telecom

Article

# Green Areas: The Secret of the Happiness of the People of Madrid

**Diegoe de Vega**, **Óscar Araque** and **Carlos Á. Iglesias** *

Intelligent Systems Group, ETSI Telecomunicación, Universidad Politécnica de Madrid, Avda. Complutense 30, 28040 Madrid, Spain
* Correspondence: carlosangel.iglesias@upm.es; Tel.: +34-910671900

**Abstract:** Green areas play an important role in people's well-being in urban areas. However, traditional survey methods hinder understanding their actual impact. Fortunately, social networking analysis provides valuable information that city planners can use to transform cities and improve city life. This research studies geolocated tweets published in parks, both urban and natural, in Madrid, for their subsequent analysis and classification with machine-learning techniques, and determines the emotional impact of green areas on citizens. The main conclusions of this study are that people express more positive sentiments and emotions (i.e., joy and trust) in urban parks in Madrid compared to the sentiments expressed in other areas of the city and a national park in the Madrid region. This positive sentiment is higher in the city's southern districts and the historical parks. People also tweeted photos more frequently in parks and differences in the topics expressed in the tweets. This analysis can provide additional information to policymakers in urban planning.

**Keywords:** Twitter; machine learning; sentiment analysis; emotion analysis; parks; well-being; natural language processing

check for updates

**Citation:** de Vega, D.; Araque, Ó.; Iglesias, C.Á. Green Areas: The Secret of the Happiness of the People of Madrid. *Telecom* **2022**, *3*, 514–525. https://doi.org/10.3390/telecom3030028

Academic Editor: Thomas Newe

Received: 29 July 2022
Accepted: 17 August 2022
Published: 24 August 2022



## 1. Introduction

More than 55% of the world's population is currently concentrated in cities, and this is expected to reach two-thirds by 2050 [1]. As a result, sustainable development has become an important issue, and governments and organizations, such as the United Nations (UN), have begun to debate its consequences and actions to reduce its impact. One of the most well-known initiatives is the definition of the UN of 17 Sustainable Development Goals (SDGs) whose aim is to improve human society, ecological sustainability, and quality of life. In particular, SDG 11 deals with sustainable cities and communities. One of the aspects that contributes to this objective is the presence of green spaces [2]. In addition, green spaces contribute to other sustainable goals, such as SDG 3 (health and well-being), as urban forests positive influence cultural aspects such as social cohesion, stress relief, and physical activity [2–4].

Research has shown that green areas and public parks contribute to social, psychological, and physical well-being [3]. However, most research studies are based on surveys, limiting the sample size and its applicability in other venues. Since social networks provide a wealth of information on user feelings, this research paper analyzes social networks to understand how green areas affect citizens' well-being.

In particular, this study analyzes citizens' feelings and emotions in the Madrid region, Spain. In 2014, a study titled "Happiness in the City" [5] analyzed perceived happiness in several cities: Argel, Barcelona, Bombay, Chicago, Chongqing, Madrid, Paris, Rio de Janeiro, and Warsaw. To this end, the authors surveyed more than 5000 citizens, and concluded that 92% of the inhabitants of Madrid were happy living in Madrid, only behind Rio de Janeiro and Bombay. The most critical aspect that Madrid's inhabitants highlighted was the presence of green areas (97%). This result is not surprising, as Madrid has a surface area of 15.9 $m^2$ green areas [6]. Some of the urban parks in Madrid have recently been created for social purposes, while others come from former palaces owned by the Spanish crown.

Some of the most relevant parks are Casa de Campo (1, 722 ha), Juan Carlos I (1, 600 ha), El Retiro (116.84 ha), Parque del Oeste (98 ha), El Capricho (22 ha), Campo del Moro (20 ha), Real Jardín Botánico (8 ha) and Jardines de Sabatini (2.6 ha) [6].

Our study aims to better understand the influence of green areas on the happiness of Madrid's inhabitants. Thus, we present the following research questions:

RQ1. Does sentiment vary across tweets written inside and outside Madrid's green areas?
RQ2. Do green areas positively affect people's sentiment?
RQ3. Do the tweeting places influence the commented topics?

To answer these questions, our analysis is based on information collected from the Twitter social network in the Madrid region, Spain, as previously discussed. To this end, geotagged tools were used. The remainder of this article is structured as follows. Section 2 introduces the related work. Then, Section 3 describes the techniques used for data collection and analysis. Next, Section 4 presents the results of this analysis. Lastly, Section 5 discusses the conclusions of this work.

## 2. Related Work

Several studies have addressed the use of social network analysis to better understand the importance of green parks for the well-being of citizens.

Plunz et al. [7] conducted a study in New York City, United States of America. Their objectives were to use Twitter as an effective measure of people's expression of feelings and geolocated Twitter data as a well-being indicator associated with urban park space. They concluded that the sentiment of urban parks varies depending on spatial context and user characteristics. They demonstrated Twitter data as a potential resource for urban design and planning. This conclusion reaffirms our goal of finding differences between tweets in the green areas of Madrid.

Bertrand et al. [8] also studied the sentiment in New York using Twitter data. They conducted temporal and spatial analysis, concluding that sentiment is generally highest in public parks and lowest in transport hubs. Additionally, they detected areas with strong sentiments, such as cemeteries, medical centers, a jail, and a sewage facility. On the other hand, Times Square and its surroundings were an area of positive sentiment. Regarding temporal analysis, the sentiment was more positive during weekends than that during the week. During the day, the sentiment was more negative between 9 a.m. and noon, and more positive at around midnight.

Abkar et al. [9] studied the role of urban green spaces in mood change. They conducted a survey among visitors to an urban park in the city of Yazd, Iran. They evaluated several aspects, such as the use of parks, barriers to access, motivations to visit, and the perceived impact on mood change. They concluded that visiting urban green spaces affects a change in mood.

Ghahramani et al. [10] used services such as TripAdvisor or Foursquare to explore the specific characteristics of different green spaces in Dublin. Due to the opinions posted by users about the parks, their objective was to develop an application to support local authorities and stakeholders in understanding and justifying future investments in urban green spaces. This demonstrates the importance of the information that users provide and their emotions when they are in a green area.

Zhu and Ju [11] analyzed the sentiment expressed in parks in Beijing during the COVID-19 pandemic. For this purpose, they collected messages from the Sina microblogging network. Their main conclusions were that people demand access to green spaces even during the COVID-19 pandemic. In addition, they observed significant differences in emotional scores inside and outside of the parks, with them being more positive inside the parks. They also observed that older people have more positive feelings. Furthermore, the main elements of positive emotions were the landscape and plants.

Mangachena et al. [12] analyzed the opinion about national parks in South Africa to improve their management. On the basis of the analysis of posts on Twitter, they concluded that most were positive and included emotions such as participation and trust.

Regarding the discussed topics, the majority of the messages were related to management (70%), followed by biodiversity (34%), tourism (27%), natural attractions (17%), and crime and safety (10%). Another study was developed in the national parks of South Africa. Hausmann et al. [13] analyzed Instagram posts and concluded that visitors' sentiment on social networks was primarily positive, with joy, anticipation, trust, and surprise being the most frequent emotions, with only a small occurrence of negative feelings expressed in posts.

Another interesting aspect comes from the research by Schwartz [14]. They analyzed tweets posted in parks of San Francisco, USA. They concluded that the sentiment was higher in parks, and that the duration of this mood remains elevated for six hours. This insight highlights the importance of green spaces for the mental well-being for people.

## 3. Materials and Methods

The objectives of this work are the following: (i) locate by geographical coordinates the different locations in which we carry out the study; (ii) collect the tweets according to the established parameters; (iii) preprocess and clean the collected tweets; (iv) analyze the tweets; and (v) draw conclusions.

### *Data Acquisition*

The data analyzed in this paper consist of tweets posted over the past three years, specifically between January 2018 and February 2021. The following criteria were considered to collect tweets:

- They are geographically restricted to users who have tweeted in parks in the Madrid region.
- They must be in Spanish.
- Retweets are excluded since they do not necessarily express the original sentiment or intention.

The Twint tool [15] was used to collect tweets. Table 1 shows the number of tweets collected in each park.

**Table 1.** Total number of tweets per data set.

| Dataset | Num. Tweets | Dataset | Num. Tweets |
|---|---|---|---|
| Campo del Moro (CM) | 189 | Juan Carlos I (JCI) | 1921 |
| Capricho (C) | 1703 | Parque del Oeste (PO) | 3 |
| Casa de Campo (CC) | 1463 | Quinta de los Molinos (QM) | 675 |
| Cerro del Tío Pío (CTP) | 16 | Retiro (R) | 4415 |
| Dehesa de La Villa (DV) | 73 | Emp. María de Austria (EMA) | 85 |
| Total urban parks (CM+C+CC+CTP+DV+JCI+PO+QM+R+EMA) | | | 10,543 |
| National park Guadarrama (G) | 12,063 | Control (Ctrl) | 59,856 |

Twint was configured to collect tweets from the parks. For this, we used Google Maps [16] to determine the geographic coordinates of the parks. First, the parks were searched by name, and we then used the integrated tool to measure the distances. Lastly, a bounding box was defined by measuring the radius length that included each park. The most crowded parks were selected to collect a significant number of tweets. The map in Figure 1 shows which green areas were studied, and Figure 2 shows the number of retrieved tweets per zone. Each circle represents a dataset with the corresponding covered area. On the basis of this analysis, we excluded parks with a low number of tweets (i.e., Oeste park) and focused on parks with a significant number (i.e., parks Retiro and Juan Carlos I).


**Figure 1.** Map of Madrid with the studied green areas.

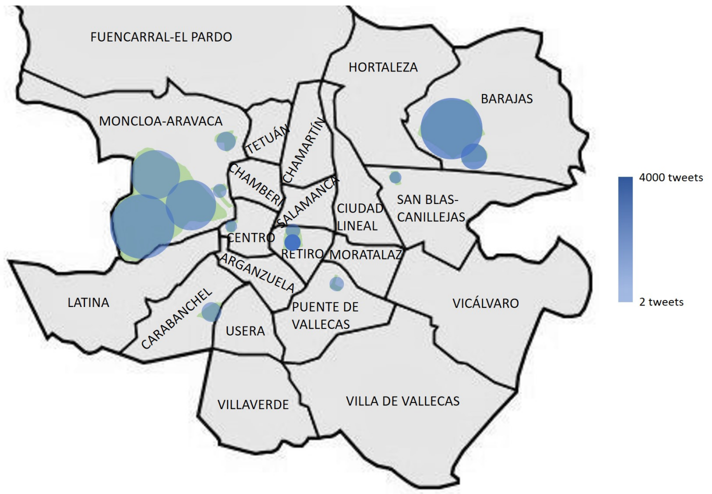


**Figure 2.** Map of Madrid with the number of collected tweets per zone.

Since we were in the Madrid region, we could also compare these results with the sentiment analysis results of Guadarrama National Park. Guadrarama is the only national park from which we collected data in this study. People who live in Madrid usually visit this natural park; therefore, they are similar population targets.

Lastly, a control dataset was also compiled to evaluate whether there were differences between tweets in parks and tweets in the city. The control dataset comprised 58,915 tweets and was recorded in the city of Madrid in random areas in the same time period as the park datasets were.

## 4. Results

### *4.1. Preprocessing*

One of the problems that we faced was that many tweets came from bots. To address this issue, we used the RapidMiner [17] filter, which detects and eliminates tweets repeated many times by the same account and if they were at the same time.

Another relevant aspect that we observed was that many posts came from automatic tweets posted when users posted photos on Instagram. These posts were discarded. After cleaning up the collected tweets, the final number of tweets per dataset is shown in Table 2.

**Table 2.** Total amount of tweets per dataset after cleaning

| Dataset | Num. Tweets | Dataset | Num. Tweets |
|---|---|---|---|
| Campo del Moro (CM) | 176 | Juan Carlos I (JCI) | 1792 |
| Capricho (C) | 1540 | Quinta de los Molinos (QM) | 573 |
| Casa de Campo (CC) | 1311 | Retiro (R) | 4050 |
| Cerro del Tío Pío (CTP) | 15 | Emp. María de Austria (EMA) | 81 |
| Dehesa de La Villa (DV) | 66 | | |
| Total Urban parks (CM + C + CC + CTP + DV + JCI + QM + R + EMA) | | | 9604 |
| National Park Guadarrama (G) | 11,244 | Control (Ctrl) | 54,112 |

### *4.2. Sentiment Analysis*

The sentiment of the tweets was analyzed using the online analysis service provided by the MeaningCloud API [18], which is available on the RapidMiner platform [17]. This sentiment analysis classifies tweets into sixdifferent categories: very positive (P+), positive (P), neutral (NEU), negative (N), very negative (N+), and none (NONE). Table 3 shows samples of tweets annotated with a sentiment label.

**Table 3.** Examples of tweets annotated with a sentiment label.

| Tweet Message | Tweet Message (Translated) | Sentiment Label |
|---|---|---|
| Estupenda mañana para pasear por el Parque del Retiro. | Wonderful morning to walk through the Retiro Park | P+ |
| En el cumple de mi nieta. | At my granddaughter's birthday. | P |
| Hoy toca... por fin. La calor impide salir tanto... | Today it's time... finally. The heat prevents you from going out so much... | NEU |
| Ya echo de menos Madrid. | I already miss Madrid. | N |
| Podría estar horas limpiando el parque, de mierda, mierda de gente irresponsable. | I could spend hours cleaning the shit of the park, shitty irresponsible people. | N+ |
| Mi bosque particular. | My private forest. | NONE |

Our first analysis aimed to understand whether there were differences in sentiment expressed in the urban parks. Figure 3 shows the distribution of tweets labeled with sentiments per urban park, considering two categories: positive (P and P+) and negative (N and N+). The urban park with the most positive tweets was El Capricho, which had the highest percentage of positive tweets (45%) and only 5% negative ones. This difference in sentiment can be explained because it is a romantic historical park with restricted access and is open only on weekends. Bertrand et al. [8] reported that sentiment is generally more positive during the weekend.

Another interesting insight is that urban parks in the city's southern districts express a more positive sentiment. In particular, we refer to parks Cerro del Tío Pío and Emperatriz María de Austria, whose positive sentiment was 34% and 35%, respectively. However, the number of negative tweets was similar for all parks.

The most negative tweets were found in the Dehesa de la Villa park. However, with 12% negative sentiment tweets, it cannot be concluded that users express more pessimistic views in this park, as it is still a relatively low percentage compared to that of positive ones, which was 26% in this case.

Lastly, the last column of Figure 3 shows that the control dataset did not stand out because of its number of positive tweets, but it did with the negative ones. Indeed, it was the dataset with the most negative tweets, at 13%.

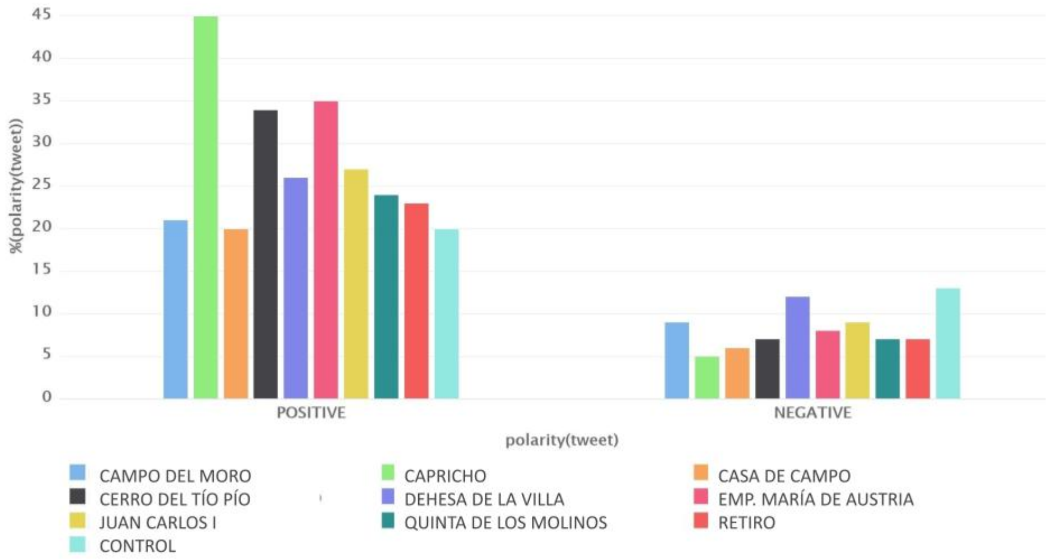


**Figure 3.** Sentiment analysis of the tweets from Madrid parks in percentages.

The second analysis that we carried out aimed to determine if there were differences in the sentiment expressed in urban and national parks. In the Madrid region, there is only one national park, Guadarrama. Figure 4 shows the sentiment distribution for urban parks, national parks, and the control dataset, and there was a significant difference. While Guadarrama National Park had the same distribution of positive and negative tweets, urban parks had three times more positive tweets.

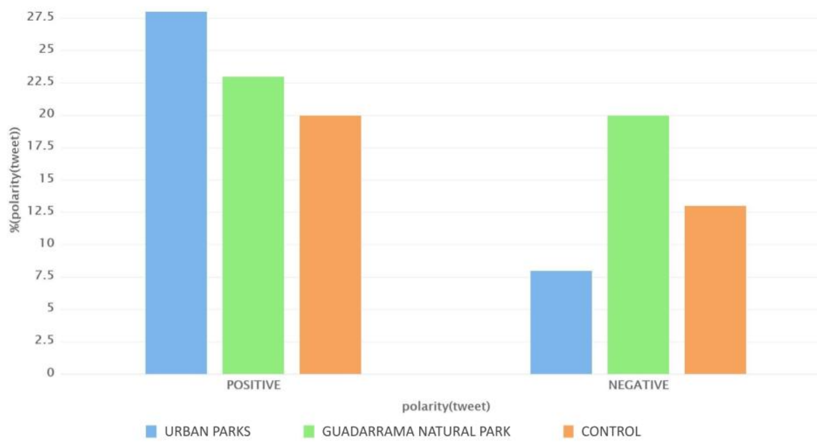


**Figure 4.** Comparison of feelings among the three main datasets.

In the case of Guadarrama National Park, the negative tweets were almost equal in percentage to the positive ones, unlike in the case of urban parks, where the positive sentiment tweets were more than triple those of negative sentiment. In this case study, we could conclude that people tweet more positively in urban parks than in natural parks.

A large number of tweets were marked as NONE. If we investigate this category more, we find that most of them were photographs or links to Instagram, as mentioned before, which also suggests that something was photographed and posted. However, this happens only in the case of parks. As shown in Table 4, the percentage of tweets with these characteristics was much higher in the case of parks. This shows that people change their behavior in places with more nature.

**Table 4.** Comparison of the number of images in the datasets.

| Dataset | $\frac{\text{Tweets with photo}}{\text{Total tweets}}$ | $\frac{\text{Tweets with photo}}{\text{NONE sentiment tweets}}$ |
|---|---|---|
| Urban parks | 44% | 97% |
| Guadarrama National Park | 31% | 83% |
| Control | 21% | 42% |

We can find a significant difference between the number of posted photos considering whether people were in a park or not. The number of photographs in urban parks was more than double that in the city. Practically all tweets without sentimentin the parks were because they were photographs, unlike the control dataset where tweets with sentiment "NONE" were much more varied. It is clear that, in the parks, more photos were posted.

In the rest of this section, we analyze whether significant differences exist between the collected datasets. To this end, we use the nonparametric Kruskal–Wallis test [19] to determine if the groups originated from the same distribution. Table 5 shows the rank assigned to the different polarities before applying the test.

**Table 5.** Integer number assigned to each type of polarity.

| Polarity | P+ | P | NEU, NONE | N | N+ |
|---|---|---|---|---|---|
| Assigned integer | 5 | 4 | 3 | 2 | 1 |

We conducted the nonparametric Kruskal–Wallis test. Table 6 shows the results of the $p$-value for the Kruskal–Wallis hypothesis test (H value), which tests the hypothesis that the medians of k groups are statistically equal. On the basis of these results, we can reject the null hypothesis, and affirm that the three datasets originated from different distributions.

**Table 6.** Results of the calculation of the Kruskal–Wallis H test.

| Dataset | H Value | $p$-Value |
|---|---|---|
| Urban parks–national parks | 246.015 | $p < 0.001$ |
| Urban parks–control | 24.1 | $p < 0.001$ |
| National parks–control | 27.95 | $p < 0.001$ |

### *4.3. Emotion Analysis*

In this section, we complement the previous sentiment analysis with an emotion analysis. For this, the tweets were labeled with emotions on the basis of Plutchik's wheel of emotions [20]. This model considers eight primary bipolar emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. This analysis was carried out, as in the previous analysis, using MeaningCloud and RapidMiner. In this case, we used the API of Deep Categorization. Table 7 shows some examples of labeling.

Our hypothesis was that positive emotions are the most common in the tweets in green areas. This would be consistent with the predominance of a positive polarity of tweets in green areas. According to Plutchik's model, joy and trust are positive emotions, while anger, disgust, fear, and sadness are negative. In contrast, anticipation and surprise emotions can be both positive and negative. Since anticipation and surprise had hardly any representation in the datasets, we excluded them from our analysis.

In emotion analysis, tweets whose sentiment was NONE or NEUTRAL did not return any results in any of the eight emotion tests carried out, which is why the number of analyzed tweets obtained in this section is much lower. A tweet could be annotated with more than one emotion. For example, a message may simultaneously show joy and surprise.

**Table 7.** Examples of tweets tagged by emotion.

| Tweet Message | Tweet Message (Translated) | Emotion Label |
|---|---|---|
| En este mundo traidor. | In this treacherous world. | ANGER |
| Acabó el finde. Amenazaba con empezar mal. | The weekend is over. It threatened to get off to a bad start. | ANTICIPATION |
| Los colores de la tauromaquia me fascinan pero me repugna 'La Fiesta Nacional' | The colors of bullfighting fascinate me but 'The National Holiday' disgusts me. | DISGUST |
| Dios que pesadilla madre mía | God what a nightmare, my mother | FEAR |
| Hoy me he levantado feliz. | Today I woke up happy. | JOY |
| Qué triste es verlo todo vacío. | How sad it is to see everything empty. | SADNESS |
| La vida es absolutamente asombrosa. | Life is absolutely amazing. | SURPRISE |
| Mi estación favorita. | My favourite season. | TRUST |

Tweets that contain photographs were discarded, which, even knowing that they were not going to have feelings or emotions, were present in the sentiment analysis for the reasons that we explained earlier.

As previously, we added the control dataset to the study to compare the results between green areas and urban environments. Figure 5 shows the results with a stacked bar chart that allowed for us to appreciate the percentages of tweets that each emotion had, and whether some areas had more of one or the other. Positive emotions were placed at the bottom of the columns, so that the comparison between positive and negative emotions could be better understood and be able to easily add those that were of the same polarity.

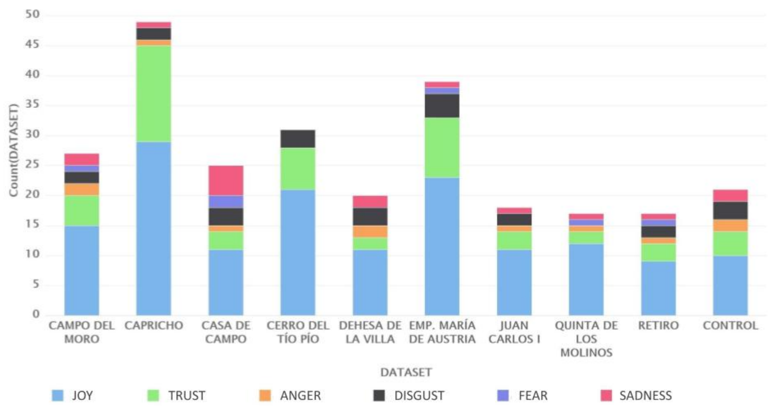


**Figure 5.** Emotion analysis of the tweets from Madrid parks in percentages.

The obtained results are consistent with those of sentiment analysis. Once again, El Capricho Park is where people tweeted the happiest. The emotions of joy and trust had their maximum there, this also being one of the parks with the least negative emotions. The 45% of the total tweets were of a positive emotion versus just 4% of negative, which represents a relevant difference.

The same differences also occurred with the parks in the south of the city. The Cerro del Tío Pío and Emperatriz María de Austria parks were the other two parks with the most positive emotions. Therefore, it is clear that people in this area of Madrid tweet more happily than others. Here, 28% and 33% of the tweets were of a positive emotion, respectively, which, compared to other parks on the graph, was a substantial difference. On the other hand, we found the most negative tweets in the Casa de Campo park. Here, positive tweets had almost the same representation as that of negative ones, with 14% and 11%, respectively.

Furthermore, the control dataset had a large number of negative emotion tweets. This confirmed the results obtained in sentiment analysis that showed that, compared to urban parks, this dataset had more negative tweets than the others did.

Once again, we compared the urban parks of Madrid with the natural park located in this community, Guadarrama National Park. As shown in Figure 6, the urban parks' positive emotions were more intense. In urban parks, there were 22% of positive emotion tweets compared to 5% of negative ones. On the other hand, the natural park had 17% positive emotions, and 12% negative emotions. As in sentiment analysis, the difference between positives and negatives was very small compared to that in urban parks. Lastly, it is once again clear that people tweet differently depending on where they do it, and that there are areas that are more positive than others.

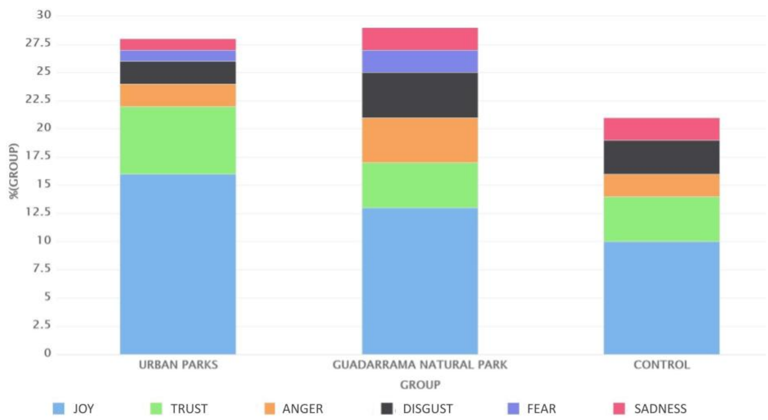


**Figure 6.** Comparison of emotions among the three main datasets.

### *4.4. Topic Analysis*

In this section, we analyze the topics expressed in the tweets in the Madrid parks. For this purpose, the analysis was carried out using MeaningCloud's Topics Extraction API, executed in RapidMiner. Figure 7 shows the results of the analysis in urban parks.

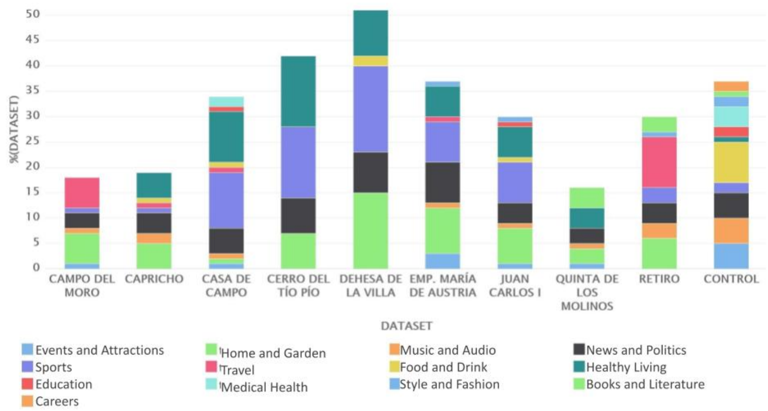


**Figure 7.** Topic analysis of the tweets from Madrid parks in percentages.

At first glance, some topics clearly stand out from others and represent a large part of the dataset. To begin with, we look at the topic of home and garden. This topic had a good percentage in most datasets; however, there was no trace of it in the control data set. Therefore, being in an urban park not only means that the user is in nature, but it also reflects it in their tweets. On average, in the urban parks in Madrid, this topic represents around 7% of the tweets, while in the control dataset, it barely had a representation.

A similar case occurred with the topic of healthy living. This topic represented only 1% of the tweets in the control dataset, while it represented almost 8% in almost all urban parks. This makes us think that people associate parks with a healthy lifestyle. However, there is an exception: there were two parks in which this topic was hardly mentioned: the Campo del Moro and Retiro parks. In these two parks, the most popular topic was "travel". This fact can be easily interpreted, as these two parks are central and touristic.

In many parks, the sports topic predominated. There was almost no representation of this topic in the parks that we previously called touristic, and in El Capricho park in which, since it is a restricted-entry park, sports are not the most common. Regarding the topic of books and literature, according to the obtained percentages, the Quinta de Los Molinos and Retiro parks are where people go to read the most or at least where they talk about reading the most.

Unsurprisingly, the only generalized topic, no matter where the user was, was news and politics. This topic was at similar percentages in all datasets, and all of them had a considerable representation. We found the same in analysis at the national level.

Lastly, we focused on the control dataset, which looks different from the rest of the datasets on the graph. Apart from having a much larger variety of topics, those that predominated were very different from those of urban parks. Topics such as medical health, education, careers, food, and drink are some of those in the control dataset that hardly had any representation in the parks. It is clear that each dataset is different, and that many factors can determine the appearance of some topics. However, similarities can be found among them, and this is why some results are obtained and not others.

Now, we compare the average of the urban parks with that of the Guadarrama National Park. Figure 8 shows this comparison with a similar graph, where we introduce the topics present in a representative number of tweets from at least one of the two groups.

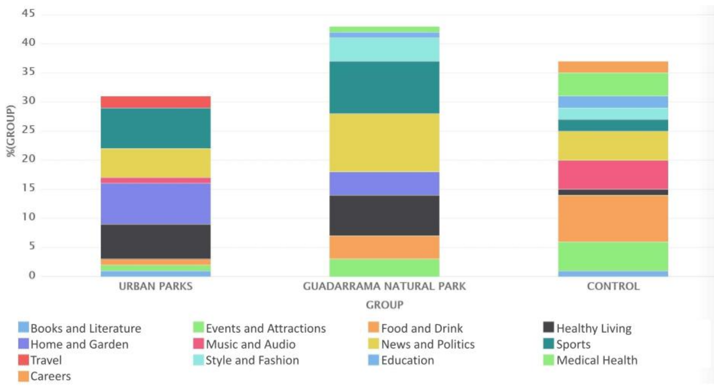


**Figure 8.** Comparison of topics among the three datasets.

There was a greater variety of topics in natural parks, with the most representative here being News and politics with a 10%, once again being the most common topic in Spain. This was closely followed by the topic of sports, with 9% here, and 7% on average in urban parks; this is the topic with the most significant presence in general. It is clear that sports practice is common in parks, and that users also share it on Twitter. Topics such as healthy living or home and garden were present in both as expected, since they are typical of natural environments. The others were less represented and varied depending on the dataset. Regarding the control dataset, the topic of home and garden related to green areas was not relevant, as expected. In contrast, we saw other topics with more relevance, such as food and drink, news and politics, events and attractions, and medical health.

## 5. Discussion and Conclusions

The first conclusion from this analysis is that people in Madrid usually tweet more frequently in a positive than in a negative way. In all studies carried out in this paper, regardless of location or analysis, there was a difference between positive and negative tweets. This result is aligned with that of other studies [21,22], though other authors [23] reported different conclusions. One potential limitation of our study is the time frame and size that we used. In future work, it would be very interesting to extend the dataset size and timeframe to understand the sentiment evolution. Emotion analysis also confirmed this conclusion. A higher number of positive emotions such as joy was found in urban parks.

Another relevant conclusion is that there was a much higher number of positive tweets in urban parks compared to the rest of the datasets. As discussed in the article, the sentiment was different when tweeting inside green areas (RQ1) and was more positive (RQ2). The relaxation that people experience in nature may be responsible for these statistics. There were thus areas of the city where tweeting was more positive than that in others. Here, the south contributed many more positive tweets than those in other areas of Madrid, the number being higher than the average in urban parks and in the control dataset. The happiest park in Madrid was El Capricho. This fact could interest city planners, since this park provides a unique concept a romantic park style based on English, French, and Italian styles.

The number of photographs increased considerably in all types of parks. Most tweets lacking polarity in these places were photograph posts. In contrast, this did not happen in the control dataset. This also gave us insight into how human behavior changes depending on the context. The relationship between emotions, and selection and publication in social networks is still an unexplored research topic [24] that requires further research to entirely understand the sentiments expressed in social media.

In addition, topic analysis provided some insights (RQ3). News and politics are frequent topics addressed in Madrid Twitter activity, and had the highest representation in all datasets. In parks, nature topics were more popular than in urban surroundings, as expected. Instead, more routine topics predominated in the city.

Social media analysis provides powerful instruments for city planners and decision makers to better understand citizen's life, needs, and sentiments. This study shows how sentiment and emotion analysis can reveal interesting facts about specific parks or city areas for urban planning, so that green areas can improve citizen wellness.

**Author Contributions:** Conceptualization, D.d.V., Ó.A. and C.Á.I.; methodology, D.d.V., Ó.A. and C.Á.I.; software, D.d.V.; validation, D.d.V.; writing—original draft preparation, D.d.V.; writing—review and editing, D.d.V., Ó.A. and C.Á.I.; supervision, C.Á.I. All authors have read and agreed to the published version of the manuscript.

**Funding:** "Gregorio Fernández" Scholarship for Research in Machine Learning and Big Data.

**Acknowledgments:** This research work was supported by the scholarship GSI "Gregorio Fernández" for Research in Machine Learning and Big Data as well as by the R&D project COGNOS (PID2019-105484RB-I00) funded by the Spanish Ministry of Science and Innovation. The authors also thank MeaningCloud for providing access to their resources for research purposes.

**Conflicts of Interest:** the authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| API | Application programming interface |
| M | Mean |
| NLP | Natural language processing |
| SD | Standard deviation |
| TWINT | Twitter Intelligence Tool |

## References

1. United Nations. 17 Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development. Available online: https://www.undp.org/sustainable-development-goal (accessed on 9 May 2022).
2. Devisscher, T.; Konijnendijk, C.; Nesbitt, L.; Lenhart, J.; Salbitano, F.; Cheng, Z.C.; Lwasa, S.; van den Bosch, M. SDG 11: Sustainable cities and communities–impacts on forests and forest-based livelihoods. In *Sustainable Development Goals: Their Impacts on Forests and People*; Cambridge University Press: Cambridge, UK, 2019; pp. 349–385.
3. Larson, L.R.; Jennings, V.; Cloutier, S.A. Public parks and wellbeing in urban areas of the United States. *PLoS ONE* **2016**, *11*, e0153211. [CrossRef] [PubMed]

4. White, M.P.; Alcock, I.; Wheeler, B.W.; Depledge, M.H. Would you be happier living in a greener urban area? A fixed-effects analysis of panel data. *Psychol. Sci.* **2013**, *24*, 920–928. [CrossRef] [PubMed]
5. Lafarge, G. Happiness in the City. 2014. Available online: https://www.esmartcity.es/2014/09/19/la-felicidad-en-la-ciudad-percepcion-ciudadana-de-sus-urbes (accessed on 9 May 2022). (In Spanish)
6. Flores-Xolocotzi, R.; González-Guillén, M.d.J. Planificación de sistemas de áreas verdes y parques públicos. *Rev. Mex. Cienc. For.* **2010**, *1*, 17–24. [CrossRef]
7. Plunz, R.; Zhou, Y.; Carrasco Vintimilla, M.; McKeown, K.; Yu, T.; Uguccioni, L.; Sutto, M.P. Twitter sentiment in New York City parks as measure of well-being. *Landsc. Urban Plan.* **2019**, *189*, 235–246. [CrossRef]
8. Bertrand, K.Z.; Bialik, M.; Virdee, K.; Gros, A.; Bar-Yam, Y. Sentiment in new york city: A high resolution spatial and temporal view. *arXiv* **2013**, arXiv:1308.5010.
9. Abkar, M.; Kamal, M.; Mariapan, M.; Maulan, S.B.; Sheyban, M. The role of urban green spaces in mood change. *Aust. J. Basic Appl. Sci.* **2010**, *4*, 5352–5361.
10. Ghahramani, M.; Galle, N.J.; Ratti, C.; Pilla, F. Tales of a city: Sentiment analysis of urban green space in Dublin. *Cities* **2021**, *119*, 103395. [CrossRef]
11. Zhu, J.; Xu, C. Sina microblog sentiment in Beijing city parks as measure of demand for urban green space during the COVID-19. *Urban For. Urban Green.* **2021**, *58*, 126913. [CrossRef]
12. Mangachena, J.R.; Pickering, C.M. Implications of social media discourse for managing national parks in South Africa. *J. Environ. Manag.* **2021**, *285*, 112159. [CrossRef] [PubMed]
13. Hausmann, A.; Toivonen, T.; Fink, C.; Heikinheimo, V.; Kulkarni, R.; Tenkanen, H.; Di Minin, E. Understanding sentiment of national park visitors from social media data. *People Nat.* **2020**, *2*, 750–760. [CrossRef]
14. Schwartz, A.J.; Dodds, P.S.; O'Neil-Dunne, J.; Danforth, C.M.; Ricketts, T.H.; Burlington, V. Exposure to Urban Parks Improves Affect and Reduces Negativity on Twitter. 2018. Available online: https://pdodds.w3.uvm.edu/research/papers/years/2018/schwartz2018a.pdf (accessed on 28 July 2022).
15. Twitter Open Source Intelligence. Twint Project. Available online: https://github.com/twintproject/twint (accessed on 9 May 2022).
16. Google Maps. Available online: https://www.google.es/maps (accessed on 9 May 2022).
17. Miner, R. RapidMiner Studio. Available online: https://rapidminer.com/products/studio/ (accessed on 9 May 2022).
18. MeaningCloud, Sentiment Analysis API. Available online: https://www.meaningcloud.com/products/sentiment-analysis (accessed on 9 May 2022).
19. Sawilowsky, S.; Fahoome, G. Kruskal-Wallis Test: Basic. In *Wiley StatsRef: Statistics Reference Online*; Wiley: Hoboken, NJ, USA, 2014.
20. Plutchik, R. *A Psychoevolutionary Theory of Emotions*; Sage Publications: Newbury Park, CA, USA, 1982.
21. Nguyen, Q.C.; Kath, S.; Meng, H.W.; Li, D.; Smith, K.R.; VanDerslice, J.A.; Wen, M.; Li, F. Leveraging geotagged Twitter data to examine neighborhood happiness, diet, and physical activity. *Appl. Geogr.* **2016**, *73*, 77–88. [CrossRef] [PubMed]
22. Hollander, J.B.; Renski, H. *Measuring Urban Attitudes Using Twitter: An Exploratory Study*; Technical Report; Lincoln Institute of Land Policy: Cambridge, MA, USA, 2015.
23. Pauken, B.; Pradyumn, M.; Tabrizi, N. Tracking happiness of different US cities from tweets. In *International Conference on Big Data*; Springer: Cham, Switzerland, 2018; pp. 140–148.
24. Robinson, P. Emediating the tourist gaze: Memory, emotion and choreography of the digital photograph. *Inf. Technol. Tour.* **2014**, *14*, 177–196. [CrossRef]