

Making Sense of the City: Exploring the Use of Social Media Data for Urban Planning and Place Branding

Ville Santala^{1,3}, Sandro Miczewski², Saulo A de Brito², Ariane Lao Baldykowski²,
Tatiana Gadda³, Nadia Kozievitch², Thiago H Silva²

¹Geography and Geology, University of Turku. Turku, Finland

²Informatics, Federal University of Technology, Paraná. Curitiba, Brazil

³Civil Engineering, Federal University of Technology, Paraná. Curitiba, Brazil

***Abstract.** Conventional planning of cities is usually based on surveys and other hard evidence collected and organized by the cities' officials and research institutions. The data is usually collected in certain time intervals and it provides a limited overview to the use of urban space. Social media, including online social networks, make available vast amounts of data which nowadays is easily obtained for various types of analyses. Cities have a chance to explore social media data as a new source to study urban dynamics and complement traditional data used for urban planning. Perhaps the new data sources could also be used to overcome limitations in traditional data used in planning. In this paper, we investigated data available in Untappd, a mobile phone application for sharing beer drinking experiences, in the context of planning and place branding in Curitiba. Curitiba recently announced the creation of a Craft Beer Street, to promote local beers. By using this as a real case study we investigate a new approach that could help creating this kind of attractions, considering the aspects related to place branding as well.*

1. Introduction

Studying urban dynamics in different spatial scale in cities has traditionally been challenging. It has usually required large social studies including a vast amount of interviews, surveys, and observations, usually resulting in only a limited presentation of the reality. In this paper, we explore the potential and possibilities to use social media data for measuring the use of urban space in the city of Curitiba. Analyzing how people use the urban space could develop new knowledge about urban dwellers, an urban area under the scope and their relationship. We claim that providing new knowledge on the use of space and characteristics of Curitiba's neighborhoods could help the city in their planning and decision making practices, including place branding and marketing.

Nowadays, the amount of data sources publicly available is diverse and widespread. Cities around the world have been creating initiatives to open public data but the most used open data sources include big social media networks, such as Instagram¹, Foursquare² and Untappd³. These examples of social media are also called location-based social networks (LBSNs), where users act as a sort of sensor using mobile devices to produce vast amounts of data related to various urban and social aspects, that may be a rich source of information supporting decision making of individuals, businesses and cities [Silva et al. 2016]. Although social media data has attracted significant interest from governments all around the world, there are still challenges and

¹<http://www.instagram.com>.

²<http://www.foursquare.com>.

³<http://www.untappd.com>.

unanswered questions related to the successful exploitation of social media data in the decision making process.

Analyses of data produced with social media applications may provide accurate real life information on the usage of space, population and activities in certain place [Cranshaw et al. 2012, Silva et al. 2014a, Noulas et al. 2011, Schwartz et al. 2013, Cerrone et al. 2015]. Identifying areas which generate specific uses and activities in the city can help in planning and guide urban development. In that way, social media can provide new data for cities to use in different management practices. In this paper, we are looking at the city management from several points of view. We elaborate the possibilities social media provides for planning and decision making, but we also pay special attention to place branding and marketing potential. Detailed data about the use of urban space could help the city to manage tourism related activities and provide detailed evidence required for place marketing in the city but also at a neighborhood level.

Sensing the city with a more detailed spatial and temporal scale can open new opportunities to monitor and manage ongoing and future development. Curitiba is a globally well-known example of urban planning, but could the city use new methodologies and approaches to sustain its position among the most well planned cities also in the future? In order to give empirical evidence, for this paper we collected and analyzed data from Untappd, a mobile phone application for sharing beer drinking experiences. Our goal is to analyze suitable locations for so-called Craft Beer Street in Curitiba, a public area that the city is planning to create to promote local craft beers. Studying beer in Curitiba's context is interesting as previous study [Silva and Graeml 2016] shows that in Curitiba and Belo Horizonte people seem to have more interest towards craft beers than in other big cities in Brazil. Further evidence of the craft beer movement, would give the city of Curitiba proof that craft beer is something the city could take into account in place branding and neighborhood marketing.

The paper proceeds next with presenting some related work (Section 2). After that, Section 3 explains traditional approaches to evidence based planning and decision making. In Section 4, the empirical case study is presented and discussed. Section 5 presents characteristics, possibilities, and limitations that the social media data has in planning and city management activities. Finally, Section 6 presents conclusions and possible future work.

2. Related Research

Exploring social media data, Curitiba has been identified as an important city from the craft beer perspective in Brazil [Silva and Graeml 2016]. Most of the global academic research on using social media data in measuring urban activities and dynamics have been conducted in Europe, North America, and Asia, and, in this direction, to the best of our knowledge, no study considers Latin American cities. With this paper, we want to raise the possibilities and issues of using new data sources to study planning, decision making, and other city management activities also in Brazil and Latin America in general.

Previous studies have shown that social media, especially location-based social networks, may be used to deepen understanding of user behavior and the city dynamics [Zhang et al. 2013, Noulas et al. 2011, Preoțiu-Pietro and Cohn 2013, Wu et al. 2014, Cranshaw et al. 2012, Kling and Pozdnoukhov 2012]. The consensus seems to be that studying the city through social media data offers a way to understand both people's activity and interests better and easier than before [Schwartz et al. 2013]. Zhang et al. [Zhang et al. 2013] studied the distribution and type of venues in New York and San Francisco together with users check-in temporal activity patterns to detect clustering phenomena. Based on their study social media

data is very valuable to study which parts of the city are active during different times of the day. Social media data also brings new valuable data especially to study human mobility patterns [Wu et al. 2014, Cheng et al. 2011].

In addition, social media data can be used to discover functional areas inside the cities [Vaca et al. 2015, Wakamiya et al. 2011]. High temporal granularity and a large volume of available data offer an opportunity to study urban areas and zoning in more detailed spatial scale than a traditional representation of the land use [Kling and Pozdnoukhov 2012]. Vaca et al. [Vaca et al. 2015] found out that, in addition to discovering functional differences, Foursquare data can also reveal seasonal changes in the use of urban space.

In place branding literature the social media is most often referred as a tool to conduct marketing activities and campaigns [Sevin 2016]. Social media provides cities a platform to reach much more people with a smaller budget and in a shorter period of time compared to traditional marketing campaigns [Zhou and Wang 2014]. As far as we know, there has not been research on using social media data to collect evidence specifically for place branding management.

Many large and mid-sized cities in the United States have adopted big data practices [Ho 2017] and the study shows that cities that can institutionalize a culture of evidence-based decision making are more likely to embrace big data. More related to our work, there are some studies on possibilities to use social media as a new source of data to urban planning and decision making [Cerrone et al. 2015, Xia et al. 2014, Afzalan and Evans-Cowley 2015, Kavanaugh et al. 2012]. This indicates that there is also a growing interest in using location-based social media data in order to understand the urban dwellers' behavior and usage of urban space. Our study differs from all the previous studies because we present an evidence based on a real life example of how social media data helps in identifying the functional areas and also how it could be used by the city to better manage itself.

3. Evidence Based Planning, Decision Making, and Place Branding

The evidence based planning and decision making is not a new topic, however, the discussion and importance of evidence and data in urban planning have been growing again recently. The discussion and terminology vary greatly among different fields of study, for example, the term *data* is often used instead of evidence or even as a synonym. Data is the most simple form of information (consisting e.g. numbers and measurements). Data becomes evidence as it is put into context and used in the base of decision making. The main reason for the increased discussion is the development of information and communication technology, which allows users to collect, measure, visualize and analyze information faster than earlier. Following the same development, there is much more information available for usage in different practices: open APIs (Application Programming Interface) and open data combined with new open source tools to analyze and visualize data, give different actors the same possibility to study and understand phenomena from their perspective.

In the planning context, the nature of evidence and its role in policy-making have changed considerably during different phases of planning ideologies [Davoudi 2006]. Traditionally cities and regional authorities have paid a lot of attention on quantitative data in monitoring and decision making. This reflects the positivist understanding of knowledge, where the hard evidence is seen as the basis for a variety of policy options. Planners have widely been using data collected with different surveys and this data has been used primarily for the plans emphasizing the physical urban structure. From the traditional evidence-based point of view of policy processes, the information and knowledge are seen as providing enough sup-

port for policymakers to choose the best or most optimal course of action among a range of options. In that case, the information is perceived in political processes as value free and objective basis for making well justified decisions [Dühr and Müller 2012]. However, the seeking of evidence or prioritizing and determining research programs demands some sort of political selection [Faludi and Waterhout 2006]. These decisions define the output of new knowledge and limit the seek of information from certain points of view.

In a social–constructivist understanding, on the other hand, data and information are never seen as objective and unbiased inputs into planning processes [Dühr and Müller 2012]. Valid information and knowledge are always dependent on the circumstances within which decision making takes place. Participatory methods were implemented to planning practices in the 1960s to give more importance for citizens views on urban development. Participatory methods usually include surveys, interviews, target group meetings and participatory hearings. Nowadays the value of social media is getting more attention [Kleinhans et al. 2015].

Conventional city planning requires lots of information on the anthropic as well as the biophysical urban system. If we consider, for example, a public transportation route planning, it requires very complex analysis based on, for instance: demographics, existing and future land use, economics, commuting patterns and stakeholders interest. Usually, gathering all the data a city needs for the decision making, with traditional measurement, is time and labor intensive, and not rarely also expensive. In recent years, cities and academics have started to look at new opportunities provided by developing information technology. New technology does not only provide new devices but also lots of new data sources. Big data, provided for instance by social media, has recently gained a considerable amount of attention, and there are some strong indications that big data is changing the future planning of the cities [Batty 2013].

Social media has more users than ever before. Social media may increase the awareness of people and it could also ease the collection of public opinions and enrich the data used in the decision making process. People’s awareness has increased the demand for more real life evidence to support the decision making process. For example, several countries in Europe and the European Union itself have established research institutions in order to follow the regional and local development, and produce evidence to support decision making [Dühr and Müller 2012]. New evidence can be used to fulfill local and regional knowledge gaps and offer the best validated information to support decision making.

Nowadays, one important part of city’s decision making and management is taking care of the city’s image. Cities pay growing amount of attention to city brand, which is considered as an important asset for urban development. In fact, developing a successful brand requires a deep understanding of the city’s real characteristics. Brand as a concept is traditionally seen as a character differing products and companies from others [Knox and Bickerton 2003], but during the last decades also place branding has increased rapidly. Nations, cities and destinations have started implementing branding and marketing theories in order to manage their image and attract people, tourists, businesses and investments. City brand is seen as an important way to differ from other cities and as a great asset for the development of the city [Ashworth and Kavaratzis 2009]. Brand tells to a target audience the characteristics of the city, how the city differs from others and how the specific city is better than others. Strategically well-executed branding can improve the attractiveness and increase regional competitiveness [Moilanen and Rainisto 2008].

There is a lot of confusion with the terms place branding and place marketing. In this paper we refer to place branding as including all the management activities, planning, and decision

making, the city uses to manage itself and its image. City brand reflects the real identity of the place [Govers 2011], so it must be based on the place's actual characteristics. Place marketing is an important part of place branding. Using marketing tools (such as logos, slogans, adverts, fliers or other materials), place marketing is about the ways a city or certain place is communicating with the selected target audiences. Place marketing requires decisions not just about the target audiences, but also about the characteristics and features that the city wants to tell to the audience.

Place marketing is conducted in order to achieve short-term goals. Branding, on the other hand, is about building the place's reputation and maintaining the image in the long run. That is how place branding is also closely related to planning and need to be taken into consideration in the city decision making. Understanding how brand management and place marketing is organized in different spatial levels (e.g. neighborhoods and cities), has not been under question in academic literature related to place branding [Syssner 2010]. In smaller spatial scale the same territory can be included with very different values, interests, and purposes. This means that a brand and branding of a nation must differ from a brand and branding of a small city or region. According to Syssner [Syssner 2010], it is important that future researchers take territoriality into account when studying place branding and brand management in different spatial scales.

4. Case study: Untappd Data in Curitiba

Curitiba is recognized in Brazil as one of the most important cities regarding the production and consumption of craft beer. Recently, the city of Curitiba announced that it is planning to create a Craft Beer Street in the city. Based on that, could the city have explored social media data to help in this planning? To investigate this question, we used the data from Untappd application for this case study.

4.1. Planning Curitiba

Curitiba, the capital of the state of Paraná, is not only widely known as the best-planned city in Brazil, but also as an international model for sustainable urban development. City won Globe Sustainable City Award 2010 as being sustainable and livable city⁴. The city's achievements are the result of strategic and well-organized urban planning, which includes social, economical and environmental programs. The most well-known feature of the city is the globally known efficient public transportation system, Bus Rapid Transit (BRT).

The main guidelines for the planning of the city were established already several decades ago. In 1940's the city responded to economic development and rapid population growth by ordering a master plan [Macedo 2004]. This first comprehensive plan suggested a radial road system departing from downtown and directing different activities to secondary centers structured by functions [IPPUC 2017]. The early 1960s the municipal administration proposed a revision of the plan to manage the growth resulting from migration and industrialization [Macedo 2004]. During the development of the new master plan, the Institute for Research and Urban Planning of Curitiba (IPPUC - Instituto de Pesquisa e Planejamento Urbano de Curitiba) was created to supervise its implementation [Macedo 2004]. One of the institute's main tasks was to promote studies and research for the integrated planning of the development of the city.

Curitiba's most recent master plans are from 2004 and 2015. Mostly the plans follow the previous master plans but updating the plan regarding the recent urban development and especially the population growth in the metropolitan area. Curitiba's urban planning is based

⁴<http://globeaward.org/winner-city-2010.html>.

on principles combining sustainable economic, social and environmental development. Taking all these factors into account in planning and decision making means that the city has constant need of new evidence on the development of the urban area. The city needs updated data and knowledge related to development in order to be able to plan the future actions. Curitiba is well known for its planning history which makes it very interesting case study related to new data sources and data-driven decision making. As the technological development provides the city more data, could the city manage its planning in a more responsive and flexible way? Studies on new data sources can provide the city new knowledge that helps it to maintain its global position of a well managed city.

4.2. Untappd

Untappd is a social media that allows users to share information about beer drinking experiences. It was launched in 2010 and defines itself as “a new way to socially share a brew you’re currently enjoying, as well as where you’re enjoying it, with your friends!” [Mather 2010]. According to [Chorley et al. 2016] there were more than one million users, just three years after launching, who were responsible for over sixty million check-ins.

There are some recent efforts exploring Untappd data. For instance, [Silva and Graeml 2016] collected and analyzed messages generated by Brazilian users of Untappd. Among other things, it was very easy to notice that users from a few cities in the country have developed a more sophisticated taste for beer, especially craft beer, and are keen to try new types of beer and tell their friends about it. According to the authors, this information could be useful for an entrepreneur who wishes to settle a new small craft beer business and needs to decide on the best place to do it. [Chorley et al. 2016] also studied data from Untappd, envisioning to understand the drinking habits of people through the lenses of social media. According to [Chorley et al. 2016] one of the advantages of using Untappd data is that a formal schema is used for all check-ins (data shared by users), which makes them easy to explore for data analysis. Anyone can collect data from that social media and the data is easy to aggregate and analyze.

4.3. Methodology of Data Collection and Processing

The data used in the analysis performed in this paper was gathered from Twitter messages (called tweets) shared by users of Untappd. To collect the data we created a program in Python⁵ language exploring TwitterAPI⁶ library, which eases the process to gather data from Twitter. More details about the procedure to collect tweets can be found in [Silva et al. 2016]. Tweets include a text message following a structured pattern, created by Untappd, from which all the data we used could easily be extracted. The following messages are examples of collected tweets from English speaking users included in our dataset: “Drinking an Anniversary Ale by @lostabbey @ The Bad Apple”; “Drinking an American IPA by Win2 Brewing”.

As we can see, the type of beer comes after “Drinking a/an” and before “by <name of a company> @ <name of a place>”. This pattern makes data parsing easy using just about any programming language. In our case, we performed this step using Python. Besides the beer type, we also extracted the location (geographical coordinates) where the beer was consumed, the date of the post, user id and the message from the tweet.

We collected data from two different periods: from 06 April 2013 to 30 April 2013 (Dataset2013); and from 03 November 2016 to 05 March 2017 (Dataset2017). After data collec-

⁵<http://www.python.org>.

⁶<http://github.com/geduldig/TwitterAPI>.

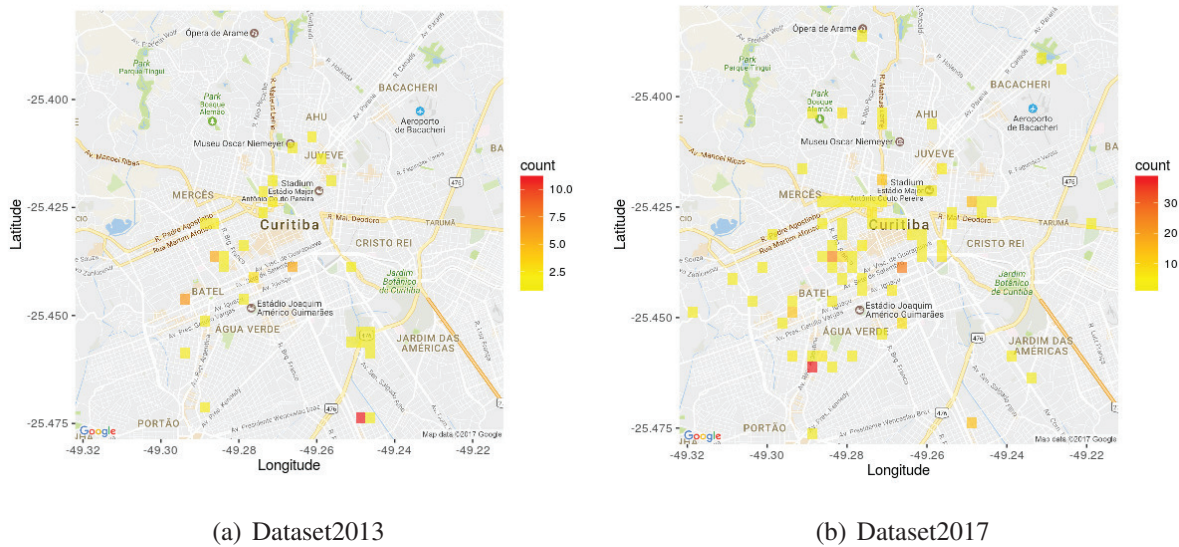


Figure 1. Heatmap of the number of Untappd check-ins in Curitiba.

tion, we performed a filtering process. We considered only one check-in per user, in a specific drink location to avoid bias in the data caused by heavy users. Besides that, we considered only messages containing geographic information about where the beer was being consumed, and also messages containing the type of beer the user was having. After that, we have 400 records for Dataset2017 and 71 records for Dataset2013.

Heatmaps, Figures 1a and 1b, present the number of Untappd check-ins in Curitiba for Dataset2013 and Dataset2017, respectively. Heatmaps provide an intuitive overview of the intensity of a phenomenon within a spatial region. Based on Figure 1 one thing that we noticed is that check-ins are spread around the city. They do not only seem to cluster in, for example, known gastronomic areas of the city. We even suspected that maybe people were drinking their beer at home, as this is a possibility for craft beer, particularly when there are many home brewers that distribute the beer they produce directly to friends and neighbors. Therefore, as done by [Silva and Graeml 2016], we also checked the names of places where those beers were being drunk to confirm that the majority of places are not homes. Another result from heatmaps is that the amount of data is more intense and spread nowadays (Figure 1b) compared to what were observed in 2013 (Figure 1a).

In order to analyze the data, we used R⁷, an open software for statistical computing and graphics available on-line, able to generate plots on maps and other mining tasks.

4.4. Results

To accomplish our proposed goal, we needed to discover popular areas in the city based on the number of check-ins observed in our datasets. In this study, we propose to use the clustering algorithm DBSCAN [Ester et al. 1996]. DBSCAN is a density-based clustering algorithm that groups closely packed points together (points with many nearby neighbors) and marks points that lie alone in low-density regions as outliers (whose nearest neighbors are too far away). It requires two parameters: *eps*, used to identify neighboring points; and the minimum number of points required to form a dense region *minPts*. Due to its characteristics, the particularities of our problem is met. In this study we used an R package available online⁸ containing the

⁷<https://www.r-project.org>.

⁸<http://github.com/mhahsler/dbscan>.

implementation of that algorithm. We consider in this work the great-circle distance, which is the shortest distance between two points on the surface of a sphere, measured along the surface of the sphere. This is calculated with the help of the *Haversine Formula*.

We set the *eps* parameter as 250 meters and considered $minPts = 10$ for all datasets. However, we also present the results considering $minPts = 5$ for the Dataset2013, since this dataset is smaller. Figure 2 show the results of the clustering. Looking first at Figure 2a, which shows the clustering results for Dataset2017, we see that the algorithm found eight clusters (all of them are labeled with one letter). Note that the colors and symbols are used only to differentiate the clusters. Among the clusters, the area where the Beer Street is going to be built is identified (see cluster *H*). As the Beer Street area was discovered by the analyses, this is the first evidence on how social media data could be used in planning.

If the decision makers of Curitiba were using the strategy we are discussing to help make better urban planning decisions, at this moment they would have those eight candidate areas to create the Beer Street. The first observation is that all these clusters are important areas regarding craft beers, and this emerged from the data considered, which implicitly express the preference of the users. As we could see in the heatmap (Figure 1b), there are other areas for craft beer consumption, however, the identified areas emerged as the most popular ones. This means that the city can concentrate efforts on more strategic areas.

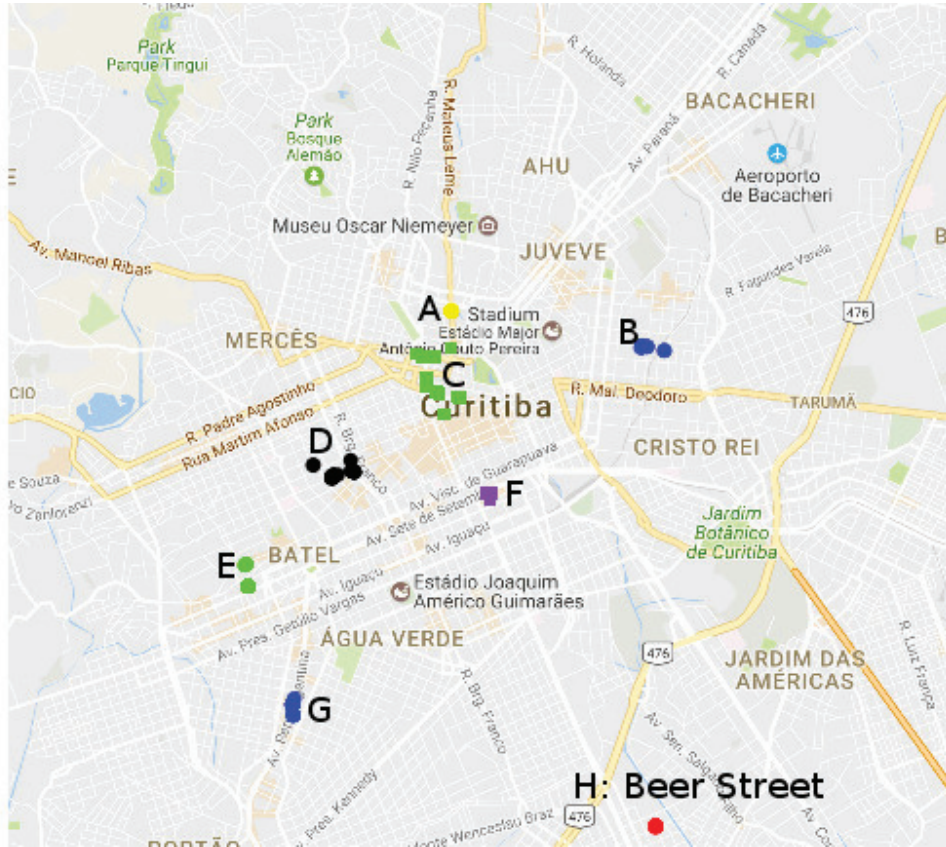
Before we proceed, it is important to mention some of the key features that a street should have to support a beer street creation: (*F1*) it has to support being blocked for several hours at night and also during the day on weekends, without causing considerable impact on traffic; (*F2*) it can not be located nearby calm zones, such as hospitals and residential areas; (*F3*) it can not affect negatively on current businesses installed on it; (*F4*) it has to be big enough to support one or more stages for concerts, space for products exposition, food and beer commercialization, and at least a public of 10k people⁹ or more, offering conditions for people to move freely.

Cluster *A* is represented by *Hop'n Roll*, a well-known pub specialized in craft beer. Cluster *B* is represented mainly by *Masmorra Cervejaria*, *Crazy4Beer* and *Mercado153*, located along Itupava Street, a gastronomic hub. Cluster *C* represents São Francisco area, a neighborhood that is famous for attracting alternative groups, such as Hipsters. Clusters *D* and *E* are located at Batel, one of the wealthiest neighborhoods in Curitiba, which hosts many restaurants, pubs, and nightclubs. Those clusters also offer several specialized pubs in craft beers, such as *The Meatpack House*, *Whatafuck* and *Clube do Malte*. Cluster *F* represents a mall that hosts, among other places that commercialize local craft beer, *Way Beer Pub*, a pub from an important craft beer brewery in Curitiba. Cluster *G* is located in the residential neighborhood Aguá Verde. That cluster is mainly represented by *Império Cervejeiro*, a pub specialized in craft beer.

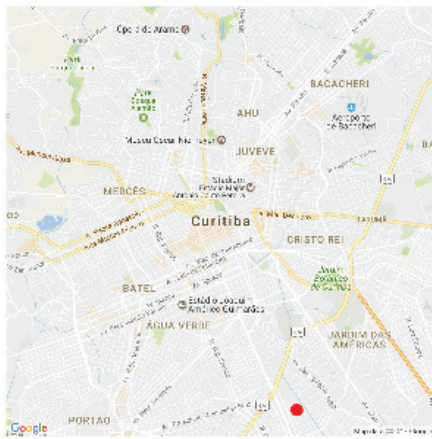
Based on feature *F1*, clusters *A*, *F* and *G* are not good candidates because they are located in busy streets, and closing them would impact considerably the traffic. Besides that, cluster *G* also do not support *F2*. Clusters *B*, *D* and *E* are located in important gastronomic areas in the city. Creating a Beer Street there would meet feature *F3*, for this reason, they are also not good candidates. Nearby Cluster *C* there are some streets with less traffic, however, they do not offer scalability, i.e. feature *F4* are not supported, the location is not suitable to accommodate larger crowds without impacting negatively to the already busy downtown area.

The best candidate according to the set of key features demanded for a Beer Street is Cluster *H*, which is mainly represented by *Bodebrown Pub*, a famous pub for craft beer in Cu-

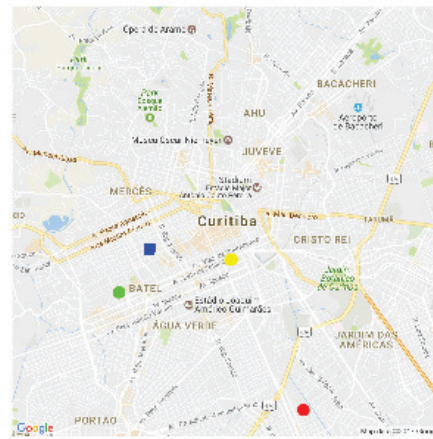
⁹Based on an average number of public of some recent craft beer festivals promoted in Curitiba: <https://goo.gl/a3OxwK>.



(a) 2016-2017 - minPts =10, eps=250 meters



(b) 2013 - minPts = 10, eps=250 meters



(c) 2013 - minPts = 5, eps=250 meters

Figure 2. Clusters found in Curitiba using Untappd data.

Curitiba. This pub appears as being the 4th most popular on our dataset. Cluster *H* is located quite far way from the city center, as well as from the other clusters we found. This area accommodates a different kind of companies, such as conveyor sheds and storage places. It is also not as developed and densely used as the center, which makes it a strategic area for development. In addition, the streets are larger and the traffic is less intense, which means that the closing of streets during specific time periods could possibly cause a comparatively less negative impact.

The same clustering process was performed for Dataset2013 (see Figure 2b). Note that,

for $eps = 250$ meters and $minPts = 10$ the only cluster found was the area where the Beer Street will be developed. As we can see, the Beer Street area has been popular for a long time, at least from 2013 and still popular today. This can be other criteria to be considered by the decision maker. Finding fewer clusters on this configuration from Dataset2013 was expected, because of the restriction of $minPts = 10$ (remember that Dataset2013 is smaller). In practice, the $minPts$ might have to be adapted for different datasets or goals, the decision maker should have this control. For this reason we also considered $minPts = 5$ for Dataset2013 (see Figure 2c). This way also other clusters, D , E and F of Figure 2a, were also found.

5. Social Media in the Context of Urban Planning and Place Branding

The rise of popularity of social media provides new data and ways to look at the city. Social media has changed the way we communicate with each other and applications are also collecting vast amounts of information on daily life activities. Academics in different parts of the world have been interested in using social media data, specially LBSNs, to measure activities and the use of urban spaces [Cheng et al. 2011, Noulas et al. 2011, Cranshaw et al. 2012, Kling and Pozdnoukhov 2012, Schwartz et al. 2013, Silva et al. 2013, Silva et al. 2014b]. The planning has traditionally been based on large datasets limited to administrative division, but new data sources give cities an opportunity to implement spatial analysis at different scales. Social media can provide more sophisticated, wider-scale and finer resolution and sometimes even real-time understanding of urban space [Kitchin 2014]. For these reasons, social media can be considered as a resource for evidence based decision making and strategic management, especially at the local level.

Recently, different social medias and mobile applications have been studied as tools for increasing public participation and engagement in urban planning [Ertiö 2015, Afzalan and Evans-Cowley 2015, Jones et al. 2015, Kleinhans et al. 2015]. The growing interest in the potential of social media and mobile technologies reflects the broad political agenda to foster citizen engagement and participation in urban planning [Kleinhans et al. 2015]. The potential of the data is often associated with its size and easiness to collect and analyze, but social media can also provide new opportunities for cities to hear the voice of user groups that do not formally participate in the planning and decision making process. Perceptions about the city or particular areas are also vital for the monitoring of the successfulness of place branding strategies [Sevin 2016].

Social media data has been used for identifying specific activities and spaces within cities. As some studies have pointed out, e.g [Wakamiya et al. 2011] and [Frias-Martinez and Frias-Martinez 2014], it is possible to use social media data, such as Twitter or Foursquare, to study and define characteristics of certain urban areas. Data shared by users in those systems can be used for analyzing and clustering selected groups and activities in the city, which eventually makes it possible to characterize urban areas and differentiate them [Silva et al. 2014a]. This kind of analysis can provide city planners detailed information on the usage of space, without being forced to analyze phenomena in the city through administrative divisions (e.g. neighborhood division).

Receiving spatially detailed information can help the city to evaluate the characteristics chosen to be used in place branding and marketing. Detailed data from social media has potential to help the city to better organize strategies for place branding in different territorial scales, which can make a substantial difference on the effectiveness of branding [Syssner 2010]. As previous study [Silva and Graeml 2016] shows, Curitiba as a city seems to be the most fruitful city for craft beer among other big cities in Brazil. Beer culture is, therefore, a feature the city

could use in city marketing. The analysis of Untappd data helps the city to determine which geographic units in the city (e.g. neighborhoods) would be suitable to develop around the beer craft activities and nightlife and rank them according to their potential. In Curitiba's southern part, where one of the most famous brewery *Bodebrown* is located, is where the Beer Street is being created. This street attracts a lot of people to spend free time and drink craft beer especially during Friday nights and Saturday mornings.

The results of the case study illustrate the place brand management potential of social media data. Using the beer drinkers' data, detailed information about the use of urban space can help the city to manage place marketing in the city but also at the neighborhood level. In other words, the data does not only help to confirm that Curitiba is an important city related to craft beer in Brazil [Silva and Graeml 2016], but also which neighborhoods deserve more attention. In the neighborhood where the Beer Street is going to be developed, brand management could help to change its current reality and bring better conditions and opportunities for its businesses and dwellers.

This case study reveals that social media data can help the city to react and adapt to changes in the use of urban space. Quick reaction time can reduce negative effects of certain phenomena and provides the city better opportunity to manage activities and the usage of urban space. The city of Curitiba has not yet taken advantage of emerging initiatives which have been transforming different areas inside the city. Therefore place marketing activities led by businesses, for instance, are yet to be officially recognized by the municipality. For example, one of the city's most upscale areas, Batel, has recently seen new development around the so-called *Batel Soho area* which has established its own website¹⁰ for marketing the area and local business.

Our case study indicates that social media data can provide information on different users and activity groups inside the city and this information could be used in analyzing and planning for example tourism related services. Despite the fact that social media data can provide detailed information to support city branding, in Curitiba's case the city could further benefit from the use of new techniques. Furthermore, by using new techniques the city could highlight its position among the best planned cities in the world.

As conventional citizen participation methods, that were incorporated into traditional planning since the 1960s, include a range of tools and tactics that require citizens to be physically present at a particular time and place [Kleinhans et al. 2015], social media data is not dependent on the place and time. Social media also gives special groups and groups that usually are silent an opportunity to express their opinions and concerns. Social media data is available anytime and data usually includes different time series. In contrast to traditional data used in urban planning, this provides cities opportunities to be more reactive in the decision making process [Townsend 2000]. Social media data could be suitable for quick analysis supporting targeted decision making especially in small spatial scales such as the neighborhood or street scale.

One important argument for using these new data sources is their availability and ability to use them without large investments. However, even if the social media and new data sources can provide comparative advantage related to the cost associated with certain planning tasks, some studies, e.g. [Kleinhans et al. 2015], have found that using social media and new technologies will not reduce the workload of professionals. New ways of communicating and collecting data from urban dwellers can even increase the workload related to analyzing and understanding complex urban issues. The analyst should be capable of understanding how the

¹⁰<http://batelsoho.com.br/>.

data is collected and what are the potential limitations of the data so that it can serve to better reflect the truth. Besides, a throwback of social media data for the public officials is the fact that even if social media data seems to be extensive and large in size, the data may be discontinuous and uncertain due to its dependence on the user [Silva et al. 2016, Wu et al. 2014].

Twitter data seems very appealing for research because of its qualities (e.g. large quantities of data, easiness to collect and analyze) [Blank 2016]. Still, some issues remain regarding the usability. One huge issue for the social media data usage is how the data is covering the population or the possible bias associated. According to [Blank 2016] there are huge demographic differences between users and not users of Twitter. In Britain and the United States of America the users are more likely to be younger, better educated, students or employed, single, and wealthier than the people not using Twitter. Our collection, explored in the case study, was based on data shared by users of Untappd on Twitter. Data may reflect the behavior of a fraction of consumers and therefore, there could be biases related to the fact that the users of such application are not necessarily representative of the population of craft beer drinkers. They are likely to be young, owners of smartphones, and urban dwellers.

Besides that, some social media applications may offer interfaces only in English, as is the case of Untappd. This would surely scare away the majority of the less educated citizens of countries that do not have English as their mother tongue. Consequently, urban areas with older and poorer populations will provide fewer data and be underrepresented in any analysis. Besides, users may not share data concerning all of their destinations, for example, because the information will be made public on Twitter. Thus, datasets obtained from social media might offer a partial view of consumers habits, which needs to be handled carefully.

6. Conclusion and Future Work

Technological development has not only changed the way how new data is collected but also changed the ways how urban dwellers communicate. This development enables also new sources to collect data for urban planning and decision making. Social media is offering new opportunities to interact, communicate and execute more data-informed decisions. Social media provide cities new ways to communicate with dwellers but also collect perceptions about the city. Some cities have also shown interest in using new data sources in policy making and planning: new data sources could be used to validate official data and provide detailed insight complying knowledge gaps that the city has regarding decision making.

Review of international research shows that benefits of the social media data, compared to conventional official datasets, include large quantity, ability to conduct pro-active problem solving, new time horizon to urban dynamics and ability to measure urban phenomena in very detailed spatial scale. The potential of social media data in measuring the use of urban space and people's mobility can, for example, provide new information about functional areas. The city could use this data to organize activities and make well-informed planning at different scales. The results of the case study show that social media data can provide information on how certain parts of the city are developing. We were able to identify the recently developed Curitiba's Beer Street from the data collected from Untappd. The city could use this data as an evidence to support their decision making. Finding characteristics describing certain areas inside the city can help the city organize urban activities and provide tools for neighborhood and city marketing. This is why social media data can be beneficial for the place branding and the building of the city's image.

Even if the case study and related research on social media data in planning context raises lots of opportunities for the future planning of cities, there are also remaining questions that fu-

ture research needs to address. The limitations of the data and social media user groups set certain concerns for large scale implementation. Use of new data sources requires vast knowledge of the data and phenomena under investigation which can set limits for large implementation by the public sector. More studies are required in order to fully understand the potential and usability of social media data in evidence based decision making.

Curitiba is well known globally for its planning, but how the city is taking new responsive techniques into consideration requires more research. Further research on the use of data in the planning of Curitiba, would help us to determine better the opportunities the different social media data offer to the city. Other future work should extend the analysis into multiple data sources and combine the new data sources to city's own official datasets. Considering Curitiba, it would be interesting to study social media data together with the city's own data such as citizen complaints, business licenses, and accidents. Combining social media data to conventional official datasets could help to fulfill the knowledge gaps the city have related to certain urban issues. In the best case, social media data could help the city to better sense and organize itself.

Acknowledgment

This work was partially financed with resources from FAPEMIG, Fundação Araucária, Capes and CNPq. This work is partially funded by EU-BR EUBra-BigSea project (MCTI/RNP 3rd Coordinated Call).

References

- Afzalan, N. and Evans-Cowley, J. (2015). Planning and social media: Facebook for planning at the neighbourhood scale. *Planning Practice & Research*, 30(3):270–285.
- Ashworth, G. and Kavaratzis, M. (2009). Beyond the logo: Brand management for cities. *Journal of Brand Management*, 16(8):520–531.
- Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, 3(3):274–279.
- Blank, G. (2016). The digital divide among twitter users and its implications for social research. *Social Science Computer Review*, pages 1–19.
- Cerrone, D., Pau, H., and Lehtovuori, P. (2015). *A Sense of Place. Exploring Analytic Potential of Location Based Social Network Data in Turku City Centre*. Turku Urban Research Programme's Research Report 1/2015.
- Cheng, Z., Caverlee, J., Lee, K., and Sui, D. Z. (2011). Exploring millions of footprints in location sharing services. *ICWSM*, 2011:81–88.
- Chorley, M., Rossi, L., Tyson, G., and Williams, M. (2016). Pub crawling at scale: tapping untapped to explore social drinking. In *Proc. of ICWSM'16*.
- Cranshaw, J., Schwartz, R., Hong, J. I., and Sadeh, N. (2012). The livelihoods project: Utilizing social media to understand the dynamics of a city. *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, pages 58–65.
- Davoudi, S. (2006). Evidence-based planning. rhetoric and reality. *disP – The Planning Review*, 42(165):14–24.
- Dühr, S. and Müller, A. (2012). The role of spatial data and spatial information in strategic spatial planning. *Regional Studies*, 46(4):423–428.
- Ertiö, T.-P. (2015). Participatory apps for urban planning—space for improvement. *Planning Practice & Research*, 30(3):303–321.
- Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. of KDD'96*, volume 96, pages 226–231.
- Faludi, A. and Waterhout, B. (2006). Introducing evidence-based planning. *disP-The Planning Review*, 42(165):4–13.
- Frias-Martinez, V. and Frias-Martinez, E. (2014). Spectral clustering for sensing urban land use using twitter activity. *Engineering Applications of Artificial Intelligence*, 35:237–245.
- Govers, R. (2011). From place marketing to place branding and back. *Place Branding and Public Diplomacy*, 7(4):227–231.
- Ho, A. (2017). Big data and evidence-driven decision-making: Analyzing the practices of large and mid-sized us cities. In *Proceedings of the 50th Hawaii International Conference on System Sciences*, pages 2794–2803.

- IPPUC (2017). *História do IPPUC*. Instituto de Pesquisa e Planejamento Urbano de Curitiba. Website: www.ippuc.org.br, Curitiba, Brazil.
- Jones, P., Layard, A., Speed, C., and Lorne, C. (2015). Maplocal: use of smartphones for crowdsourced planning. *Planning Practice & Research*, 30(3):322–336.
- Kavanaugh, A. L., Fox, E. A., Sheetz, S. D., Yang, S., Li, L. T., Shoemaker, D. J., Natsev, A., and Xie, L. (2012). Social media use by government: From the routine to the critical. *Government Infor. Quar.*, 29(4):480–491.
- Kitchin, R. (2014). The real-time city? big data and smart urbanism. *GeoJournal*, 79(1):1–14.
- Kleinhans, R., Van Ham, M., and Evans-Cowley, J. (2015). Using social media and mobile technologies to foster engagement and self-organization in participatory urban planning and neighbourhood governance. *Planning Practice & Research*, 30(3):237–247.
- Kling, F. and Pozdnoukhov, A. (2012). When a city tells a story: urban topic analysis. In *Proc. of the 20th international conference on advances in geographic information systems*, pages 482–485. ACM.
- Knox, S. and Bickerton, D. (2003). The six conventions of corporate branding. *European journal of marketing*, 37(7/8):998–1016.
- Macedo, J. (2004). Curitiba. *Cities*, 21(6):537–549.
- Mather, T. (2010). Untappd connects beer drinkers everywhere. untappd - drink socially. In *Untappd Blog*. <http://blog.untappd.com>.
- Moilanen, T. and Rainisto, S. (2008). *How to brand nations, cities and destinations: a planning book for place branding*. Springer.
- Noulas, A., Scellato, S., Mascolo, C., and Pontil, M. (2011). An empirical study of geographic user activity patterns in foursquare. In *Proc. of ICWSM'11*, volume 11, pages 570–573.
- Preoțiu-Pietro, D. and Cohn, T. (2013). Mining user behaviours: a study of check-in patterns in location based social networks. In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 306–315. ACM.
- Schwartz, R., Naaman, M., and Matni, Z. (2013). Making sense of cities using social media: Requirements for hyper-local data aggregation tools. In *Proc. of ICWSM'13*, pages 15–22.
- Sevin, E. (2016). Branding cities in the age of social media: a comparative assessment of local government performance. In *Proc. of Social Media and Local Governments*, pages 301–320. Springer.
- Silva, T., Celes, C., Neto, J., Mota, V., Cunha, F., Ferreira, A., Ribeiro, A., Vaz de Melo, P., Almeida, J., and Loureiro, A. (2016). Users in the urban sensing process: Challenges and research opportunities. In *Pervasive Computing: Next Generation Platforms for Intelligent Data Collection*, pages 45–95. Academic Press.
- Silva, T., Vaz De Melo, P., Almeida, J., and Loureiro, A. (2014a). Large-scale study of city dynamics and urban social behavior using participatory sensing. *Wireless Communications, IEEE*, 21(1):42–51.
- Silva, T. H. and Graeml, A. R. (2016). Exploring collected intelligence from untappd to support the location decision for new smes. In *Proc. of Brazilian Symposium on Multimedia and the Web*.
- Silva, T. H., Vaz de Melo, P. O. S., Almeida, J. M., Salles, J., and Loureiro, A. A. F. (2013). A comparison of foursquare and instagram to the study of city dynamics and urban social behavior. In *Proc. of ACM SIGKDD Int. Workshop on Urban Computing (UrbComp'13)*, Chicago, USA.
- Silva, T. H., Vaz de Melo, P. O. S., Almeida, J. M., Salles, J., and Loureiro, A. A. F. (2014b). Revealing the city that we cannot see. *ACM Trans. Internet Technol.*, 14(4):26:1–26:23.
- Syssner, J. (2010). Place branding from a multi-level perspective. *Place Branding and Public Diplomacy*, 6(1):36–48.
- Townsend, A. M. (2000). Life in the real-time city: Mobile telephones and urban metabolism. *Journal of urban technology*, 7(2):85–104.
- Vaca, C., Quercia, D., Bonchi, F., and Fraternali, P. (2015). Taxonomy-based discovery and annotation of functional areas in the city. In *Proc. of ICWSM'15*.
- Wakamiya, S., Lee, R., and Sumiya, K. (2011). Crowd-based urban characterization: extracting crowd behavioral patterns in urban areas from twitter. In *Proc. of 3rd ACM SIGSPATIAL international workshop on location-based social networks*, pages 77–84. ACM.
- Wu, L., Zhi, Y., Sui, Z., and Liu, Y. (2014). Intra-urban human mobility and activity transition: Evidence from social media check-in data. *PloS one*, 9(5):e97010.
- Xia, C., Schwartz, R., Xie, K., Krebs, A., Langdon, A., Ting, J., and Naaman, M. (2014). Citybeat: real-time social media visualization of hyper-local city data. In *Proc. of WWW'14*, pages 167–170. ACM.
- Zhang, K., Jin, Q., Pelechris, K., and Lappas, T. (2013). On the importance of temporal dynamics in modeling urban activity. In *Proc. of ACM SIGKDD International Workshop on Urban Computing*, page 7. ACM.
- Zhou, L. and Wang, T. (2014). Social media: A new vehicle for city marketing in china. *Cities*, 37:27–32.