Mapping Political Extremism on Twitter in Brazil

Geisa Tamara Bugs[1(✉)], Agnes Silva de Araujo[1], Diego Saez-Trumper[2], and Rodrigo Firmino[1]

[1] Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba PR, Brazil
gestaourbana@pucpr.br
http://www.pucpr.br
[2] Universitat Pompeu Fabra, Barcelona, Spain
dsaez-trumper@acm.org

Abstract. Data from Social Media Networks may contain individual level information that is rarely available in institutional databases. It can reveal subjective and immaterial aspects that can help us to understand social behavior and, when geolocated, it may also enable understanding of spatial patterns and relationships. There is a wide range of studies that address the relationships between Twitter use and political campaigns and how it has been used in several countries to drive debates, news agenda and influence public opinion. This paper focuses on analyzing the relationship between Brazil's 2022 election results, the most polarized in recent years, and Twitter geolocated data regarding the January 8th events in Brazil, when the invasion of the National Congress and the Supreme Court in Brasília took place. The goal was to understand what the territory might reveal about this relationship. The state of Paraná was selected because it is from where many demonstrators left for Brasilia. We performed visual exploratory data analysis and spatial data analysis at municipality scale. As a result, we can cite a strong relationship between the municipalities that have more votes in each of the candidates and the number of tweets in favor or against the far-right extremist activities, and that a significant number of individuals turned against the act, especially after the media showed the damage caused to public property.

Keywords: location based social networks · Brazilian elections · January 8th

1 Introduction

The omnipresence of digital technologies, the ubiquity of geolocated urban data, and media activism of social movements, among other phenomena have emerged from the age of information and the network society along with the geospatial

Supported by Pontifícia Universidade Católica do Paraná (PUCPR).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023 Workshops, LNCS 14107, pp. 439–454, 2023.
https://doi.org/10.1007/978-3-031-37114-1_30

revolution. For a while now we have been experiencing a massive use of new information and communication technologies as digital social networks in which geographic localization plays a central role in the context of dynamic data production and appropriation [1].

On digital social networks users share all sorts of information and professionals from different fields are paying great attention to them [2]. From these data treatment and modeling, besides visualizing the users' manifestations, it is possible to make several inferences about them, such as: what identity elements they carry, what they are posting about, what keywords they use, what their positions are on certain controversies, with which other users they interact, or even how they articulate themselves in a network to consume and pass on the news or other updates [3,4].

The information transmitted in online platforms is often attached to a geolocation related component, which in most cases means the physical locale where we are sending information from [5]. Digital geographical footprints left on the so-called Location Based Social Networks - LBSN [6] bring locative information allowing other readings about the indirect relationships that are formed [3]. A broad definition of LBSN includes any social media services with geo tagged media content, such as Twitter for geo tagged Tweets, or Flickr for geo tagged photos [7]. Places become informational layers by connecting different actors, allowing computational processes to generate inferences about profiles and behaviors [3].

But according to Martí et al. [8] obtaining meaningful information from these sources represents both a challenge and an opportunity. On one hand, the amount, quality, and usability of data are growing, and automatic retrieval represents a technological advance. Moreover, the information contained in LBSN data enables the exploration of intangible aspects of urban life, offering a multiperspective approach to the study of cities. On the other hand, there is a lack of consistency in providing acceptable amounts of geolocated data, and LBSN data retrieved about specific locations may not be transferable to other studies. Also, no personal details are retrieved, and some users are not private and the data in this case is driven by public relations.

Nevertheless, online social networks such as Twitter have been the subject of scientific investigation with respect to citizen-government interactions, spatial convergence and segregation, modes of individual and collective performativity, urban analysis, social activity, and other approaches [7,9–14]. Twitter offers particularly attractive conditions as it makes its metadata bank public through the principle of anonymity. The set of variables provided by the Twitter API includes user IDs along with the timestamp and geographical coordinates of each tweet [10]. Besides being a popular service in terms of political content postings [9,10].

There are several studies on how political demonstrations can happen in line with the use of social networks, especially Twitter. These studies aim to understand how manifestations of political inclination are expressed in social networks, as well as to understand the strategies used by political and social

agents to mobilize and influence public opinion. One of the main approaches used in these studies aims to identify the emotions and attitudes expressed by users on social networks in relation to certain topics, which makes it possible to understand how public opinion is divided in relation to political issues, and to identify sources of polarization and conflict. Another common approach seeks to understand how people relate on social networks through the measurement of hubs and connections in the spread of information or opinions.

These understandings of social behavior on social networks have been heavily used by political campaigns in majoritarian elections in several countries, the most recent examples being the US presidential elections in 2016 and the plebiscite for the Brexit in the UK in the same year, as well as the general elections in Brazil in 2018 [15–19]. Most of these studies highlight the polarization and partisan orientation of users, with emphasis on the role of emotions in the expression of opinions.

There is, however, an important gap in studies on political engagement through social media, including Twitter, which concerns the possibility of understanding the spatial relationships between the use of networks and political positions, as well as the cultural and social context of users. Spatializing engagement data with geotagging may help to understand layers of network behavior, linking geographic position and positioning on policy issues.

In this context, the present article uses part of the Twitter database to analyze the posts about the January 8th events in Brazil, when the invasion of the National Congress and the Supreme Court (STF) in Brasília took place, and the repercussions compared to the result of the presidential elections. The objective is to verify if there is a relationship between the political positioning of the event, the votes for each candidate and the geolocation of Twitter users in the state of Paraná, located in the southern region of the country (Fig. 1). In terms of demographic, economic and social data, Paraná stands out as an interesting case because it is sometimes homogeneous and stands out clearly from the rest of the country, and at other times presents internal contrasts that resemble part of its territory to other regions of Brazil [20]. This state is commonly referred to as the land of one of the most politically conservative populations in the country. Paraná was also one of the states of origin of most of the demonstrators arrested after the antidemocratic acts. According to Agência Pública [21], most coaches hired to transport demonstrators to Brasilia came from the states of São Paulo and Paraná, thirty-one and twenty-three respectively, followed by fourteen from Minas Gerais, and less than four from each of the other twenty-three states. Therefore, proportionally to their populations (44 million people live in São Paulo and 11 million in Paraná) Paraná was the most prominent state in the acts of January 8th.

The acts occurred 8 days after the inauguration of President-elect Luiz Inácio Lula da Silva. The 2022 presidential election was one of the most polarized in recent decades in Brazil. The candidates Lula (PT) and Bolsonaro (PL) finished the second round with a very narrow margin of votes: 50.90% percent and 49.10%, respectively. According to Nagy and Thal's [22] study a geographical

polarization can be observed, with each of the two candidates having strengths where they were particularly strong and areas of weakness where the other dominated. Lula won in fourteen states, mostly in the Northeast, and Bolsonaro won in twelve states and the Federal District. In the state of Paraná Lula received 37.60% of the valid votes and Bolsonaro 62.40%.

It was the first time in Brazil's history that an incumbent President was not re-elected. The result generated strong complaints from supporters of the defeated candidate, for not accepting the results. Many people took to the streets in Brazilian cities in protest, voicing against Lula's victory and alleging electoral fraud in the electronic voting system [23]. Others have camped out in front of headquarters, calling for military intervention [24]. Everything leads one to believe that the attack on democratic powers was organized by alleged supporters of former President Jair Bolsonaro - with the former President himself flashing doubts about the electoral process through his social media accounts.

Although he claims to have no responsibility over the attacks and yet defended the attackers [25], for 58 percent of Brazilian voters [26] Bolsonaro had some degree of responsibility. Bolsonaro's campaign admittedly made use of digital social media to disseminate his ideas and interact with voters, being regularly associated with fake news dissemination [27,28]. The association between Bolsonaro's political performance with the production and distribution of misinformation is not a novelty in the Brazilian electoral scenario. According to [29]:102, the "strategy of spreading misinformation through social media is a powerful tool, used by Bolsonaro's supporters during and following the 2018 general election". Among the tools, WhatsApp and Telegram groups stand out, in which messages circulated with calls for anti-democratic acts in Brasilia [30].

2 Methodology

2.1 Data Collection

In this study we focused on Twitter information about the invasion of the Brazilian National Congress and Supreme Court, generated within the state of Paraná (Fig. 1). Our goal was to obtain geo located tweets, looking for a municipality level granularity.

There are two main channels to obtain Twitter's data: the Application Programming Interface - API [31] and the User Interface - UI. The former allows programmers to hit the Twitter database and receive information in machine readable format (JSON). However, the type and number of queries are limited. Moreover, its free access has been cut since February 2023. On the other hand, the UI is built for humans that access Twitter through phone or desktop frontends. One interesting feature on the UI, that does not exist on the API, is the ability to query tweets using an approximate location, that Twitter infers using a set of heuristics [32], while the API considers only tweets that have an explicit location (through the phone's GPS location).

Considering that just around 1% of the tweets contain exact coordinates [33], the approximated location becomes a powerful tool to perform location-aware

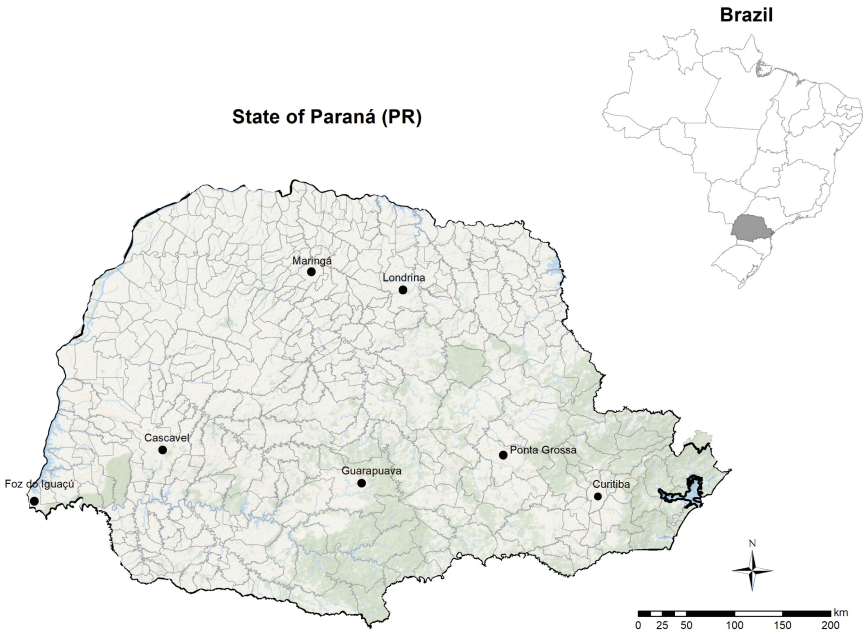


Fig. 1. Parana state municipalities

studies. The main limitation of Twitter UI is its format, coming in HTML that is easy to read for humans, but not directly readable for machines. In order to overcome these problems, there are a set of efforts to build wrappers. In this context, wrappers refer to a piece of code used to gather human-readable data and transform it into structured data. In this study, we have used the popular Python scraper package [34] that allows us to perform approximate location searches and transform them into JSON, CSV, or Excel readable files.

Our data collection process was built by searching for a set of keywords and hashtags (see Figs. 2 and 3) and limiting the search space by location. Specifically, we took the geo center coordinates of all 401 Parana state municipalities, using a radius of 20km. We also temporarily restricted our search, looking for tweets produced between 7 to 11 of January 2023. A total number of 3299 tweets compose the initial dataset.

This query returns several fields, including the Tweet content, the approximate geolocation computed by Twitter, tweet ID, and the author's username. Given that the 20km produces some overlaps between regions, we removed all the duplicated data from our dataset, and used the location assigned by Twitter. Also, to preserve users' privacy, we anonymize the username, using a hashing function [35]. This technique allows us to assign a random identification number of each user in our dataset, as exemplified in Fig. 4.

'Eleição OR Democracia OR "Forças armadas" OR "Atos antidemocráticos" OR "Três poderes" OR "Ataque" OR Armas OR Golpistas OR "Palácio do planalto" OR Invasão OR Terrorista OR Povo OR Poder OR Direita OR Esquerda OR Comunista OR Acampamento OR Bolsonarista OR Militares OR Exército OR "Bolsonaro é culpado" OR semanistiaparagolpista OR "Estado democrático de direito" OR "pátria e família" OR "3 poderes" OR "Faz o B" until:2023-01-15 since:2023-01-07 near:" -23.7358, -53.5920," within:20km'

Fig. 2. Query Example for the "Xambrê" municipality (in Portuguese)

Eleição , Democracia , "Forças armadas" , "Atos antidemocráticos" , "Três poderes" , "Ataque" , Armas , Golpistas , "Palácio do planalto" , Invasão , Terrorista , Povo , Poder , Direita , Esquerda, Comunista, Acampamento , Bolsonarista , Militares , Exército , "Bolsonaro é culpado" , semanistiaparagolpista, "Estado democrático de direito", "pátria e família", "3 poderes", "Faz o B"

Fig. 3. Keywords used for Search (in Portuguese)

Next, we manually evaluate each tweet, classifying them with three labels: "Support the act", "Reject the act" and "Not sure". During this process, some tweets were removed because they were out of context (such as the use of the word 'povo', which means people in Portuguese, to other issues). Our final dataset has 2.968 tweets.

2.2 Data Limitations

Our dataset has some limitations. First, the list of keywords was built by reviewing the trending topics for the selected day, and might be biased towards popular trends, ignoring other - less popular, but relevant - keywords. This is an intrinsic limitation of this kind of study. However, the number of tweets not considered in our search should be small in comparison with the analyzed dataset. Unfortunately, it is not possible to estimate the coverage of the analyzed dataset. This limitation should be considered when analyzing our results.

The second limitation is the potential errors in the Twitter approximate location estimation. This process is usually based on an IP address that does not have some precision from GPS coordinates. Anyhow, we opted for using this approximation to capture as many tweets as possible, focusing on recall instead of the precision for our data.

	username_anon	renderedContent	lat	lon	clean_place
91	500518638337186926	eu queria ver bolsonarista sendo tratado igual...	-25.993234	-48.995433	Guaratuba
899	-6271232442799670818	@eas_info_br @exercitooficial @NBCNews @cbstv ...	-25.644752	-49.391643	Curitiba
0	-3021135016698098404	Os caras não conhecem história e querem tomar ...	-23.375411	-50.812575	Cornélio Procópio
241	8257822057118756708	Deixou 2 salários mínimos em uma peixaria, iss...	-25.644752	-49.391643	Curitiba
40	3178254838579307705	@rafaelsatie Por quê o impeachment do luladrao...	-25.709038	-53.990943	Capanema

Fig. 4. Dataset sample

A third problem was the stacking of tweets in the Centroid of the Municipality. To overcome this limitation, a technique of random distribution of points within the municipal limits was applied. This distribution was later validated manually, comparing whether the data referring to a tweet in the city of "Curitiba", for example, was within the limits of the municipality.

Another relevant aspect refers to the classification of Tweet content. This process cannot be performed automatically, due to the use of the same words in positive (in favor of the act) and negative (against the act) contexts. To get around this limitation, the points were classified manually by the authors based on the content analysis of the Tweets.

2.3 Data Analysis

Three main methodologies are used in this paper: 1st) analysis of the statistical relationship between the number of tweets in favor and against the act of January 8th and the number of votes for Lula and Bolsonaro in the second round of the 2022 elections by municipality, obtained in the Superior Electoral Court (2023); 2nd) spatial analysis between the clusters of municipalities that most voted for each of the candidates and the density of points of tweets for and against the act and 3rd) statistical bivariate maps to display of association between the variables of the most voted candidate and number of Tweets.

Regarding statistical analysis, a simple linear regression model was applied. It determines the relationship between two variables: the dependent (y) and the independent or explanatory (x) variables. Equation 1 describes the statistical relationship between the variables in this study:

$$Y = \beta\theta + \beta_1 X + \epsilon \tag{1}$$

where: Y represents the number of Tweets in favor or against the act (independent variable); β are the intercept and slope parameters of the regression model; X represents the number of votes for Bolsonaro or Lula in the second round of the 2022 elections (explanatory variables); ϵ is the sum of the residuals in the model. It is important to mention that the variables were analyzed in pairs: number of votes for Lula and Tweets against the act and number of votes for Bolsonaro and Tweets in favor of the act.

Regarding the spatial analysis, a kernel density estimator (k) was used to transform the tweets point data into a density surface which allows the analysis of the locations where there were more tweets in favor or against the act. The density surface was then visually compared to hotspot maps of votes concentration in Bolsonaro or Lula, prepared using the Getis-Ord (Gi*) algorithm. The Getis-Ord (Gi*) algorithm allows spatial analysis of hotspots in area vector layers, which is not possible using a kernel density estimator. Nonetheless, Manepalli et al. [36] discuss that both methods can be used comparatively since they present similar spatial results.

To analyze and demonstrate the spatial relationship between two variables (votes for each candidate and number of tweets in favor or against the act) in a

single map, we adopted bivariate maps. The idea was to analyze where the two variables tend to be in agreement or disagreement. To develop a bivariate map, we counted the number of tweets per municipality which allowed both variables to be aggregated in the same vector data of municipalities, then we broke the variables into 3 classes each and we based the choice of color scheme on Tumbo's [37] recommendations.

3 Results and Discussion

Almost 43% of the tweets were in favor of the act. However, there is an increase in the Tweets with disapproval content over time (Fig. 5). It is inferred that as the media presented the damage caused to public property during the act, part of the population changed its opinion, even of individuals who can be considered extreme right. Most of the far-rightist disinformation and extreme opinions are focused on the fact that "there were leftists infiltrated in the act" and this was the reason for the destructive result (Fig. 6).

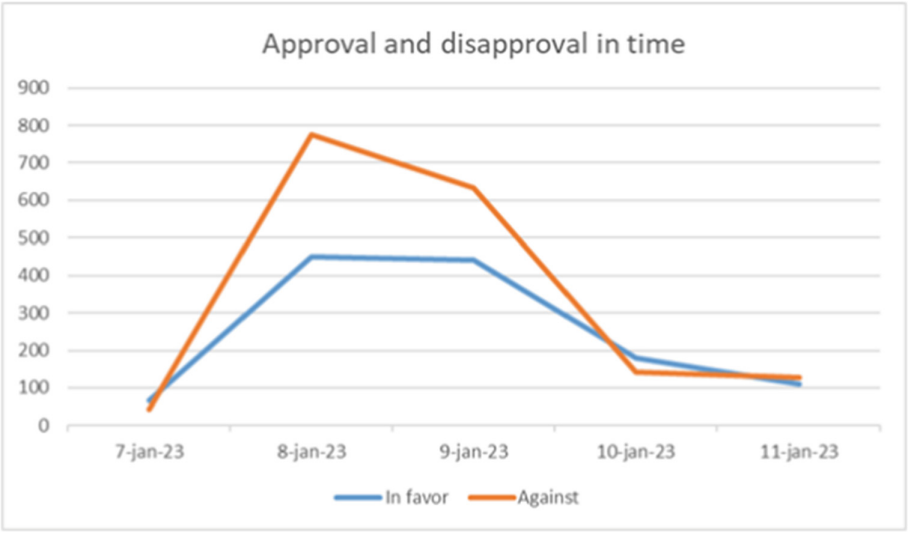


Fig. 5. Analysis of Tweet content over time

Content analysis shows that extreme rightist opinions are centered on the opinion that "there were leftists infiltrated in the act". The second most used word in tweets in favor of the act is "esquerda" (the original in Portuguese, means "left") and the word "infiltrados" (the original in Portuguese, means "infiltrators") is the one that appears more frequently (right corner of the word cloud). The second most used word in tweets against the act is "golpista" (the original in Portuguese, means "scammer"). What's more interesting, however,

is the fact that for both categories of tweets (in favor and against) the word "terrorist" appears more frequently (Fig. 6). It reflects a position change of the left and the extreme right in the legal debate around the term. Since 2016 (on the occasion of the Olympic Games) there has been a specific law in Brazil that classifies certain crimes as terrorism and the STF Minister classified the 8th of January acts as terrorism against democracy and the Brazilian institutions. While the left has previously mobilized to exclude political motivation from the characterization of terrorism, the extreme right is now using the law to prevent the attacks perpetrators from being punished as such.



Fig. 6. Content analysis of Tweets (against on left side and in favor on right side)

The results of the linear regression are presented in Table 1. The statistical analysis showed R^2 greater than 0.85 for both pairs of variables with a positive and statistically significant relationship (p-value = 0.0000). In this case, it demonstrates a relationship between the political opinion displayed in Social Media Networks and election results. Two recent examples of this type of social media influence are the Facebook/Cambridge Analytica scandal in 2018 and the invasion of the US Capitol building in 2021.

Table 1. Linear regression results

Regression outputs	**Votes for Bolsonaro and teet in favor**	**Votes for Lula and tweets against**
R^2	0.867443	0.870003
R^2 adjusted	0.867111	0.869678
F-statistics	2611.03	2670.31
AIC	2966.28	3282.45

The Facebook/Cambridge Analytica case involves the power of large corporations to change political opinions and behavior. Cambridge Analytica used data from Facebook users' clicks and likes to produce political messages tailored to the interests and beliefs detected by users' behavior. This allowed them to manipulate people's political opinions, often without their prior knowledge or consent. The case was particularly worrying because it demonstrated the ease with

which large organizations can manipulate public opinion and influence election results [38].

Similarly, in the case of the invasion of the Capitol building, social media platforms such as Facebook, Twitter, and YouTube played a critical role in organizing and disseminating the message of the people and groups involved [39,40]. Social networks were used to plan and publicize the event, using false information that fueled the anger and frustration of the participants. This type of use of networks also allowed everything to be documented and shared in real time, further amplifying the impact of the event, and contributing to inciting violence and encouraging others to participate in the insurrection. Ironically, both in the US and in Brazil, the publicized material also allowed prosecutors and the courts to form proof against the demonstrators, facilitating their indictment.

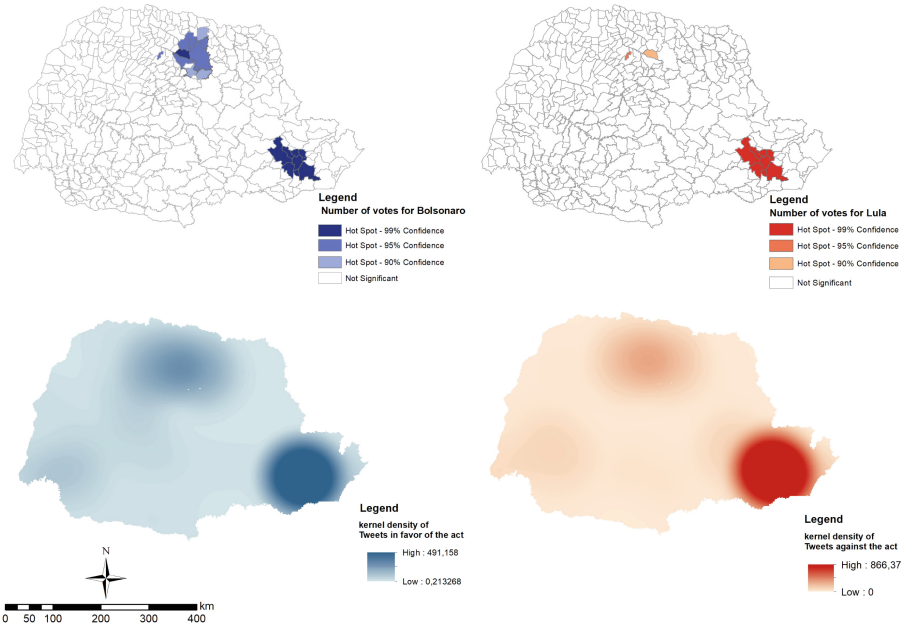


Fig. 7. Spatial patterns between municipalities with the highest number of votes for Bolsonaro and Lula and number of Tweets in favor and against the act

In addition to quantitative data, we sought to understand the spatial relationship between variables. According to Théry et al. [20] in Paraná there are two main population concentrations, one coastal around Curitiba (the state capital, and the only city with more than one million inhabitants in the state), and another in the north around Maringá and Londrina, although pole-cities and medium-sized cities, such as Guarapuava and Cascavel, stand out in the interior of the state, and in the far west Foz do Iguaçu (Fig. 1). As expected, the number of tweets collected per city follows the demography distribution. The top 5 cities

with the most tweets collected are Curitiba (1288), Maringá (214), Londrina (158), Cascavel (82) and Foz do Iguaçu (77). In other words, one can observe clusters of tweets in municipalities with a larger number of inhabitants. Both the Metropolitan Region of Curitiba and the axis in the northeast of Paraná (Londrina and Maringá) had the highest number of Tweets in favor and against the act (Fig. 7).

The capital and its surroundings have a high industrial concentration, the center of the state and part of the north are dominated by agriculture, while the secondary cities depend mainly on services. The state has more than a third of the land with priority use for agriculture. Within this activity, in the southern part family labor predominates and in the northern part hired labor, as it is the main agribusiness territory [20]. The highest Human Development Index - HDI of the state is concentrated in the regions of greater industrial activity and agribusiness-oriented agricultural activities in the western and northwestern portions, also marked by social inequality, whereas the central region of Paraná is the least developed [20]. The HDI of the smaller municipalities tends to be lower than that of the municipalities in the Metropolitan Region of Curitiba and the medium-sized cities in northern Paraná (Londrina and Maringá) [40].

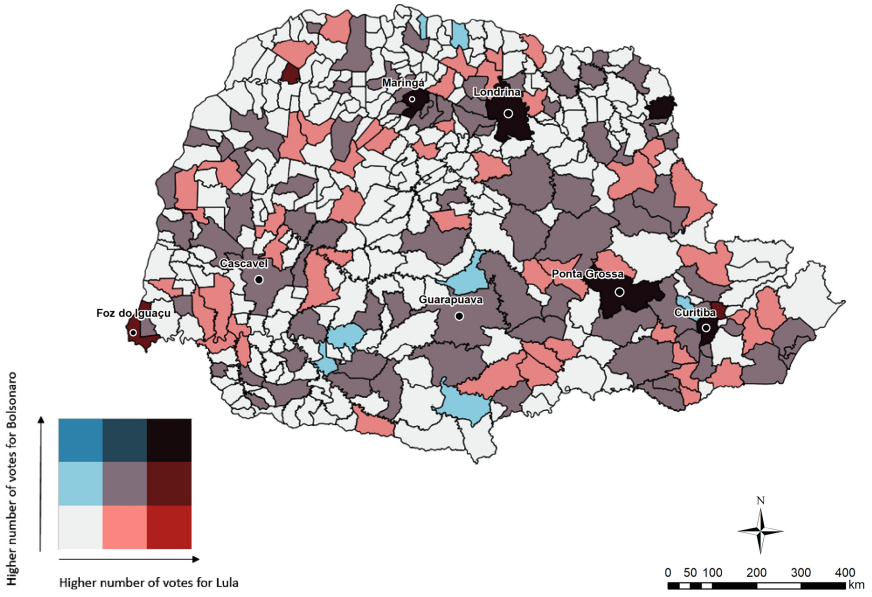


Fig. 8. Results of the second round of the 2022 elections

The voting results in the second round (Fig. 8) can be read considering socioeconomic characteristics. The map clearly demonstrates the polarization of these elections. Bolsonaro won in large cities in the interior with positive indicators,

a result of the good performance of the countryside in recent years. Lula won in a much smaller number of cities and in the smallest towns in Paraná, a state considered politically conservative, as mentioned earlier.

Figures 9 and 10 show the relationship between the number of tweets in favor and against the act and the number of votes for Bolsonaro or Lula. One aspect that must be considered when analyzing these maps is that in small municipalities, the number of tweets was very small or non-existent, for this reason, only the variables of the result of the second-round elections stand out. Despite this limitation, spatial relationships can be observed.

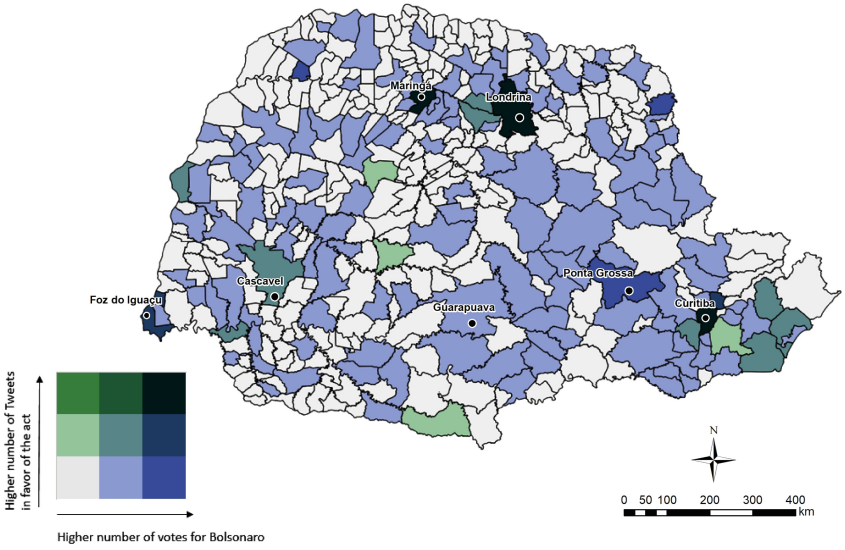


Fig. 9. Spatial relationship between votes for Bolsonaro and number of tweets in favor of the act.

One can observe that there is a greater number of Tweets in favor of the act in Londrina and Maringá, in line with the election results in those cities that had the highest number of votes for Bolsonaro. The Metropolitan Region, on the other hand, stands out for the greater concentration of tweets against the act, despite having a higher number of votes for Bolsonaro as well. Smaller municipalities did not show significant results, even though it was in smaller municipalities where the percentage of votes for candidate Lula was higher. This fact highlights the problem of representativeness of the social media data which is exacerbated when spatial patterns are discussed.

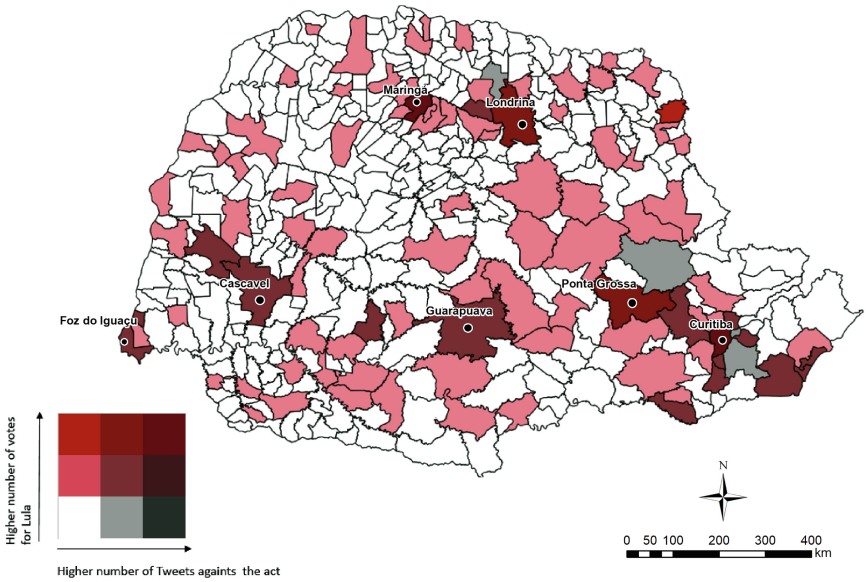


Fig. 10. Spatial relationship between votes for Lula and number of tweets against the act

4 Final Remarks

The results revealed a strong relationship ($R^2 = 0.86$) between the municipalities that have more votes in each of the candidates (Bolsonaro or Lula) and the number of tweets in favor or against the act, respectively. As expected, some municipalities that had more votes for Bolsonaro, also had a greater number of tweets in favor of the act. That is the case of the municipalities in the north of the state such as Londrina and Maringá, agribusiness territory. On the other hand, the Metropolitan Region, in which the capital Curitiba is inserted, had a greater number of tweets against the act and a greater number of votes for Bolsonaro as well. In the municipalities where the votes in favor of Lula prevailed, smaller ones, the number of tweets, both in favor of the act and against it, were very small or non-existent.

The results show that, firstly, a significant number of individuals were against the act, even rightists, especially after the media showed the damage caused by demonstrators to public property and, second, emphasizes the problem of representativeness of the social media data which is exacerbated when spatial patterns are discussed because we have big voids with no tweets related to small towns and rural areas, this does not mean that this population has no opinion on the subject, just that they do not use social media.

A contribution that should be highlighted is that this article is part of a range of research that seeks to assess the use of Location Based Social Networks

data to analyze their concurrence or divergence from established methodologies and thereby increase our understanding regarding the use of this kind of data.

References

1. Florentino, P.V., Pereira, G.C.: Structures of interactions and data in urban networks: the case of PortoAlegre.cc. In: Oliveira, J., Farias, C.M., Pacitti, E., Fortino, G. (eds.) BiDU 2018. CCIS, vol. 926, pp. 156–170. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11238-7_10
2. Rocha, M.C.F., Pereira, G.C., Loiola, E., Murgante, B.: Conversation about the city: urban commons and connected citizenship. In: Gervasi, O., et al. (eds.) ICCSA 2016. LNCS, vol. 9790, pp. 608–623. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42092-9_46
3. Sousa, P.V., Florentino, P.V.: Rastreamento e visualização de redes sociais: Investigando redes indiretas a partir de dados georreferenciais. In: 3o Simpósio Internacional LAVITS: Vigilância, Tecnopolíticas, Territórios (2015)
4. Quercia, D., Saez, D.: Mining urban deprivation from foursquare: implicit crowdsourcing of city land use. IEEE Pervasive Comput. **13**(2), 30–36 (2014)
5. Lemos, A., Firmino, R.: I connect, therefore I am!: places, locales, locations and informational territorialization. Revista Estudos do Século XX **15**, 17–34 (2015)
6. Zheng, Y.: Tutorial on Location-Based Social Networks. In: Proceedings of the 21st international conference on World Wide Web, p. 1078. ACM (2012)
7. Yang, D., Qu, B., Cudre-Mauroux, P.: Location-centric social media analytics: challenges and opportunities for smart cities. IEEE Intell. Syst. **36**(5), 3–10 (2021)
8. Martí, P., Serrano-Estrada, L., Nolasco-Cirugeda, A.: Social media data: challenges, opportunities and limitations in urban studies. Comput. Environ. Urban Syst. **74**, 161–174 (2019)
9. Flores, C.C., Rezende, D.A.: Twitter information for contributing to the strategic digital city: towards citizens as co-managers. Telemat. Inform. **35**(5), 1082–1096 (2018)
10. Rosales Sánchez, C., Craglia, M., Bregt, A.K.: New data sources for social indicators: the case study of contacting politicians by Twitter. Int. J. Digit Earth **10**(8), 829–845 (2017)
11. Netto, V.M., Pinheiro, J.M., Meirelles, J.V., Leite, H.: Digital footprints in the cityscape. In: International Conference on Social Networks on Proceedings, Publisher, Athens, USA (2015)
12. Matzner, T.: Beyond data as representation: the performativity of big data in surveillance. Surveill. Soc. **14**(2), 197–210 (2016)
13. Bernabeu-Bautista, A., Serrano-Estrada, L., Perez-Sanchez, V.R., Martí, P.: The geography of social media data in urban areas: representativeness and complementarity. ISPRS Int. J. Geo Inf. **10**(11), 747 (2021)
14. Adelfio, M., Serrano-Estrada, L., Martí-Ciriquián, P., Kain, J.H., Stenberg, J.: Social activity in Gothenburg's intermediate city: mapping third places through social media data. Appl. Spat. Anal. Policy **13**(4), 985–1017 (2020)
15. Waisbord, S., Amado, A.: Populist communication by digital means: presidential Twitter in Latin America. Inf. Commun. Soc. **20**(9), 1330–1346 (2017)
16. Wignell, P., Tan, S., O'Halloran, K.L., Chai, K.: The twittering presidents. J. Language Polit. **20**(2), 197–225 (2021)

17. Tromble, R.: The great leveler? Comparing citizen-politician Twitter engagement across three Western democracies. Eur. Polit. Sci. **17**(2), 223–239 (2018)
18. Jungherr, R.: Twitter use in election campaigns: a systematic literature review. J. Inf. Technol. Polit. **13**(1), 72–91 (2016)
19. Vergeer, M.: Twitter and political campaigning. Sociol. Compass **9**(9), 745–760 (2016)
20. Théry, H., Nagy, A., Junior, R. N.: O Paraná no Brasil: uma contextualização em treze imagens (e meia), Confins 27 (2016)
21. Agência Pública Homepage. https://apublica.org/2023/01/maioria-dos-onibus-que-transportaram-terroristas-sao-do-parana-e-sao-paulo/. Accessed 11 Apr 2023
22. Nagy, A., Thal, F.: Geografia das eleições de 2022: o Brasil cortado em quatro, Confins **57** (2022)
23. Correia, E.F.C.R.D., Tavares, F., Soares, A.M., Assis, I.B., Eleições presidenciais no Brasil. Polis **26** (2022)
24. The Guardian Homepage. https://www.theguardian.com/world/2022/dec/22/brazil-bolsonaro-militants-lefist-lula-president. Accessed 09 Apr 2023
25. O Globo Homepage. https://oglobo.globo.com/politica/noticia/2023/02/nos-eua-bolsonaro-defende-presos-no-dia-8-de-janeiro-chefes-de-familia.ghtml. Accessed 09 Apr 2023
26. Datafolha Homepage. https://www1.folha.uol.com.br/poder/2023/04/datafolha-58-veem-alguma-responsabilidade-de-bolsonaro-no-ataque-de-81.shtml. Accessed 09 Apr 2023
27. Viscardi, J.M.: Fake news, true and lie according to Jair Bolsonaro's twitter account. Trabalhos em Linguística Aplicada **592**, 1134–1157 (2020)
28. Lisbôa, J.P., Oliveira, V.: Fake News enquanto Instrumento de Política Externa: Uma Perspectiva Construtivista do Caso Bolsonaro. Trabalhos em Linguística Aplicada **592**, 1134–1157 (2020)
29. Evangelista, R., Firmino, R.: Modes of pandemic existence: territory, inequality, and technology. In: Data Justice and COVID-19: Global Perspectives, pp. 100–107, Meatspace Press (2020)
30. G1 Homepage. https://g1.globo.com/politica/noticia/2023/01/08/mensagens-bolsonaristas-terroristas-brasilia.ghtml. Accessed 09 Apr 2023
31. Twitter API Developer Platform. https://developer.twitter.com/en/docs/twitter-api. Accessed 09 Apr 2023
32. Filtering Tweets by location - API Developer Platform. https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location. Accessed 09 Apr 2023
33. Hoang, T.B.N., Mothe, J.: Location extraction from tweets. Inf. Process. Manag. **54**(2), 129–144 (2018)
34. Snscrape Github Homepage. https://github.com/JustAnotherArchivist/snscrape. Accessed 09 Apr 2023
35. Hash function Wikipedia Homepage. https://en.wikipedia.org/wiki/Hashfunction. Accessed 09 Apr 2023
36. Manepalli, U.R.R., Bham, G.H., Kandada, S.: Evaluation of hotspots identification using kernel density estimation (K) and Getis-Ord (Gi*) on I-630. In: 3rd International Conference on Road Safety and Simulation on Proceedings, pp. 14–16 (2011)
37. Tumbo, B.E.: A theory for coloring bivariate statistical maps. Am. Stat. **35**(4), 220–226 (1981)
38. Isaak, J., Hanna, M.J.: User data privacy: Facebook, Cambridge Analytica, and Privacy Protection. Comput. (Long Beach Calif.) **51**(8), 56–59 (2018)

39. Chirwa, C., Manyana, Z.: The rise of Fake news: surveying the effects of social media on informed democracy. The Thinker **88**(3), 59–66 (2021)
40. Calderón, C. et al.: Relatório do Desenvolvimento Humano 2021–2022: Tempos incertos, vidas instáveis: A construir o nosso futuro num mundo em transformação. Report UN, Brazil (2022)