nationwide with high precision. This paper aims at an open data solution to discover the hidden hierarchical structure of the for-profit music venues and their dynamic relationship with urban economies. Data collected from the largest two public ticketing websites are used for clustering-based ranking modeling and spatial pattern discovery of music venues in 28 cities as recorded. The model is based on a multi-stage hierarchical clustering algorithm to level those cities into four groups according to the website records which can be used to describe the total music industry scale and activity vitality of cities. Data collected from the 2018 China City Statistical Year Book, including the GDP per capita, disposable income per capita, the permanent population, and the number of patent applications, are used as socio-economic indicators for the city-level potential capability of music industry development ranking. The Spearman’s rank correlation coefficient and the Kendall rank correlation coefficient are applied to test the consistency of the above city-level rankings. The results are 0.782 and 0.744 respectively, which means there is a relatively significant correlation between the scale level of current music venue configuration and the potential to develop the music industry. Average nearest neighbor index (ANNI), quadrate analysis, and Moran’s I are used to identify the spatial patterns of music venues of individual cities. The results indicate that music venues in urban centers show more spatial aggregation, where the spatial accessibility of music activity services takes the lead significantly, while a certain amount of venues with high service capacity distribute in suburban areas. The findings can provide decision support for urban planners to formulate effective policies and rational site-selection schemes on urban cultural facilities, leading to smart city rational construction and sustainable economic benefit.

Keywords: music venues; spatial layout pattern; internet-accessible data; hierarchical clustering

1. Introduction

Social economists including Edward Banfield and Daniel Bell argued the value of culture to produce incentives for economic growth [1,2]. In many cities, a cultural and creative industry cluster is viewed as a panacea for economic and environmental survival and prosperity [3,4]. The music industry, which belongs to one branch of creative industries, directly or indirectly produces music cultural products and includes commercial and artistic enterprises as well as public and non-profit organizations. Comparing to traditional industries with the manufacturing industry included, the music industry has significant advantages: (1) lower dependence on other industries resulting from the shortness of the music industrial chain, (2) less necessity to build hardware facilities due to the liquidity of musical functions, (3) ability to receive economic benefit both rapidly and effectively.

Since the 1990s, China’s national economy has maintained rapid and stable development. With the rapid growth in the GDP per capita, the consumption pattern of urban
residents transformed from subsistence to a comfortable life. Therefore, the demand of Chinese consumers for cultural products began to increase rapidly [5], and the music cultural products are no exception, including concerts, instruments, and so on. The music industry is a broad concept, and the concerts belong to the music industry with other forms of musical events. Live music is performed in a range of venues and settings: concert halls and other performing arts venues; mass entertainment and sports venues; festivals of all sizes; local halls and community centers; through to busking on the streets [6].

However, not all cities have the foundation to develop the music industry, and conditions and opportunities for developing the creative and cultural industry are uneven. Studies on cultural consumption mainly focus on the empirical analysis of the influencing factors of cultural consumption, and the conclusions are varied. Some scholars conclude that residents’ income has a significant positive effect on the demand of cultural consumption [7], Colbert declares that education and occupational status are positively related to cultural consumption [8], while other scholars find that consumers’ age, gender, social status, quality of cultural products and services, and geographical factors all have certain influences on cultural consumption [9,10]. On the other hand, a large number of research studies focus on regional differences in cultural industry development in China. The condition to develop the music industry of different cities is affected by the quality of living, openness, population density, geographical conditions as well as the urban development strategy, historical conditions, and cultural deposits [11–14]. Therefore, the potential of developing the music industry shows a great difference in different China’s cities. Only when decision-makers and urban planners understand the otherness of this potential, they could propose an urban macro strategy of industrial development scientifically and rationally. Based on appropriate urban policies, the spatial systematic planning scheme can meet the real urban demand and gain the benefits both effectively.

A profound understanding of the current status of the music industry and the entity of space for it is necessary. This study focuses on the otherness of these dimensions among China’s cities. Moreover, the interactive relationship between the music industry and the urban economy is verified in this study. Therefore, commercial music activities benefit a city both directly and indirectly. Studies on commercial music activities and the music venues are valuable because it is a reflection of the music industry with the market size and maturity. Hidden spatial layout pattern discovery is one of the main targets of this study. The disparity of the patterns among different city levels is further extracted and discussed. By discovering the hierarchical difference in the patterns of music venues, the pattern of demand for music activities may be found out. The contribution could deepen the understanding of the music industry market and the supply–demand relationship of China’s cities. A more rational decision can be made on urban music industry development for smart city construction and sustainable economic benefit with a better understanding.

2. Literature Review

Studies on urban cultural facilities are copious. The discovery of spatial layout patterns and verification of the influencing factors of urban cultural facilities is always attractive. One of the focuses on the spatial layout patterns of urban service facilities is the spatial aggregation analysis, for the reason that this type of spatial pattern is a reflection of spatial self-organization, which impacts the urban economy, residents’ life, employment, transportation, and other important socioeconomic aspects. However, spatial aggregation is mainly studied as a complete subject on commercial facilities, due to the relevance to direct economic benefits and the quality of life of urban residents [15]. As for urban cultural facilities, the spatial clustering method is mainly applied to measure the equity and effectiveness to meet citizens’ needs. Spatial autocorrelation analysis with Moran’s I, Getis-Ord Gi* and K-function are frequently-used methods to study the spatial aggregation pattern with quantifying the degree of aggregation [15–17].

Another focus on spatial layout pattern of urban cultural facilities is mainly on spatial accessibility, for the reason that the achievement of equity in the geographical distribution
of urban public facilities is a goal of paramount importance to urban planners, who must analyze whether and to what degree their distribution is equitable [18]. Therefore, spatial accessibility is widely used in emergency services, transportation, education, medical, and other types of urban facilities for spatial assessment besides cultural facilities. A variety of methods are commonly applied to measuring spatial accessibility, such as interviews, supply-to-demand ratio, distance to the closest facility, Kernel Density, gravity model, and floating catchment area method. Park et al. analyze the accessibility of the public library with the descriptive and statistical method and network-based distance measure based on GIS to discover the determining factors for library use in Lake County [19]. Donnelly uses geographic information systems to study variations in library accessibility by state and by the socio-economic group at the national level [20].

However, research on the spatial layout pattern of urban music venues is insufficient compared to other types of cultural facilities such as parks and libraries. Shanshi Li et al. selected Beijing as the study area for the spatial distribution pattern of music tourism resources with spatial analysis methods in which they find out music resources in Beijing spatially attach urban parks and commercial centers. Besides this, the music venues in Beijing show similarity to the catering industry in spatial distribution [21]. Ying Jing et al. used point pattern analysis and cluster analysis with multi-sources geospatial data in Wuhan, Central China to examine the spatial distribution of leisure venues including music venues, and explored its underlying dominating factors. They draw a conclusion that the clusters formed by these venues are mostly distributed in urban centers, and a small part of them are in suburban areas [22].

These studies provide useful insights into the spatial layout pattern of urban cultural facilities. The data sources and methods being used also have great value for posterity research. In terms of data acquisition for urban facility analysis in previous studies, a portion of data comes from government records including facility lists at an early stage, with on-the-spot investigation additional to increase the creditability of data materials. In recent years, geoscience big data based on Internet spatiotemporal mega data have brought new ideas and methods, helping us to understand and quantitatively analyze the spatial and temporal pattern characteristics of complex natural regions and socioeconomic systems [23]. Point of interest (POI) data, which contain information of coordinates, are able to satisfy the demand of large scale and high precision in urban research, and it is available to obtain from web mapping companies freely and effectively. However, the data sources of the studies above have some limitations: (1) Data from the government are not available to everyone, and data from field investigations are costly and biased. (2) POI data may lack sufficient and necessary attributes for specific research purposes. POIs are collected by data companies with the main purpose of navigation. Therefore, POI data have the attribute of geographic coordinates. However, in this study, other dimensions of information are expected in research data: the attributes which can reflect the socioeconomic activities held in/around those POIs are essential. With the limitations of the research data, input parameters of spatial accessibility analysis of cultural facilities are mainly based on experience, which is not sufficient to describe the real demand for urban cultural products. Therefore, what information the data can provide to depict cultural demands and configuration is essential to research conclusions. On the other hand, studies on spatial layout patterns of music cultural facilities are mainly focused on the urban scale, and the conclusions are pertinent to one city separately, leaving global-perspective patterns of the whole country out of consideration.

This paper aims to study the spatial layout pattern of music venues based on open data. Open data have the advantages of totally free and wide accessibility, which expands the possibilities of the urban study area. Music activities ticket transaction websites are used as an open data source in this paper. The records can provide information of related commercial activities in different cities, whose attribute is highly relative to the realistic demand. The commercial live music activities are hosted by for-profit music venues, which are the main research subjects of our research.
3. Materials and Methods

This study proposed a methodological workflow that included three essential steps: clustering-based analysis of city-level music activity ranking and urban socioeconomic indicator ranking, correlation tests of the two rankings, and spatial layout patterns discovery of individual cities. In city-level music activity ranking, cities were clustered according to websites’ records of the music activities. Based on city-level music activities ranking, the spatial layout pattern of music venues was further discovered with descriptive and statistical methods. Spatial aggregation patterns and spatial accessibility patterns are studied. With the comparison of cities’ spatial patterns at different levels, similarities and differences can be summarized. Figure 1 provides the flow chart of the analytical framework and logical procedure. The methods will be deeply explained according to the procedure in the following sections.

![Analytical framework and logical procedure.](image)

The open data sources for the two rankings are different. Data of indicators for socioeconomic ranking can be extracted from the 2018 China City Statistical Year Book. However, website data are used for city-level music activity ranking, and both data capturing and data preprocessing work are more complicated. As Figure 2 shows, data preprocessing contains data fusion and normalization. The technical process is further illustrated in the following section.
3.1. Data Capturing: Data Source Selection and Data Scraping

Open data is increasingly important to urban researches. Multiple types of services are covered by ticketing websites, including commercial music events and activities, such as music festivals, concerts, musicals, and operas. Two consumer-oriented music events ticketing websites were used in this research: Show Start [24] (https://www.showstart.com/ accessed on 1 June 2021) and Sky Wheel [25] (https://www.moretickets.com/ accessed on 1 June 2021). Both websites are the most popular ticketing platforms for music activities. The website operation is stable and mature with new records updated every day. China’s cities are all listed on the homepage of the websites, but some of them only have a blank web page. Cities with blank pages may have no activities in the short term, and they were not studied in this paper. These two websites have historical activity records tracing back to around two years. Data from the two websites are complementary to be reliable and abundant. Attributes of music activities, including the name of activities, music venue, showtime, ticket price, and popularity (how many times have been viewed and how many viewers liked to attend) are provided on both websites. These attributes can directly reflect the vitality of commercial music activities, which is the ranking standard in the next step.

Figure 3 provides the diagram of data capturing work. Library rvest was used to access HTML and collect data in R language. DOM Parsing method was applied during capturing. R program can retrieve the dynamic content from client-side scripts. Relying on the CSS selectors of the web page, the relevant fields containing the target information were found more quickly. R scripts were created separately for data collection from two websites. Both scripts were made into executable files. By scheduling a task through the operating system, the executable files run every day from 1 October 2019 to 30 December 2019 for data collection. Due to the small size of data, all the records for music activities were local-saved in the CSV format. Each city has its own folder for data, and the city for music activities is identified by a unique id. The attribute of popularity was removed on account of the difficulty during data fusion, even though it is valuable to describe the vitality of activities. Some records have no ticket price information, which means these activities are free of charge. Since the proportion of records with a missing value (including ticket price) was too small to affect the total sample number, the processed data were returned with incomplete records removed. A number of the activities recorded from both websites...
overlapped, and the music activities were uniquely identified with the following strategy: the records were regarded to represent the same activity if the name of venues and holding time both agreed. If a discrepancy in ticket price appeared from the two websites for the same activity, the mean value was calculated to represent the final price of this record.

A total of 101 cities have activities information according to Appendix A, Table A1, and the activities range from 21 September 2016 September to 30 December 2019. The maximum sample size reaches 2311. However, not all cities participate in city-level ranking work. The rule for the further ranking work was defined as follows: cities should have more than 10 records on either website during the scraping work to avoid bias from the small sample size. In total, 28 cities were filtered for music activity ranking work. Appendix A, Table A2 shows the number of records and music venues of these cities. Data on websites updates every day.

An official report on 2015 Beijing music activities is published online by the China Association of Performing Arts [26] (http://www.capa.com.cn/ accessed on 1 June 2021). The report shows that the average value of music activities held in Beijing in 2015 is 4.70. According to the data captured from two websites, 1719 music activities in Beijing are recorded. The average value of activities number updated per day is 11.75. Data in October, November, and December were used to calculate the average value of music activities number held per day due to the sufficiency of records in these months. The average value was 3.79 from the two websites, and the activities coverage percentage in Beijing was estimated by dividing the two average values above. The result shows that the data have a high coverage rate for music activities in Beijing at 80.63% (Table 1). Activities on both websites are displayed nationwide, which leads to the assumption that music activities in different cities are recorded in a unified standard. Therefore, 80.63% was estimated to represent the overall coverage rate of all the music activities. The data sample from both websites is shown in Appendix A, Tables A3 and A4.
3.2. Data Preprocessing: Normalization of Index Representing Current Configuration

Indicators representing the scale of music venues can be divided into two categories: the number of venues and the scale of corresponding music activities, where the latter is comprised of indicators of frequency and ticket price of music activities. A massive discrepancy in the magnitude of variables may exist: for instance, the venue number remains in single digits in some cities with relatively backward development, while it is common that the price of a concert costs more than 100 yuan. Therefore, rescaling data into values between 0 and 1 before other possible calculations are necessary. The normalization formula of the numbers of music venues is given by:

\[
\text{Numbers\_of\_venues\_Normalization\_Value}(n) = \frac{\text{Total\_numbers}(n)}{\sum_{n=1}^{N} \text{Total\_numbers}(n)}
\]

where \(n\) represents each city while \(N\) represents all of the cities on record. The total frequency of activities holding time is obtained through numerical statistics of information from websites.

Similarly, the way to normalize musical activity holding frequency which pertains to the discrete variable is to use division as well, where the numerator is the total frequency of activities holding time within each city, and the denominator represents the total frequency of activities holding time of all the cities.

\[
\text{Frequency\_Normalize\_Value}(n) = \frac{\text{Total\_Frequency}(n)}{\sum_{n=1}^{N} \text{Total\_Frequency}(n)}
\]

Concerning the price of activities, we use the following formula for normalization:

\[
\text{Fare\_Normalize\_Value}(n) = \frac{\text{Sum\_of\_unique\_Prices}(n)}{\sum_{n=1}^{N} \text{Sum\_of\_unique\_Prices}(n)}
\]

The total sum of the unique prices of all the music events in city \(n\) from June 2016 to June 2020 is used to represent the music activities price level of city \(n\), while the scaling denominator is the sum to add up the total sum of the unique prices of each city. The variable, Sum\_of\_unique\_Prices is calculated by adding up the unique values of the prices, and here’s an example to illustrate this variable: if there are three activities someday in a city, and the price of admission is 100 yuan, 100 yuan, and 150 yuan separately, then the sum of 100 yuan and 150 yuan, which equals 250 yuan, is regarded to be the sum of unique prices of music events in this city during the day.

3.3. Clustering-Based Ranking Procedure

Machine learning algorithms are suitable to rank cities according to their music activities from the websites’ data and socioeconomic conditions from the yearbooks. Generally, there are two branches of machine learning methods: classification and clustering. Classification methods aim at identifying the category of a new observation among a set of categories on the basis of a labeled training set [27]. Logistic regression, Random Forest, Support Vector Machine, and Neural Networks are classic classification algorithms. However, the classification method is not applicable for urban classification for the following reasons: (1) Urban classification work aims to grade the cities based on index values reflecting the demand of commercial music activities, which is at variance with the supervised learning feature of classification methods: pre-defining training sets is necessary. (2) The

### Table 1. The figure featuring open data (Beijing for instance).

| City     | Average Value of Activities Number Added Per Day | Average Value of Activities Number Per Day from Websites | 2015 Average Value of Activities Number per Day | Activities Coverage Rate |
|----------|--------------------------------------------------|--------------------------------------------------------|-----------------------------------------------|--------------------------|
| Beijing  | 11.75                                            | 3.79                                                   | 4.70                                          | 80.63%                   |
number of groups is unknown in advance, which is essential to determine according to the results for division. However, the number of groups should be clear while using classification algorithms. On the contrary, clustering methods need no prior knowledge input. Moreover, clustering methods are the most efficient data processing methods to identify homogeneous aggregates inside a heterogeneous group [28]. Clustering methods can be summarized into three categories: hierarchical clustering, partitional clustering, and Bayesian clustering.

In order to propose the grading rule considering both factors: conciseness and validity synthetically hierarchical clustering algorithms, one of the most classical clustering methods is applied in this research. Hierarchical clustering is a method of cluster analysis, attempting to create a hierarchy of clusters by grouping points into their specific clusters according to their distances [29]. Hierarchical clustering starts with \( k = N \) clusters and proceeds by merging the two closest ones into one cluster, obtaining \( k = N - 1 \) clusters. The process of merging two clusters to obtain \( k-1 \) clusters is repeated until we reach the desires number of clusters \( K \) [30]. The Euclidean distance is feasible here to find which clusters to merge. In this study, three variables (musical activities holding frequency, the total price of musical activities, and the number of venues) are selected for hierarchical clustering computation.

K-means algorithm is applied to perform cross-validation of the urban division result obtained by hierarchical clustering. K-means is one of the classic partitional clustering algorithms, which can aid the investigator in obtaining a qualitative and quantitative understanding of large amounts of N-dimensional data by providing reasonably good similarity groups [31]. However, the ranking result from the hierarchical clustering algorithm, rather than K-means, is adopted for further study for the reason that the description of the result by the hierarchical method is more intuitive and graphical, which makes it easier to form a judgment on the rationality of the result. Even though the clusters obtained have no ranking information, it is quite simple to determine the order of city groups, as the result of the great gaps in variable values among different clusters.

In order to validate the effectiveness of the classification results, this study also put forward the city classification according to the potential to develop the music industry based on the authoritative indicators (GDP per capita, disposable income per capita, the permanent population, and the number of patent applications) from the yearbooks.

3.4. Correlation Test of the Ranking

The Spearman’s Rank Correlation Coefficient and The Kendall Rank Correlation Coefficient are two measures applied to support the relationship between music activities ranking and urban socioeconomic indicator ranking. Two ordinal, interval, or ratio variables are expected for both methods, therefore they are suitable for the test of these two city-level rankings. Both methods have their specific characters and advantages, while Spearman’s Rank Correlation Coefficient assesses how well an arbitrary monotonic function can describe a relationship between two variables, without making any assumptions about the frequency distribution of the variables [32], and The Kendall Rank Correlation Coefficient can easily be generalized to other combinatorial structures such as weak orders, partial orders, or partitions [33]. Comparing to Spearman’s Rank Correlation, Kendall Rank Correlation has a smaller gross error sensitivity and a smaller asymptotic variance, which makes the correlation test more robust and efficient. The correlation test result of ranking will be more reliable with the complementary of both methods. The range of correlation coefficients is \(-1\) to \(+1\), where the correlation coefficient of \(+1\) indicates a great positive association of ranks, and the value of \(-1\) indicates a perfect negative association of ranks. The closer coefficient is to 0, the weaker the association between the ranks.
3.5. Spatial Layout Pattern Discovery

3.5.1. Spatialization

Due to the verification of urban classification in the previous step, we can obtain valid information that ranks the cities into different groups based on the vitality of music activities. With the ranking result, it is natural to analyze the characteristics of each group to further excavate the similarities and differences, and to get the spatial layout pattern of urban music venues.

On account of descriptions representing addresses to hold music activities on the Internet that can be applied in space, it is necessary to spatialize nonspatial data in advance. Characters of addresses’ names may represent stadiums, music restaurants, or music bars. Therefore, the map API (Application Program Interface) is necessary for obtaining longitude and latitude information with provided geocoding services. The current large-scale Internet companies in China, such as Baidu, Gaode, and Tencent, provide map API services. The data used in this study are derived from the API services of AMAP of Gaode company. Gaode’s AMAP service is the alternative to Google’s for Apple company’s map service being used in China. Furthermore, the musical activities scale can be reflected by variables including frequency and price level, therefore the weight of activities-scale factor of each site lying in urban space is accessible, which plays an important role in mining local distribution patterns on music space.

3.5.2. Spatial Aggregation Pattern Discovery

ANNI, quadrat analysis, and Global Moran’s I were applied to discover the spatial aggregation pattern of for-profit music venues at the city level. ANNI and quadrat analysis methods can describe the distribution pattern of the venues. Variables for city-level music activities ranking reflect the vitality level of music activities. By mapping variables of activities frequency and ticket price to music venues in space, the vitality of each venue can be represented by summing up the activities frequency and price level normalized values. Moran’s I was applied to discover the spatial distribution pattern of music activities vitality. The Average Nearest Neighbor Index (ANNI) was a complicated tool to measure precisely the spatial distribution of a pattern and see if it is regularly dispersed, randomly dispersed, or clustered. This method was developed by P. Clark and F. Evans in their research [34] to depict the pattern of distribution of the population of plants or animals. The result of the average nearest neighbor result is a rational number no less than 0. It is interpreted that if the result is less than 1, the tendency of cluster occurs, otherwise spatial points are more likely to be distributed discretely with outcome over 1. ANNI is widely applied in spatial point pattern analysis, and the extensiveness in application leads to its high efficiency during calculation as the extension program in common software, including ArcGIS and R. Quadrat analysis is a kind of variance analysis. It uses the measures of the number of points inside quadrats located inside the region. It is firstly developed by Greig Smith for ecological investigation [35]. The spatial pattern of points is estimated by Variance Mean Ratio (VMR). If VMR is greater than 1, an aggregation distribution is shown. The size of unit space has a great effect on results. The quadrate should be twice as large as the average area of each point [36]. Urban space was partitioned into quadrants by administrative boundary. Spatial autocorrelation was applied to detect patterns of spatial association, namely, to present locational similarity, and Moran’s I was the indicator for depicting spatial autocorrelation, involving global Moran’s I and local Moran’s I [22]. As a single value, Global Moran’s I was employed to the whole study area [37]. The expected Global Moran’s I value ranges from $-1$ to $+1$ after normalization. Values of I exceed 0 indicate positive spatial correlation, and values below 0 indicate negative spatial correlation. If values of I are close to 0, the spatial distribution is random. In general, the clustering results are more statistically accurate when the number of input features increases. Referring to the input number threshold proposed by Arcgis official document, the cities with more than 30 music venues recorded were applied with ANNI and Global Moran’s I.
Furthermore, the association between the spatial clustering degree of urban music venues and the vitality of music activities can be studied with correlation analysis. Pearson’s $r$ was selected for correlation analysis of ANNI results and variables including the number of venues, the total frequency, and the price level of music events in each city. The Pearson product-moment correlation coefficient is a measure of the strength of the linear relationship between two variables. Pearson’s $r$ can range from $-1$ to $1$. An $r$ of $-1$ indicates a perfect negative linear relationship between variables, an $r$ of $0$ indicates no linear relationship between variables, and an $r$ of $1$ indicates a perfect positive linear relationship between variables [38].

3.5.3. Spatial Accessibility Pattern Discovery

In addition, to analyze the spatial aggregation level, measuring the spatial accessibility of the music venues within city limits is an intuitive approach to highlight the spatial layout patterns of urban music venues. The two-step floating catchment area (2SFCA) method, pioneered by Luo and Wang is one of the effective tools for spatial accessibility measurement [39,40]. 2SFCA is commonly applied for healthcare. However, the demand for music activities is not as urgent as which of medical or education service. The commercial live music activities recorded on the website are available for audiences all over the country, and the tourism demand is regarded to be similar to music activities demand to some level. Y. Niu developed a formula to describe the tourism demand of the tourism-generating region, indicating that tourism demand is positively associated with per capita income and population [41]. Referring to the tourism demand research mentioned above, the new method based on the original 2SFCA method is improved, which specifically targets the entertainment services with external parameters in this paper. The formula is as follows:

$$A_i^F = \sum_{j \in \{t_{ij} \leq t_0\}} R_j = \frac{\sum_{j \in \{t_{ij} \leq t_0\}} S_j}{\sum_{j \in \{t_{ij} \leq t_0\}} \frac{P_i}{H_iN}}$$

(4)

where $A_i^F$ represents the accessibility $F$ of each spatial unit $i$, $S_j$ represents the service capacity of music venue $j$, $P_i$ stands for the population within spatial unit $k$, $H_i$ represents the house price level of spatial unit $i$, and $N$ is the empirical parameter. However, by applying this formula, the paper can get access to the relative relation of spatial accessibility of city-wide spatial units. It is not necessary to set the value of the parameter $N$ precisely, therefore it is set to constant 1. The research unit $i$ refers to the city township (subdistrict offices) division, and the population data come from the sixth census, while the data of house price are from a house selling website (https://lianjia.com/ accessed on 1 June 2021) [42]. The website records the housing transaction price of communities all over the city, therefore we implemented the price into spatial units and averaged it as the argument $H_i$ during the calculation, whose value depends on experience reference as well. In order to simplify the calculation process, instead of time distance, Euclidean distance—the second quartile of the sum of spatial units centroids—was regarded as the threshold.

4. Results

4.1. Ranking of Scale on Urban For-Profit Music Venues and Cross-Validation

The input variables are the number of sites, total price, and total frequency of musical events through a hierarchical clustering for the vitality of music activities. Figure 4a shows the results obtained by the hierarchical clustering method, on behalf of the vitality of music activities. The cities were clustered into four groups, which is the most appropriate group number of categories educed with Bayesian Information Criterion (BIC) via the R Language Model-based clustering package “mclust” [43]. Meanwhile, the K-means method was applied for the cross-validation of the city clustering result. The maximum number of iterations was set to be 10,000 times for algorithm convergence.
From the hierarchical result, we can see that: Beijing and Shanghai possess the greatest capacity to hold more music-related activities, which leads them to the first level. Meanwhile, Guangzhou, Chengdu, Nanjing, Shenzhen, Wuhan, and Hangzhou come in the second level, while Changsha, Tianjin, Chongqing, Xi’an, Ningbo, Kunming, Zhengzhou, Suzhou, and Hefei lie in the third level and Shijiazhuang, and Dalian, Harbin, Fuzhou, Changchun, Nanning, Zhuhai, Shenyang, Jinan, Qingdao, and Xiamen are in the last level. It is not surprising to see that the capacity of carrying music activities in Beijing and Shanghai far exceeds that in other cities, considering the intuitive massive gap in the scale of music venues. On the other hand, the K-means result has a high similarity with the city rankings described above. Cities were clustered into four groups as well. Table 2 shows the cities in different levels divided by two clustering methods. The city ranking result by K-means agrees completely with which obtained by hierarchical algorithm.

Table 2. Cities in different levels clustered by hierarchical and K-means methods.

| Level Number | Cities Divided by Hierarchical Method | Cities Divided by K-Means Method |
|--------------|--------------------------------------|----------------------------------|
| 1            | Beijing, Shanghai                    | Beijing, Shanghai                |
|              | Guangzhou, Chengdu, Nanjing, Shenzhen, Wuhan, and Hangzhou | Guangzhou, Chengdu, Nanjing, Shenzhen, Wuhan, and Hangzhou |
| 2            | Changsha, Tianjin, Chongqing, Xi’an, Ningbo, Kunming, Zhengzhou, Suzhou, Hefei | Changsha, Tianjin, Chongqing, Xi’an, Ningbo, Kunming, Zhengzhou, Suzhou, Hefei |
| 3            | Dalian, Harbin, Fuzhou, Changchun, Nanning, Zhuhai, Shenyang, Jinan, Qingdao, Xiamen | Dalian, Harbin, Fuzhou, Changchun, Nanning, Zhuhai, Shenyang, Jinan, Qingdao, Xiamen |

The clustering result reveals the disparity in the demand for commercial music activities, which represents the difference in the scale of for-profit music venues. In addition, music industry markets of cities at the top level have a larger size and higher maturity. Cities at the last levels may face the problem of supply and demand imbalances of music activities. In the premise of a clear investigation on the Chinese music industry market maturity, the grade difference on the demand of commercial music activities is directly reflected by the clustering result, which can provide decision supports on the future music industry developing strategies of cities with potential demand.
However, the ranking result above will be unconvincing without support from statistical examination or authoritative conclusions. For the purpose of confirming the result’s validity and verifying the relationship between economic development and demand for commercial music activities, a number of variables, which are highly related to the urban economy, were selected from the statistical yearbook, for the new urban ranking work. If the new urban grading result is verified to be correlated with the result early got, the purpose can be achieved.

4.2. Ranking of City-Level Social Capability of Music Industry Development

First, America’s non-profit organization, Partners for Livable Communities, summarizes the tasks of developing cultural planning into four aspects: human, economic, environmental, and social development [44]. Therefore, GDP per capita, disposable income per capita, the permanent population, and the number of patent applications are selected as variables to cluster the cities according to their potential of music industry development, on account that the first two variables represent the urban economic foundation, the permanent population partly reflects urban cultural demand, and the number of patent applications represents urban innovative capacity, which is essential for developing cultural industry with music industry included as well. The data of variables are available in the 2018 China City Statistical yearbook. By using a hierarchical clustering method, cities are clustered into three classes as Figure 4b shows, where Chongqing, Beijing, Shanghai, Tianjin, Chengdu, Shenzhen, Guangzhou, and Suzhou belong to the first level, Changsha, Ningbo, Nanjing, Hangzhou, Xi’an, Hefei, Zhengzhou, Wuhan, and Tsingtao are in the second level, while Zhuhai, Xiamen, Fuzhou, Jinan, Nanning, Shijiazhuang, Shenyang, Harbin, Kunming, Changchun, and Dalian are in the third level.

4.3. Results of Rank Correlation Test

Afterward, the Spearman’s Rank Correlation Coefficient and The Kendall Rank Correlation Coefficient were applied to test the consistency of both rankings above, and the results were 0.782 and 0.744 respectively, which means there is a relatively significant correlation between the scale level of current music venues configuration and the potential to develop music industry.

4.4. Result of Comparison with SARFT Ranking

The China State Administration for Radio, Film, and Television (SARFT) planning corporation carried out research on city rankings for the music industry in 2017. The rankings are divided into multiple works to assess different aspects of music development including music industry foundation, music education resources, music industry talent, and music industry capital. The gradation basis in this research has similarities with urban rankings for the music industry foundation. As a consequence, the ranking result for the music industry foundation by SARFT can be used as a reference for classification results achieved in this step. Urban per capita GDP, urban per capita disposable income, the number of permanent urban residents, and the number of urban music infrastructure were selected as variables for the classification process, and there are multiple channels for data acquisition including from the administrative department and statistic yearbook. The ranking results are shown in Figure 5.

The ranking shows that Beijing is in the first place in municipalities, followed by Shanghai, Tianjin, and Chongqing. Ranking in the sub-provincial city is Guangzhou, Shenzhen, Hangzhou, Wuhan, Chengdu, and Nanjing orderly. At the same time, the ranking for provincial cities has Changsha, Fuzhou, and Kunming in the top three. The rest of the results are hidden due to the secrecy provisions by SARFT, thus the complete ranking result is unavailable online. However, the authoritative ranking results can be seen to have a certain degree of similarity comparing with our clustering result of musical activities vitality.
4.5. Results of Spatial Layout Pattern Analysis

4.5.1. The Geospatial Distribution of Different Levels of Urban For-Profit Music Venues Configuration

Figure 6 depicts the geographical distributions of cities in distinct levels. The cities we did not show lack records on the ticket sales platform, which means activities in those cities are not held frequently enough to be listed on websites. From the figure, it is easy to see that all the cities with abundant music events for the record are located in the east of China. The distribution just meets the rule, in which China is divided into two parts geographically by Aihui-Tengchong Line proposed by geographer Hu [45]. A great disparity also embodies the population density and economic development between the west and east parts according to this rule. Therefore, it can be inferred that factors such as the population quantity and economic development conditions have an effect on the cultural demands of cities from another side. Cities in the northeast are all on the fourth level, from which influence factors related to the music industry can be extrapolated that the economic openness of cities in northeast China is slightly behind in comparison with that in cities from southeast China. Besides, the extreme weather in northeast China may also cause a lack of innovative talents and a suitable outdoor environment for music activities.
4.5.2. The Scale and Spatial Layout Pattern of Urban For-Profit Music Venues

By combining data from both ticket sales websites, the number of for-profit music venues in each city is obtained. The sample size of spatial aggregation pattern discovery should be over 30 to get statistically significant results. The 12 cities that participate in music activities ranking have more than 30 music venues recorded. ANNI, quadrat analysis, and Global Moran’s I were applied to these cities. ANNI and quadrat analysis were utilized to find out whether the for-profit music venues in each city show spatial aggregation in space. Moran’s I, however, was calculated to testify if the vitality of music activities in venues is more likely to aggregate. Table 3 shows the number of music venues, the ANNI, VMR, and Moran’s I of these 12 cities. As shown in Table 3, it is obvious that the number of music venues in Beijing and Shanghai significantly exceeds those in other cities. ANNI value indicates that music venues in these two cities show the most obvious spatial aggregation effect, which leads to our preliminary estimation that the number of music venues is related to their spatial aggregation level. Music venues in other cities of which ANNI value is greater than 1 are distributed randomly. However, the VMR value by quadrat analysis of all cities is manifestly greater than 1, which indicates that music venues show a great aggregation tendency in all 12 cities. On one hand, spatial points are distributed mainly in the city center and suburbs, and it leads to more blank spatial units during grid partition. On the other hand, the optimal size of quadrats is large due to the small number of music venues, which can distinctly increase the VMR value. However, both methods can reflect the common regularity that the clustering result according to music venue configuration is highly related to the actual number of music venues. Besides this, the aggregation level of music venues in the city may be significantly related to its rank number. Moreover, it may be reasonable to presume that the music venues are scattered all over the urban space, and as the number of venues for music activities increases, the venues begin to present the trend of aggregation. The configuration in the city center may increase, and the venues in the suburb are still spatially isolated. If this assumption is proved later, it can provide instructions on the real live music industrial construction. Nevertheless, further work needs to be accomplished to verify this assumption and will not be conducted in this paper. Despite missing that part, this study can confirm that indicators representing the scale of music venues are related to each other through the correlation analysis method. Furthermore, the results of the Global Moran’s Index are close to 0, which suggests that the vitality of music activities tends to be distributed in the city randomly rather than concentrating in the city center. We can put forward the hypothesis that some of the music venues in suburban areas have enough space to hold massive music shows due to the abundance of unused construction land compared with the city center, despite the traffic inconveniences the suburbs may face. The results and hypothesis can provide future construction guidance for cities in China to develop the music industry by evaluating the potential of target cities and analyzing cities with high music prosperity.

Table 3. The number and ANNI result of music cultural venues in 12 cities.

| City     | Number of Music Venues | ANNI | VMR   | Moran’s Index | Rank of Cities |
|----------|------------------------|------|-------|---------------|----------------|
| Shanghai | 117                    | 0.521| 26.369| −0.010        | 1              |
| Beijing  | 102                    | 0.664| 21.440| 0.115         | 1              |
| Chengdu  | 51                     | 0.968| 22.818| 0.107         | 2              |
| Hanzhou  | 45                     | 1.051| 15.216| −0.053        | 2              |
| Shezhen  | 44                     | 0.693| 9.595 | 0.017         | 2              |
| Guazhou  | 42                     | 0.768| 8.544 | −0.015        | 2              |
| Nanjing  | 41                     | 0.977| 13.444| 0.026         | 2              |
| Tianjin  | 34                     | 1.236| 16.124| 0.100         | 3              |
| Chongqing| 33                     | 1.159| 15.610| −0.060        | 3              |
| Xi’an    | 32                     | 1.008| 7.267 | −0.091        | 3              |
| Ningbo   | 31                     | 0.979| 9.735 | −0.097        | 3              |
| Wuhan    | 30                     | 1.963| 19.436| −0.063        | 2              |
With the normalized value of extracted information including the number of music venues, frequency of musical events, and price of musical events as well as the ANNI of each city, it is feasible to figure out if there are related-to relationships between variables determined by Pearson’s r on a national scale.

Firstly, a diagram is depicted where we regard the ANNI value as the independent variable and the other three variables as dependent variables along the y-axis. The reason for the allocation for independent and dependent variables is that the expected independent variable may reflect space attributes representing the degree of spatial aggregation, through which we can discover if there is a relevant relationship between scale and spatial pattern of music cultural facilities preliminary. Figure 7 is the scatter diagram that demonstrates the relevant relationships between different variables in an intuitive form. It can, therefore, be assumed that spatial aggregation has an impact on the total price and total frequency of music venues in each city. As for the number of music venues, it may be reasonable to assume that it has the interaction with spatial aggregation level. The result also demonstrates the presumption that for cities, those who possess more venues for music activities also show a higher aggregation.

![Figure 7. The Scatter diagram of correlation analysis.](image)

Pearson correlation determines the extent to which the values of the two variables are linear related [46]. The value range (−1, 1) indicates whether there is a linear correlation between variables. Besides this, different absolute values signify different strengths of correlation. Table 4 elaborates on the Pearson correlation coefficients between different variables. The result shows negative correlated behavior between the average nearest neighbor value and the number of sites, total price, and total frequency of musical activities in each city, with Pearson correlation coefficients of −0.613, −0.532, and −0.501 respectively. This result proves again what the first figure shows quantitatively that as the aggregation of sites increases, the quantity of sites, total price, and frequency of each city increase likewise, and the aggregation level has a moderate correlation with other variables.

|                  | ANNI Value | Site Number | Ticket Price | Events Frequency |
|------------------|------------|-------------|--------------|------------------|
| ANNI Value       | 1.000      | −0.613      | −0.532       | −0.501           |
| Site number      | −0.613     | 1.000       | 0.828        | 0.964            |
| Ticket price     | −0.532     | 0.828       | 1.000        | 0.839            |
| Events frequency | −0.501     | 0.964       | 0.839        | 1.000            |
Moreover, there is a significant positive correlation between the number of venues and total price (Pearson correlation coefficient of 0.828), total price, and total frequency (Pearson correlation coefficient of 0.839) as well as the number of venues and total frequency (Pearson correlation coefficient of 0.964). The evidence suggests an obvious result that if the city has more venues to hold music events, it is available to hold more music activities, and the average price level increases.

Furthermore, the modified spatial accessibility formula based on 2SFCA is applied to the cities in the first two grades, and the results are shown in Figure 8. It can be inferred that space in urban centers is more available for residents to obtain music cultural service. A fraction of spatial units in suburbs in some cities have higher spatial accessibility as well, due to some music venues far away from downtown with higher service capacity.

![Spatial Accessibility Maps](image)

**Figure 8.** (a–h) The spatial accessibility distribution of cities in the top two levels of classification.

### 4.5.3. The Analysis Based on the Comparison and Difference among Groups

After leveling 28 cities into four groups according to the scale of for-profit music venues, it is able to calculate the mean normalized value of each group. According to Table 5, cities in different grades varies in both aspects of music venues scale and music events scale, and the relative disparity among the average normalized value of each category on all three indicators: the number of music venues, frequency, and the ticket price of music events is similar. For instance, the average value of the three indicators of Beijing and Shanghai in the first level is 0.130, 0.162, and 0.113 respectively, while that of cities in the second level is 0.050, 0.056, and 0.063. The table indicates that at current situation city in the first level has more than twice music venues as those in the city from the second level on average, and the ratio of total frequency for holding music events remains close to that of music venues numbers, which indirectly reflect that the average frequency of each music venue differs slightly even the cities are from different levels during the study period, regardless of the diversity in the scale of the site. However, the difference in average ticket price level is within one time of that in the first level, which implies that the price of a part of music events is consistent through different cities, still, the disparity in price may reflect the gap on consumption level.
Table 5. The Average Normalized Scale Index of four levels of cities on music venues.

| Classification Level | Cities | Average Normalized Number of Music Venues | Average Normalized Frequency of Music Events | Average Normalized Ticket Price of Music Events |
|----------------------|--------|------------------------------------------|---------------------------------------------|-----------------------------------------------|
| 1                    | Beijing, Shanghai                           | 0.130                                     | 0.162                                       | 0.113                                         |
| 2                    | Guangzhou, Chengdu, Nanjing, Shenzhen, Wuhan, Hangzhou | 0.050                                     | 0.056                                       | 0.063                                         |
| 3                    | Hefei, Suzhou, Zhengzhou, Kunming, Ningbo, Xi’an, Chongqing, Tianjin, Changsha | 0.030                                     | 0.026                                       | 0.029                                         |
| 4                    | Shijiazhuang, Dalian, Harbin, Fuzhou, Changchun, Nanning, Zhuhai, Shenyang, Jinan, Qingdao, Xiamen | 0.015                                     | 0.010                                       | 0.012                                         |

Furthermore, the results in Table 5 may provide instructions in the arrangement of new music venues, since they indicate urban needs of cultural consumption. For instance, if we select a city in the fourth level as the case for music cultural facility construction in the future, corresponding strategies can be proposed on the scale of construction. In consideration of growth in actual demand, the number of new music venues should maintain within one time of the original scale, while a larger level of music activities can be introduced with a higher frequency and a higher ticket price for greater benefits.

5. Discussion and Conclusions

In this study, an open data based approach is proposed for urban music cultural facilities classification and spatial pattern discovery. Music venues for commercial music activities are the selected type of cultural facilities for instance, and two nationwide city-level rankings are introduced, which are based on the scale of the current configuration of music venues, and socioeconomic indicators selected from the yearbook of the corresponding cities, respectively. For the ranking based on music activities, two music events ticketing websites (Show start and Sky wheel) are selected as data sources, where three variables: the total number of music venues in the city, the frequency, and the ticket price level of music activities are selected to cluster and level the 28 recorded cities into four groups, by using a hierarchical clustering method and it is cross-validated by a k-means clustering method.

Ranking of the city-level potential capability of music industry development uses data collected from the 2018 China City Statistical Year Book including GDP per capita, disposable income per capita, the permanent population, and the number of patent applications are used as socioeconomic indicators. Results of correlation tests (the Spearman’s rank correlation coefficient and the Kendall rank correlation coefficient) are 0.782 and 0.744, respectively. The results indicate there is a relatively significant correlation between the scale level of current music venue configuration and the potential to develop the music industry of the 28 Chinese cities. As expected, Beijing and Shanghai are located at the first level, and cities in northeast China are mainly on the fourth level. Afterward, the music venues in each city are located in space via an API-based geocode technology.

Along with the ranking results conclusions can be drawn on both the scale and spatial aggregation pattern of music venues: (1) cities with more venues to hold music activities have much more possibilities to hold more events with higher ticket price; (2) the spatial aggregation level is related with the clustering level; (3) as the number of music venues increases, those venues are more likely to aggregate, resulting in the growth of frequency and ticket price level for music events as well. The gaps of music venues among the cities in different groups are analyzed specifically.

This study has both theoretical and practical significance. Open data are utilized widely for urban researches recently. With the open data, the music venues’ status can be quantitatively analyzed by breaking through constraints from traditional data limitations.
There are also practical significances for urban construction guidance. With the ranking results embodying music cultural demands, the scale characteristics can be used to propose a nationwide strategic plan on music venues incremental. It can also provide site selection advice on layout planning for music industry-related venues or firms qualitatively. Although the instructions are only applicable for cities in China since spatial patterns discovered in this study are the reflection of the music industry status quo of China, the approach proposed in this paper can be extended to other countries and regions.

This study suffers from some other limitations as well. Firstly, due to the imperfection of websites for data extraction, the number of cities participating in ranking is incomplete despite the related factors interpreted from this incompleteness including economic development and population density. Besides, the data may have certain limitations as well, since the acquisition time may cause contingency. In our future work, this limitation can be solved by using data from more websites to improve the comprehensiveness of music events through data fusion methods.

Secondly, the ranking method proposed in this paper is subjective inherently to some extent, which is mainly based on the urban capacity of holding commercial music activities, without evaluating music education and creation level. However, the specific indicators for leveling can be modified with the future research target. Thirdly, there are multiple types of facilities that perform functions as the complete music chain, including musical creation, live activities, music education, and so on [47], which can be located in urban space as a music culture industrial park. Nevertheless, the result obtained from this study so far is not convinced enough to guide implement of music culture industrial park construction both quantitatively and qualitatively, not to mention the instruction for site selection [48–53].

Therefore, in future work, it is necessary to further discover the spatial pattern of music venues and that of other classes of functional music cultural facilities. In addition, excavating spatial relationships among different types is significant as well. The co-location mining method will be applied to discover which types of POIs are more likely related to the venues for music activities, in other words, to filter out the types on the music industry chain in a spatial statistics way, and discover co-location distribution pattern afterward. It should be noted that factors hard to quantify are also important to be taken into consideration during the actual construction, for instance, historical issues, cultural connotation, or policy reason, and to compromise for the optimal planning scheme.

**Author Contributions:** Conceptualization, Z.Z.; Data curation, X.W.; Formal analysis, X.W.; Methodology, Z.Z.; Project administration, Z.Z.; Supervision, Z.Z.; Validation, X.W.; Visualization, X.W.; Writing—original draft, X.W.; Writing—review & editing, Z.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Educational Reform Program of Harbin Institute of Technology (XYZ2020038); the Fundamental Research Funds for the Central Universities (HIT.HSS.202108, XNAUEA575000120) and the Graduate Education and Teaching Reform Project of Harbin Institute of Technology (21HX0303).

**Data Availability Statement:** Publicly available datasets were analyzed in this study. The data can be found here: https://www.showstart.com/ accessed on 1 June 2021, https://www.moretickets.com/ accessed on 1 June 2021 and https://lianjia.com/ accessed on 1 June 2021.

**Acknowledgments:** All authors express their sincerest gratitude to website programmers involved in this study.

**Conflicts of Interest:** The authors declare no conflict of interest.
Appendix A

Table A1. The general summary of data sample from two websites.

| Original Number of Cities | Activities Time Range | Activities Number of the Most Active City | Activities Number of the Least Active City |
|--------------------------|-----------------------|-------------------------------------------|-------------------------------------------|
|                          | 2016/09/21–2019/12/30 | 2311                                      | 1                                         |

Table A2. The overall number of music activities and venues of selected cities.

| City     | Overall Number of Music Activities in Record | Overall Number of Music Venues in Record |
|----------|---------------------------------------------|----------------------------------------|
| Beijing  | 1719                                        | 102                                    |
| Shanghai | 2311                                        | 117                                    |
| Wuhan    | 705                                         | 30                                     |
| Nanjing  | 727                                         | 41                                     |
| Guangzhou| 771                                         | 42                                     |
| Shenzhen | 746                                         | 44                                     |
| Tianjin  | 330                                         | 34                                     |
| Suzhou   | 294                                         | 22                                     |
| Hangzhou | 714                                         | 45                                     |
| Ningbo   | 286                                         | 31                                     |
| Zhengzhou| 231                                         | 20                                     |
| Changsha | 270                                         | 24                                     |
| Chengdu  | 666                                         | 51                                     |
| Chongqing| 555                                         | 33                                     |
| Fuzhou   | 61                                          | 11                                     |
| Shenyang | 148                                         | 14                                     |
| Tsingtao | 104                                         | 16                                     |
| Nanning  | 112                                         | 12                                     |
| Kunming  | 234                                         | 20                                     |
| Xi’an    | 460                                         | 32                                     |
| Zhuhai   | 93                                          | 11                                     |
| Xiamen   | 169                                         | 18                                     |
| Hefei    | 211                                         | 15                                     |
| Jinan    | 222                                         | 15                                     |
| Changchun| 56                                          | 10                                     |
| Shijiazhuang | 63                      | 15                                    |
| Dalian   | 70                                          | 9                                      |
| Harbin   | 102                                         | 11                                     |

Table A3. The data sample from Sky Wheel website (scraped in 2019/10/1, Beijing as case).

| Name of Music Activities | Music Venue                           | Activities Time | Ticket Price       | Popularity          |
|--------------------------|---------------------------------------|-----------------|--------------------|---------------------|
| Classic Concert “Old memories” | Beijing Forbidden City Concert Hall | 2019.10.03 19:30 | Minimum Price: ¥86 | Viewed 17,355, Like to attend 63 |
| [Ronghao Li, Ying Na] 2019 China Great Voice | National Stadium (Bird’s Nest) | 2019.10.07 19:00 | Minimum Price: ¥205 | Viewed 277,335, Like to attend 314 |
| Yuheng Jiang “me and old friends” Concert | Cadillac Center | 2019.10.09 18:30 | Minimum Price: ¥256 | Viewed 18,675, Like to attend 43 |
| Jiayun Yu “Since 1994” Music Tour | Tango Live | 2019.10.12 20:00 | Minimum Price: ¥390 | Viewed 1485, Like to attend 17 |
| 2019 Weibo Pan “Coming home” Music Tour | Cadillac Center | 2019.10.13 19:00 | Minimum Price: ¥418 | Viewed 411,630, Like to attend 443 |
| 2019 PHARAOH Music Tour | Tango Live | 2019.10.13 20:30 | Minimum Price: ¥483 | Viewed 60, Like to attend 2 |
Table A3. Cont.

| Name of Music Activities            | Music Venue              | Activities Time       | Ticket Price          | Popularity                      |
|-------------------------------------|--------------------------|-----------------------|-----------------------|---------------------------------|
| Carly Rae Jepsen 2019 Music Tour    | Century Theater          | 2019.10.16 20:00      | Minimum Price: ¥586   | Viewed 4530, Like to attend 47  |
| Beam! Beam! Pro.3                   | Beijing Livehouse OMNISPACE | 2019.10.18 18:00      | Minimum Price: ¥354   | Viewed 45, Like to attend 0     |
| Xu Huang 2019 Sun and Iron Concert  | Beijing Livehouse OMNISPACE | 2019.10.19 20:30      | Minimum Price: ¥404   | Viewed 1455, Like to attend 17  |

Table A4. The data sample from Show Start website (scrapped in 2019/10/01, Beijing as case).

| Name of Music Activities            | Activities Time       | Ticket Price          | Music Venue              | Popularity                      |
|-------------------------------------|-----------------------|-----------------------|--------------------------|---------------------------------|
| Piano Concert by Taiwan Pianist Zhihao, Gao | Time: 2019/10/03 19:30 | Minimum Price: ¥180   | [Beijing] Beijing Concert Hall | Like to attend 63               |
| 2019LIL JET Zhengting, Lu Music Tour | Time: 2019/10/04 20:00 | Minimum Price: ¥168   | [Beijing] MAO Livehouse   | Like to attend 705              |
| Ghibli Music symphony concert       | Time: 2019/10/05 19:30 | Minimum Price: ¥280   | [Beijing] Beijing Beizhan Theatre | Like to attend 86               |
| “Ultraman: Cosmic Live Show” Music drama | Time: 2019/10/06 15:00 | Minimum Price: ¥126   | [Beijing] Beijing Theatre | Like to attend 25               |
| [Rock Brother] “Our era”            | Time: 2019/10/07 19:30 | Minimum Price: ¥120   | [Beijing] Beijing MAO Livehouse | Like to attend 540              |
| Yuheng Jiang “me and old friends” Concert | Time: 2019/10/09 18:30 | Minimum Price: ¥380   | [Beijing] Cadillac Center | Like to attend 17               |
| Harvey Mason Chameleon              | Time: 2019/10/10 19:30 | Minimum Price: ¥360   | [Beijing] Blue Note Beijing | Like to attend 50               |
| Jiayun Yu “Since 1994” Music Tour   | Time: 2019/10/12 20:00 | Minimum Price: ¥180   | [Beijing] Tango Live     | Like to attend 1348             |
| Chen Hong Concert “sound in memory 2019” | Time: 2019/10/13 20:00  | Minimum Price: ¥260   | [Beijing] Beijing Laifu Livehouse | Like to attend 30               |

References

1. Zheng, J. ‘Creative Industry Clusters’ and the ‘Entrepreneurial City’ of Shanghai. Urban Stud. 2011, 48, 3561–3582. [CrossRef]
2. Zheng, J.; Chan, R. The impact of ‘creative industry clusters’ on cultural and creative industry development in Shanghai. City Cult. Soc. 2014, 5, 9–22. [CrossRef]
3. David Bell, K.O. Cultural Policy; Routledge: Third Avenue, NY, USA, 2015.
4. Jayne, M. Culture that works? Creative industries development in a working-class city. Cap. Cl. 2004, 28, 199–210. [CrossRef]
5. Cai, R.; Wang, Y. Research on foreign and domestic policies in developing cultural creative industries. China Soft Sci. 2009, 8, 77–84.
6. Performing, A.; Association, R. Economic contribution of the venue-based live music industry in Australia. Retrieved Jan. 2011, 20, 2014.
7. Brito, P.; Barros, C. Learning-by-Consuming and the Dynamics of the Demand and Prices of Cultural Goods. J. Cult. Econ. 2005, 29, 83–106. [CrossRef]
8. Willis, K.G. A latent class model of theatre demand. J. Cult. Econ. 2012, 36, 113–139. [CrossRef]
9. Milner, A. A New Framework for Building Participation in the Arts; Rand: Santa Monica, CA, USA, 2001.
10. Grisolia, J.M.; Willis, K.; Wymer, C.; Law, A. Social engagement and regional theatre: Patterns of theatre attendance. Cult. Trends 2010, 19, 225–244. [CrossRef]
11. Gu, J.; Wu, J.; Hu, H. The Regional Characteristics and Its Influential Factors of China’s Cultural Industry: Based on the 5th and 6th census data GU. Econ. Geogr. 2013, 33, 89–114. [CrossRef]
12. Resea, R.; An, R.C.H. A Comparison of the Innovation Collaboration Network between Electronic Information Industry and Biotechnology and Pharmaceutical Industry in Pudong New Area of Shanghai City. Areal Res. Dev. 2005, 24, 5–9.
13. Zhao, J. Review on Urban Transformation Development and Cultural and Creative Industry Research. Academics 2014, 198, 223–228.
14. Chen, H. Chinese Culture Creative Industry: Developmental Actuality and Prospect. Econ. Geogr. 2008, 28, 728–733. [CrossRef]
15. Wang, T.; Wang, Y.; Zhao, X.; Fu, X. Spatial distribution pattern of the customer count and satisfaction of commercial facilities based on social network review data in Beijing, China. *Comput. Environ. Urban Syst.* 2018, 71, 88–97. [CrossRef]

16. Gao, J.; Zhou, C.; Wang, Y.; Jiang, H. Spatial analysis of Guangzhou’s public services during the transitional period. *Geogr. Res.* 2011, 30, 424–436.

17. Jana, M.; Sar, N. Modeling of hotspot detection using cluster outlier analysis and Getis-Ord Gi* statistic of educational development in upper-primary level, India. *Modeling Earth Syst. Environ.* 2016, 2, 1–10. [CrossRef]

18. Tsou, K.W.; Hung, Y.T.; Chang, Y.L. An accessibility-based integrated measure of relative spatial equity in urban public facilities. *Cities* 2005, 22, 424–435. [CrossRef]

19. Park, S.J. Measuring public library accessibility: A case study using GIS. *Libr. Inf. Sci. Res.* 2012, 34, 13–21. [CrossRef]

20. Donnelly, F.P. The geographic distribution of United States public libraries: An analysis of locations and service areas. *J. Librariansh. Inf. Sci.* 2014, 46, 110–129. [CrossRef]

21. Li, S.; Liu, J.; Huang, W. Distribution of Music Tourism Attractions in Beijing. *Resour. Sci.* 2012, 34, 381–392.

22. Liu, Y.; Jing, Y.; Cai, E.; Cui, J.; Zhang, Y.; Chen, Y. How leisure venues are and why? A geospatial perspective in Wuhan, Central China. *Sustainability* 2017, 9, 1865. [CrossRef]

23. Hu, Y.; Han, Y. Identification of urban functional areas based on POI Data: A case study of the Guangzhou economic and technological development zone. *Sustainability* 2019, 11, 3885. [CrossRef]

24. Available online: https://www.showstart.com/ (accessed on 26 May 2021).

25. Available online: https://www.moretickets.com/ (accessed on 26 May 2021).

26. Available online: https://www.showstart.com/ (accessed on 26 May 2021).

27. Jimenez-Del-Toro, O.; Otárola, S.; Andersson, M.; Euren, K.; Hedlund, M.; Rousson, M.; Müller, H.; Atzori, M. *Analysis of Histopathology Images: From Traditional Machine Learning to Deep Learning*, 1st ed.; Elsevier Ltd.: Amsterdam, The Netherlands, 2017; pp. 281–314. ISBN 9780128121337.

28. Kaufman, L.; Rousseew, P. *Finding Groups in Data an Introduction to Cluster Analysis*; John Wiley and Sons: Hoboken, NJ, USA, 1990.

29. Wang, S.; Shi, W. *Data Mining and Knowledge Discovery*; Springer: New York, NY, USA, 2012; pp. 123–142, ISBN 9783540726807.

30. Teichgraeber, H.; Brandt, A.R. Systematic Comparison of Aggregation Methods for Input Data Time Series Aggregation of Energy Systems Optimization Problems. *Comput. Aided Chem. Eng.* 2018, 44, 955–960. [CrossRef]

31. MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Oakland, CA, USA, 21 June–18 July 1967; Volume 1, pp. 281–297.

32. Hauke, J.; Kosowski, T. Comparison of values of pearson’s and spearman’s correlation coefficients on the same sets of data. *Quast. Geogr.* 2011, 30, 87–93. [CrossRef]

33. Abdi, H. Kendall Rank Correlation Coefficient. *Concise Encycl. Stat.* 2007, 278–281. [CrossRef]

34. Clark, P.J.; Evans, F.C. Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. *Ecology* 1954, 35, 445–453. [CrossRef]

35. Greig Smith, P. *Quantitative Plant Ecology*; University of California Press: Berkeley, CA, USA, 1983.

36. Taylor, P.J. *Quantitative Methods in Geography: An Introduction to Spatial Analysis*; Houghton Mifflin Company: Boston, MA, USA, 1977; pp. 146–147.

37. Moran, P.A.P. Notes on Continuous Stochastic Phenomena. *Biometrika* 1950, 37, 17. [CrossRef] [PubMed]

38. Nettleton, D. Selection of Variables and Factor Derivation. In *Commercial Data Mining*; Elsevier: Waltham, MA, USA, 2014; pp. 79–104.

39. Mcgrail, M.R. Spatial accessibility of primary health care utilising the two step floating catchment area method: An assessment of recent improvements. *Int. J. Health Geogr.* 2012, 11, 1–12. [CrossRef]

40. Wang, F.; Luo, W. Assessing spatial and nonspatial factors for healthcare access: Towards an integrated approach to defining health professional shortage areas. *Health Place* 2005, 11, 131–146. [CrossRef]

41. Niu, Y. The study on spatial linkage between the supply and demand of tourism. *J. Mater. Process. Technol.* 2018, 1, 1–8.

42. Available online: https://lianjia.com/ (accessed on 26 May 2021).

43. Gaussian Mixture Modelling for Model-Based Clustering, Classification, and Density Estimation Description. Available online: https://mlust-org.github.io/mclust/ (accessed on 26 May 2021).

44. Tu, Q.; Lin, L. Some Thoughts on Cultural Planning and Planning Culture. *Soc. Sci.* 2012, 11, 50–58.

45. Tien, H.Y. *China’s Population Struggle: Demographic Decisions of the People’s Republic of China, 1949–1969*; Ohio State University Press: Columbus, OH, USA, 1973.

46. Ratner, B. The correlation coefficient: Its values range between 1/1, or do they? *J. Target. Meas. Anal. Mark.* 2009, 17, 139–142. [CrossRef]

47. Liu, G. Study on Chinese Music Industry Value Chain. Ph.D. Thesis, China University of Geosciences, Wuhan, China, 2006.

48. Wang, D.; Qiao, C.; Liu, S.; Wang, C.; Yang, J.; Li, Y.; Huang, P. Assessment of spatial accessibility to residential care facilities in 2020 in Guangzhou by small-scale residential community data. *Sustainability* 2020, 12, 3169. [CrossRef]

49. Brown, A.; O’Connor, J.; Cohen, S. Local music policies within a global music industry: Cultural quarters in Manchester and Sheffield. *Geoforum* 2000, 31, 437–451. [CrossRef]
50. Small, M.L.; Harding, D.J.; Lamont, M. Reconsidering culture and poverty. *Ann. Am. Acad. Political Soc. Sci.* 2010, 629, 6–27. [CrossRef]

51. Xu, H.; Zhu, J.; Wang, Z. Exploring the spatial pattern of urban block development based on POI analysis: A case study in Wuhan, China. *Sustainability* 2019, 11, 6961. [CrossRef]

52. Rahman, K.M.A.; Zhang, D. Analyzing the Level of Accessibility of Public Urban Green Spaces to Different Socially Vulnerable Groups of People. *Sustainability* 2018, 10, 3917. [CrossRef]

53. Cherbo, J.M. Creative synergy: Commercial and not-for-profit live theater in america. *J. Arts Manag. Law Soc.* 1998, 28, 129–143. [CrossRef]