The Grass is Greener on the Other Side: Understanding the Effects of Green Spaces on Twitter User Sentiments

Kwan Hui Lim[†], Kate E. Lee[‡], Dave Kendal[#], Lida Rashidi[†], Elham Naghizade[§ †], Stephan Winter[§], and Maria Vasardani[§]
[†] School of Computing and Information Systems, The University of Melbourne, Australia
[‡] School of Ecosystem and Forest Sciences, The University of Melbourne, Australia
[§] Department of Infrastructure Engineering, The University of Melbourne, Australia
[#] School of Technology, Environments and Design, The University of Tasmania, Australia
{kwan.lim, kate.lee, rashidi.l, e.naghi, winter, maria.vasardani}@unimelb.edu.au, dave.kendal@utas.edu.au

ABSTRACT

Green spaces are believed to improve the well-being of users in urban areas. While there are urban research exploring the emotional benefits of green spaces, these works are based on user surveys and case studies, which are typically small in scale, intrusive, time-intensive and costly. In contrast to earlier works, we utilize a non-intrusive methodology to understand green space effects at large-scale and in greater detail, via digital traces left by Twitter users. Using this methodology, we perform an empirical study on the effects of green spaces on user sentiments and emotions in Melbourne, Australia and our main findings are: (i) tweets in green spaces evoke more positive and less negative emotions, compared to those in urban areas; (ii) each season affects various emotion types differently; (iii) there are interesting changes in sentiments based on the hour, day and month that a tweet was posted; and (iv) negative sentiments are typically associated with large transport infrastructures such as train interchanges, major road junctions and railway tracks. The novelty of our study is the combination of psychological theory, alongside data collection and analysis techniques on a large-scale Twitter dataset, which overcomes the limitations of traditional methods in urban research.

CCS CONCEPTS

• **Information systems** → **Social networking sites**; **Social networks**; *Location based services*; *Data mining*; • **Applied computing** → Psychology; Sociology;

KEYWORDS

Green Spaces; Urban Areas; Empirical Study; Twitter

ACM Reference Format:
Kwan Hui Lim[†], Kate E. Lee[‡], Dave Kendal[#], Lida Rashidi[†], Elham Naghizade[§ †], Stephan Winter[§], and Maria Vasardani[§]. 2018. The Grass is Greener on the Other Side: Understanding the Effects of Green Spaces on Twitter User Sentiments. In *WWW '18 Companion: The 2018 Web Conference Companion, April 23–27, 2018, Lyon, France.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3184558.3186337

WWW '18 Companion, April 23–27, 2018, Lyon, France

ACM ISBN 978-1-4503-5640-4/18/04.
https://doi.org/10.1145/3184558.3186337

1 INTRODUCTION

Half of the world's population is living in urban areas today, and this figure is projected to increase to two-thirds of the world's population by 2050 [44]. With this rapid urbanization of today's cities, there is an increased emphasis on ensuring the well-being of people living in urban areas [15]. In response, city planners have sought to incorporate green spaces in urban areas, as green spaces are believed to improve the physical and mental health of people residing in these urban areas. The importance of incorporating green spaces is also reflected in the UN's recent Sustainable Development Goals, which have a specific target to provide accessible green space for all urban residents [45]. Given the importance of green spaces, this topic has also garnered the interest of researchers, increasingly so in recent years [12, 15, 38, 48].

Existing research has explored the emotional benefits of green spaces in urban areas [16, 21]. These are largely based on user surveys, questionnaires and case studies, which are typically either small in scale or involve the explicit participation of users. Moreover, these traditional methods are often intrusive, time-intensive and costly for researchers to perform a longitudinal study or fine-grained analysis involving participants. For example, to track users with a fine-grained resolution, personal tracking devices have to be used and worn by the participants. Similarly, to study sentiment change across the different days or months, surveys would need to be regularly administered over the course of the study, which is time consuming for both researchers and participants. To overcome these challenges of traditional methods, we apply sentiment analysis techniques on geo-tagged tweets posted by Twitter users, which serves as an unintrusive way of exploring sentiment expressed in user-generated content and is also easily available on a large scale.

Research Objectives and Contributions. In this empirical study, we aim to examine the effects that visits to green spaces have on people's sentiments and the implications of these findings for urban planning. The novelty is the combination of psychological theory and analysis of digital traces left by Twitter users, which, as we will demonstrate, overcomes the limitations of traditional methods in urban research. In particular, we will answer the following research questions (RQ).

- RQ1: How do sentiments and emotions differ in green spaces compared to urban areas?
- RQ2: How does the time of day and season of visit to green spaces affect these sentiments and emotions?

- RQ3: How does the proximity of green spaces affect these sentiments and emotions?

We apply this proposed methodology on a large set of 21.2 million tweets to better understand the relationship between green space and user sentiments, and our main findings are:

- Tweets in green spaces exhibit higher levels of joy, anticipation and trust (positive emotions), and lower levels of anger and fear (negative emotions), compared to tweets in urban areas.
- While tweets in green spaces are generally more positive than tweets in urban areas, the season (spring, summer, autumn, winter) when a tweet was posted affects the various emotion levels differently in green space and urban areas.
- In addition, we observe interesting changes in sentiments based on the hour, day and month that a tweet was posted, which reflect trends in real-life.
- We also find a positive correlation between the sentiment polarity of tweets in urban areas and their proximity to green spaces.

Structure and Organization. This paper is structured as follows. Section 2 provides an overview of literature on Twitter-related analytics and green space studies. Section 3 describes our dataset collection and analysis framework. Sections 4, 5 and 6 highlight the results from our experiments using Twitter, while Section 7 discusses the implications of our main findings. Finally, Section 8 concludes and summarizes the paper.

2 RELATED WORK

There are two streams of research that are related to our work, namely research on general Twitter-related analytics and research that examines the emotional benefits of green spaces.

General Twitter-related Analytics. Twitter is a popular microblogging social networking site that allows users to post short messages of 140 characters and share these tweets with their followers. In recent years, researchers have made extensive use of Twitter to understand many social phenomena and behaviours such as identifying popular topics [19] or witness accounts [42] associated with various places, studying correlations between user mobility patterns and happiness levels [14], predicting levels of happiness, food preferences and physical activities [28], recommending friends [5], predicting flu outbreaks [1], constructing interest profiles [7] and topical expertise [46, 53] of Twitter users, and numerous other applications in politics [10], academic conferences [49], community detection [23, 24], travel trends [13], crisis management [18], crowd sensing [36], event detection [11, 32, 52], among others. Although Twitter has been extensively used for these purposes, Twitter has not been used for the study of green spaces and their effects on user sentiment, to the best of our knowledge.

Analysis of Psycho-social Response to Green Spaces. The study of green spaces in urban areas have garnered strong interest in recent years [12, 15, 38, 48], ranging from determining the appropriate levels of green spaces [50] to understanding the usage patterns of urban green spaces [38]. Among these works, we are most interested in works that study the effects of green spaces on people in urban areas. Many of these works utilized surveys or questionnaires to understand how green spaces affect personal well-being [9], thermal comfort [48], life expectancy of residents [41], and prevalence of myopia [12]. Researchers [35, 43] augmented these surveys with clinical measurements to study the correlations between green spaces and stress level, via measurements of blood pressure and salivary cortisol levels. Others [4] have also used wearable biosensors to study the physiological response of users to different types of environments. However, green space research typically rely on traditional methods based on surveys, questionnaires, case studies or wearable sensors, and has not previously used twitter data analytics to explore user sentiment. This observation is supported by recent comprehensive literature surveys of existing work on the benefits of green spaces [16, 21].

Discussion. While previous research has examined interesting aspects of Twitter and green spaces separately, we note two key differences with our study, namely: (i) while the works using Twitter-related analytics present interesting and useful understanding of some social phenomena, none of these earlier works examine topics related to green spaces or its effect on people's sentiments; and (ii) while earlier studies of green space examine how green spaces are associated with various health and well-being outcomes, they are based on surveys, questionnaires or case studies, which are typically small in scale, intrusive, time-intensive, costly, and difficult to replicate. In contrast, our study utilizes a big data driven framework based on implicit digital traces left by Twitter users, which is large in scale and non-intrusive, to study how green spaces affect user sentiments across different time periods and spatial areas.

3 EXPERIMENTAL DESIGN

In this section, we describe our Twitter data collection approach and how we calculate the sentiment level related to a tweet.

3.1 Dataset and Data Collection

Our dataset comprises a set of 21.2 million tweets (2.2 million geo-tagged) generated by 10,510 users in Melbourne, Australia. We also have access to a green space dataset, comprising the locations and coverage of 482 green spaces (e.g., parks, gardens, green fields and other open areas) in the same city.

Twitter Dataset. We first describe our data collection methodology for the Twitter dataset, which was collected from Nov 2016 to Jan 2017 using the Twitter REST API. For this dataset, we employed a two-stage collection as follows:

1. **Stage 1 Collection**: This initial stage involves collecting all geo-tagged tweets (i.e., tagged with latitude/longitude coordinates) that are posted within a 5km × 5km grid in central Melbourne, Australia. This 5km × 5km grid is centered on approximately the Melbourne GPO building.
2. **Stage 2 Collection**: Based on the set of retrieved geo-tagged tweets (from Stage 1), we then proceed to extract the list of unique Twitter users who have posted these tweets, i.e., a set of seed users who have posted tweets in Melbourne, Australia. Thereafter, we retrieve the most recent 3,200 tweets of these users, as per Twitter API constraints, to build a tweeting profile for these users.

Green Space Dataset. We also have access to a green space dataset, provided by the City of Melbourne, which is the local government authority in charge of urban planning and regulations for the central Melbourne area. This dataset is in the form of a GeoJSON file that comprises 482 green spaces in Melbourne. These green spaces are represented by polygons, which encompasses the entire and exact area of each green space.

Mapping Tweets to Green Spaces. We also identified: (i) if a tweet was posted in a green space, which park was it posted from; and (ii) if this tweet was not posted in a green space, how far was it from the nearest green space. Using our collected tweets and green space dataset, we then labelled each tweet with the ID of the green space that these tweets were posted from. For tweets that were not posted from a green space (i.e., posted from an urban area), the distance to the nearest green space and the ID of this green space was identified.

3.2 Data Preprocessing

Prior to performing our sentiment analysis on the tweets, we perform a number of pre-processing steps on the collected tweets. We restrict our work to using tweets that are explicitly geo-tagged as such tweets allow us to determine where they are posted from. These steps include the following:

- Filtering tweets that are explicitly geo-tagged with latitude and longitude coordinates and within the 5km × 5km grid in Melbourne, Australia.
- Selecting tweets that are written in English, based on the "language" field provided by the Twitter API. We chose to only consider English tweets as English is the main language spoken in the Australia and more importantly, focusing on one language allows us to abstract away the nuances associated with sentiment analysis based on different languages.[1]
- Tokenizing each tweet into individual words based on separation by white-spaces.
- Converting all tweets and tokenized words into lower-case.

3.3 Sentiment Analysis

We utilize a commonly used sentiment analysis technique [8, 20], which involves first splitting each tweet into a series of tokens/words, then comparing each token/word to determine the sentiment category in which they belong to. Similar to these earlier work, we calculate sentiment score $Senti_t^S$ of a tweet t based on the word usage frequency of each sentiment category S. To account for different tweet lengths, we normalize each sentiment score $Senti_t^S$ by the number of words in each tweet. Based on this definition, the calculated sentiment level will take on a value in the range of [0,1], with 0 and 1 representing the weakest and strongest levels of the sentiment, respectively.

For these sentiment categories, we utilize the NRC Word-Emotion Association Lexicon (EmoLex) [26, 27], which is a widely used emotion word lexicon that has been used in many other works [3, 29, 33]. The EmoLex lexicon comprises 10,170 words that are associated with the emotions of anger, anticipation, disgust, fear, joy, sadness, surprise and trust, introduced in Plutchik's theory of emotions [31].

[1]Although we focus on English tweets in this work, this work can also be easily extended to any text-based social media written in other languages by using a sentiment dictionary of that language. In this work, we focus on the text/words used in tweets and future work can also consider the embedded links, photos and videos using image recognition techniques.

As pointed out in [27], the emotions of anger, disgust, fear and sadness are generally associated with negative sentiments, while the emotions of anticipation, joy and trust are generally associated with positive sentiments. The emotion of surprise is neutral, i.e., can belong to either category, and hence is used independently but not for the calculation of positive or negative sentiments. Thus, we define another two sentiment categories of positive (comprising the emotions of anger, disgust, fear and sadness) and negative (comprising the emotions of anticipation, joy and trust). Similar to [20], we define the polarity of a tweet based on the difference between the positive and negative sentiment scores of a tweet.

4 RQ1: GREEN SPACE EFFECTS

In this section, we aim to address RQ1 on the effects that green spaces have on the sentiments and emotions in such green spaces, compared to tweets posted in urban areas.

4.1 Comparison of Tweet Sentiments in Green Spaces Versus Urban areas

We first examine the presence of any significant difference in mean sentiment (positive, negative, polarity) between tweets posted in green spaces and those posted in urban areas. Table 1 shows the average sentiment level of tweets posted in green space and urban area, and associated p-values. In particular, the column "difference" indicates the increase in a specific sentiment level of tweets in green space over that of urban areas, the reported p-values are based on a two-sided Student's t-test.[2]

Table 1: Comparison of tweet sentiments in green spaces and urban areas. The bold/blue numbers indicate a statistically significant difference.

Sentiment Type	Green-space	Urban Area	Difference	p-value
negative	.0300	.0318	**-5.60%**	<.0001
positive	.0815	.0764	**6.79%**	<.0001
polarity	.0515	.0446	**15.62%**	<.0001

Table 1 shows that there is a statistically significant increase ($p < .0001$) of more than 15% in the polarity of tweets posted in green spaces, compared to those in urban areas. Similarly, there is a statistically significant decrease ($p < .0001$) of more than 5% in the negativity of tweets posted in green spaces, and also an increase of more than 6% in the positivity of tweets. These results show that green spaces generally benefit from higher positivity and lower negativity, compared to urban areas, and we examine more specific emotions in the next section.

4.2 Comparison of Tweet Emotions in Green Spaces Versus Urban areas

Similar to Section 4.1, we performed a two-sided Student's t-test to compare if there is any difference in each emotion level between

[2]This "difference" is calculated by dividing the mean sentiment levels in green spaces over that of urban areas, and the reported values are based on the exact (non-rounded) sentiment levels for a higher precision, whereas the values reported in the tables are rounded to the nearest 4 decimal points for brevity.

tweets posted in green spaces and those in urban areas. The results are shown in Table 2, and the columns are similarly defined as those in Section 4.1. In contrast to Section 4.1 that examines how positive or negative the tweets are, this section examines a more detailed breakdown of the sentiments into specific emotions, which are discussed later.

Table 2: Comparison of tweet-level emotions in green spaces and urban areas. The bold/blue numbers indicate a statistically significant difference.

Sentiment Type	Green-space	Urban Area	Difference	p-value
anger	.0071	.0079	**-9.75%**	<.0001
anticipation	.0264	.0256	**2.95%**	<.0001
disgust	.0051	.0053	-3.13%	.05689
fear	.0085	.0095	**-10.27%**	<.0001
joy	.0300	.0271	**10.62%**	<.0001
sadness	.0093	.0092	1.39%	.24705
surprise	.0129	.0122	**5.60%**	<.0001
trust	.0252	.0236	**6.54%**	<.0001

Based on our analysis, we find that there is an increase of more than 10% in joy and decrease of approximately 10% in the fear and anger emotions, for tweets posted in green spaces compared to their counter-part in urban areas. There is also an increase of 6.5%, 5.6% and 2.95% for the trust, surprise and anticipation emotions, respectively. The reported difference in the emotions of joy, fear, anger, trust, surprise and anticipation are also statistically significant, with p-values of less than 0.0001. While there are differences in the emotions of disgust and sadness, these differences are not statistically significant with p-values of more than 0.05.

Section 4.1 shows that green spaces generally display higher positivity and lower negativity than urban areas, and there are higher levels of positive emotions of joy, trust and anticipation, and lower levels of negative emotions of fear and anger in green spaces than urban areas. For the negative emotions of disgust and sadness, there is insufficient evidence to indicate any differences between green spaces and urban areas. We now explore how these sentiments and emotions change over various time periods.

5 RQ2: IMPACT OF TIME

In this section, we perform a longitudinal study of sentiments and emotions across time periods of different seasons and, in finer-grained time steps, of time of day and month.

5.1 Comparison of Sentiments and Emotions across Seasons

For our analysis of sentiments and emotions across the four seasons, we label a tweet as belonging to a particular season if this tweet was posted within the three months of the season, as widely used in Melbourne: Spring (Sep-Nov), Summer (Dec-Feb), Autumn (Mar-May), Winter (June-Aug).

5.1.1 Comparison of Sentiments Across Seasons. We start our longitudinal study on tweet sentiments by first examining the level of positive, negative and polarity of tweets posted across the four seasons in green spaces and urban areas, as shown in Figure 1. When examining levels of positive (Figure 1a) and negative sentiments (Figure 1b), we note that positive sentiments are higher and negative sentiments are lower in green spaces compared to urban areas, across all seasons of spring, summer, autumn and winter. For both tweets in green spaces and urban areas, we also observe that negative sentiments are the highest in autumn and winter, a trend that resembles the seasonal affective disorder where "depressive symptoms occur during the winter months"[3] [34, 37].

Recall that a tweet can contain both positive and negative sentiments (as described in Section 3.3), hence we use the polarity of a tweet to better measure the positivity or negativity of a tweet on its own. Figure 1c shows that the polarity levels of tweets posted in green spaces are higher (more positive) compared to those in urban areas, regardless of the season when a tweet is posted. In particular, we observe that tweets posted in green spaces are the most positive in summer, followed by spring, autumn and winter, in an order corresponding to the temperatures associated with each season. The polarity of these tweets gives us an overview of the positivity and negativity of tweets, and we examine the emotions associated with these tweets in the following sections.

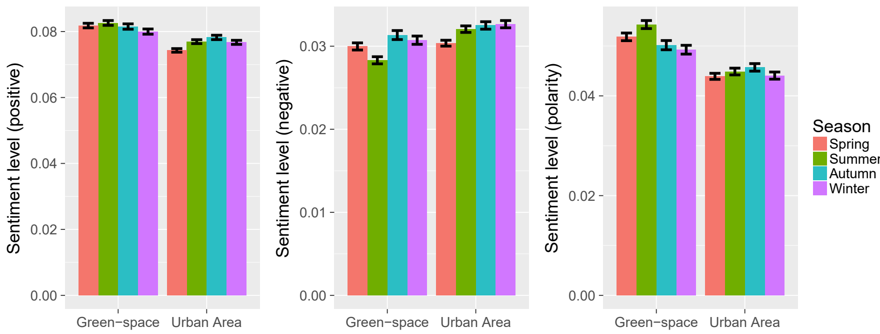
5.1.2 Comparison of Emotions Across Seasons. Figure 2 shows the average level of emotions for anticipation, joy, surprise and trust for tweets posted across the four seasons in green spaces and urban areas. The results for these positive emotions are similar to that in Section 5.1.1, as tweets in green spaces show higher levels of anticipation, joy, surprise and trust, compared to those in urban areas in the same season. Figure 2b shows that the emotion of joy is most prevalent out of all four emotions, with the highest levels for tweets in both green spaces and urban areas. In general, the results show that tweets in green spaces evoke more positive emotions of anticipation, joy, surprise and trust.

Next, we examine the average levels of emotions for anger, disgust, fear and sadness, as shown in Figure 3. In terms of the emotions of anger (Figure 3a) and fear (Figure 3c), tweets in green spaces show lower levels of these negative emotions, compared to its counterparts in urban areas in the same season. In terms of the emotions of disgust (Figure 3b) and sadness (Figure 3d), we observed mixed results where there are no clear "winners" between tweets posted in green spaces and urban areas, i.e., green spaces exhibit lower levels of these emotions in some seasons but not others. In all cases for tweets in green spaces, we note that the lowest levels of anger, disgust, fear and sadness are found during summer, an observation similar to that of the seasonal affective disorder where depressive symptoms are less likely during summer months [34, 37].

5.2 Comparison of Sentiments Across Hours, Days and Months

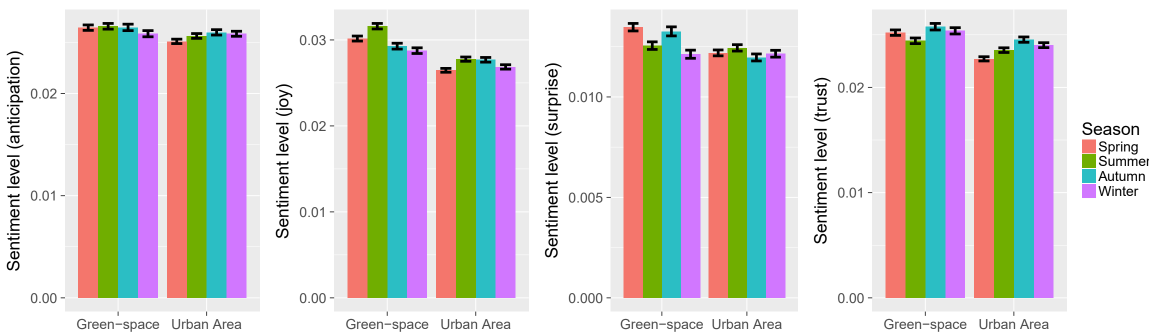
After examining how sentiments change across the seasons, we now examine how these sentiments change across finer time periods of hours, days and months, as shown in Figure 4.

[3] In the study by Rastad et al. [34], they consider that "winter was defined as the combination of autumn and winter seasons".



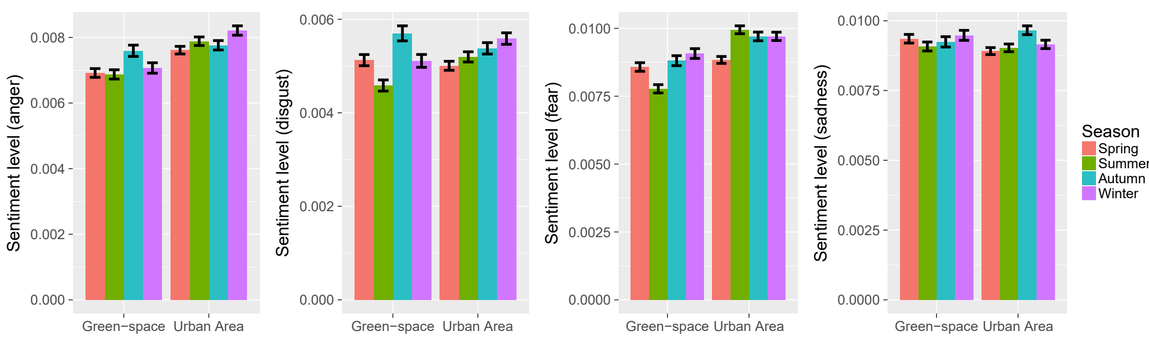
(a) (b) (c)

Figure 1: Longitudinal study of tweet sentiments across the four seasons.



(a) (b) (c) (d)

Figure 2: Longitudinal study of tweet emotions (positive only) across the four seasons.

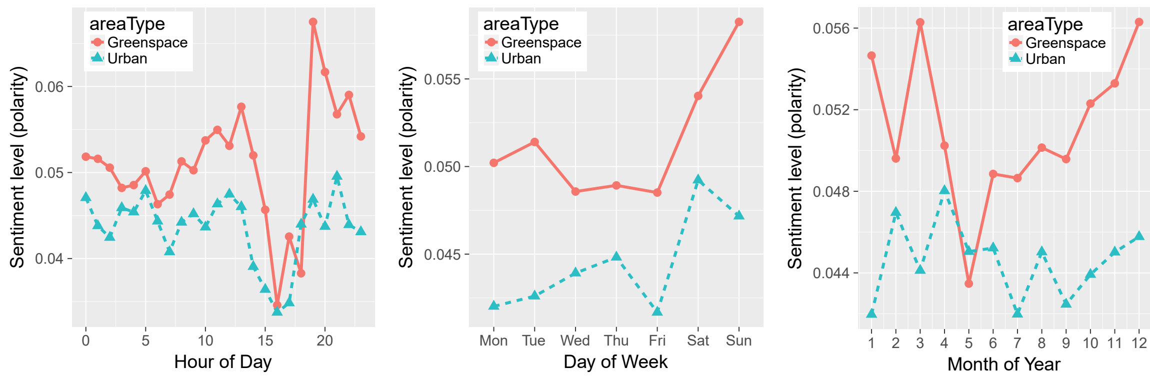


(a) (b) (c) (d)

Figure 3: Longitudinal study of tweet emotions (negative only) across the four seasons.

In terms of sentiments change across the hours of the day (Figure 4a), e.g., 12am, 1am, 2am, etc, we notice that the sentiment polarity of tweets becomes lower (less positive) from approximately 12-1pm onward until reaching a trough at 4-5pm, before increasing drastically thereafter. While this trend applies to both green spaces and urban areas, the change during this time period is more pronounced for tweets in green spaces. We attribute this due to the fact that most people are either at work (or school) from 8am to 5pm, and they become more negative towards the end of this work cycle, i.e, 12pm to 4 pm. However, the recovery period (work detachment and relaxation) takes place at the end of this work cycle [39] and sentiment of the person improves through the evening, i.e., 5 pm onwards. Similarly, social scientists have also noted that "positive emotion runs high in the morning, declines throughout the day, and rebounds in the evening" [25].

Figure 4b shows the change in sentiment based on the day of the week. Psychological studies have shown that people tend to be happier during weekends [40] and our Twitter-driven study shows the same observation, as indicated by higher levels of sentiment polarity during Sat and Sun for both green spaces and urban areas. While there is a trend of more positivity during weekends, we also



(a) (b) (c)

Figure 4: Longitudinal study of tweet sentiments based on the hour of day (left), day of week (middle) and month of year (right) that a tweet was posted. Scales do not start from zero for a clearer comparison.

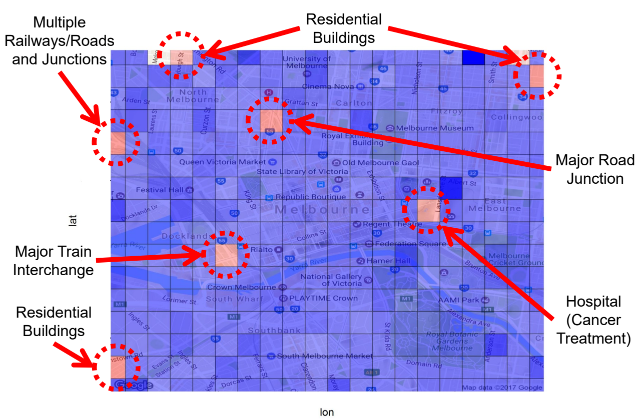


Figure 5: Grid-based Sentiment Analysis.

observe that tweets are consistently more positive in green spaces compared to urban areas, regardless of the day a tweet was posted.

The sentiment changes across the months (Figure 4c) show that sentiments in green spaces are the lowest (most negative) in May, i.e., the end of Autumn, before gradually increasing to a peak in December, i.e., the start of summer. While there are some variations in tweet sentiments in urban areas, we note that there are no obvious trends in sentiment change for urban areas. These results shows a finer grained analysis of how sentiments change across the months, while displaying the same general trends of how sentiments change across the broader seasons (as discussed in Section 5.1.2).

6 RQ3: GREEN SPACE PROXIMITY EFFECT

In this section, we investigate the effects of green space proximity by performing a high-level study of sentiments in broad city grids, and studying the correlation between sentiments in urban areas and their proximity to green spaces.

6.1 Grid-based Analysis

For a broader-scale understanding of sentiments in Melbourne, we perform a grid-based analysis of sentiment polarity within the same city, where each 250m grid comprises the aggregated sentiment polarity of all tweets within that grid. Figure 5 shows the result of this analysis where blue grids indicate positive sentiments and red grids indicate negative sentiments, while deeper colours indicate a higher level of that sentiment.

Figure 5 shows that most of the grids with negative sentiments are relating to areas that contain large transport infrastructures (train stations, road junctions, railway tracks) or residential areas. Most of the grids containing green spaces exhibit positive sentiments with the exception of one grid that contains a hospital (which has since shifted), where most tweets mention about visiting patients or going for their cancer treatments.

6.2 Proximity of Green Spaces and Urban Sentiments

To understand how the proximity of green spaces affects user sentiments in urban areas, we calculate the Pearson correlation coefficients between the sentiment levels of urban tweets and their distance to the nearest green space. Tables 3 and 4 shows the results of this correlation test in terms of sentiments (positive, negative, polarity) and emotions (anger, anticipation, disgust, fear, joy, sadness, surprise and trust), respectively.

Table 3: Pearson Correlation of Sentiments and Distance to Nearest Green Space. The bold/blue numbers indicate a statistically significant correlation.

Sentiment	Correlation	p-value
negative	**-0.0150**	<.0001
positive	0.0003	.90918
polarity	**0.0091**	.00017

The results (Table 3) show a significant negative correlation between negative sentiments and green space proximity ($p < .0001$), and a significant positive correlation between sentiment polarity and green space proximity ($p = .00017$) but none for positive sentiments. Table 4 shows that anger, anticipation, fear, sadness and trust are negatively correlated with green space proximity ($p < .0001$ for all, $p = .00655$ for sadness), while joy is positively correlated ($p < .0001$). These results show that while green spaces have an effect on urban areas, this effect is significant in terms of a reduced negative sentiment, but not significant in terms of an increase in positive sentiments.

Table 4: Pearson Correlation of Emotions and Distance to Nearest Green Space. The bold/blue numbers indicate a statistically significant correlation.

Sentiment	Correlation	p-value
anger	**-0.0111**	<.0001
anticipation	**-0.0103**	<.0001
disgust	-0.0011	.64589
fear	**-0.0244**	<.0001
joy	**0.0191**	<.0001
sadness	-0.0066	.00655
surprise	0.0031	.19721
trust	**-0.0094**	<.0001

7 DISCUSSION OF MAIN FINDINGS

In this section, we first highlight the main findings of our study, then discuss some implications of these findings in urban planning. Our main findings of how green spaces affect user sentiment are:

- RQ1: In general, tweets in green spaces are more positive and less negative than those in urban areas. When we examine these changes in terms of specific emotions, green spaces exhibit higher levels of joy, anticipation and trust (positive emotions), and lower levels of anger and fear (negative emotions), compared to urban areas.
- RQ2: While green spaces are generally more positive than urban areas, the season when a tweet was posted affects the various emotions differently. We observe that green spaces display higher polarity (more positive) than urban areas across the four seasons, with warmer seasons (spring and summer) being more positive and colder seasons being less negative.
- RQ2: Breaking down our analysis in terms of hours and days, the results show sentiment changes that reflect the general lifestyle of users. For example, sentiment polarity is the lowest at the end of a work day (early evening) before gradually increasing through the evening after work. Similarly, sentiments are more positive during weekends than weekdays, with green spaces being more positive than urban areas across all days.
- RQ3: Our grid-based analysis show that areas containing major transport-related infrastructures and residential areas are more likely to show negative sentiments, while almost all areas with green spaces exhibit positive sentiments (with the exception of an area that contained both a green space and a hospital).
- RQ3: Examining urban tweets, we find a correlation between the sentiment polarity of urban tweets and its distance to the nearest green space. The results show a significant negative correlation with tweets in urban areas and distance to green spaces for negative sentiments but no significant correlation for positive sentiments.

These findings have some important implications for urban planning authorities [15, 38, 54] and smart city applications [2]. They provide supporting evidence for policies aiming to improve well-being outcomes through urban greening interventions; people express more positive emotions and less negative emotions in green spaces or in close proximity to one. In Melbourne, this effect is particularly notable in warmer months and on weekends. At some times of the year, e.g., autumn, more negative sentiment is expressed in parks than in urban areas. Further research could explore whether these seasonal changes can be mitigated (e.g. negative sentiment could be related to mess from falling leaves, which could be mitigated through additional maintenance), or whether park use can be promoted at optimal times. In addition to urban planning, we can also improve existing tour recommendation and route planning systems [6, 17, 22, 47] by using our sentiment analysis approach to identify and recommend Points-of-Interest that elicit more positive sentiments.

8 CONCLUSION

In this paper, we studied the effects of green spaces on user sentiments based on digital traces left by Twitter users in the form of geo-tagged tweets, and presented our main findings in Section 7. As far as we are aware, our work is the first to utilize a big data driven approach to understand how green spaces are related to user sentiments across different time periods and spatial areas. In contrast to earlier works that utilizes surveys, questionnaires and case studies, our approach utilizes a large amount of Twitter data which can be easily collected and is neither intrusive nor time-consuming for the users (as the tweets are publicly available). These properties allow an unprecedented capacity for fine-grained analysis, such as capturing all green spaces at once, studying local effects, size effects, time effects, and range effects, thus also allowing to identify gaps. Moreover, our study methodology can be easily extended to examine other research questions, and thus this type of analysis is relevant for social researchers and psychologists who are currently using independent studies and traditional methods. For example, instead of administering surveys to understand how a specific crisis or natural disaster affects people's emotional well-being, we can perform sentiment analysis on a large amount of tweets that are posted in close proximity to the natural disaster or by users residing near the natural disaster. In future, we intend to extend our study to utilize image recognition techniques alongside sentiment analysis on photo-sharing sites, similar to the studies on pet ownership and alcohol consumption using Instagram [30, 51].

Acknowledgments. This work is partly supported by the Melbourne Networked Society Institute and the Clean Air and Urban Landscape hub of the National Environmental Science Program. The authors thank IBM Research - Australia and the City of Melbourne for the preliminary discussions, and the anonymous reviewers for their useful comments.

REFERENCES

[1] Harshavardhan Achrekar, Avinash Gandhe, Ross Lazarus, Ssu-Hsin Yu, and Benyuan Liu. 2011. Predicting flu trends using twitter data. In *Proc. of INFOCOM'11*.

[2] Dirk Ahlers, Patrick Driscoll, Erica Lofstrom, John Krogstie, and Annemie Wyckmans. 2016. Understanding smart cities as social machines. In *Proc. of WWW'16*.

[3] Luca Maria Aiello, Rossano Schifanella, Daniele Quercia, and Francesco Aletta. 2016. Chatty maps: constructing sound maps of urban areas from social media data. *Royal Society Open Science* 3, 3 (2016), 150690.

[4] Luluah Al-Husain, Eiman Kanjo, and Alan Chamberlain. 2013. Sense of space: mapping physiological emotion response in urban space. In *Proc. of UbiComp'13*.

[5] Nicola Barbieri, Francesco Bonchi, and Giuseppe Manco. 2014. Who to follow and why: link prediction with explanations. In *Proc. of KDD'14*.

[6] Idir Benouaret and Dominique Lenne. 2016. A Composite Recommendation System for Planning Tourist Visits. In *Proc. of WI'16*.

[7] Christoph Besel, Jorg Schlotterer, and Michael Granitzer. 2016. Inferring semantic interest profiles from Twitter followees: does Twitter know better than your friends?. In *Proc. of SAC'16*.

[8] Johan Bollen, Huina Mao, and Alberto Pepe. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Proc. of ICWSM'11*.

[9] Anna Chiesura. 2004. The role of urban parks for the sustainable city. *Landscape and Urban Planning* 68, 1 (2004), 129–138.

[10] Michael Conover, Jacob Ratkiewicz, Matthew R. Francisco, Bruno Goncalves, Filippo Menczer, and Alessandro Flammini. 2011. Political Polarization on Twitter. In *Proc. of ICWSM'11*.

[11] Anqi Cui, Min Zhang, Yiqun Liu, Shaoping Ma, and Kuo Zhang. 2012. Discover breaking events with popular hashtags in twitter. In *Proc. of CIKM'12*.

[12] Payam Dadvand, Jordi Sunyer, Mar Alvarez-Pedrerol, Albert Dalmau-Bueno, Mikel Esnaola, Mireia Gascon, Montserrat De Castro Pascual, Xavier Basagana, Ian G. Morgan, and Mark J. Nieuwenhuijsen. 2017. Green spaces and spectacles use in schoolchildren in Barcelona. *Environmental Research* 152 (2017), 256–262.

[13] Emilio Ferrara, Onur Varol, Filippo Menczer, and Alessandro Flammini. 2013. Traveling trends: social butterflies or frequent fliers?. In *Proc. of COSN'13*.

[14] Morgan R. Frank, Lewis Mitchell, Peter S. Dodds, and Christopher M. Danforth. 2013. Happiness and the patterns of life: A study of geolocated tweets. *Scientific Reports* 3, 2625 (2013).

[15] Terry Hartig and Peter H. Kahn. 2016. Living in cities, naturally. *Science* 352, 6288 (2016), 938–940.

[16] Terry Hartig, Richard Mitchell, Sjerp De Vries, and Howard Frumkin. 2014. Nature and health. *Annual Review of Public Health* 35 (2014), 207–228.

[17] Haosheng Huang, Silvia Klettner, Manuela Schmidt, Georg Gartner, Sven Leitinger, Andreas Wagner, and Renate Steinmann. 2014. AffectRoute - considering peopleâĂŹs affective responses to environments for enhancing route-planning services. *International Journal of Geographical Information Science* 28, 12 (2014), 2456–2473.

[18] Andrea L. Kavanaugh, Edward A. Fox, Steven D. Sheetz, Seungwon Yang, Lin Tzy Li, Donald J. Shoemaker, Apostol Natsev, and Lexing Xie. 2012. Social media use by government: From the routine to the critical. *Government Information Quarterly* 29, 4 (2012), 480–491.

[19] Guy Lansley and Paul A. Longley. 2016. The geography of Twitter topics in London. *Computers, Environment and Urban Systems* 58 (2016), 85–96.

[20] Huyen Le, G. R. Boynton, Yelena Mejova, Zubair Shafiq, and Padmini Srinivasan. 2017. Bumps and Bruises: Mining Presidential Campaign Announcements on Twitter. In *Proc. of HT'17*.

[21] Andrew C. K. Lee and R. Maheswaran. 2011. The health benefits of urban green spaces: a review of the evidence. *J. of Public Health* 33, 2 (2011), 212–222.

[22] Kwan Hui Lim, Jeffrey Chan, Shanika Karunasekera, and Christopher Leckie. 2017. Personalized Itinerary Recommendation with Queuing Time Awareness. In *Proc. of SIGIR'17*.

[23] Kwan Hui Lim and Amitava Datta. 2013. A topological approach for detecting twitter communities with common interests. In *Ubiquitous Social Media Analysis*. Springer Berlin Heidelberg, 23–43.

[24] Kwan Hui Lim and Amitava Datta. 2016. An Interaction-based Approach to Detecting Highly Interactive Twitter Communities using Tweeting Links. *Web Intelligence* 14, 1 (2016), 1–15.

[25] Greg Miller. 2011. Social scientists wade into the tweet stream. *Science* 333, 6051 (2011), 1814–1815.

[26] Saif M. Mohammad and Peter D. Turney. 2010. Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. In *Proc. of NAACL HLT 2010 Workshop on CAAGET*.

[27] Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a word - emotion association lexicon. *Computational Intelligence* 29, 3 (2013), 436–465.

[28] Quynh C. Nguyen, Dapeng Li, Hsien-Wen Meng, Suraj Kath, Elaine Nsoesie, Feifei Li, and Ming Wen. 2016. Building a National Neighborhood Dataset From Geotagged Twitter Data for Indicators of Happiness, Diet, and Physical Activity. *JMIR Public Health and Surveillance* 2, 2 (2016), e158.

[29] Claudia Orellana-Rodriguez, Ernesto Diaz-Aviles, and Wolfgang Nejdl. 2015. Mining affective context in short films for emotion-aware recommendation. In *Proc. of HT'15*.

[30] Ran Pang, Agustin Baretto, Henry Kautz, and Jiebo Luo. 2015. Monitoring adolescent alcohol use via multimodal analysis in social multimedia. In *Proc. of BigData'15*.

[31] Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. *Emotion: Theory, research, and experience* 1, 3 (1980), 3–33.

[32] Ana-Maria Popescu and Marco Pennacchiotti. 2010. Detecting controversial events from twitter. In *Proc. of CIKM'10*.

[33] Daniele Quercia, Luca Maria Aiello, and Rossano Schifanella. 2016. The Emotional and Chromatic Layers of Urban Smells. In *Proc. of ICWSM'16*.

[34] Cecilia Rastad, Jan Ulfberg, and Per-Olow Sjoden. 2006. High prevalence of self-reported depressive mood during the winter season among Swedish senior high school students. *J. of the American Academy of Child and Adolescent Psychiatry* 45, 2 (2006), 231–238.

[35] Jenny J. Roe, Catharine Ward Thompson, Peter A. Aspinall, Mark J. Brewer, Elizabeth I. Duff, David Miller, Richard Mitchell, and Angela Clow. 2013. Green space and stress: Evidence from cortisol measures in deprived urban communities. *Intl. J. of Environmental Research and Public Health* 10, 9 (2013), 4086–4103.

[36] Haggai Roitman, Jonathan Mamou, Sameep Mehta, Aharon Satt, and L. V. Subramaniam. 2012. Harnessing the crowds for smart city sensing. In *Proc. of CrowdSens'12*.

[37] N. E. Rosenthal, C. J. Carpenter, S. P. James, B. L. Parry, S. L. Rogers, and T. A. Wehr. 1986. Seasonal affective disorder in children and adolescents. *The American J. of psychiatry* 143, 3 (1986), 356–358.

[38] Sophie Schetke, Salman Qureshi, Sven Lautenbach, and Nadja Kabisch. 2016. What determines the use of urban green spaces in highly urbanized areas? - Examples from two fast growing Asian cities. *Urban Forestry and Urban Greening* 16 (2016), 150–159.

[39] Sabine Sonnentag, Carmen Binnewies, and Eva J. Mojza. 2008. "Did you have a nice evening?" A day-level study on recovery experiences, sleep, and affect. *J. of Applied Psychology* 93, 3 (2008), 674–684.

[40] Arthur A. Stone, Susan M. Hedges, John M. Neale, and Maurice S. Satin. 1985. Prospective and cross-sectional mood reports offer no evidence of a "blue Monday" phenomenon. *J. of Personality and Social Psychology* 49, 1 (1985), 129–134.

[41] Takehito Takano, Keiko Nakamura, and Masafumi Watanabe. 2002. Urban residential environments and senior citizensâĂŹ longevity in megacity areas: the importance of walkable green spaces. *J. of Epidemiology and Community Health* 56, 12 (2002), 913–918.

[42] Marie Truelove, Maria Vasardani, and Stephan Winter. 2014. Towards credibility of micro-blogs: characterising witness accounts. *GeoJournal* 80, 3 (2014), 339–359.

[43] Liisa Tyrvainen, Ann Ojala, Kalevi Korpela, Timo Lanki, Yuko Tsunetsugu, and Takahide Kagawa. 2014. The influence of urban green environments on stress relief measures: A field experiment. *J. of Environmental Psychology* 38 (2014), 1–9.

[44] United Nations. 2014. World Urbanization Prospects: The 2014 Revision. Internet. (2014). https://esa.un.org/unpd/wup/publications/files/wup2014-highlights.pdf.

[45] United Nations. 2017. UN Sustainable Development Goals. Internet. (2017). http://www.un.org/sustainabledevelopment/cities/.

[46] Claudia Wagner, Vera Liao, Peter Pirolli, Les Nelson, and Markus Strohmaier. 2012. It's not in their tweets: Modeling topical expertise of twitter users. In *Proc. of PASSAT'12*.

[47] Xiaoting Wang, Christopher Leckie, Jeffery Chan, Kwan Hui Lim, and Tharshan Vaithianathan. 2016. Improving Personalized Trip Recommendation to Avoid Crowds Using Pedestrian Sensor Data. In *Proc. of CIKM'16*.

[48] Yafei Wang, Rudolf de Groot, Frank Bakker, Heinrich Wortche, and Rik Leemans. 2017. Thermal comfort in urban green spaces: a survey on a Dutch university campus. *International J. of Biometeorology* 61, 1 (2017), 150–159.

[49] Xidao Wen, Yu-Ru Lin, Christoph Trattner, and Denis Parra. 2014. Twitter in academic conferences: Usage, networking and participation over time. In *Proc. of HT'14*.

[50] Jennifer R. Wolch, Jason Byrne, and Joshua P. Newell. 2014. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landscape and Urban Planning* 125 (2014), 234–244.

[51] Yuchen Wu, Jianbo Yuan, Quanzeng You, and Jiebo Luo. 2016. The effect of pets on happiness: A data-driven approach via large-scale social media. In *Proc. of BigData'16*.

[52] Wei Xie, Feida Zhu, Jing Jiang, Ee-Peng Lim, and Ke Wang. 2016. Topicsketch: Real-time bursty topic detection from twitter. *IEEE TKDE* 28, 8 (2016), 2216–2229.

[53] Yu Xu, Dong Zhou, and Seamus Lawless. 2016. Inferring Your Expertise from Twitter: Integrating Sentiment and Topic Relatedness. In *Proc. of WI'16*.

[54] Peter Zeile, Bernd Resch, Jan-Philipp Exner, and Gunther Sagl. 2015. Urban emotions: benefits and risks in using human sensory assessment for the extraction of contextual emotion information in urban planning. In *Planning support systems and smart cities*. Springer International Publishing, 209–225.