Tweeting the High Line Life: A Social Media Lens on Urban Green Spaces

Jisoo Sim [1],*, Patrick Miller [2] and Samarth Swarup [3]

[1]   Korea Research Institute of Human Settlement, Urban Division, Sejong-si 30147, Korea
[2]   Landscape Architecture Program, College of Architecture and Urban Studies, Virginia Tech, Blacksburg, VA 24061, USA; pmiller@vt.edu
[3]   Biocomplexity Institute and Initiative, University of Virginia, Charlottesville, VA 22904, USA; swarup@virginia.edu
*   Correspondence: jisoosim@krihs.re.kr

check for updates

Abstract: The objective of this study is to investigate elevated parks as urban green spaces using social media data analytics. Two popular elevated parks, the High Line Park in New York and the 606 in Chicago, were selected as the study sites. Tweets mentioning the two parks were collected from 2015 to 2019. By using text mining, social media users' sentiments and conveyed perceptions about the elevated parks were studied. In addition, users' activities and their satisfaction were analyzed. For the 606, users mainly enjoyed the free events at the park and worried about possible increases in housing prices and taxes because of the 606. They tended to participate in physical activities such as biking and walking. Although the 606 provides scenic observation points, users did not seem to enjoy these. Regarding the High Line, users frequently mentioned New York City, which is an important aspect of the identity of the park. The High Line users also frequently mentioned arts and relaxation. Overall, this study supports the idea that social media analytics can be used to gain an understanding of the public's use of urban green spaces and their attitudes and concerns.

Keywords: urban green space; elevated park; big data; social media analytics

1. Introduction

Social media is emerging as a key data source and is playing an increasingly important role in the urban design and planning process. For example, social media data about urban spaces can help researchers understand urban use, urban area characteristics, and user perceptions [1–3]. In particular, social media data have the potential for enhancing the understanding of human behavior [4,5] and attitudes of a place [6]. Since social media provides a forum for communicating and sharing thoughts [7], it can be used in the design and planning process as a tool for understanding user characteristics. Martí et al. describe how social media and other recent technologies have transformed people's interactions in urban spaces [8]. As researchers pointed out, the importance of social media analytics in urban design and planning is that social media analytics provide a means to collect data efficiently, gather users' thoughts directly, and collect past data.

Traditional research methods in the design and planning fields fall into two categories. One is a quantitative approach, such as surveys, and the other is qualitative, such as interviews and observational studies [9]. The quantitative approach gathers data directly from the user in the form of surveys and certain types of observational studies. Qualitative studies rely on interviews, community meetings, and self-reporting methods [9]. Field observation has been used in many studies related to urban areas [10–12]. Surveys of anticipated park users are often used to obtain data about the survey participant. With surveys, it easy to verify who the participant is and the extent to which they are

representative of park users. In many studies, a survey is used to understand people's attitudes and their behavior in urban open spaces such as a park. Peters et al. used a survey to find that urban parks can promote social cohesion [13]. Whiting et al. used a survey of park visitors to identify motivations and preferences for outdoor recreation [14]. While traditional assessment tools can be used to evaluate on-site benefits of parks through surveys, interviews, and observations [10,15–17], the weakness of these methods is that preconceptions of the researcher can be reflected in the questions being asked or, in some cases, not asked. If a researcher asks a question with a positive framing, a respondent tends to answer positively. Although traditional methods have strengths, these approaches also have limitations compared to the new method as follows. They are often cost and labor intensive. In addition, they are static in time [18].

Social media has some advantages when compared to these traditional approaches. Though social media analytics have become quite common in several fields of research in the last ten years, its use in landscape planning and design is less prevalent. Social media, as a means of accessing users' thoughts and attitudes toward a place, can be used in landscape architecture to figure out how people react to and interact with a landscape or places. Historically, landscape architecture has attempted to investigate the relationship between humans and landscape. To do this, many studies have developed methods for understanding how people react to the landscape through survey, observation, and other tools. Social media data, since they are user-generated, offer a new mode of observation. Users post their thoughts and behavior on a social media platform to share with other users, not just with the researchers.

The objective of this study is to understand elevated parks as urban green spaces through social media. The research questions are (1) What do people use elevated parks for? and (2) What are people's opinions about elevated parks? The objective of the first question is to understand how the public perceives elevated parks through their social media postings. Although social media analytics have limitations, they also have the potential to scan and understand broad swathes of public opinion. In terms of understanding elevated parks, this approach allows the researcher how the main themes of public concern relate to the elevated parks.

The second question is related to public attitudes toward elevated parks and the way social media data can provide useful information on user satisfaction. They do this by examining the results from activity detected on a short media post and user sentiment about the activity. This study identifies a way to interpret social media data and provide useful information for designers and planners.

This article is organized into five sections: Introduction, Literature Review, Methods, Results, and Conclusion. The literature review answers three questions: (1) what constitutes social media, (2) what social media analytics are, and (3) how researchers obtain insights from social media. After reviewing previous studies, data collection and data analytics are introduced in the Methods section. This section presents the different analytics in detail, as well as the findings from the analytics. From these findings, the Conclusion draws implications for the future use of social media as a lens to understand urban green spaces.

2. Social Media

2.1. Defining Social Media

The emergence of Big Data has opened up new possibilities for urban studies. Among Big Data sources, social media comprises internet-based applications that provide services to users to create and share user-generated content [19,20]. Social media has reshaped our way of thinking about user data and become an important data source in multiple academic fields, such as urban studies, humanities, geography, and sociology. Social media has grown dramatically because of the emergence of the "Web 2.0" features [21], the development of internet accessibility, and increase in smartphone use [22]. Web 2.0 is a collection of technologies which allow users to create content, interact with other users, and share information [20]. These developments are particularly well-suited to landscape

architecture, because these new data sources can offer insights into how local environments are used and experienced.

Kaplan and Haenlein classify social media into several categories [20]. The first type is collaborative media, such as Wikipedia. The second one is designed to share content, such as YouTube and Flickr. The third one is social networking platforms, such as Twitter and Facebook. The fourth one is blogging platforms, such as WordPress and Tumblr.

Social media improves public engagement tools in two ways: social media contributes to augmenting public engagement, and social media works as public communication channels. Regarding public engagement, many researchers agree that social media use increases public engagement [23], though a well-known limitation is that the population of social media users skews heavily towards younger age groups. On the other hand, this bias contributes to increasing the young generation's participation in urban planning [24]. Fredericks and Foth describe how social media can be used as a means of augmenting public participation in planning [25]. Social media becomes a virtual public sphere by providing supplementary channels to traditional ones. In terms of a new public sphere, social media also provides a new communication space for the public. According to Jendryke et al., social media helps us to understand many issues in a city, such as disasters, emergencies, and others [26]. Brkovic and Stetovic also describe the importance of social media in terms of communication within communities. The authors explain that social media platforms act as the outreach channels of urban planners to communicate with residents for gathering information, announcing news, and sharing resources [27]. Sui and Goodchild mention the use of social media platforms to support communication among citizens [28].

2.2. Social Media Analytics

Text mining, also known as text analytics, allows researchers to reveal the user attitudes behind textual data. This field, which is at the intersection of information retrieval, data mining, statistics, and linguistics, extracts meaningful information from unstructured textual data [29]. The emergence of social media applications has contributed to the growth of text mining usage. Text mining borrows techniques from several areas, such as Natural Language Processing (NLP), Machine Learning (ML), Information Retrieval (IR), and Data Mining (DM).

Hu and Liu describe the text mining process in three phases: text pre-processing, text representation, and knowledge discovery [30]. Text pre-processing, also known as data processing, makes the textual data into a consistent form to enable text analysis. This step includes stop word removal and stemming. Stop word removal deletes all stop words which are considered meaningless in textual data. Stemming turns words to their stems based on root form [30]. By inflecting words to their root form, the representation has reduced variability. Other text pre-processing steps are lower-casing, normalization, and noise removal. Lower-casing changes all capitalization to lowercase to prevent the sparsity issue that the same words with difference cases map to the same form [31]. Normalization transforms text into a standard form. For example, it maps multiple forms of one word to a single normalized form, such as "stopwords", "stop-words", and "stop words" to "stopwords". It also changes a misspelled word to a standard form. For example, the words "wooooow" and "wuw" to "wow". Text pre-processing may differ according to the research design and purpose. In the case of analyzing tweets, most researchers eliminate hashtag, url, html, ID, and etc. before conducting text analytics.

The next step in text mining is text representation. This is a way to represent words into sparse vectors by word, known as the "Bag of Words (BOW)" representation. The basic assumption of the model is to slice textual information by word and count each word's frequency to visualize it [32]. For this mode, a clustering algorithm is widely used to visualize words. In the Bag of Words model, a sentence or paragraph is the bag of its words (Table 1).

Table 1. Example of BOW.

Sentences	Bag of Words (BOW)
(1) I walk on the High Line	"I", "walk", "on", "the", "High", Line"
(2) and you also walk on the High Line	"and", "you", "also", "walk", "on", "the", "High", Line"

The Bag of Words model contributes to detecting and identifying topics and activities in urban green spaces. Lim et al. examine park visitor sentiments and activities using text analytics [33]. In this study, the authors use Latent Dirichlet Allocation (LDA) to detect topics and activities in tweets about urban green spaces. The authors identify two main findings. One is that common activities occurred in larger parks that had food, drinks, and events and another is that park visitors liked to take a photo at prominent landmarks or historical places. These results may seem like a predictable finding that is commonly mentioned in studies [34]. However, the main contribution of this study is to examine the use of text analytics to detect behaviors in parks.

A refinement of the simple Bag of Words approach is the N-gram language model. N-gram is a sequence of N words in the textual data [35], which means a set of adjacent sub-sequences of words within a given text. For example, for the sentence "I walk in the High Line", when N is 2 (bigram), the sentence can be divided into "I walk", "walk in", "in the", "the High", and "High Line". When N is 3 (trigrams), the results are "I walk in", "walk in the", "in the High", and "the High Line". N-grams are widely used in statistical NLP for calculating the joint probability of a sentence [36]. For using vector scores and the probability of a sequence, many studies use the N-gram model to detect issues [37], behavior [38], sentiment [39], and social unrest [40]. Agrawal et al. use the term frequency model to monitor a subset of the stream elements [41].

In the case of detecting behavior, Myslin et al., in 2013, conducted a study to understand and develop a machine learning classifier to detect behaviors and attitudes to tobacco [38]. They collected over 7000 tweets related to tobacco from 2011 to 2012 and manually classified it. Using the data, they developed machine learning classifiers that were trained to detect sentiment [38]. In this study, a bigram model was used to calculate and predict the probability of sequential words for the machine learning classification algorithm.

For the last processing phase in text analytics, several methods such as classification, clustering, event detection, and sentiment analysis are used [30]. Among these tools, sentiment analysis is commonly used in text analytics, especially for analyzing social media data. According to Mukherjee, sentiment analysis is a means of understanding a user's feelings posted in positive or negative tweets by analyzing texts. It can be considered the study of people's opinions, sentiments, emotions, and attitudes [42]. In this study, we use the lexicon-based approach for sentiment analysis. The lexicon-based approach uses positive and negative sentiment terms by using a dictionary based-approach [43]. The dictionary-based approach collects a small set of opinion words and then grows this set by searching in a large collection of text [44].

There are two ways to detect the sentiment in textual information: polarity analysis and emotion analysis. Polarity analysis classifies sentiment into three categories, negative, positive, and neutral. Emotion analysis categorizes feelings as happy, sad, fearful, angry, surprised, and disgusted [45]. Although sentiment analysis is helpful to understand feelings and emotions expressed by social media users, this kind of analysis can be challenging because of the brevity of the postings and a tendency to use abbreviated words [46]. To reduce its difficulties and increase its validity, Hutto and Gilbert created a model known as Valence Aware Dictionary for sEntiment Reasoning (VADER) [46]. They examined the validity of a sentiment lexicon and then combined lexical features with five rules embodying words. After the authors developed the VADER, they examined its accuracy in social media data by comparing the results from the VADER with ground truth. The authors claim the strengths of VADER as follows: (1) it works well on short and abbreviated text, (2) it is constructed from a human-created sentiment lexicon, and (3) it is fast enough to analyze streaming data [46].

They developed a new sentiment lexicon list from existing sentiment lists such as Linguistic Inquiry and Word Count [47], Affective Norms for English Words, and General Inquirer and added human-created lists. To make a human-created lexicon list, they used a wisdom-of-the-cloud (WotC) to estimate a point for the sentiment for each feature. Each feature was rated with a score from extremely negative, −4, to extremely positive, 4, and zero for neutral by human raters. By doing so, they identified 7500 lexical features with validated score indicating the sentiment polarity. With these lexicons, VADER measures each word's sentiment score and sums the sentiment score over the text. Then, VADER normalize scores to be between −1 and +1. Positive sentiment refers to this score being larger than 0.05, and negative sentiment is under −0.05; otherwise, it is considered neutral (when the score is larger than −0.05 and smaller than 0.05). Compared to other sentiment analysis lexicons, VADER does well at estimating human sentiment, especially in tweets [46].

2.3. Related Studies

Based on the attributes of social media data, researchers use it in detecting landscape and land uses, supporting public engagement, and identifying the relationship between people and place. In this part, we introduce studies that used social media data as a major data source in urban studies and landscape architecture. Since studies in landscape architecture are comparably fewer than urban studies, social media use in urban studies is mainly discussed. Many studies aim to (1) detect hidden functions of urban areas and (2) detect hidden relationship between human and landscape.

First, social media allows characterization of the use of urban areas, green spaces, and protected areas. Tu et al. discover urban functions by using social media data [48]. From the data, the authors extracted the behaviors of users and identified the time of posting. By doing so, they reveal the use of urban areas by functions such as working and time [48]. In a similar study, Chen et al. depict urban functions by using hot place analysis with social media data [49]. The authors identify the urban functions and places that emerge as central places. Silva et al. show that using social media data allows the detection of characteristics of each urban area [50]. In urban planning, many studies use social media to find functional areas. Studies in landscape architecture also use social media to detect the ecosystem. Oteros-Rozas et al. examine over a thousand photos from Flickr and Panoramio to identify the relationship between landscape diversity and ecosystem services [51]. Landscape diversity is an important aspect to assess the quality of the landscape and cultural ecosystem services refer to the benefits of green spaces to people. The authors find a weak and positive relationship between the two factors [51].

Second, social media provides a chance to understand how people interact with the landscape. Tieskens et al. used photos in Flickr and Panoramio to examine the relationship between landscape attributes and preferences [52]. Many researchers in landscape architecture have similarly focused on the relationship between landscape attributes and people's preferences [52]. However, most studies use surveys with photographs, which means there is a gap between the photos as representative media and real places. These studies may reduce the gap between the photos and reality by using visitors' uploaded photos. Contents and multi-media information posted on social media platforms allow researchers to detect a new use of urban place and the relations between human and nature. In particular, social media data contain geo-location information, which contributes to understanding how people use a place. Social media studies which use geotagged information often provide meaningful information for the design and planning. Wu examined the use of urban trails by examining geo-location information from Flickr and Twitter [53]. The author aimed to examine the usefulness of social media as a proxy for demand. By mapping geotagged postings along the trails, they make heat maps. At the same time, they calculate trail traffic to compare two datasets. As a result, the authors find a weak meaningful statistical correlation between photos and trail traffic [53]. Fisher et al. also attempted to predict trail traffic by using social media [54]. They measured the traffic from various techniques including geotagged photos and compared the traffic. They found that correlations between official statistics and

geotagged photos are between 55% and 95%. Based on the results, they assert that user-generated data can let urban planners and landscape architects manage and monitor trail use [54].

3. Methods

3.1. Study Sites

Two well-known and popular elevated parks, the 606 in Chicago and the High Line in New York City, were selected for this study. Their location in urban areas and the large number of people who visit these parks makes them ideal for a social media-based study.

The 606 was the Bloomingdale line that was an elevated railroad crossing Chicago in the east–west direction. The railroad was constructed in 1873 and was elevated in the 1910s to reduce pedestrian fatalities. The line was used for passenger trains but stopped its operation in 2001. After this, from 2002 to 2004, the greenway concept was introduced to the public. A group known as Friends of the Bloomingdale Trail launched in 2003 for advocacy [55]. In 2015, the 606 opened as the longest elevated park in the U.S. (see Figure 1)

| (**a**) | (**b**) |
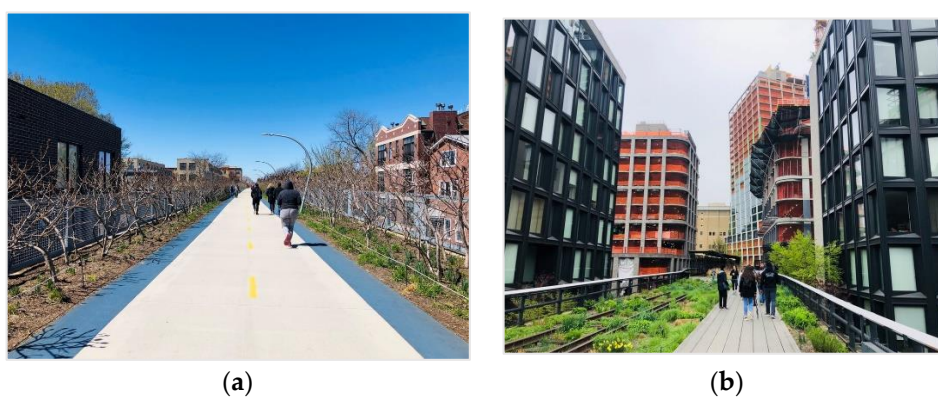
Figure 1. (**a**) The 606, (**b**) The High Line (copy right Jisoo Sim).

The High Line was the street-level railroad track along Tenth and Eleventh on Manhattan's West Side in 1847. In 1929, when safety issues were raised, the city, the state, and New York Central agreed on the West Side Improvement Project led by Robert Moses [56] and the street-level tracks turned into the elevated railway. The invention of standard container cargo fostered the growth of the interstate trucking system during the 1950s [57] and it also led to shrinkage of rail traffic throughout the U.S. The owner Conrail decided to disconnect the High Line elevated railway from the national rail system in 1980 for a year. While the railway was reconnected in 1981, there was no demand to use this line. The tracks ceased operation at this point. The High Line park opened in 2009 and has achieved economic success, becoming an iconic elevated park. In 2015, seven years after its opening, 7.6 million people visited the park, which was nearly six times the number of visitors in the first year, 1.3 million [58]. It is also expected to generate around USD 1 billion in tax revenues to the city over the next 20 years [59]. The success of the High Line has led other cities to want to build their own elevated parks to revive struggling communities.

3.2. Data Collection

We chose Twitter for social media analytics in this study due to several reasons: (1) Twitter users generate short messages, limited to 280 characters (until late 2017, the limit was 140 characters), (2) Twitter is convenient to access and collect data, and (3) many studies use Twitter as the main source. We introduce Twitter in general, its users, and its data structure. Following the approach used by Liu [42], we codified keyword-based queries related to each elevated park. We submitted each keyword

to Twitter's search interface to access full historical tweets instead of using the streaming API. By doing so, we collected tweets posted from January 2015 to June 2019 for each elevated park (Table 2).

Table 2. Total number of tweets.

	Keywords	Number of Tweets (Before Text Pre-Processing)	Total Number of Tweets for the Study
The 606	"The 606", "the Bloomingdale trail"	14,340	12,952
The High Line	"The High Line"	206,229	165,347

The data only include tweets posted in English. Since most analytics in text mining including sentiment analysis are developed for the English language, using non-English language data in the analysis derives inaccurate results [60]. To reduce the inaccuracy, we only include English tweets. Among the total tweets including keywords, we exclude all tweets posted in other languages except English. For the 606 (*n* = 14,340), we deleted 497 tweets (English tweets = 13,843). For the High Line, after the keyword search (*n* = 206,229), all tweets using other languages other than English were deleted.

After pre-processing for text mining, we reviewed the data for each park and removed unrelated tweets for reducing inaccuracy. First, we plotted the number of tweets by day for each elevated park for detecting abnormal dates. For the abnormal dates, we reviewed all tweets posted on these dates and identified keywords to delete unrelated tweets. Second, we found heavy users to detect unrelated tweets. After finding heavy users who posted more than the average level, we reviewed the profiles of these users on Twitter. For example, unrelated tweets use the elevated park name as their business in other cities. In the case of the 606, there is a jazz bar, "the 606 Club", in London. One of the model names of a product uses "High Line". These are examples of unrelated tweets. If a tweet included unrelated keywords such as the club name or product name, these tweets were eliminated before the analysis. Through these filtering processes, we deleted tweets which were not related to the elevated park. The total number of tweets of the 606 is 12,952 and the High Line is 165,347.

3.3. Data Analytics

After the text pre-processing, we conducted three main analyses: (1) topic analysis, (2) sentiment analysis, and (3) text analysis. Topic analysis is conducted to discover the main issues being discussed about each elevated park by month. Sentiment analysis aims to detect the public sentiment towards the elevated parks. For sentiment analysis, we used the VADER model, which was developed especially for social media data [46], because of its accuracy, especially for a short text, compared to other sentiment analysis models. Text analysis, at a more basic level, looks at the term frequency of selected keywords and the relationship between corpora. We selected keywords related to physical activities, social activities, nature, and scenery based on the survey questionnaire, then compared the sentiment of each activity. Through the analysis, the relationship between activity and sentiment in the virtual environment can be revealed.

4. Results

4.1. Public Attitudes about the Elevated Park

4.1.1. The 606

Regarding the 606, the number of tweets and the main topics of tweets relate to the residential location of the park and its proximal neighborhoods. The dates with the top five numbers of tweets show some of the main issues people are interested in. One of the abnormal dates was an event day for residents—on the first anniversary of the 606—and another one was the announcement of a block party in the 606. Two days were related to crime in the 606, such as theft and robbery. One of the dates was a protest of rising rents, property prices, and taxes. These show that the interests of residents

and frequent park visitors tend towards their neighborhoods and their communities. In terms of the number of tweets, we find that users mentioned the 606 mostly in May, July, September, and October. The number of tweets by year shows a decrease by year. Since the 606 opened in June 2015, the collected tweets include the tweets from August 2015 to July 2019 (Figure 2).
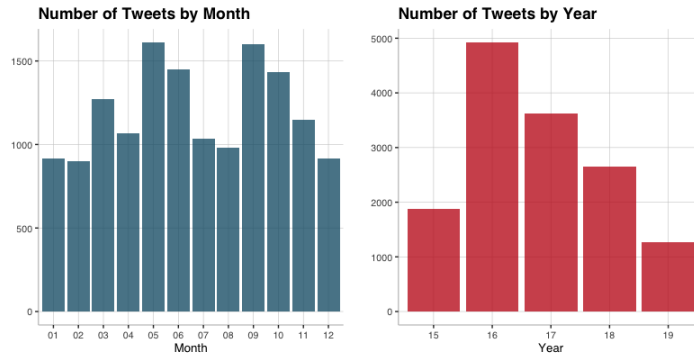

Figure 2. (**a**) Number of tweets by month, (**b**) Number of tweets by year.

The topic analysis reveals the changes in topic by month (Figure 3). During the winter season, from December to February, park users frequently mentioned "light" and "walk" (or "walking"). Compared to the other season, "jazz" was also mentioned in January and February. Some of the main topics of the spring were worries such as "robberies", "crime", and "gentrification". From March to May, people tweeted worries liked "robberies", "police", "crime", "night", "gentrification", and "housing". They also mentioned "tensions" and "buzzing" to express their feelings. Among the spring months, April shows a different result compared to other months. In April, people mentioned activities such as "bike" and "walk" frequently, and events such as "free jazz night". In the summer season, park users mentioned physical activities such as "walk", "bike", "run", with their status such as "happiness". For the last, during the fall season, people liked to take "photos". Compared to other seasons, people mentioned "photos" in the fall season more often. They also tweeted "beautiful" in November, which was the first time that they described their feelings about the scenery. Overall, park users worried about "housing", "prices", and "crimes" in the 606. Generally, they liked to do physical activities during all seasons; however, the intensity of physical activity may differ from season to season, from light walking in the winter to running in summer.
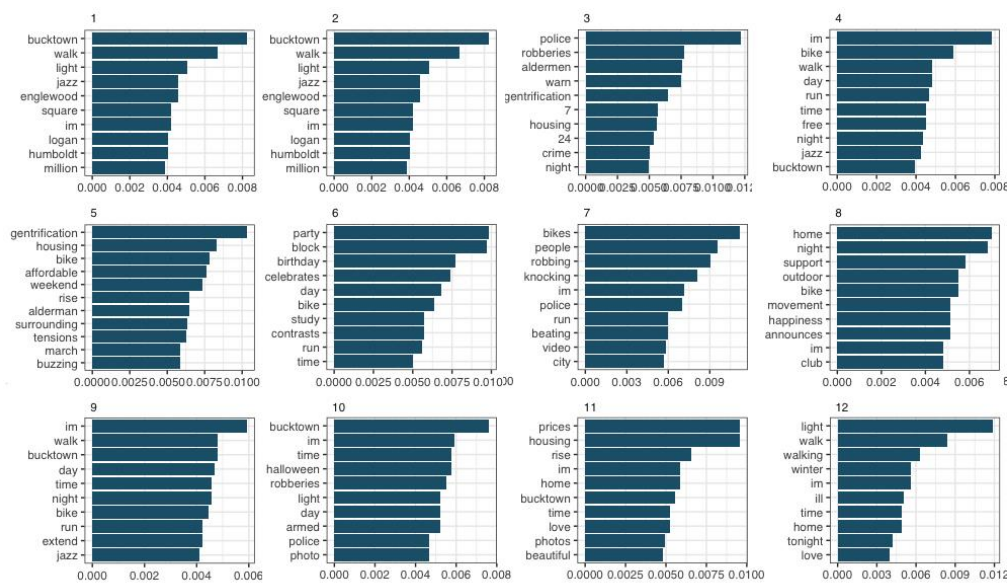

Figure 3. Topic changes by month (x: %).

For sentiment analysis, users tend to mention the 606 positively. When we compared the sentiment by year, the histogram shows that positive sentiments increased by year. Although the total number of tweets decreased by year, users tended to post about the 606 more positively than negatively (Figure 4a). An interesting aspect is the sentiment of summer and winter (Figure 4b). The difference in sentiment by month is slight, but the positive sentiment percentage of July and December is higher than other months (Figure 4c). These may reflect occasions such as Independence Day and Christmas.
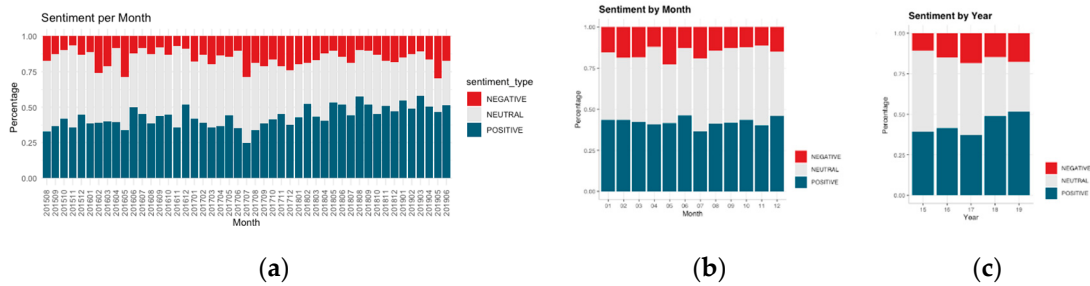

Figure 4. (**a**) Sentiment per month, (**b**) Sentiment by month, (**c**) Sentiment by year of the 606.

To scrutinize the sentiment analysis, we conducted bigram network analysis for each sentiment tweet. Users positively posting about the 606 were celebrating birthdays, free and cheap events, and a block party (Figure 5a). For the negative tweets, users mentioning the 606 focused on protest and gentrification, affordability and housing, robberies and police, surroundings and tensions, and bike robbing (Figure 5b). These tweets show issues which users worried about, with two main themes: one is gentrification, and the other one is crime.
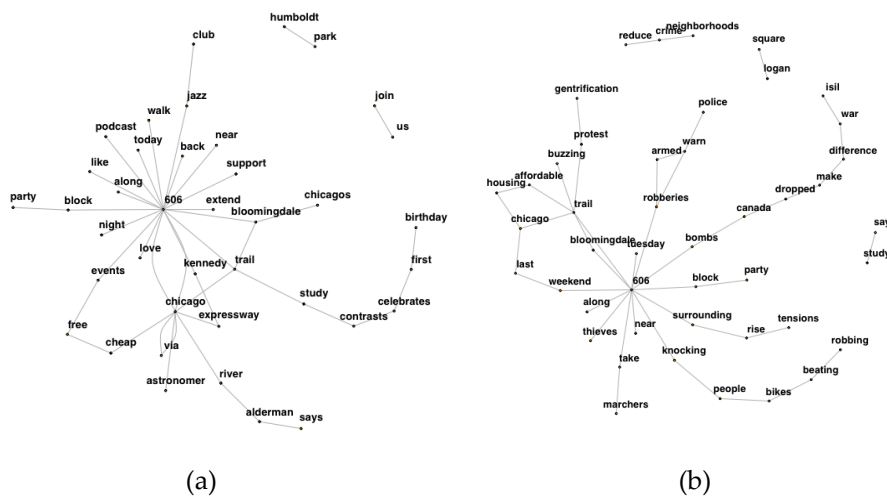

Figure 5. (**a**) Network diagram for positive sentiments, (**b**) Network diagram for negative sentiments.

4.1.2. The High Line

The High Line had the most tweets (*n* = 165,347) compared to other elevated parks in our dataset. After text preprocessing, a total of 165,347 tweets were identified for the analysis and this was almost eleven times more than the 606 tweets (*n* = 14,340). Like the 606, we reviewed the abnormal dates to identify big events or issues in the High Line. However, there was no special event on the top 10 dates of total number of tweets, except for the 5th Annual Coach and Friends of the High Line summer party on 25 July 2015. Other dates were in 2015, which was the top tweeted year of the study period (Figure 6).
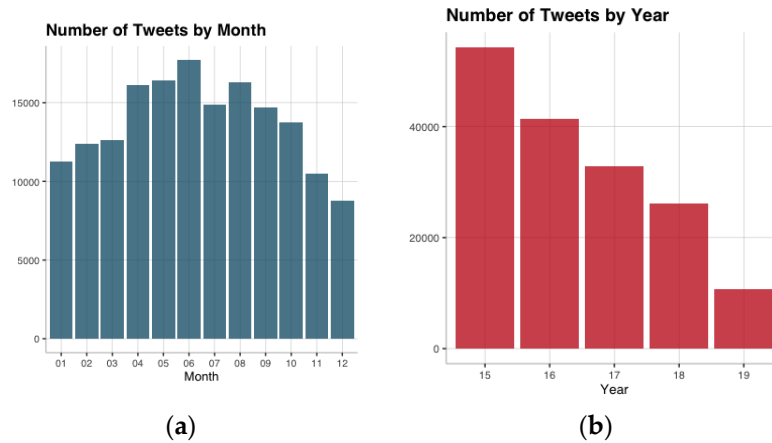
Figure 6. (**a**) Number of tweets by month, (**b**) Number of tweets by year.

Regarding the topic changes by month, the High Line shows more stable patterns in topic compared to the 606. Figure 7 shows the results from the topic analysis. Generally, people tweeting about the High Line mentioned "art", "walk", "photo", "people", and "city" frequently. The High Line users mentioned "photo" in all seasons, which indicates that the High Line is a place to memorize and post on social media, not a daily visit place, for the majority of users. The High Line users mentioned "building" and "tower" frequently also.
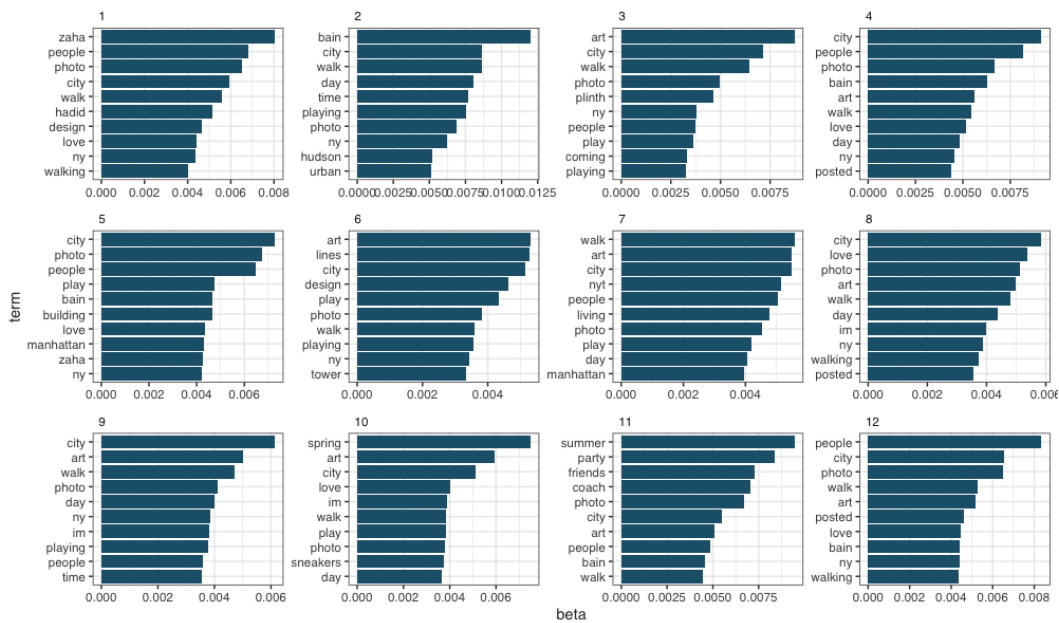

Figure 7. Topic changes by month.

For the number of tweets by month and year, the High Line users tended to post their experiences on social media platforms from summer to fall. Moreover, the number of tweets by year decreased from 2015 to 2019. Compared to the 606, the High Line users were more concentrated in summer and fall seasons (Figure 8).
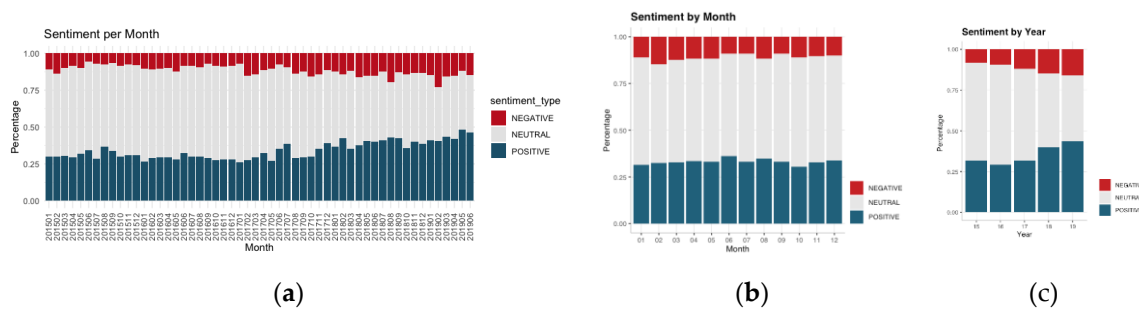
Figure 8. (**a**) Sentiment per month, (**b**) Sentiment by month, (**c**) Sentiment by year of the High Line.

Sentiment analysis of the High Line tweets reveals that the proportion of both positive and negative tweets has increased over the years (the proportion of neutral tweets has decreased). When we look at the sentiment changes by month, there are only small differences between months. In the High Line case, like the 606, the total number of tweets is decreasing; however, users tend to post about the High Line more positively than negatively.

The bigram network of positive sentiment represents related keywords (Figure 9a). For example, one of the clusters includes "im" (I'm), "pretty", and "good". Some clusters show the events in the High Line, including "friends", "coach", "annual", and "5th". The main cluster shows a number of keywords related to the city. "Central" and "Park" are also frequently mentioned with the High Line. From these results, one of the biggest reasons of the High Line's success can be considered a scenic urban view and the city itself. Regarding the negative sentiments (Figure 9b), the bigram network analysis shows some keywords like "abandoned" with "railroad" and "chelsea" with "site". An important caveat here is that lexicon-based sentiment analysis methods cannot always extract the full sentiment of a complex sentence. Thus, the sentiment analysis here should be considered a starting point for a deeper understanding of user opinions.
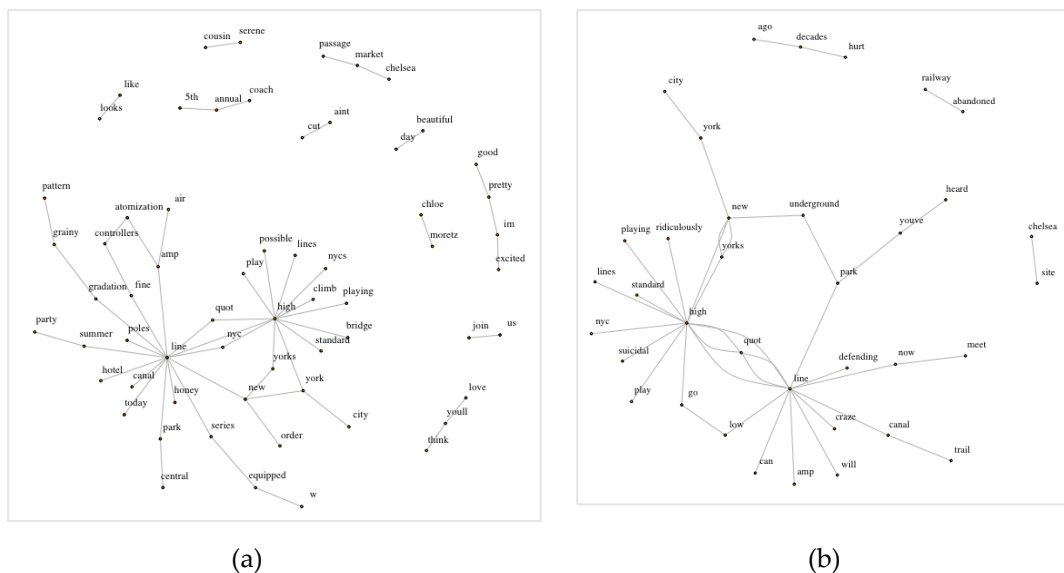

Figure 9. (**a**) Network diagram for positive sentiments, (**b**) Network diagram for negative sentiments.

4.2. Visitors' Activities and Satisfaction

This section addresses the relationship between users' activities and their sentiments through filtering a specific activity in the tweets and comparing the sentiment. Main activities in the elevated park fall into physical activities, viewing art, social activities, having picnics, scenic overlooks, and contact with nature (Table 3). Related keywords of each category were used to filter the relevant

tweets. Then, sentiment analysis was conducted to compare the sentiment between activity-related tweets and unrelated tweets. Results are as follows: (1) the 606 users posted positively about physical activities, enjoying art, social activities, and contact with nature, and (2) the High Line users posted positively in terms of having picnics and scenic overlooks. These results suggest that the 606 supports residents to use the trail as their daily activity place and the High Line allows users to have a picnic in the elevated park and have a special experience in the center of the city, overlooking the urban view.

Table 3. Number of tweets related to the activity.

Activity	Keywords	Number of Tweets	
		The 606	**High Line**
Physical activity	"walk", "jog", "run", "bike"	N = 730 (5.64%)	N = 10,509 (6.36%)
Art	"art","gallery", "museum"	987 (7.62)	15,021 (9.08)
Social	"friend", "family", "mom", "dad", "people", "father", "mother"	177 (1.37)	2408 (1.46)
Picnic	"lunch", "snack", "sandwich", "food", "coffee", "yum"	45 (0.35)	445 (0.27)
Overlooking	"overlook", "view", "see", "watch"	4 (0.03)	275 (0.17)

First, in the case of the 606, users were actively involved in physical activities, enjoying arts, social interaction, and contact with nature in the elevated park (Table 4). Users who were involved physical activity in the 606 were satisfied with the activity (positive = 51.78%) compared to the High Line (positive = 30.31%). Additionally, the percentage of physical activity-related tweets to the total tweets reinforces that the 606 users were more involved in physical activity (24.73%) than the High Line users (6.36%). For the activities involving viewing art, the result shows that the 606 users mentioned arts (33.43%) much more than the High Line users (9.08%) and more positively (48.02%) than the High Line users (39.74%). In terms of social interaction, users of both elevated parks positively tweeted about social interaction, which involves experiences with their friends or family. The 606 users positively posted about the social interaction (87.00%) slightly more than the High Line (84.50%). For contact with nature, the 606 users posted more positively (50.00%) than the users of the High Line (40.58%). However, when we consider the differences between nature-related tweets and other tweets in positive sentiment, both parks show similar differences of around 8%. Overall, the 606 users posted positively about physical activities, viewing art, social interaction, and contact with nature.

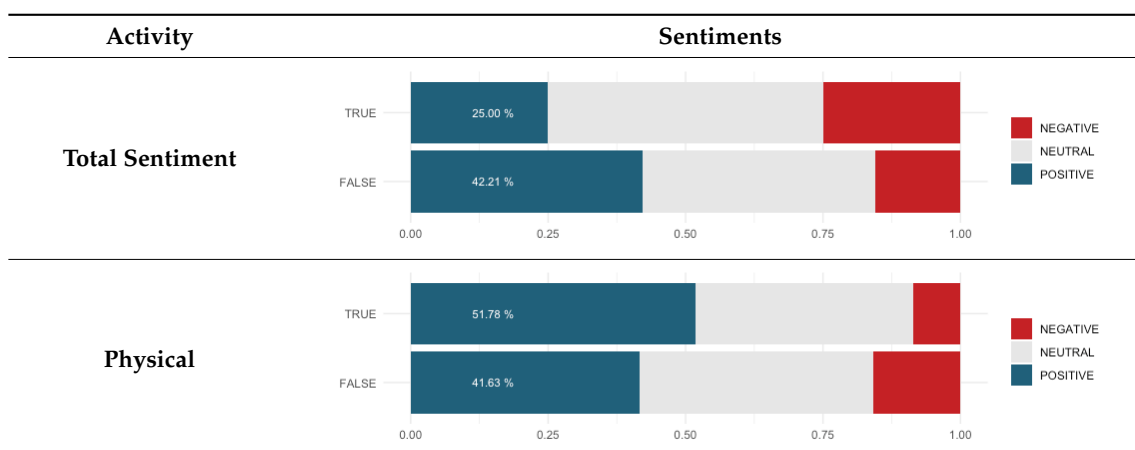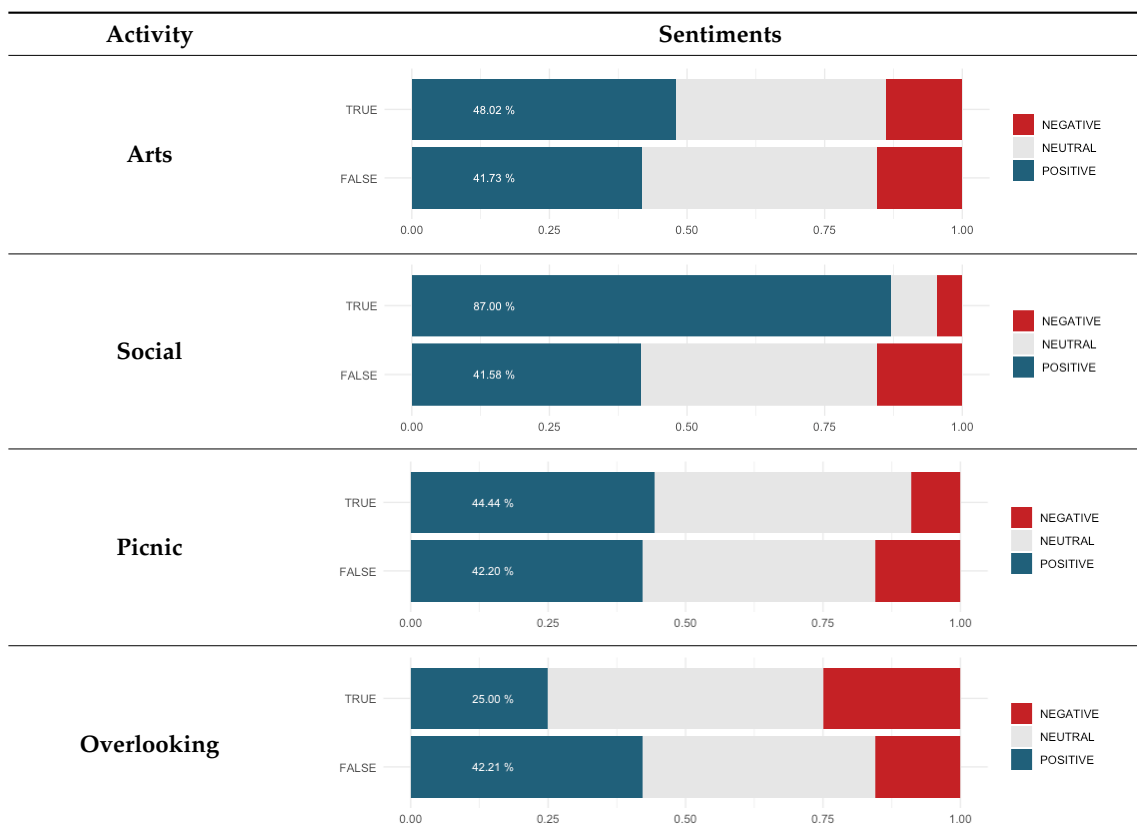**Table 4.** Sentiments for each activity in the 606.

Table 4. *Cont.*

Activity	Sentiments
Arts	TRUE 48.02 % / FALSE 41.73 % — NEGATIVE / NEUTRAL / POSITIVE
Social	TRUE 87.00 % / FALSE 41.58 % — NEGATIVE / NEUTRAL / POSITIVE
Picnic	TRUE 44.44 % / FALSE 42.20 % — NEGATIVE / NEUTRAL / POSITIVE
Overlooking	TRUE 25.00 % / FALSE 42.21 % — NEGATIVE / NEUTRAL / POSITIVE

Second, the High Line users posted positively about having picnics and overlooking experience rather than the 606 (Table 5). Users of the High Line enjoyed picnics slightly more positively (45.61%) than the 606 users (44.44%). Considering the differences between picnic-related tweets and other tweets in positive sentiment, the 606 users mentioning the picnic-related keywords were slightly more (44.44%) than unrelated tweets (42.40%). However, the result of the High Line shows the larger differences between them (45.61% vs. 32.96%). This result suggests that picnicking is a more common activity at the High Line than at the 606. The surprising result is about overlooking. Regarding overlooking, more High Line users positively tweeted about the views on the High Line (36.36%) rather than the 606 (25.00%). The 606 users were rarely satisfied with the views (25.00%) compared to other tweets about the 606 (42.21%) and the overlooking-related tweets of the High Line (36.36%). Negative tweets were also 25.00% in the 606 about overlooking. This suggests that the 606 users were not drawn to the park for the views. The overall results of the High Line show that the High Line users liked to have a picnic and enjoy the view on the elevated park more than the 606 visitors.

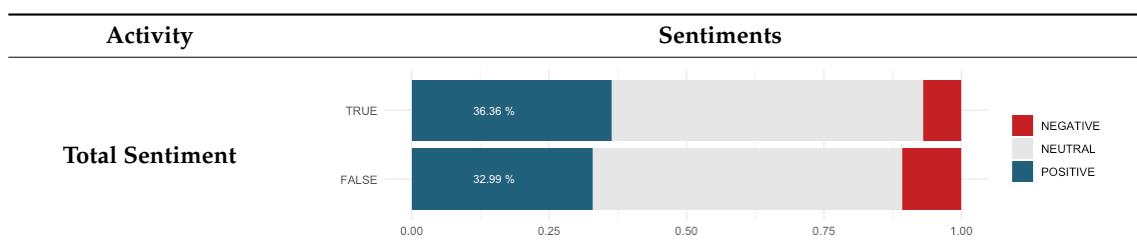**Table 5.** Sentiments for each activity in the High Line.

Activity	Sentiments
Total Sentiment	TRUE 36.36 % / FALSE 32.99 % — NEGATIVE / NEUTRAL / POSITIVE

Table 5. *Cont.*

Activity	Sentiments
Physical	
Arts	
Social	
Picnic	
Overlooking	

5. Discussion

In this section, we discuss results in detail based on our research questions.

5.1. What do People Use Elevated Parks for?

The 606 is well-integrated into the daily lives of nearby residents. The users liked to bike and walk along the park, were interested in events, and worried about housing prices and crimes. The clustering analysis highlights the main activities in the park and issues related to the park. For the main activity, biking and walking were the most frequently mentioned in the social media platform. This result supports the idea that the 606, which is designed for biking and walking [61], is well-designed and provides for the needs of residents. Interestingly, the cluster analysis points out issues related to gentrification and crime. The sentiment analysis also supports the idea that the 606 users worried about these issues. Another interesting finding from the sentiment analysis is that the 606 users rarely mentioned the park during the middle of summer and winter. Since the winter in Chicago is too cold to participate in outdoor activity, the low volume of tweets over the winter is not surprising. However, it is interesting that the volume of tweets is also low in the middle of the summer. These results broadly support the assertion that the 606, which is designed for residential and mixed use, is closely related to the daily life of residents.

The High Line users are more drawn to special events at the park. To them, the purpose of visiting the park is to play or to participate in events like "summer", "party", "coach", and "friend". A large

proportion may be tourists, since they also mentioned "New York City" frequently. The topic changes by month also support the same argument that the High Line users are happy to visit the park which is located in NYC. The keyword "NY" was mentioned during almost all months. Compared to the 606 users, the High Line users tended to take photos throughout the whole year according to the topic changes, which also could be due to tourism and/or special events.

5.2. What are People's Opinions about Elevated Parks?

Surprisingly, appreciating art in the park has a lower satisfaction level at the High Line than at the 606. This is despite the fact that there are many galleries and museums around the High Line, and the High Line regularly exhibits artworks. This result is difficult to directly compare and is one of the analyses that can only be achieved through social media analysis. In the case of the 606, compared to the High Line, there are many cases of appreciating art only during certain seasons. This may be due to 606 visitors perceiving artwork at the park as an unexpected feature and thus more novel and pleasing.

Social activities are viewed positively in both parks. This result is presumed to be due to the characteristics of the place as a park and a positive reaction to the activities of being with friends and family. Since this study was targeted at elevated parks, we did not compare whether these characteristics occurred in particular in these parks. Comparisons with other types of parks need to be conducted to see if this reaction is a characteristic of elevated parks.

In general, elevated parks consider scenic overlooks as an important design feature, and in many cases, viewing points are selected and benches from which to view the landscape are placed. Even in the case of the High Line Park, viewpoints were created and benches were placed at many points. The same goes for the 606. Many elevated parks are designed with the intention of wanting visitors to enjoy the pleasure of extended views. Surprisingly, however, the tweets relating to scenic overlooks have the lowest proportion of positive sentiment. This is contrary to the conventional design concept that has been considered so far, and it will be important to consider this result when designing elevated parks in the future.

6. Conclusions

In conclusion, we recommend to urban planners and landscape architects who want to make an elevated park that every elevated park has to be used in its own context. For example, the High Line, which is designed for art exhibitions and special events, is used by tourists who want to enjoy New York City. The 606, which is designed for the neighborhoods, sometimes reflects the neighborhoods' gentrification instead of economic vitality. This study offers an opportunity to show how every elevated park can be used in a different way. Based on this study, urban planners and landscape architects may consider the side effects of building elevated parks in their city.

This study addresses one of the reasons for the High Line's success through social media analytics. As a more well-known place, not only in New York City but also in the world, the volume of tweets indicates that users mentioned the High Line during all seasons except in November and December. Moreover, the number of tweets was over ten times more than the 606 tweets. The results of cluster analysis show that users frequently mentioned New York City and arts with respect to the High Line. Furthermore, the result of sentiment analysis and bigram network support the notion that the city and urban scenery in the park are the most frequently mentioned topics. These results suggest that tourism contributes to the success of the park. One of the other reasons for its success is arts. Users frequently mentioned arts along the elevated park with positive sentiment, which suggests that arts in the park contributes to the success of the High Line.

This study shows how social media data can be used to study urban parks. By using social media analytics, we gained some insight into how people think about and use elevated parks. One of the limitations of this study is that it only considered Twitter among social media. Additionally, this study does not consider the population bias that social media users tend to be more young, white, and male [5]. In the future, other social media platforms can be added to enrich the results.

References

1. Frias-Martinez, V.; Frias-Martinez, E. Spectral clustering for sensing urban land use using Twitter activity. *Eng. Appl. Artif. Intell.* **2014**, *35*, 237–245. [CrossRef]
2. Kelley, M.J. The emergent urban imaginaries of geosocial media. *GeoJournal* **2013**, *78*, 181–203. [CrossRef]
3. Lee, R.; Wakamiya, S.; Sumiya, K. Urban area characterization based on crowd behavioral lifelogs over Twitter. *Pers. Ubiquitous Comput.* **2013**, *17*, 605–620. [CrossRef]
4. Afzalan, N.; Muller, B. The Role of Social Media in Green Infrastructure Planning: A Case Study of Neighborhood Participation in Park Siting. *J. Urban. Technol.* **2014**, *21*, 67–83. [CrossRef]
5. Georgiou, T. Understanding the Real World through the Analysis of User Behavior and Topics in Online Social Media. Ph.D. Thesis, University of California, Santa Barbara, CA, USA, 2017.
6. Hollander, J.B.; Shen, Y. Using Social Media Data to Infer Urban Attitudes about Bicycling: An Exploratory Case Study of Washington DC. In *City Networks*; Springer: Cham, Switzerland, 2017; pp. 79–97.
7. Kankanhalli, A.; Tan, B.C.; Wei, K.-K. Contributing knowledge to electronic knowledge repositories: An empirical investigation. *Mis Q.* **2005**, *29*, 113–143. [CrossRef]
8. Martí, P.; Serrano-Estrada, L.; Nolasco-Cirugeda, A. Using locative social media and urban cartographies to identify and locate successful urban plazas. *Cities* **2017**, *64*, 66–78. [CrossRef]
9. Roberts, H.V. Using Twitter data in urban green space research. *Appl. Geogr.* **2017**, *81*, 13e20. [CrossRef]
10. Birch, E.L. *The Observation Man, A Study of William H. Whyte*; Planners Press: Chicago, IL, USA, 1986.
11. Gehl, J.; Svarre, B. *How to Study Public Life*; Island Press: Washington, DC, USA, 2013.
12. Jacobs, J. *The Death and Life of Great American Cities*; Vintage Books: New York, NY, USA, 1961.
13. Peters, K.; Elands, B.; Buijs, A. Social interactions in urban parks: Stimulating social cohesion? *Urban For. Urban Green.* **2010**, *9*, 93–100. [CrossRef]
14. Whiting, J.W.; Larson, L.R.; Green, G.T.; Kralowec, C. Outdoor recreation motivation and site preferences across diverse racial/ethnic groups: A case study of Georgia state parks. *J. Outdoor Recreat. Tour.* **2017**, *18*, 10–21. [CrossRef]
15. Carmona, M. *Public Places, Urban. Spaces: The Dimensions of Urban. Design*; Abingdon-on-Thames: Routledge, UK, 2010.
16. Ewing, R. Using a visual preference survey in transit design. *Public Work. Manag. Policy* **2001**, *5*, 270–280. [CrossRef]
17. More, T.A.; Stevens, T.; Allen, P.G. Valuation of urban parks. *Landsc. Urban. Plan.* **1988**, *15*, 139–152. [CrossRef]
18. Sim, J.; Miller, P. Understanding an Urban Park through Big Data. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3816. [CrossRef] [PubMed]
19. Boyd, D.M.; Ellison, N.B. Social Network Sites: Definition, History, and Scholarship. *J. Comput. Mediat. Commun.* **2007**, *13*, 210–230. [CrossRef]
20. Kaplan, A.M.; Haenlein, M. Users of the world, unite! The challenges and opportunities of Social Media. *Bus. Horiz.* **2010**, *53*, 59–68. [CrossRef]
21. Lin, Y.; Geertman, S. Can Social Media Play a Role in Urban Planning? A Literature Review. In *Computational Urban Planning and Management for Smart Cities*; Geertman, S., Zhan, Q., Allan, A., Pettit, C., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 69–84.
22. Kitchin, R. The real-time city? Big data and smart urbanism. *GeoJournal* **2014**, *79*, 1–14. [CrossRef]
23. Kleinhans, R.; Ham, M.V.; Evans-Cowley, J. Using Social Media and Mobile Technologies to Foster Engagement and Self-Organization in Participatory Urban Planning and Neighbourhood Governance. *Plan. Pract. Res.* **2015**, *30*, 237–247. [CrossRef]
24. Schroeter, R.; Houghton, K. Neo-planning: Location-based social media to engage Australia's new digital locals. *Aust. Plan.* **2011**, *48*, 191–202.

25. Fredericks, J.; Foth, M. Augmenting public participation: Enhancing planning outcomes through the use of social media and web 2.0. *Aust. Plan.* **2013**, *50*, 244–256. [CrossRef]

26. Jendryke, M.; Balz, T.; Liao, M. Big location-based social media messages from China's Sina Weibo network: Collection, storage, visualization, and potential ways of analysis. *Trans. Gis* **2017**, *21*, 825–834. [CrossRef]

27. Brkovic, M.; Stetovic, B. Harnessing social media for urban planning: An overview. In Proceedings of the 49th international society of city and regional planners congress, Brisbane, Australia, 1–4 October 2013.

28. Sui, D.; Goodchild, M. The convergence of GIS and social media: Challenges for GIScience. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 1737–1748. [CrossRef]

29. He, W.; Zha, S.; Li, L. Social media competitive analysis and text mining: A case study in the pizza industry. *Int. J. Inf. Manag.* **2013**, *33*, 464–472. [CrossRef]

30. Hu, X.; Liu, H. Text Analytics in Social Media. In *Mining Text Data*; Aggarwal, C.C., Zhai, C., Eds.; Springer US: Boston, MA, USA, 2012; pp. 385–414.

31. Ganesan, K. Getting Started with Text Preprocessing for Machine Learning & NLP. *Kavita Ganesan*, 23 February 2019. Available online: https://kavita-ganesan.com/text-preprocessing-tutorial/ (accessed on 5 August 2019).

32. Zhang, Y.; Jin, R.; Zhou, Z.-H. Understanding bag-of-words model: A statistical framework. *Int. J. Mach. Learn. Cyber.* **2010**, *1*, 43–52. [CrossRef]

33. Lim, K.H.; Lee, K.E.; Kendal, D.; Rashidi, L.; Naghizade, E.; Feng, Y.; Wang, J. Chapter 6—Understanding sentiments and activities in green spaces using a social data–driven approach. In *Smart Cities: Issues and Challenges*; Visvizi, A., Lytras, M.D., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 77–107.

34. Sather-Wagstaff, J. Picturing experience: A tourist-centered perspective on commemorative historical sites. *Tour. Stud.* **2008**, *8*, 77–103. [CrossRef]

35. Broder, A.Z.; Glassman, S.C.; Manasse, M.S.; Zweig, G. Syntactic clustering of the Web. *Comput. Netw. Isdn Syst.* **1997**, *29*, 1157–1166. [CrossRef]

36. Martin, J.H.; Jurafsky, D. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*; Pearson/Prentice Hall Upper Saddle River: Upper Saddle River, NJ, USA, 2009.

37. Laniado, D.; Mika, P. Making Sense of Twitter. In *The Semantic Web—ISWC 2010*; ISWC: Shanghai, China, 2010; pp. 470–485.

38. Myslín, M.; Zhu, S.-H.; Chapman, W.; Conway, M. Using Twitter to Examine Smoking Behavior and Perceptions of Emerging Tobacco Products. *J. Med. Internet Res.* **2013**, *15*, e174. [CrossRef]

39. Tang, D.; Wei, F.; Yang, N.; Zhou, M.; Liu, T.; Qin, B. Learning sentiment-specific word embedding for twitter sentiment classification. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Florence, Italy, 22 June 2014; pp. 1555–1565.

40. Compton, R.; Lee, C.; Lu, T.; Silva, L.D.; Macy, M. Detecting future social unrest in unprocessed Twitter data: 'Emerging phenomena and big data. In Proceedings of the 2013 IEEE International Conference on Intelligence and Security Informatics, Seattle, WA, USA, 4–7 June 2013; pp. 56–60. [CrossRef]

41. Agrawal, D.; Budak, C.; el Abbadi, A.; Georgiou, T.; Yan, X. Big Data in Online Social Networks: User Interaction Analysis to Model User Behavior in Social Networks. In *Databases in Networked Information Systems*; Springer: Aizu-Wakamatsu, Japan, 2014; pp. 1–16.

42. Liu, B. *Sentiment Analysis, Opinion Mining, Emotion Classification*; Cambridge University Press: Cambridge, UK, 2015.

43. Maynard, D.; Funk, A. Automatic Detection of Political Opinions in Tweets. In *The Semantic Web: ESWC 2011 Workshops*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 88–99.

44. Parrott, W.G. *Emotions in Social Psychology: Key Readings*; Psychology Press: East Sussex, UK, 2001.

45. Zhai, C.; Massung, S. *Text. Data Management and Analysis: A Practical Introduction to Information Retrieval and Text. Mining*; Morgan & Claypool: Williston, ND, USA, 2016.

46. Hutto, C.J.; Gilbert, E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, May 2014. Available online: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8109 (accessed on 6 August 2019).

47. Pennebaker, J.W.; Francis, M.E.; Booth, R.J. Linguistic inquiry and word count: LIWC 2001. *Mahway Lawrence Erlbaum Assoc.* **2001**, *2001*, 71.

48. Tu, W.; Cao, J.; Yue, Y.; Shaw, S.L.; Zhou, M.; Wang, Z.; Li, Q. Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 2331–2358. [CrossRef]

49. Chen, J.; Ma, B.; Cao, H.; Chen, J.; Fan, Y.; Xia, T.; Li, R. Attention region detection based on closure prior in layered bit Planes. *Neurocomputing* **2017**, *251*, 16–25. [CrossRef]

50. Silva, M.D.F.M.; Calijuri, M.L.; de Sales, F.J.F.; de Souza, M.H.B.; Lopes, L.S. Integration of technologies and alternative sources of water and energy to promote the sustainability of urban landscapes. *Resour. Conserv. Recycl.* **2014**, *91*, 71–81. [CrossRef]

51. Oteros-Rozas, E.; Martín-López, B.; Fagerholm, N.; Bieling, C.; Plieninger, T. Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecol. Indic.* **2018**, *94*, 74–86. [CrossRef]

52. Tieskens, K.F.; van Zanten, B.T.; Schulp, C.J.E.; Verburg, P.H. Aesthetic appreciation of the cultural landscape through social media: An analysis of revealed preference in the Dutch river landscape. *Landsc. Urban. Plan.* **2018**, *177*, 128–137. [CrossRef]

53. Wu, X.; Wood, S.A.; Fisher, D.; Lindsey, G. Photos, tweets, and trails: Are social media proxies for urban trail use? *J. Transp. Land Use* **2017**, 10. [CrossRef]

54. Fisher, D.M.; Wood, S.A.; White, E.M.; Blahna, D.J.; Lange, S.; Weinberg, A.; Lia, E. Recreational use in dispersed public lands measured using social media data and on-site counts. *J. Environ. Manag.* **2018**, *222*, 465–474. [CrossRef]

55. Gobster, P.H.; Sachdeva, S.; Lindsey, G. Up on the 606: Understanding the Use of a New Elevated Pedestrian and Bicycle Trail in Chicago, Illinois. *Transp. Res. Rec.* **2017**, *2644*, 83–91. [CrossRef]

56. Bighorse, A. The Highline: Past and Present. 13 May 2010. Available online: https://web.archive.org/web/20141023213640/http://www.gsd.harvard.edu/pbcote/courses/archive/2010/gsd6447/bighorse/ (accessed on 14 March 2019).

57. Waldheim, C. *Landscape as Urbanism: A General Theory*; Princeton University Press: Princeton, NJ, USA, 2016.

58. Ganser, A. High Line Magazine: B1G DA+A and Parks | the High Line Blog. *The High. Line,* 18 January 2017. Available online: http://www.thehighline.org/blog/2017/01/10/high-line-magazine-b1g-da-a-and-parks (accessed on 8 April 2018).

59. Bliss, L. The High Line's Biggest Issue—And How Its Creators are Learning from Their Mistakes. *CityLab* **2017**. Available online: http://www.citylab.com/cityfixer/2017/02/the-high-lines-next-balancing-act-fair-and-affordable-development/515391/ (accessed on 8 April 2018).

60. Wang, Z.; Joo, V.; Tong, C.; Chan, D. Issues of Social Data Analytics with a New Method for Sentiment Analysis of Social Media Data. In Proceedings of the 2014 IEEE 6th International Conference on Cloud Computing Technology and Science, Singapore, 15–18 December 2014; pp. 899–904. [CrossRef]

61. Sinha, A. Slow landscapes of elevated linear parks: Bloomingdale Trail in Chicago. *Stud. Hist. Gard. Des. Landsc.* **2014**, *34*, 113–122. [CrossRef]