Article

Categorization of Green Spaces for a Sustainable Environment and Smart City Architecture by Utilizing Big Data

Qi Liu [1,2], **Hidayat Ullah** [1,3,*], **Wanggen Wan** [1,3], **Zhangyou Peng** [1], **Li Hou** [2], **Sanam Shahla Rizvi** [4], **Saqib Ali Haidery** [1,3], **Tong Qu** [1,3] and **A.A.M. Muzahid** [1,3]

[1] School of Communication & Information Engineering, Shanghai University, Shanghai 200444, China; liuqi@shu.edu.cn (Q.L.); wanwg@staff.shu.edu.cn (W.W.); zypeng@mail.shu.edu.cn (Z.P.); alisaqib@shu.edu.cn (S.A.H.); qu_tong@shu.edu.cn (T.Q.); muzahid@shu.edu.cn (A.A.M.M.)
[2] School of Information Engineering, Huangshan University, Huangshan 245041, China; houli@shu.edu.cn
[3] Institute of Smart City, Shanghai University, Shanghai 200444, China
[4] Raptor Interactive (Pty) Ltd., Eco Boulevard, Witch Hazel Ave, Centurion 0157, South Africa; sanam.shahla@raptorinteractive.com
[*] Correspondence: hidayat@shu.edu.cn

check for updates

Abstract: Urban green spaces promote outdoor activities and social interaction, which make a significant contribution to the health and well-being of residents. This study presents an approach that focuses on the real spatial and temporal behavior of park visitors in different categories of green parks. We used the large dataset available from the Chinese micro-blog Sina Weibo (often simply referred to as "Weibo") to analyze data samples, in order to describe the behavioral patterns of millions of people with access to green spaces. We select Shanghai as a case study because urban residential segregation has already taken place, which was expected to be followed by concerns of environmental sustainability. In this research, we utilized social media check-in data to measure and compare the number of visitations to different kinds of green parks. Furthermore, we divided the green spaces into different categories according to their characteristics, and our main findings were: (1) the most popular category based upon the check-in data; (2) changes in the number of visitors according to the time of day; (3) seasonal impacts on behavior in public in relation to the different categories of parks; and (4) gender-based differences. To the best of our knowledge, this is the first study carried out in Shanghai utilizing Weibo data to focus upon the categorization of green space. It is also the first to offer recommendations for planners regarding the type of facilities they should provide to residents in green spaces, and regarding the sustainability of urban environments and smart city architecture.

Keywords: spatiotemporal analysis; smart city; environment; KDE; green parks; quality of life

1. Introduction

Urban green spaces, parks in particular, offer numerous advantages for people's well-being, such as enhancing their psychological and social well-being, as well as their physical health [1–3]. Visits to green parks provide the opportunity to directly experience the extra benefits of "natural" environments, especially for people with limited contact with nature. Consequently, quantifying the visits of residents to urban green spaces, and recognizing the factors that influence their visits, are vital for urban park management and planning. It is important to measure visitation to parks and urban green spaces to recognize people's recreational interests and to define the factors that influence them. Visitor polls, on-site counters and direct observation are the traditional methods for estimating

the number of visitors [4,5]. These methods of organized observation usually pick a representative sample of urban green spaces and collect information regarding the use of green parks, for example, visitation numbers and features, as well as the activities and behavior of green park visitors. However, these approaches are typically time-consuming and site-specific, and therefore have restricted spatial coverage [6]. The advent of free social media data offers new and different methods to analyze urban green park visits. Data collected from social media, and other "big data" as reliable sources, grow in size every year, and can be utilized to investigate how the public interact with actual environments, and to accurately assess individual preferences through time and space [7].

Green areas or parks are obviously the best way to encourage people to take part in physical exercise. In this research, we analyzed visitors' spatiotemporal behavior in green parks, taking into account 122 green parks with nearly 250,000 visitor check-ins, using the location-based social network (LBSN) service named Sina Weibo, also known as Weibo [8]. Weibo was created in August 2009 and is one of China's largest social media sites, as well as a popular microblogging platform that allows people to post and share their daily activities with their circle of friends. The term "check-in" refers to a user willingly sharing their location on the LBSN, by publicizing it on a location-sensing smartphone or tablet while participating in certain activities [9].

In this research, we examined the effect of visits to green spaces on users' check-in behaviors. We utilized a sample of data to research green park-related check-in behaviors, including likes and dislikes, as a summary of the behavior or preferences of the general public in terms of various activities while using park resources. We also implemented the linear regression model by including variables that have a mild correlation with the dependent variable. Similarly, the aim of the study was to respond directly to the following questions: (1) What type of parks do people choose to visit, and what are the characteristics of each category of park? (2) How often does the time of the day influence the behavior of the general public in relation to green parks? (3) How does the season influence the behavior of the general public in terms of green park visits? (4) Does gender have an impact on visits to green park? To properly grasp the link between green parks and visitor behavior, we analyzed users' spatiotemporal trends in reference to urban green spaces by using a large-scale dataset that covered the period of July 2014 to June 2017. The research was conducted within Shanghai city in south-east China. Our main results indicate the distribution of users' check-ins in different green parks, and exciting behavioral changes based on the time of day (i.e., a 24-h period) and the day of the week. We also investigated seasonal impact on the behavior of the public in relation to green spaces to show real-life patterns. Through examining these factors, we were able to analyze the use of urban green parks and the subsequent variance across a range of temporal scales. This method of capturing visitor activity in urban green spaces is an effort to solve the drawbacks of previous studies, and it paves the way for socio-ecological research using crowd-sourced and social network data, instead of relying on the results of subjective coverage and observational data.

2. Literature Review

Urban areas are generally marked by downgraded ecosystems, rising environmental damage, higher temperature and decreasing urban green parks. Urban green spaces (UGS) are important for improving the quality of life in metropolitan areas, for balancing thermal comfort and providing an appropriate thermal budget [10], and for helping the public heal from the physical and emotional pressures of their daily lives [11]. Green spaces in urban environments are vital for improving living conditions in urban areas by improving the quality of the air and the area's esthetics, which eventually leads to higher real estate values and a reduction in energy use for cooling. UGS can also be applied as a cumulative resource for the development of sustainable urban areas [12]. UGS also provides children with much-needed space to play, which is significant for their social, cognitive and mental functioning [13].

Urbanization has become one of the worldwide agendas for development. Based upon the United Nations' Sustainable Development Goal 11 (SDG 11)—which seeks to expand human settlements and

urban communities and to ensure that they are safeguarded, strong and feasible—the United Nations expanded the sustainable development goal (SDG) Agenda by implementing the New Urban Agenda in 2016 [14]. Urban sustainable development and social justice rely on different requirements, including equitable spatial distribution, planning, environmental facilities, urban strategic planning, quality of green space, and socio-economic facilities. The social benefits issued by UGS to urban residents are crucial for maintaining and increasing the personal choices of urban residents [15]. Public urban green spaces (PUGS) are important for the mitigation of high summer surface temperatures [16], and are also essential for the elimination of pollution and for decreasing noise levels [17]. The temperature difference among urban and green areas is high in summer and low in winter. Moreover, in summer, the difference is greater during the day than during the night, while in winter, the opposite is true.

In recent years, several analytical efforts have been made to use social media data for the fields of application and urban planning, the purposes of which are dynamic, and range from basic tasks to quite complex analysis, for example, urban form and feature verification. In reality, Facebook, Weibo, Twitter, and the other big data networks are utilized to assess the mobility and behavior of people, with calculations ranging from the inter-urban to the global scales [18]. However, it would be extremely difficult to determine the finer spatial and temporal stages of these two phenomena by utilizing traditional methods, such as questionnaires or on-site observations. Furthermore, LBSN data can be utilized for socio-spatial analysis [19] using techniques such as the absorption of user's tweet components, or by exploring their feelings and disparities over space and time [20], also taking into consideration health factors such as physical exercise or diet. Campagna [21] implemented the idea of "Social Media Geographic Information" as a means to investigate people's insights, opinions and interests in space and time in order to promote spatial planning and geo-design by using space–time analysis. By using various sources of data (except for survey data), such as cellular data, it is feasible to better define the spatiotemporal elements of urban environments [22]. Meanwhile, other authors have worked on smart city transportation, technologies and applications [23–25].

The kernel density estimation (KDE) method has been used to spatially model geolocation data, and it offers an additional particular and comfortable structure for the spatial estimation of density. KDE has also been used for the assessment of environmental attributes, for example, healthcare properties [26], the diet environment [27] and green space access. It boosts up the rate of access, or measures distance by transforming point data into something similar to a nonstop surface, thus allowing the density of the characteristic to be measured at any point onto the map's surface [28]. The researchers in [29] investigated the relationships among location, walking speed and adequate levels of physical activity by utilizing KDE. KDE was also utilized to research sites such as food stores in connection with factors such as overweight, body mass index (BMI) and dietary intake [30]. In [31], the authors availed the KDE approach to analyze the relative importance of external factors correlated with temporal and spatial user distribution in urban green spaces. The identification and analysis of accidents or catastrophes and their effect on daily urban development procedures also offers another interesting potential usage of social media network data [32]. The authors compared KDE and point density in their research [33].

3. Materials

3.1. Study Area

As part of the alluvial plain, Shanghai and the Yangtze River Delta have an average elevation of around 4 m. From east to west, there is a slight terrain gradient. The land is plain, except for some south-western foothills. Shanghai is part of the climate zone for the north subtropical humid monsoon. There are four seasons, with abundant sunshine and rainfall. Therefore, in such a metropolitan area, extreme events and urbanization have had a substantial influence on public health promotion services and the country's economy. Shanghai is indeed the world's 10th, and most prominent, agglomeration region. Therefore, the process of urbanization is relatively quick. In the United Nations' (UN) future

forecasts of urbanization, it was stated that Shanghai's urban population ranks as the world's second largest, and China's first [34].

In 2016, Shanghai was divided into 16 regions: 15 districts (i.e., Baoshan, Changning, Fengxian, Hongkou, Huangpu, Jiading, Jing'an, Jinshan, Minhang, Pudong New District, Putuo, Qingpu, Songjiang, Xuhui and Yangpu) and one county (i.e., Chongming) [35]. Seven regions (i.e., Changning, Hongkou, Huangpu, Jing'an, Putuo, Xuhui and Yangpu) are located in Puxi (literally, Huangpu West). These seven areas are known as Shanghai's downtown or city center [36]. The study area in this research includes 10 districts (i.e., Baoshan, Changning, Hongkou, Huangpu, Jing'an, Minhang, Pudong New District, Putuo, Xuhui and Yangpu), as shown in Figure 1. The locations of the green parks can also be seen within the study area.
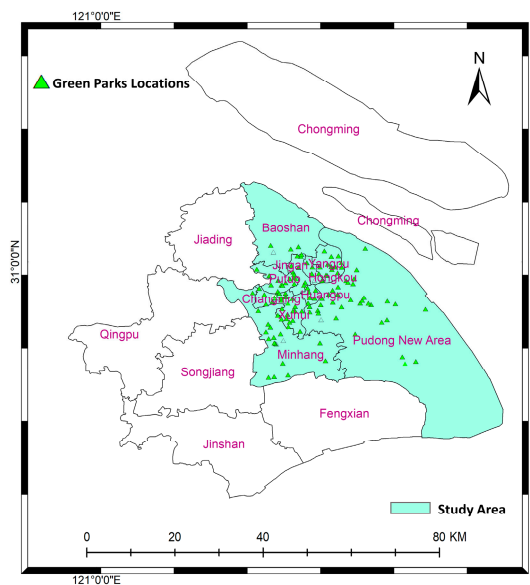

Figure 1. Location of the study area and its green parks.

3.2. Dataset

The dataset that we used to examine the number of visits to green parks came from the Chinese popular micro-blog Weibo. Weibo is considered to be comparable to China's Twitter, which is the major social media blog in China. Weibo has a massive number of customers, representing the biggest dataset of obtainable geotags. According to the latest Weibo annual report, it was announced that above 500 million active users were registered on the platform in 2018, and it reached 462 million monthly active users in December 2018 [37]. The demographics of Weibo users are inconsistent with the total population; thus, it was decided to launch the public interface of the Weibo Location Based Service (LBS) on 28 May 2012. Weibo users have since been able to share their location on the internet in real-time. As a kind of scalable and available large-scale crowdsource dataset, Weibo sign-in data were the most appropriate that we could obtain to estimate actual park visitors. However, an earlier study of 87 city parks situated in Shanghai, China showed that there is an important link between data from Jiepang's social media platforms and official visitor numbers [38].

In addition, earlier research has already revealed that Weibo data offer a perfect representation of the interests and behaviors of the individuals in urban areas [39]. Using the check-in data of Weibo [40], a new model was proposed, which integrates the parameters of urban environmental function networks, thereby enriching the definition of the structure of urban networks. However, while using Weibo check-ins as a proxy for visits is still uncommon, earlier research utilizing data from similar social media sites or platforms (for example, Facebook, Flickr and Instagram) found important beneficial ties amongst official visit statistics and the number of visitors reported on these sites [6,41]. Figure 2 represents the criteria that we deployed for collecting the check-in data.
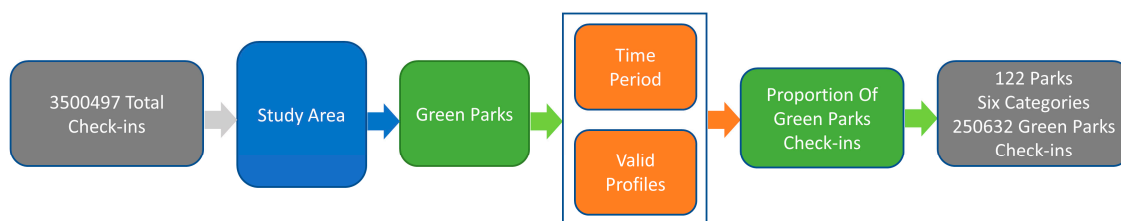
Figure 2. Criteria for the collection of check-in data.

The Weibo application program interface (API) helps in the processing of data collection. The dataset was obtained from Weibo for a time frame of 3 years, from July 2014 to June 2017. After retrieving this location information, we found that some locations that we had added were not green parks (e.g., sidewalks, former homes of famous people, and sculptures). We checked these locations one after another and removed those that were not in the green park category, but we still counted the parking lots or park areas linked to green parks. Some larger parks have more than two location IDs, for example, garden area, kids' play areas and barbecue areas were all combined into one location ID [42]. After pre-processing, filtration and cleaning, a total of 250,632 geo-tagged visits to green parks in 10 districts of Shanghai were included. The data were collected with the help of the programming language of Python (version 2.7.12), and were filtered for exclusions such as invalid records and fake users, including:

- The geographical location of data should exist only in Shanghai;
- The lowest number of check-ins per green park should be 100 within the time period of the study;
- Every record should have a geo-location (latitude and longitude), user id, time, gender, day, month and year;
- Parks that are separated into several geo-locations inside the green spaces were combined into one geo-location.

3.3. Park Type Classification

The green parks in Shanghai were divided into six commonly selected categories: (1) community parks (*n* = 30), (2) cultural relic parks (*n* = 8), (3) large urban parks (*n* = 12), (4) natural parks (*n* = 09), (5) neighborhood parks (*n* = 46), and (6) recreational parks (*n* = 17) (Table 1). These parks were classified on the basis of their dominant functions and different kinds of administration [43,44]. Figure 1 reveals the distribution of the different types of urban green parks throughout the study area. Amongst the 122 green parks in our study area, the greatest proportion belonged to neighborhood parks, followed by community parks.

Table 1. Description of the green parks.

Park Type (*N* = 122)	Description
Recreational park (*n* = 17)	Recreational green parks consist of botanical gardens, children's parks, zoos and sports fields, categorized by variable locations and sizes.
Cultural relic park (*n* = 8)	Cultural relic green parks are home to ancient valuable relics and are places that are situated in urbanized areas, which are of value in terms of education and tourism.
Large urban park (*n* = 12)	Large urban green parks are community spaces that serve a wide range of inhabitants, categorized by adequate services, various activities and large sizes.
Natural park (*n* = 9)	Natural green parks are categorized by a natural background and have environmental value, situated in suburban areas.
Community park (*n* = 30)	Community green parks are spaces that are situated in residential areas for local inhabitants' recreational purposes.
Neighborhood park (*n* = 46)	Neighborhood green parks are tiny or small residential urban green spaces that are situated in a certain residential district, serving a much smaller population in comparison to community parks.

4. Methodology

4.1. Data Preparation

The data that we collected cover all check-ins from July 2014 to June 2017 made within Shanghai's boundary. The data downloaded were included in many JSON files. Figure 3 represents the data preparation process.
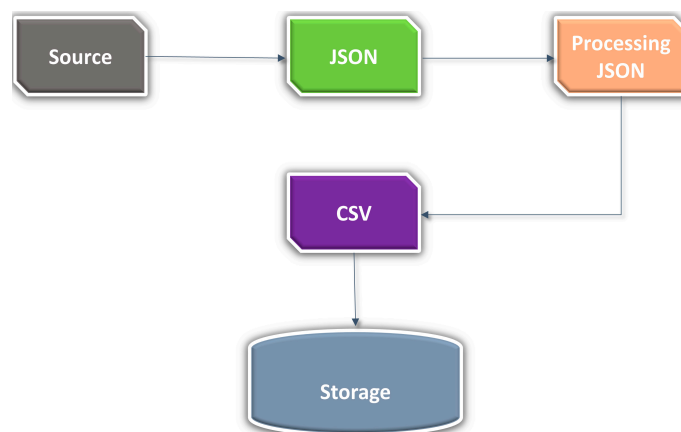

Figure 3. The data preparation process.

In this study, the Weibo dataset included information such as a unique user ID and check-in date and time. In addition, information about the geographic location (latitude and longitude) and gender was collected through the Weibo API. The LBSN dataset therefore assumes that daily trends are evidence of users' daily activities, behaviors on social media, and spatiotemporal patterns [45]. A typical Weibo "check-in" is represented as: check-in (B2094554D064ABF44293) = {1758115961, ####, B2094554D064ABF44293, Mon July 25 14:47:41 +0800 2016, m, 121.484566, 31.270601}, where B2094554D064ABF44293 denotes "location_id," 1758115961 denotes "user_id," #### denotes the "user_name," Mon July 25 14:47:41 +0800 2016 denotes "day, month, date, time, and year," m denotes "gender," and 121.484566, 31.270601 denotes the geo-location. JSON is a Java platform programming platform format, which is the most commonly used data format though Java, and is considered to be the main programming language with open source accessible reader and writer modules. Using the selected software [46], the data were filtered into a CSV (comma-separated values) file format, so that all user data, which include geo-locations, could be identified and stored in the database regardless of the publication date. Table 2 shows an example of a "check-in" in CSV format.

Table 2. Example of a Weibo check-in in CSV (comma-separated values) format.

Building_id	User_id	Month	Date	Day	Time	Year	Gender	Lon	Lat	Address
B2094554D0 64ABF4429	###*	08	24	Wed.	0:00:03	2016	F	121.6650	31.14363	ZhaoHang_Park
B2094757D0 6FA0F8409D	###*	11	11	Fri.	1:24:45	2016	F	121.5244	31.265614	Jiangpu_Park
B2094757D0 6FA6FB409B	###*	05	29	Mon.	3:02:19	2017	M	121.3787	31.34228	Gucun_Park

In view of the heterogeneity problem, only green parks that include more than 100 check-ins were chosen to establish the user sample, in order to confirm a fairly high level of representativeness.

4.2. Social Media Data Analytics

In this research, we analyzed Weibo-based geo-location datasets in 10 districts of Shanghai, China (July 2014 to June 2017). Figure 4 shows a check-in behavior analysis framework in which the LBSN data analysis method includes the framework, the data pre-processing and cleaning, the temporal and spatial analysis of the LBSN data, and the statistics that indicate the worth of the LBSN data.
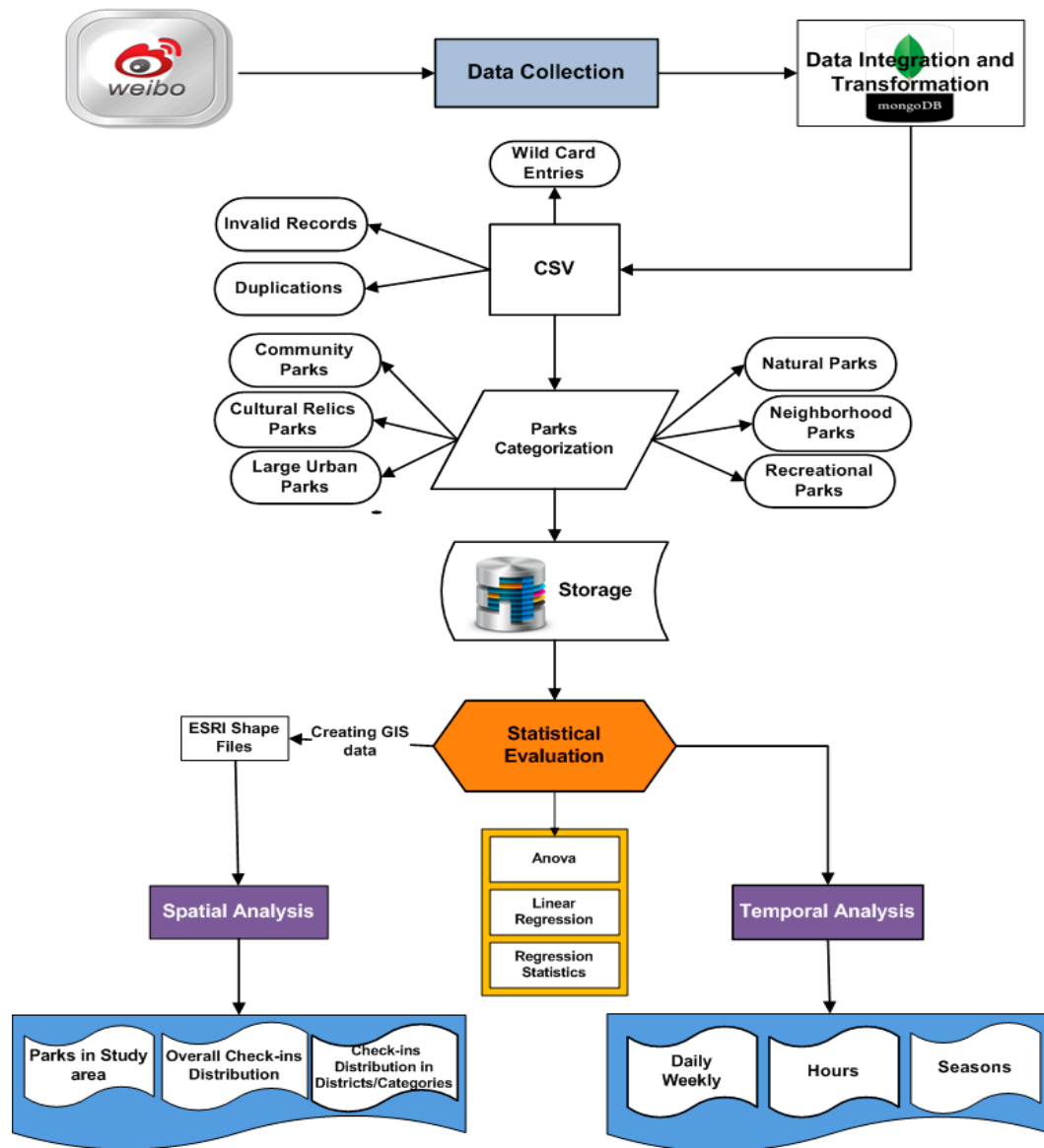

Figure 4. The methodology.

Across the three-year period, the number of visitors to the urban green spaces was obtained as statistical data. A spatiotemporal assessment, utilizing statistical graphs and tables, was performed based on the results, and the distribution of check-ins was examined using the KDE method.

4.3. Temporal Analysis

With the purpose of tracking variations in user behavior, we divided the check-in time stamps into different time classifications—daily, weekly and seasonal. The daily pattern shows the hourly distribution of check-ins during the day, and the weekly pattern shows the weekly distribution of check-ins. Seasonal trends and climate factors were also taken into account, as they may influence

the characteristics of green parks. Winter trends could provide significant data regarding park usage throughout the colder months. We did not define seasons by different dates in our research, but as a substitute used simple categories based upon months: March–May is spring; June–August is summer; September–November is autumn; and December–February is winter.

SPSS v25 was used for statistical analysis. SPSS is a frequently-used program in social science for performing statistical analysis to solve various study problems, and it is used by industry and health researchers, survey agencies, government education specialists, data mining companies, marketing agencies and more. It provides various methods, including hypothesis checking and reporting, ad-hoc analysis and data management to facilitate analysis. We used Table 2019.2 for its visualization techniques to explore and analyze relational databases and data cubes.

4.4. Statistical Analysis

To assess the significance of the explanatory variables, it was necessary to statistically explore the predictors (i.e., explanatory variables) and their effect on the response variable (i.e., the number of check-ins). For this model, we used the following regression equation:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_k x_k + \epsilon \tag{1}$$

Table 3 displays the parameters and the explanatory variables used in our regression model.

$$\begin{aligned} Y = \beta_0 &+ \beta_1 Baoshan + \beta_2 Changning + \beta_3 Huangpu + \beta_4 Jingan + \beta_5 Yangpu + \beta_6 Mon \\ &+ \beta_7 Tue + \beta_8 Wed + \beta_9 Fri + \beta_{10} Sat + \beta_{11} July + \beta_{12} Aug + \beta_{13} Feb \\ &+ \beta_{14} Mar + \beta_{15} Apr + \beta_{16} May + \beta_{17} Jun + \epsilon \end{aligned} \tag{2}$$

Table 3. Final multiple linear regression model interpretation.

Min.	1Q	Median	3Q	Max.	
−15.489	−2.308	0.325	2.184	183.555	-
Coefficients:	**Estimate**	**Std. Error**	***t*-Value**	**Pr (>|t|)**	-
(Intercept)	0.545664	0.389241	1.402	0.161213	-
Baoshan	0.802558	0.438164	1.832	0.067252	.
Changning	2.739148	0.283543	9.66	<2E−16	***
Huangpu	1.847305	0.260773	7.084	2.38E−12	***
Jingan	1.384735	0.340477	4.067	5.07E−05	***
Yangpu	1.666539	0.145104	11.485	<2E−16	***
Mon.	−0.17896	0.078598	−2.277	0.022966	*
Tue.	−0.185066	0.06221	−2.975	0.00299	**
Wed.	−0.182712	0.094293	−1.938	0.052891	.
Fri.	−0.083539	0.063926	−1.307	0.191526	-
Sat.	0.003532	0.059012	0.06	0.952281	-
July	0.871949	0.22472	3.88	0.00011	***
Aug.	0.837961	0.202606	4.136	3.78E−05	***
Feb.	1.866674	0.177615	10.51	<2E−16	***
Mar.	0.612028	0.166562	3.674	0.000249	***
Apr.	0.940388	0.126142	7.455	1.71E−13	***
May	0.35173	0.140485	2.504	0.012422	*
Jun.	0.829923	0.137177	6.05	1.93E−09	***
July 2014–June 2015	0.061276	0.077947	0.786	0.431952	-
July 2015–June 2016	0.320808	0.082899	3.87	0.000115	***

Significance codes: '***' 0.001; '**' 0.01; '*' 0.05; '.' 0.1; and '-' 1.

Afterward, by applying the linear regression model, our fitted value equation converted to:

$$\hat{y} = b_0 + b_1 Baoshan + b_2 Changning + b_3 Huangpu + b_4 Jingan + b_5 Yangpu + b_6 Mon$$
$$+ b_7 Tue + b_8 Wed + b_9 Fri + b_{10} Sat + b_{11} July + b_{12} Aug + b_{13} Feb+ \quad\quad (3)$$
$$b_{14} Mar + b_{15} Apr + b_{16} May + b_{17} Jun + \epsilon$$

We implemented the linear regression model by including variables that have a mild correlation with the dependent variable; in our scenario, the dependent variable was the number of check-ins, with correlation values from 0.10 to 0.50. As for the model's inference, the *p*-value of the F-statistic model showed that the model is itself significant. It must be remembered that not all predictors have a significant *p*-value, since the model was built using Table 4's highest adjusted R^2. Table 3 defines the model coefficients, in which the modification shows that the number of check-ins increased by an average of approximately 2.73% for each unit, with a very small *p*-value. Likewise, for each unit of Huangpu, Jingan and Yangpu that increased, the average check-in period increased, with very small *p*-values, by approximately 1.84%, 1.38% and 1.66%, respectively. Analysis of ANOVA has been represented in Table 5.

Table 4. Regression statistics summary.

Residual Standard Error	Degrees of Freedom	Multiple R-Squared	Adjusted R-Squared	F-Statistic	*p*-Value
6.512	1205	0.4873	0.4792	60.28	<2.2E−16

Table 5. Analysis of variance (ANOVA). Response: Number of Check-ins.

	Df	Sum Sq	Mean Sq	F-Value	Pr (>F)	
Baoshan	1	379	378.7	8.93	0.002861	**
Changning	1	14,815	14,814.7	349.3	<2.2E−16	***
Huangpu	1	8031	8030.6	189.4	<2.2E−16	***
Jingan	1	2060	2059.6	48.56	5.25E−12	***
Yangpu	1	10,885	10,884.8	256.7	<2.2E−16	***
Mon.	1	45	44.8	1.057	0.3042	-
Tue.	1	216	215.7	5.087	0.024288	*
Wed.	1	95	94.8	2.235	0.135139	-
Fri.	1	77	76.9	1.814	0.178312	-
Sat.	1	1	1.5	0.034	0.853309	-
July	1	639	638.5	15.06	0.00011	***
Aug.	1	814	814.4	19.2	1.28E−05	***
Feb.	1	4674	4674	110.2	<2.2E−16	***
Mar.	1	339	338.6	7.985	0.004795	**
Apr.	1	2959	2958.8	69.77	<2.2E−16	***
May	1	143	142.5	3.361	0.067015	.
Jun.	1	1162	1162.4	27.41	1.94E−07	***
July 2014–June 2015	1	138	138.2	3.258	0.071333	.
July 2015–June 2016	1	1106	1105.9	26.08	3.81E−07	***
Residuals	1205	51,105	42.4	-	-	-

Significance codes: '***' 0.001; '**' 0.01; '*' 0.05; '.' 0.1; and '-' 1.

All independent variables are significant predictors that rely on the *p*-values [47], as revealed in Table 4. For the statistical analysis, we used the statistical programming language R [48] and the program RStudio [49] to perform basic descriptive and regression analysis.

4.5. Spatial Analysis

We utilized the KDE method to create a smooth surface density for check-in hotspots in the geographic area. The KDE method is a non-parametric estimation technique for determining the density of a random sample of data [50]. KDE smooths each data point into small density bumps, and then all of these small bumps are combined to make a final estimation of the density. KDE is widely accepted for spatial distribution [51–53], and it describes the spatial density distribution combined with the distance–decay effect, and projects hotspots by transforming scatter point data into a continuous density surface [54]. KDE is an evolving spatiotemporal technique that has been used previously [55,56] to inspect several features of social media (but not restricted to LBSN) data analytics, such as users' online activity and movement trends [57], check-in behavior [58], city boundary descriptions [59,60] and point-of-interest recommendations [61]. It also explores the distribution of destinations in communities, enabling researchers to see where there are densely scattered destinations, and where they are more sparsely scattered. Eventually, this method seeks to create a smooth surface of density within the geographical space of spatial point cases [46]. The authors utilized the KDE method for the analysis of spatiotemporal patterns in green parks [62,63].

KDE can effectively calculate the visitor density spatial structure within an area of study. KDE is a statistical method for estimating a smooth and continuous distribution from a small number of observations [64]. The data taken into account in our analysis were in the form of geo-tagged check-ins. Let E be a collection of historical data for check-in, i.e., $E = \{e^1, \ldots, e^n\}$, where $e^i = <x, y>$ is a check-in geo-location $1 < i < n$, of individual i and at time t, where E represents the dataset we used. The total of the kernel's functions was scaled to construct a smooth curve, i.e., a unit field. This resulted in a bivariate of KDE in the following form:

$$f_{KD}(e|E,h) = \frac{1}{n} \sum_{i=1}^{n} K_h\left(e, e^i\right) \tag{4}$$

$$K_h = \frac{1}{2\pi h} \exp\left(-\frac{1}{2}(e, e^i)^t \sum_{h}^{-1} (e, e^i)\right) \tag{5}$$

where e denotes the check-in location in dataset E, along with bandwidth h. h is supposed to be reliant on the estimated density f_{KD}, producing a smooth density surface around E at the data point e^i.

ArcGIS 10.0 was used to evaluate the spatial distribution of the check-ins in space. In particular, ArcGIS 10.0 (Environmental Systems Research Institute, Inc., Redlands, CA, USA) software, with a 2016-developed Shanghai map using the WGS 1984 geodesic coordinate system, was used. The base map also included the major transportation lines (i.e., the line layer) and the administrative districts (i.e., the polygon layer and the district layers) with the new OpenStreetMap subway lines and entries (i.e., the point layer).

5. Results

Shanghai city is among the fastest growing metropolises in the world, with a population of 22,125,000 per 4015 km^2 [65]. The total number of green parks within the city of Shanghai is 366 [66], which encourages the city's inhabitants to participate in various healthy activities. For this analysis, 122 of these green parks were chosen after processing the data collected from Weibo. The distribution of the different categories of green parks included in this research is shown in Figure 5, where the different colors reveal different categories of parks.
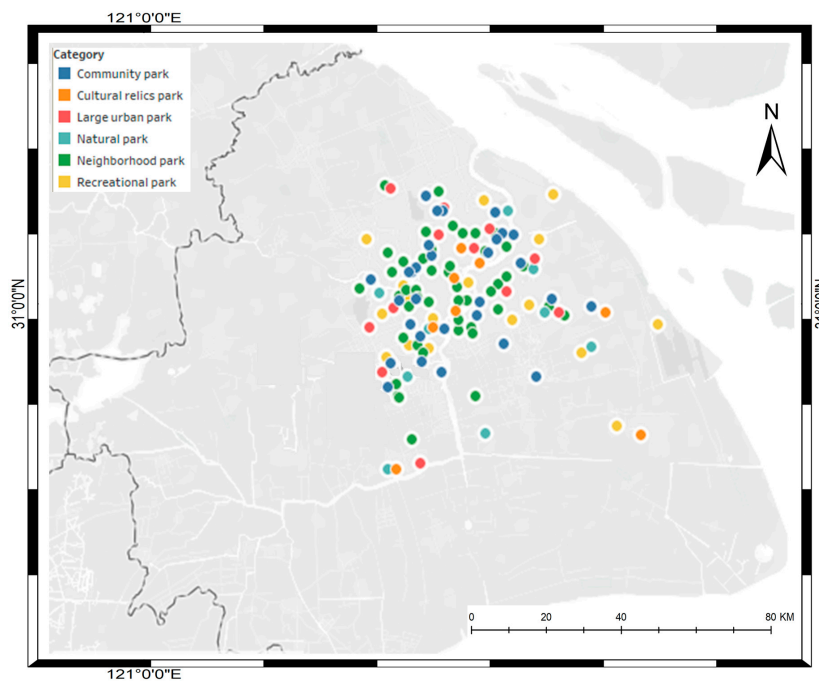
Figure 5. The different categories of green parks.

We utilized KDE to examine the spatial distribution of the check-in data, and ArcGIS for visualization, to investigate the Weibo geo-location check-in data. Figure 6 indicates the overall check-in density in Shanghai between July 2014 and June 2017, where the areas colored in red display a higher human density, a higher level of activity, and a higher percentage of social media usage. It is therefore no surprise that the downtown green parks have large clusters of activity.
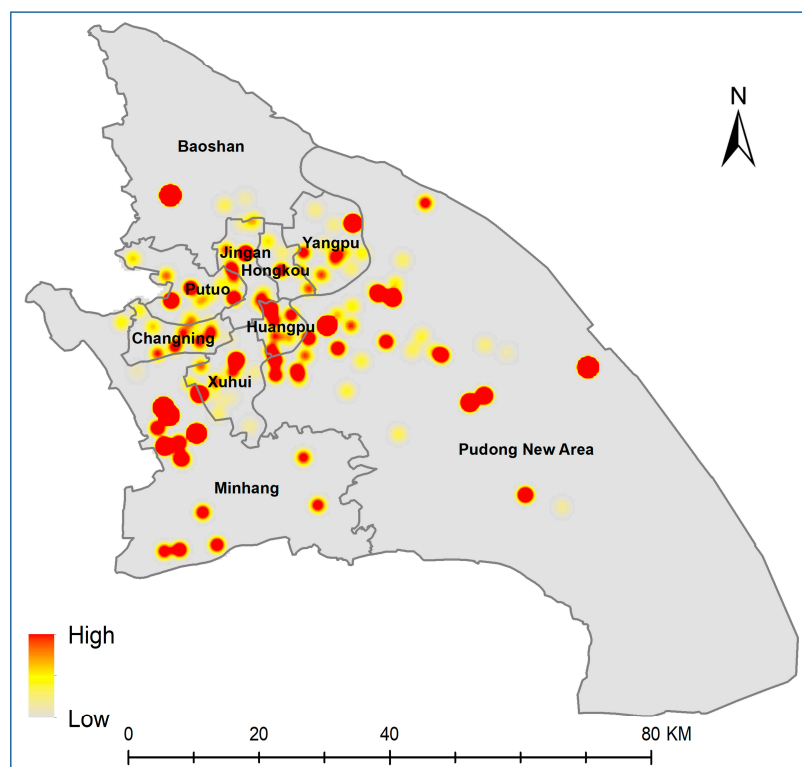

Figure 6. The distribution of the check-ins.

Figure 7 presents the overall check-in density in all of the different categories of parks; the parks were divided into six categories, and our results reveal that the most check-ins were found in neighborhood parks and recreational parks as the research question was raised, and so the density is higher in these categories. The reason behind this is that the number of neighborhood parks is higher compared to the other types of parks, and they are located near residential areas.
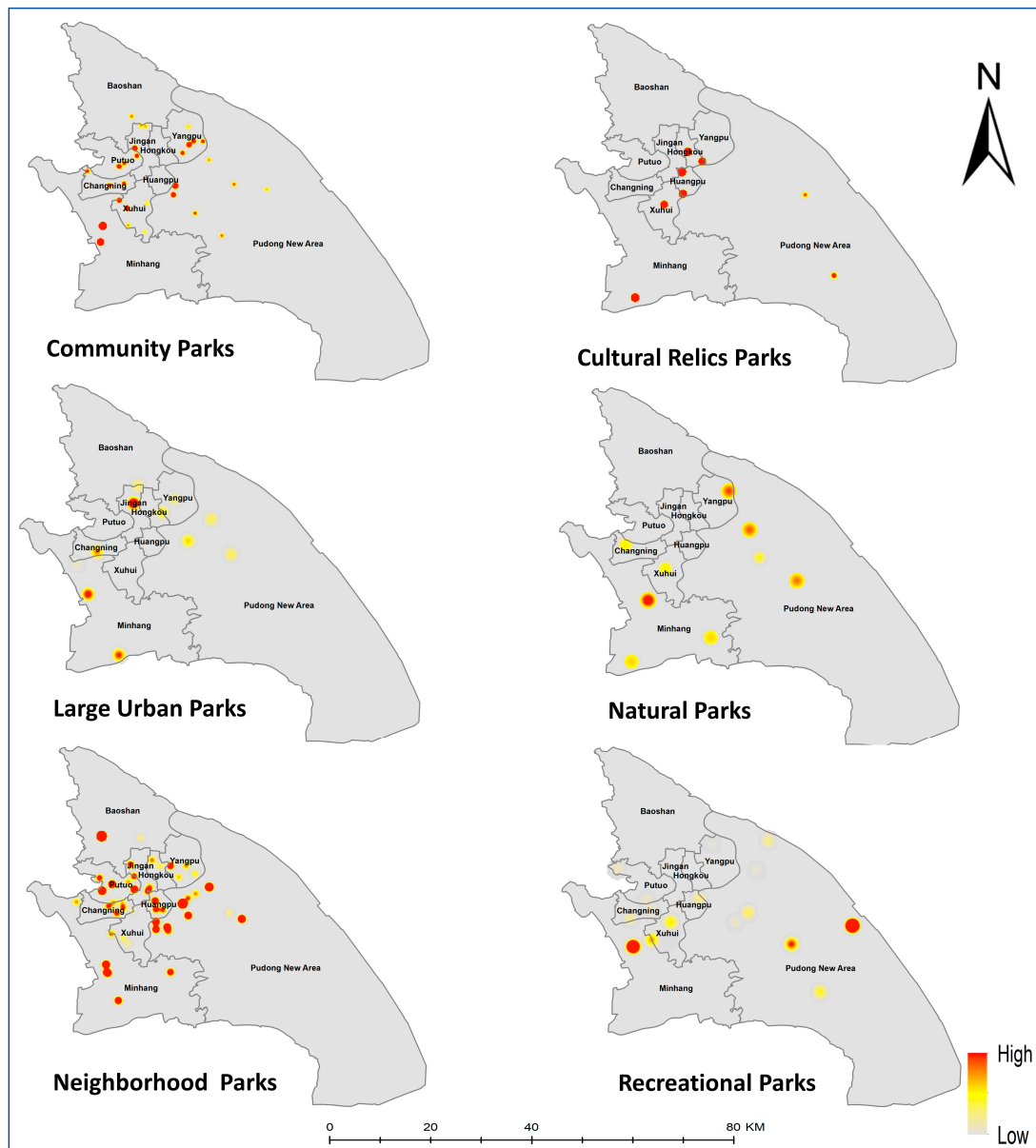

Figure 7. The density of the different categories of parks.

Figure 8a shows the temporal differences in the number of visitors during a 24-h period. Although the visitors accessed the parks during all periods of the day, the maximum number of check-ins were made at 4:00 p.m.–6:00 p.m. and at 10 p.m. for all of the parks considered in the research. This trend continues to increase until midnight. These results indicate that Shanghai's green parks are fast becoming popular recreational destinations for the people of Shanghai. Figure 8b shows the percentage of the total number of check-ins, including a record of every hour, and it also testifies that most people like to visit neighborhood parks up until midnight.
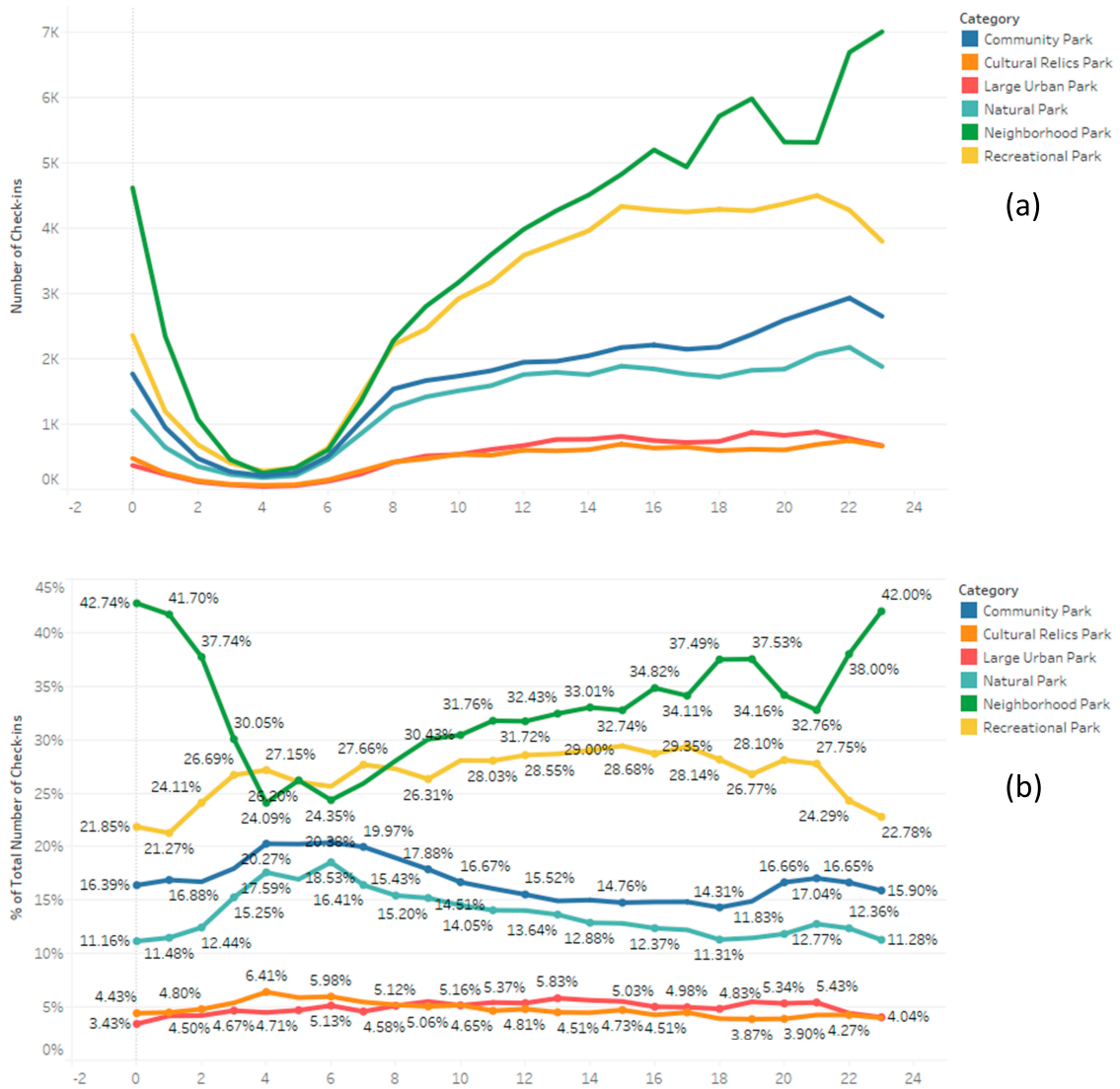
Figure 8. Daily check-ins: (**a**) Number of check-ins; (**b**) percentage of the total number of check-ins.

Although a nearly equal number of check-ins were made every weekday, more check-ins were made on Saturdays and Sundays, and this weekly pattern is consistent across all categories of parks. Figure 9a describes the general pattern of weekly check-ins, while Figure 9b reveals the weekly trend in all categories of green parks, further highlighting that the most check-ins were made on weekends across all park types.
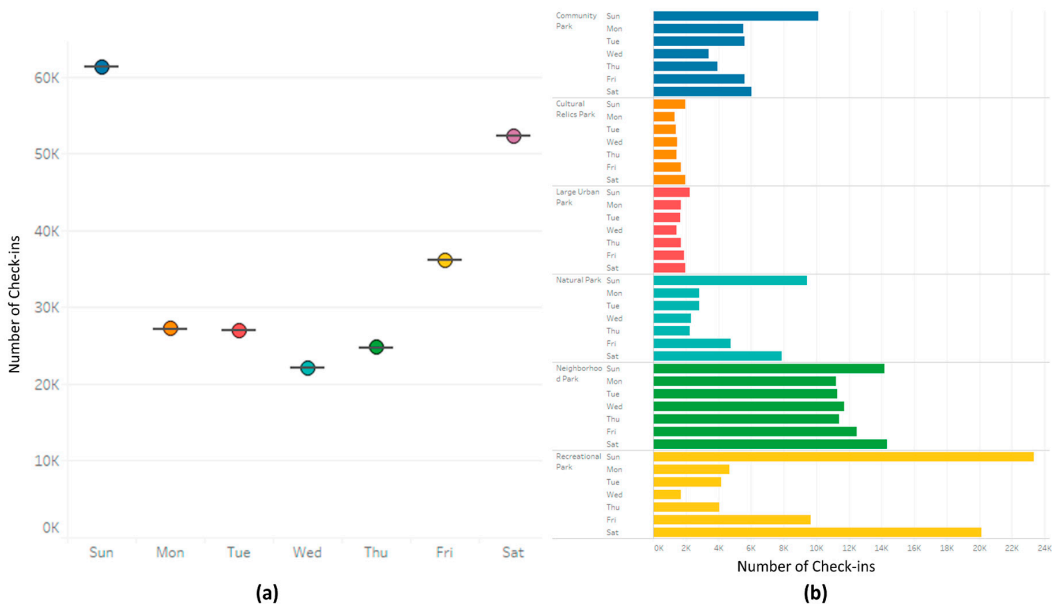
Figure 9. Weekly check-ins: (**a**) Number of check-ins during the week; (**b**) number of weekly check-ins in different categories.

We investigated the data based on gender (i.e., male and female) in 10 districts of Shanghai, to examine the check-in rates and behavior. Figure 10a, displays the overall check-in pattern among the districts. Moreover, Figure 10b reveals the number of check-ins across the different categories of parks. The important thing to note here is that female visitors were more active users of Weibo in contrast to male visitors.
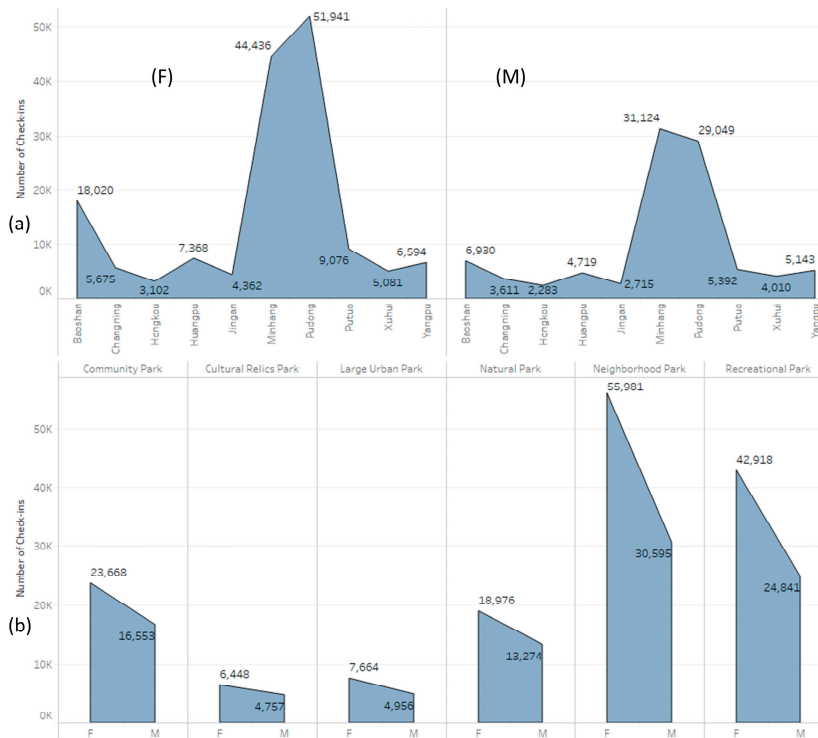

Figure 10. Gender differences: (**a**) In districts; (**b**) in categories.

A comparison of genders in terms of their check-ins (regarding both frequency and behavior), across week days and districts, was used to investigate the differences among male and female visitors

in Shanghai. Tables 6 and 7 represent the outcomes of this comparison across the days of the week, seasons, districts, and categories of parks.

Table 6. Gender differences among districts.

Gender	Season	Baoshan	Changning	Hongkou	Huangpu	Jingan	Minhang	Pudong	Putuo	Xuhui	Yangpu
F	Autumn	1.94%	0.50%	0.27%	0.61%	0.32%	3.92%	4.09%	0.75%	0.47%	0.58%
	Spring	1.59%	0.66%	0.37%	1.01%	0.54%	5.25%	6.36%	1.06%	0.58%	0.78%
	Summer	2.10%	0.67%	0.34%	0.77%	0.56%	5.00%	6.40%	1.02%	0.56%	0.75%
	Winter	1.55%	0.43%	0.25%	0.55%	0.32%	3.56%	3.87%	0.80%	0.42%	0.52%
M	Autumn	0.90%	0.35%	0.18%	0.42%	0.20%	2.71%	2.32%	0.45%	0.39%	0.44%
	Spring	0.55%	0.42%	0.27%	0.54%	0.33%	3.78%	3.52%	0.65%	0.44%	0.61%
	Summer	0.76%	0.39%	0.26%	0.52%	0.34%	3.34%	3.51%	0.61%	0.44%	0.55%
	Winter	0.55%	0.28%	0.20%	0.40%	0.22%	2.58%	2.24%	0.44%	0.32%	0.44%

Table 7. Gender differences among categories.

Gender	Day	Community Parks	Cultural Relic Parks	Large Urban Parks	Natural Parks	Neighborhood Parks	Recreational Parks
F	Sun.	2.34%	0.43%	0.57%	2.21%	3.62%	5.82%
	Mon.	1.29%	0.30%	0.42%	0.66%	2.89%	1.18%
	Tue.	1.32%	0.32%	0.39%	0.67%	2.90%	1.06%
	Wed.	0.84%	0.35%	0.35%	0.57%	2.99%	0.47%
	Thu.	0.95%	0.33%	0.41%	0.53%	2.93%	1.02%
	Fri.	1.30%	0.39%	0.45%	1.12%	3.26%	2.47%
	Sat.	1.42%	0.44%	0.49%	1.83%	3.74%	5.11%
M	Sun.	1.71%	0.35%	0.33%	1.56%	2.03%	3.49%
	Mon.	0.91%	0.22%	0.26%	0.46%	1.58%	0.69%
	Tue.	0.92%	0.23%	0.27%	0.45%	1.60%	0.61%
	Wed.	0.52%	0.24%	0.23%	0.36%	1.68%	0.21%
	Thu.	0.61%	0.23%	0.28%	0.37%	1.62%	0.60%
	Fri.	0.94%	0.28%	0.31%	0.78%	1.72%	1.39%
	Sat.	0.98%	0.34%	0.30%	1.32%	1.97%	2.92%

The seasonal differences in visit check-ins to green parks were investigated for autumn, winter, spring and summer. In accordance with similar studies, significant seasonal differences in user check-ins were identified [62,67]. According to the research question, the seasonal pattern shows a higher percentage of check-ins in green parks throughout the summer and spring. It is worth noting that check-ins throughout the winter were slightly lower than in autumn. Figure 11b represents the number of check-ins across the different categories of green parks during the different seasons, and shows that neighborhood parks and recreational parks dominate the number of check-ins in all seasons.
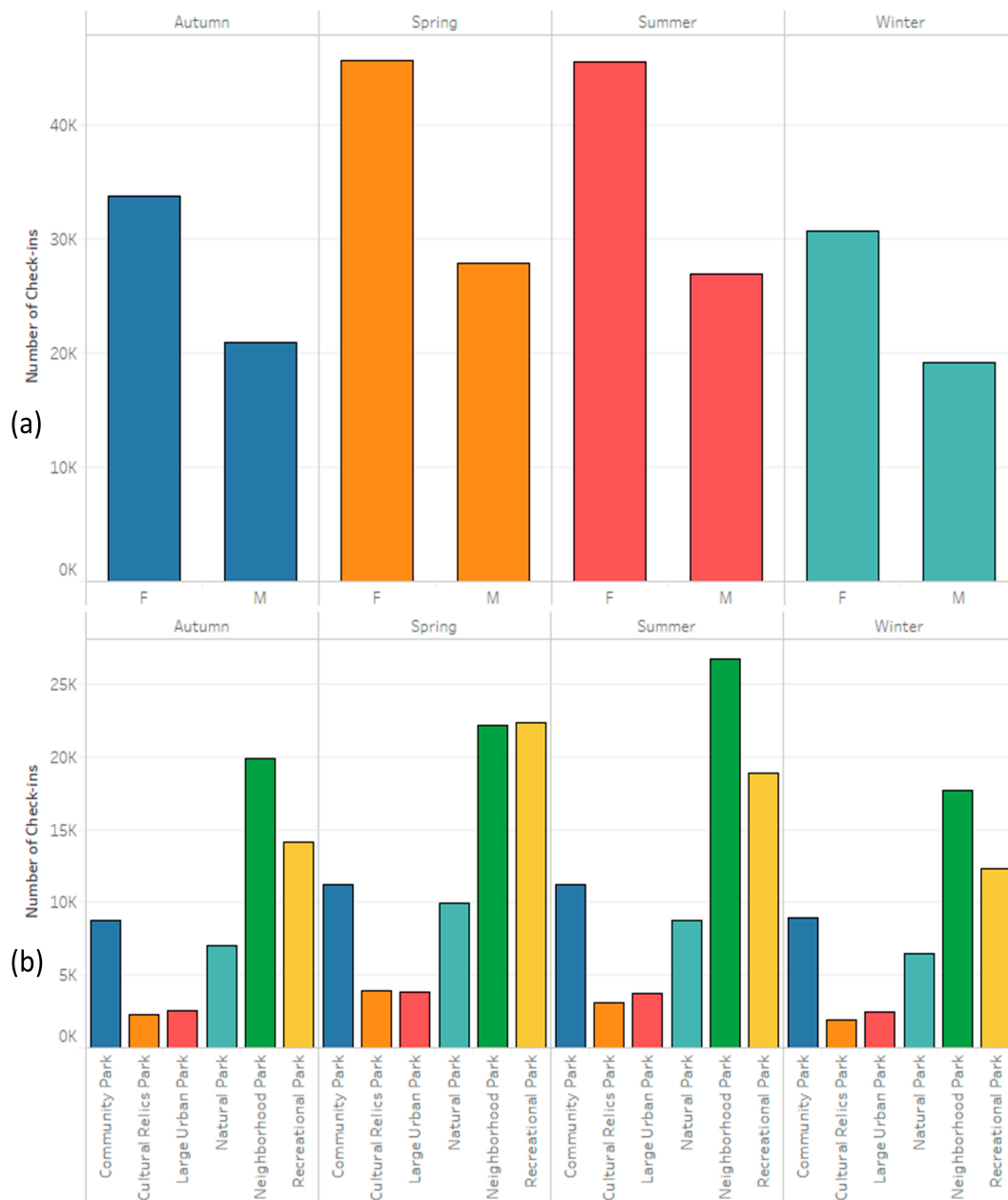
Figure 11. Seasonal check-ins: (**a**) Number of check-ins during the seasons; (**b**) number of check-ins in the different categories of parks among the seasons.

Finally, the overall statistics regarding the number of check-ins are shown in Table 8, separated by seasons, districts and daily trends.

Table 8. Daily, seasonal, and district check-in trends.

Season	District	Sun.	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.
Autumn	Baoshan	0.44%	0.38%	0.39%	0.40%	0.39%	0.40%	0.44%
	Changning	0.11%	0.12%	0.13%	0.12%	0.12%	0.14%	0.11%
	Hongkou	0.09%	0.04%	0.04%	0.05%	0.05%	0.08%	0.09%
	Huangpu	0.16%	0.14%	0.14%	0.16%	0.14%	0.15%	0.15%
	Jingan	0.08%	0.08%	0.07%	0.08%	0.07%	0.07%	0.08%
	Minhang	1.73%	0.88%	0.82%	0.41%	0.75%	0.95%	1.10%
	Pudong	1.94%	0.45%	0.45%	0.41%	0.40%	0.95%	1.83%
	Putuo	0.22%	0.14%	0.14%	0.16%	0.15%	0.17%	0.21%
	Xuhui	0.23%	0.05%	0.07%	0.07%	0.07%	0.14%	0.23%
	Yangpu	0.31%	0.06%	0.05%	0.07%	0.07%	0.15%	0.31%
Spring	Baoshan	0.33%	0.30%	0.30%	0.29%	0.30%	0.30%	0.33%
	Changning	0.16%	0.15%	0.16%	0.15%	0.15%	0.15%	0.17%
	Hongkou	0.13%	0.07%	0.07%	0.07%	0.08%	0.10%	0.12%
	Huangpu	0.24%	0.21%	0.21%	0.24%	0.20%	0.23%	0.22%
	Jingan	0.12%	0.12%	0.13%	0.12%	0.13%	0.13%	0.12%
	Minhang	2.37%	1.20%	1.14%	0.60%	0.95%	1.29%	1.47%
	Pudong	3.03%	0.71%	0.72%	0.59%	0.59%	1.40%	2.83%
	Putuo	0.30%	0.21%	0.20%	0.23%	0.21%	0.25%	0.31%
	Xuhui	0.26%	0.09%	0.10%	0.10%	0.10%	0.14%	0.24%
	Yangpu	0.42%	0.11%	0.09%	0.11%	0.09%	0.19%	0.39%
Summer	Baoshan	0.45%	0.38%	0.39%	0.39%	0.37%	0.41%	0.46%
	Changning	0.15%	0.14%	0.15%	0.16%	0.15%	0.15%	0.15%
	Hongkou	0.12%	0.06%	0.07%	0.06%	0.06%	0.09%	0.14%
	Huangpu	0.19%	0.17%	0.19%	0.18%	0.18%	0.19%	0.19%
	Jingan	0.13%	0.13%	0.13%	0.12%	0.13%	0.12%	0.13%
	Minhang	2.17%	1.11%	1.10%	0.55%	0.84%	1.21%	1.37%
	Pudong	2.87%	0.77%	0.77%	0.73%	0.70%	1.44%	2.63%
	Putuo	0.28%	0.22%	0.20%	0.22%	0.20%	0.23%	0.29%
	Xuhui	0.23%	0.09%	0.10%	0.11%	0.10%	0.15%	0.22%
	Yangpu	0.39%	0.08%	0.09%	0.10%	0.08%	0.19%	0.37%
Winter	Baoshan	0.31%	0.30%	0.27%	0.29%	0.31%	0.30%	0.32%
	Changning	0.11%	0.10%	0.09%	0.10%	0.10%	0.10%	0.10%
	Hongkou	0.09%	0.06%	0.04%	0.05%	0.05%	0.08%	0.08%
	huangpu	0.11%	0.13%	0.13%	0.14%	0.14%	0.14%	0.16%
	Jingan	0.08%	0.08%	0.08%	0.08%	0.07%	0.08%	0.07%
	Minhang	1.57%	0.83%	0.79%	0.42%	0.68%	0.84%	1.01%
	Pudong	1.86%	0.41%	0.43%	0.39%	0.39%	0.87%	1.74%
	Putuo	0.21%	0.16%	0.17%	0.15%	0.16%	0.17%	0.21%
	Xuhui	0.18%	0.07%	0.06%	0.07%	0.07%	0.12%	0.19%
	Yangpu	0.28%	0.07%	0.06%	0.06%	0.07%	0.14%	0.28%

6. Conclusions and Recommendations

In this research, we applied big data methodology to revisit the issue of environmental justice linked to the spatial provision of urban green parks in Shanghai. We utilized geo-tagged Weibo check-in data as a park-visit indicator within 10 districts of Shanghai, and investigated the number of check-in visits accordingly. We focused upon the different categories of parks in Shanghai because, to the best of our knowledge, we are the first ones to address this problem in such a highly crowded and metropolitan area. We achieved an in-depth experimental analysis of check-in behavior in the current research, which used intensity maps and patterns from LBSN data. The conclusions reveal the distribution of users in parks by analyzing the check-in data, and the findings show that neighborhood green parks are much more crowded than other green spaces. Analysis of the seasons' impacts upon people's behavior toward green spaces in different categories of parks shows that the number of check-ins is much higher in summer and spring, compared to autumn and winter. The total number of

Weibo users, depending on the hour of the day, shows that the peak time to visit green parks is from midday to midnight. Lastly, gender-based variations were measured in relation to green parks visits, and the findings reveal that female visitors are more involved in their use of social media services when visiting green parks.

Kernel density estimation is a technique for defining the probability density function, and is a must-have method that allows the user to examine the studied probability distribution more effectively than when using an old-style histogram. The core technique, different from the histogram, provides a uniform estimate, uses the locations of all the sample points, and, more importantly, involves multimodality. KDE is a function in which events are balanced according to their distances and the two necessary parameters. The first of these is bandwidth, the control distance. Bandwidth selection has a great influence on performance. The second parameter is the K-weighting function, more often a normal function. The kernel bandwidth is a free parameter that indicates a strong effect on the resulting estimate. Comparing to the commonly used histogram, the kernel density estimator offers many advantages. It is a smooth curve, and therefore shows details better, and uses the locations of all sample points, so the information contained in the sample should be better represented. KDE estimates smooth distributions by not including local noise to a particular degree, which minimizes flaws by providing a non-parametric probability distribution with optimal bandwidth.

The wide spatial coverage of this research offers valuable information, which can improve urban green space development and planning in other major cities. The results indicate that planners must start paying more attention to the importance of small neighborhood green parks in urban green space arrangements. This article provides encouragement for the entertainment value of these parks, by defining their high check-in visitation intensity. The presence of accessible, well-maintained, small green spaces in the urban park system is also crucial for meeting the recreational needs of local residents. More support services, including the area and location of the green park, can charm more visitors, as this research reveals that green urban parks in the city center or the downtown area attract more visitors than parks in the other parts of the city. These findings can also be useful in the urban development of smart cities for green spaces, by considering visitors' preferences. In addition, similar research must be carried out in other megalopolis cities to assess the common issues affecting the utilization of urban green spaces.

7. Limitations and Future Work

Based on the outcomes of the present study, LBSN data have the ability to deliver a new outlook, in addition to providing observations of gender differences and check-in intensity. LBSN check-in data have some huge benefits, such as high spatial accuracy and minimal cost. However, some constraints are associated with this type of data, such as low sample size frequency, gender bias and location category bias. In conclusion, LBSN data is more likely to be supplementary to, than a replacement for, traditional sources of data. When implementing similar techniques, social media data that are shared or posted by green park visitors can be further understood to help visitors realize their feelings or sentiments, as well as to evaluate the numerous benefits offered by urban green parks.

References

1.　Hand, K.L.; Freeman, C.; Seddon, P.J.; Recio, M.R.; Stein, A.; van Heezik, Y. The importance of urban gardens in supporting children's biophilia. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 274–279. [CrossRef] [PubMed]

2. Richardson, E.A.; Pearce, J.; Mitchell, R.; Kingham, S. Role of physical activity in the relationship between urban green space and health. *Public Health* **2013**, *127*, 318–324. [CrossRef] [PubMed]

3. Zhou, W.; Wang, J.; Cadenasso, M.L. Effects of the spatial configuration of trees on urban heat mitigation: A comparative study. *Remote Sens. Environ.* **2017**, *195*, 1–12. [CrossRef]

4. Cohen, D.A.; Marsh, T.; Williamson, S.; Derose, K.P.; Martinez, H.; Setodji, C.; McKenzie, T.L. Parks and physical activity: Why are some parks used more than others? *Prev. Med.* **2010**, *50*, S9–S12. [CrossRef] [PubMed]

5. Wendel, H.E.W.; Zarger, R.K.; Mihelcic, J.R. Accessibility and usability: Green space preferences, perceptions, and barriers in a rapidly urbanizing city in Latin America. *Landsc. Urban Plan.* **2012**, *107*, 272–282. [CrossRef]

6. Tenkanen, H.; Di Minin, E.; Heikinheimo, V.; Hausmann, A.; Herbst, M.; Kajala, L.; Toivonen, T. Instagram, flickr, or twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Sci. Rep.* **2017**, *7*, 17615. [CrossRef]

7. Wood, S.A.; Guerry, A.D.; Silver, J.M.; Lacayo, M. Using social media to quantify nature-based tourism and recreation. *Sci. Rep.* **2013**, *3*, 2976. [CrossRef]

8. Weibo. Available online: https://www.weibo.com (accessed on 25 June 2019).

9. Zhen, F.; Cao, Y.; Qin, X.; Wang, B. Delineation of an urban agglomeration boundary based on Sina Weibo microblog 'check-in'data: A case study of the Yangtze River Delta. *Cities* **2017**, *60*, 180–191. [CrossRef]

10. Cetin, M. Determining the bioclimatic comfort in Kastamonu City. *Environ. Monit. Assess.* **2015**, *187*, 640. [CrossRef]

11. Lee, A.C.; Maheswaran, R. The health benefits of urban green spaces: A review of the evidence. *J. Public Health* **2011**, *33*, 212–222. [CrossRef]

12. Haq, S.M.A. Urban green spaces and an integrative approach to sustainable environment. *J. Environ. Prot.* **2011**, *2*, 601. [CrossRef]

13. Amoly, E.; Dadvand, P.; Forns, J.; López-Vicente, M.; Basagaña, X.; Julvez, J.; Alvarez-Pedrerol, M.; Nieuwenhuijsen, M.J.; Sunyer, J. Green and blue spaces and behavioral development in Barcelona schoolchildren: The BREATHE project. *Environ. Health Perspect.* **2014**, *122*, 1351–1358. [CrossRef] [PubMed]

14. Klaufus, C.; Van Lindert, P.; Van Noorloos, F.; Steel, G. All-inclusiveness versus exclusion: Urban project development in Latin America and Africa. *Sustainability* **2017**, *9*, 2038. [CrossRef]

15. Kothencz, G.; Kolcsár, R.; Cabrera-Barona, P.; Szilassi, P. Urban green space perception and its contribution to well-being. *Int. J. Environ. Res. Public Health* **2017**, *14*, 766. [CrossRef] [PubMed]

16. Bowler, D.E.; Buyung-Ali, L.M.; Knight, T.M.; Pullin, A.S. A systematic review of evidence for the added benefits to health of exposure to natural environments. *BMC Public Health* **2010**, *10*, 456. [CrossRef] [PubMed]

17. Escobedo, F.J.; Kroeger, T.; Wagner, J.E. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. *Environ. Pollut.* **2011**, *159*, 2078–2087. [CrossRef]

18. Hasan, S.; Ukkusuri, S.V. Urban activity pattern classification using topic models from online geo-location data. *Transp. Res. Part C Emerg. Technol.* **2014**, *44*, 363–381. [CrossRef]

19. Longley, P.A.; Adnan, M. Geo-temporal twitter demographics. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 369–389. [CrossRef]

20. Frank, M.R.; Mitchell, L.; Dodds, P.S.; Danforth, C.M. Happiness and the patterns of life: A study of geolocated tweets. *Sci. Rep.* **2013**, *3*, 2625. [CrossRef]

21. Campagna, M. The geographic turn in social media: Opportunities for spatial planning and geodesign. In Proceedings of the International Conference on Computational Science and Its Applications, Guimarães, Portugal, 30 June–3 July 2014; pp. 598–610.

22. Sagl, G.; Resch, B.; Hawelka, B.; Beinat, E. From social sensor data to collective human behaviour patterns: Analysing and visualising spatio-temporal dynamics in urban environments. In Proceedings of the GI-Forum, Berlin, Germany, 36 July 2012; pp. 54–63.

23. Ghani, A.; Zubair, M.; Saeed, M.I.; Singh, D. SOS: Socially Omitting Selfishness in IoT for smart and connected communities. *arXiv* **2020**, arXiv:08948.

24. Nouh, R.M.; Singh, D. Introducing blockchain for smart city technologies and applications. In *Blockchain Technology for Smart Cities*; Springer: Seoul, Korea, 2020; pp. 1–17.

25. Agrahari, A.; Singh, D. Smart city transportation technologies: Automatic no-helmet penalizing system. In *Blockchain Technology for Smart Cities*; Springer: Seoul, Korea, 2020; pp. 115–132.

26. Smiley, M.J.; Roux, A.V.D.; Brines, S.J.; Brown, D.G.; Evenson, K.R.; Rodriguez, D.A. A spatial analysis of health-related resources in three diverse metropolitan areas. *Health Place* **2010**, *16*, 885–892. [CrossRef]

27. Thornton, L.E.; Pearce, J.R.; Macdonald, L.; Lamb, K.E.; Ellaway, A. Does the choice of neighbourhood supermarket access measure influence associations with individual-level fruit and vegetable consumption? A case study from Glasgow. *Int. J. Health Geogr.* **2012**, *11*, 29. [CrossRef] [PubMed]

28. Thornton, L.E.; Pearce, J.R.; Kavanagh, A.M. Using Geographic Information Systems (GIS) to assess the role of the built environment in influencing obesity: A glossary. *Int. J. Behav. Nutr. Phys. Act.* **2011**, *8*, 71. [CrossRef] [PubMed]

29. King, T.L.; Thornton, L.E.; Bentley, R.J.; Kavanagh, A.M. The use of kernel density estimation to examine associations between neighborhood destination intensity and walking and physical activity. *PLoS ONE* **2015**, *10*, e0137402. [CrossRef] [PubMed]

30. Buck, C.; Börnhorst, C.; Pohlabeln, H.; Huybrechts, I.; Pala, V.; Reisch, L.; Pigeot, I. Clustering of unhealthy food around German schools and its influence on dietary behavior in school children: A pilot study. *Int. J. Behav. Nutr. Phys. Act.* **2013**, *10*, 65. [CrossRef] [PubMed]

31. Li, F.; Zhang, F.; Li, X.; Wang, P.; Liang, J.; Mei, Y.; Cheng, W.; Qian, Y. Spatiotemporal patterns of the use of urban green spaces and external factors contributing to their use in central Beijing. *Int. J. Environ. Res. Public Health* **2017**, *14*, 237. [CrossRef]

32. Kovacs-Gyori, A.; Ristea, A.; Havas, C.; Resch, B.; Cabrera-Barona, P. London2012: Towards citizen-contributed urban planning through sentiment analysis of twitter data. *Urban Plan.* **2018**, *3*, 75–99. [CrossRef]

33. Haidery, S.A.; Ullah, H.; Khan, N.U.; Fatima, K.; Rizvi, S.S.; Kwon, S.J. Role of big data in the development of smart city by analyzing the density of residents in shanghai. *Electronics* **2020**, *9*, 837. [CrossRef]

34. Department of Economic and Social Affairs. *World Urbanization Prospects the 2014 Revision*; United Nations: New York, NY, USA, 2014.

35. Xiong, X.; Jin, C.; Chen, H.; Luo, L. Using the fusion proximal area method and gravity method to identify areas with physician shortages. *PLoS ONE* **2016**, *11*, e0163504. [CrossRef]

36. Shen, J.; Kee, G. *Development and Planning in Seven Major Coastal Cities in Southern and Eastern China*; Springer: Shanghai, China, 2017.

37. Weibo Statistics. Available online: http://ir.weibo.com/news-releases/news-release-details/weibo-corporation-report-fourth-quarter-and-fiscal-year-2018 (accessed on 10 June 2019).

38. Shen, Y.; Sun, F.; Che, Y. Public green spaces and human wellbeing: Mapping the spatial inequity and mismatching status of public green space in the central city of shanghai. *Urban For. Urban Green.* **2017**, *27*, 59–68. [CrossRef]

39. Ebrahimpour, Z.W.; Wanggen, W.; Cervantes, O.; Luo, T.; Ullah, T. Comparison of Main Approaches for Extracting Behavior Features from Crowd Flow Analysis. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 440. [CrossRef]

40. Shen, Y.; Karimi, K. Urban function connectivity: Characterisation of functional urban streets with social media check-in data. *Cities* **2016**, *55*, 9–21. [CrossRef]

41. Keeler, B.L.; Wood, S.A.; Polasky, S.; Kling, C.; Filstrup, C.T.; Downing, J.A. Recreational demand for clean water: Evidence from geotagged photographs by visitors to lakes. *Front. Ecol. Environ.* **2015**, *13*, 76–81. [CrossRef]

42. Ullah, H.; Wan, W.; Haidery, S.A.; Khan, N.U.; Ebrahimpour, Z.; Muzahid, A.A.M. Spatiotemporal patterns of visitors in urban green parks by mining social media big data based upon WHO reports. *IEEE Access* **2020**, *8*, 39197–39211. [CrossRef]

43. Xiaoli, T.; Mingxing, C.; Wenzhong, Z.; Yongping, B. Classification and its relationship with the functional analysis of urban parks: Taking Beijing as an example. *Geogr. Res.* **2013**, *32*, 1964–1976.

44. Zhang, S.; Zhou, W. Recreational visits to urban parks and factors affecting park visits: Evidence from geotagged social media data. *Landsc. Urban Plan.* **2018**, *180*, 27–35. [CrossRef]

45. Wang, Y.; Wang, T.; Ye, X.; Zhu, J.; Lee, J. Using social media for emergency response and urban sustainability: A case study of the 2012 Beijing rainstorm. *Sustainability* **2016**, *8*, 25. [CrossRef]

46. Xie, Z.; Yan, J. Kernel density estimation of traffic accidents in a network space. *Comput. Environ. Urban Syst.* **2008**, *32*, 396–406. [CrossRef]

47. Abidi, S.; Hussain, M.; Xu, Y.; Zhang, W. Prediction of confusion attempting algebra homework in an intelligent tutoring system through machine learning techniques for educational sustainable development. *Sustainability* **2019**, *11*, 105. [CrossRef]

48. Language, R. Available online: https://cran.rproject.org/ (accessed on 12 May 2019).

49. Studio, R. Available online: https://www.rstudio.com (accessed on 12 May 2019).

50. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Routledge: London, UK, 2018.

51. Maroko, A.R.; Maantay, J.A.; Sohler, N.L.; Grady, K.L.; Arno, P.S. The complexities of measuring access to parks and physical activity sites in New York City: A quantitative and qualitative approach. *Int. J. Health Geogr.* **2009**, *8*, 34. [CrossRef]

52. Yu, W.; Ai, T.; Shao, S. The analysis and delimitation of Central Business District using network kernel density estimation. *J. Transp. Geogr.* **2015**, *45*, 32–47. [CrossRef]

53. Yu, W.; Ai, T.; He, Y.; Shao, S. Spatial co-location pattern mining of facility points-of-interest improved by network neighborhood and distance decay effects. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 280–296. [CrossRef]

54. Ying, L.; Shen, Z.; Chen, J.; Fang, R.; Chen, X.; Jiang, R. Spatiotemporal patterns of road network and road development priority in three parallel rivers region in Yunnan, China: An evaluation based on modified kernel distance estimate. *Chin. Geogr. Sci.* **2014**, *24*, 39–49. [CrossRef]

55. Wu, C.; Ye, X.; Ren, F.; Wan, Y.; Ning, P.; Du, Q. Spatial and social media data analytics of housing prices in Shenzhen, China. *PLoS ONE* **2016**, *11*, e0164553. [CrossRef]

56. King, T.L.; Bentley, R.J.; Thornton, L.E.; Kavanagh, A.M. Using kernel density estimation to understand the influence of neighbourhood destinations on BMI. *BMJ Open* **2016**, *6*, e008878. [CrossRef] [PubMed]

57. Hasan, S.; Zhan, X.; Ukkusuri, S.V. Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing, Chicago, IL, USA, 11–14 August 2013; p. 6.

58. Li, L.; Goodchild, M.F.; Xu, B. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartogr. Geogr. Inf. Sci.* **2013**, *40*, 61–77. [CrossRef]

59. Sun, Y.; Fan, H.; Li, M.; Zipf, A. Identifying the city center using human travel flows generated from location-based social networking data. *Environ. Plan. B Plan. Des.* **2016**, *43*, 480–498. [CrossRef]

60. Yuan, N.J.; Zheng, Y.; Xie, X.; Wang, Y.; Zheng, K.; Xiong, H. Discovering urban functional zones using latent activity trajectories. *IEEE Trans. Knowl. Data Eng.* **2014**, *27*, 712–725. [CrossRef]

61. Li, H.; Ge, Y.; Hong, R.; Zhu, H. Point-of-interest recommendations: Learning potential check-ins from friends. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 975–984.

62. Ullah, H.; Wan, W.; Haidery, S.A.; Khan, N.U.; Ebrahimpour, Z.; Luo, T. Analyzing the spatiotemporal patterns in green spaces for urban studies using location-based social media data. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 506. [CrossRef]

63. Liu, Q.; Ullah, H.; Wan, W.; Peng, Z.; Hou, L.; Qu, T.; Haidery, S.A. Analysis of green spaces by utilizing big data to support smart cities and environment: A case study about the city center of shanghai. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 360. [CrossRef]

64. Maia, M.; Almeida, J.; Almeida, V. Identifying user behavior in online social networks. In Proceedings of the 1st Workshop on Social Network Systems, Glasgow, UK, 1–4 April 2008; pp. 1–6.

65. Demographia World Urban Area. 2019. Available online: http://www.demographia.com/db-worldua (accessed on 14 June 2019).

66. Xiao, Y.; Wang, Z.; Li, Z.; Tang, Z. An assessment of urban park access in Shanghai–implications for the social equity in urban China. *Landsc. Urban Plan.* **2017**, *157*, 383–393. [CrossRef]

67. Roberts, H.; Sadler, J.; Chapman, L. Using twitter to investigate seasonal variation in physical activity in urban green space. *Geo Geogr. Environ.* **2017**, *4*, e00041. [CrossRef]