IDENTIFYING THE URBAN SPACE FOR LOCALS AND TOURISTS THROUGH "FOURSQUARE" DATA IN BARCELONA

IDENTIFICACIÓN DEL ESPACIO URBANO POR RESIDENTES Y TURISTAS, A TRAVÉS DE DATOS DE "FOURSQUARE" EN BARCELONA

YANG, Liya
Department of Architectural Technology (TA), Centre of Land Policy and Valuations (CPSV)
Technical University of Catalonia (UPC)
PhD student
Av. Diagonal, 649, 4th floor, Barcelona, C.P. 08028, Spain
E-mail: li.ya.yang@upc.edu
Telephone: +34 934054385

MARMOLEJO DUARTE, Carlos
Department of Technology of Architecture (TA), Center for Land Policy and Valuations (CPSV)
Polytechnic University of Catalonia (UPC)
Associated Professor
Av. Diagonal, 649, 4th floor, Barcelona, C.P. 08028, Spain
E-mail: carlos.marmolejo@upc.edu
Telephone: +34 934054385

MARTÍ CIRIQUIÁN, Pablo
Department of Building and Urbanism
University of Alicante (UA)
Associated Professor
Carretera de San Vicente del Raspeig s/n, San Vicente del Raspeig, Alicante, C.P. 03690. Spain
E-mail: pablo.marti@ua.es
Telephone: +34 965903982

Key words: tourism; urban space; LBSN; land uses

Palabras clave: turismo; espacio urbano; LBSN; usos del suelo

Abstract

Barcelona is an important touristic city in the world. According to Annual Report of Tourism of Barcelona (2014), more than 7.5 million tourists visited here in that year. The studies related to tourism of Barcelona are numerous; however, the comparison of activities and land uses between tourists and locals is scarcely analyzed. In fact, tourism may be a dominant factor of urban development as well as a source of social conflict. Therefore, it is crucial to understand the co-living situation of tourists and residents in a touristic city. The main objective of the study is to identify touristic users and local users through their Foursquare behaviors. Furthermore, it explores the difference of geospatial activities and POIs' usages between the two groups. The analytical period is from April of 2012 to September of 2013, based on the monitoring span of Foursquare data. After filtration, the total check-ins during this period is 80,936 coming from 4,250 Foursquare users. The POIs of Foursquare are 13,887 in Barcelona. The geographic range of data roughly covers the central conurbation of the Metropolitan area of Barcelona.

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

The methodology includes four parts. The first step is to select indicators of behavior and standardization. The second step consists of selecting two short-period samples and classifying them into tourists and locals by K-means clustering. After the manual examination of the initial result, a threshold of classification is introduced to improve the result. Finally, the same method of identification is applied to the whole dataset.

According to the result, the difference of POI usages verifies that the identification is effective. It reflects the typical activities of tourists and locals separately in the city. The most visited POIs of tourists are: outdoor resorts, transport, restaurants, hotel, and store. The corresponding rank of locals is restaurants, workplaces, outdoor resorts, educational places, and transport.

Moreover, the two groups appear different Foursquare behaviors, regardless of the length of analyzing period. In general, behaviors of tourists -- the stay duration, number of check-ins, and total travel distance, are smaller than the local group. K-means clustering can effectively identify users who possess the extreme values of attributes. However, it is unavoidable to introduce artificial intervention for users without extreme-characteristics.

Besides, the geospatial distribution and active time also embody differences between locals and tourists. In terms of movement scale, tourists seem more concentrated than the residents. With regard to the active time, tourists' active period is similar every day. On the contrary, locals show an evident periodic variation daily and weekly.

It is undeniable that this paper has several limitations. Firstly, Foursquare data has bias. The high proportion of check-ins is restaurants because Foursquare aims to provide practical information about places for users. What's more, the lack of demographic information of users also limits the scope of the study, due to the privacy policy.

In sum, this study demonstrates that it is possible to distinguish tourists from locals via Foursquare data, though the uncertainty of data is recognized. How to improve the accuracy of the unsupervised identification and cooperate with other datasets will be the object of further investigation. Furthermore, whether the identification model can be universally applied is another issue that is worth to test in the future.

Resumen

Barcelona es una importante ciudad turística en el mundo. Según el Informe Anual de Turismo de Barcelona (2014), más de 7,5 millones de turistas la visitaron este año. Los estudios relacionados con el turismo en Barcelona son numerosos, sin embargo, la comparación de actividades y usos del espacio entre turistas y residentes es poco analizada. De hecho, el turismo puede ser un factor dominante del desarrollo urbano, así como una fuente de conflicto social. Por lo tanto, es crucial comprender la situación de convivencia de turistas y residentes en una ciudad turística. El objetivo principal del estudio es identificar usuarios turísticos y usuarios locales a través de sus comportamientos de Foursquare. Además, explora la diferencia entre las actividades geoespaciales y los usos de los puntos de interés *(POIs)* entre los dos grupos. El período analizado abarca desde abril de 2012 a septiembre de 2013, según el intervalo de monitoreo de los datos de Foursquare. Después de la filtración, el total de los registros durante este período son 80,936 provenientes de 4,250 usuarios de Foursquare. Los

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

POIs de Foursquare son 13,887 en Barcelona. El rango geográfico de los datos cubre aproximadamente la conurbación central del área metropolitana de Barcelona.

La metodología incluye cuatro partes. El primer paso es seleccionar indicadores de comportamiento y estandarización. El segundo paso consiste en seleccionar dos muestras de corto período y clasificarlas en turistas y locales por agrupación de K-means. Después del examen manual del resultado inicial, se introduce un umbral de clasificación para mejorar el resultado. Finalmente, el mismo método de identificación se aplica a todo el conjunto de datos.

De acuerdo con el resultado, la diferencia de uso de *POIs* verifica que la identificación sea efectiva, reflejando las actividades típicas de turistas y residentes por separado en la ciudad. Los *POIs* más visitados de los turistas son: complejos turísticos al aire libre, transporte, restaurantes, hoteles y tiendas. El rango correspondiente de los residentes es: restaurantes, lugares de trabajo, centros turísticos al aire libre, lugares educativos y transporte.

Además, independientemente de la duración del período de análisis, los dos grupos tienen diferentes comportamientos de Foursquare. En general, los comportamientos de los turistas: la duración de la estadía, el número de registros y la distancia total de viaje son menores que los del grupo de locales. El cluster de K-means puede identificar efectivamente a los usuarios que poseen los valores extremos de los atributos. Sin embargo, es inevitable introducir una intervención artificial para usuarios sin características extremas.

Además, la distribución geoespacial y el tiempo activo también representan diferencias entre los lugareños y los turistas. En términos de escala de movimiento, los turistas parecen más concentrados que los residentes. Con respecto al tiempo activo, el período activo de los turistas es similar todos los días. Por el contrario, los residentes muestran una evidente variación periódica diaria y semanal.

Es innegable que este trabajo presenta limitaciones. En primer lugar, los datos de Foursquare tienen sesgo. La alta proporción de check-ins en restaurantes es producto de que Foursquare tiene como objetivo proporcionar información práctica sobre los lugares para los usuarios. Además, la falta de información demográfica de los usuarios también limita el alcance del estudio, debido a su política de privacidad.

En resumen, este estudio demuestra que es posible distinguir a los turistas de los residentes a través de los datos de Foursquare, aunque se reconoce la incertidumbre de los datos. Cómo mejorar la precisión de la identificación no supervisada y cooperar con otros conjuntos de datos será objeto de investigación adicional. Además, si el modelo de identificación puede aplicarse universalmente es otro tema que vale la pena probar en el futuro.

1. Introduction

With the increasing mobility among cities, visitors are becoming an important part of the city because cities actually provide permanent services for them, such as hospitality, tourist information center. From the perspective of activities, visitors usually "occupy" some areas of a city. As Page and Hall (2003, pp. 49) note that “(…) tourism is subsumed and integrated into the postmodern city …it is one aspect of the form of the city.” However, most researches either only focus on the tourists (Vu, Huy Quan *et al.* 2015; Kádár, B. 2014) or the residents (Sun, Y.

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through “Foursquare” data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. “Ciudades y Territorios Inteligentes”. UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

2016). Their behaviors (i.e. moving patterns, use of services, etc.) in the same city are scarcely compared. Besides, the specific land uses of tourists are not discussed completely. Many studies of touristic activities just center on the moving patterns (García-Palomares, J.C. *et al.* 2015; Mckercher, B. & Lau, G 2008), lack the investigation of tourists' land uses. Therefore, this paper chooses Barcelona as the case study, to investigate differences between locals and tourists in terms of behaviors and activities performed through Foursquare data.

Created in 2009, Foursquare is a local search-and-discovery service application. It provides practical living information about places for users. Globally it has over 50 million users and cumulates more than 12 billion check-ins. The global distribution of Foursquare check-ins was mainly in America, Europe, and Southeast Asia in 2012 (Pontes Tatiana *et al.* 2012).

As a location-based social network (LBSN), the main components of Foursquare data include venue (i.e. a place), Foursquare users and their check-ins on the platform. Because of the accessibility of huge dataset, researchers have opportunity to examine its relation with urban activities, and to compare human behaviors across different datasets. For example, Silva, T. H. *et al*. (2013) compared Foursquare and Instagram datasets in three different cities. They suggested that both datasets "might be compatible in finding popular regions of cities". Agryzkov, T. *et al.* (2017) utilize foursquare data to build a network to measure urban activities in Murcia.

The main objective of the study is to identify touristic users and local users through their Foursquare behaviors. Furthermore, it explores the variation of tourists and locals on their geospatial activities and land uses. The remain of the paper develops as follows: section two reviews the studies of touristic behaviors and location-based social network data; section three delimitates the scope of research; section four explains the methodology and the data; section five displays the results of identification; and the final part is conclusions and discussions.

2. Literature review

With regard to the identification of tourists, field survey is a traditional approach. For example, in 2017, Barcelona government estimated the proportion of touristic pedestrians in tourist season around one of the main touristic street --Passeig de Gràcia, based on a field survey. Mckercher, B. & Lau, G (2008) found out tourists though making questionnaires on hotel lobbies. However, it is difficult to make surveys in large areas or a large number of tourists.

LBSN data can overcome this limitation and provide more precise human tracks in urban areas. It records the specific timestamps and locations automatically when people make check-ins and share their locations in social network applications. Therefore, LBSN data combine temporal information with spatial information. It has become an important data source to understand and forecast human movements (Hasan, S.*et al.* 2013).

Moreover, many studies have proved that human movement is not a random process (Gonzalez, M. C. et al. 2008), and different groups of users exhibit different characteristics on LBSN data. For example, the "locality of social media behaviors" has been mentioned in several pieces of research (Sun, Y. 2016; Yin Zhihong, 2014; Jue, J., & Xiaolu, G. 2012). It indicates that most of our movements concentrate in a certain range, instead of random distribution. Based on three different LBSN datasets, Cho *et al.* (2011) found that periodic behaviors

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

(moving between home and workplaces) account for 50%-70% of all human movement. Gao, Qi *et al.* (2012) proved that behaviors on Weibo and Twitter are related to the cultural background. Cranshaw *et al.* (2012) states that there are differences between the behavior of residents and visitors.
Da Rugna *et al.* (2012) showed that geotagged photos on Flickr could identify the original country of tourists. They counted the number of countries that each user visited from 2010 to 2011 and calculated users' total length of stay in those countries. The country that a user stayed longest was considered as his/her original country. However, as the paper noticed, the method would fail if users did not make enough check-ins in their home countries.

Vu, Huy Quan *et al.* (2015) combined GPS data and Flickr to show tourists' main routes in Hongkong. They identify tourists through the user's demographic information on Flickr and their locations of publishing photos of tourist attractions. Luo *et al.* (2016) distinguished residential Twitter users from visitors through their locations during the night in Chicago. For each user, the most frequently visited place at night is defined as "home place". Users are identified as locals if most of their check-ins are located in residental areas during nights. However, this method is effective only if hospitality services are segregated from residential areas. It is hard to apply the approach into a compact city which tends to mix-use land. Noulas, A. *et al.* (2011) defined a local active user as a user whose total number of check-ins were above 30 and the majority activities were within their monitoring areas in New York and London. The limitation of the method is that it has to drop inactive users who usually account for a large volume in datasets.

Kádár, B. (2014) adopted a threshold of 5 consecutive days to distinguish tourists in Vienna, Prague and Budapest. Girardin *et al.* (2008) used a period of 30 days to separate tourists from locals in Province of Florence of Italy through Flickr data. If a user took pictures in the region beyond 30 days, he/she was considered as a resident. Eric Fischer (2012) [1] identified them by Flickr and Picasa data in 2012. He defined "locals" as: "people who have taken pictures in the same city dated over a range of a month or more". If the users who have not taken pictures anywhere for over a month, they are classified as unknown. García-Palomares, J.C. *et al.* (2015) used a similar method to identify tourists via Panoramio data. The longer time threshold could be more reliable; however, it means that huge dataset is indispensable.

In sum, the above methods of classification mainly rely on the geo-location or the time threshold, rather than the social media behaviors. Moreover, the time threshold is usually derived from empirical experiences (Kádár, B. 2014; Girardin *et al.* 2008; Noulas, A. *et al.* 2011) or advices of tourism experts (Luo *et al.* 2016). It is hard to judge what length of time span is the "correct" one. Thus, this work tries to classify them by user's behaviors on Foursquare and a threshold based on statistic results of the data.

3. Study scope

Barcelona has a long history of development of tourism. The government established Commission for the Attraction of Foreigners and Tourists in 1906 and aimed to build the city as the "Pearl of Mediterranean". Especially, it turned into an important touristic city in the world after the Olympic Games of 1992. According to the Annual Tourism Sector of Barcelona Report

1 Source: https://www.flickr.com/photos/walkingsf/sets/72157624209158632/

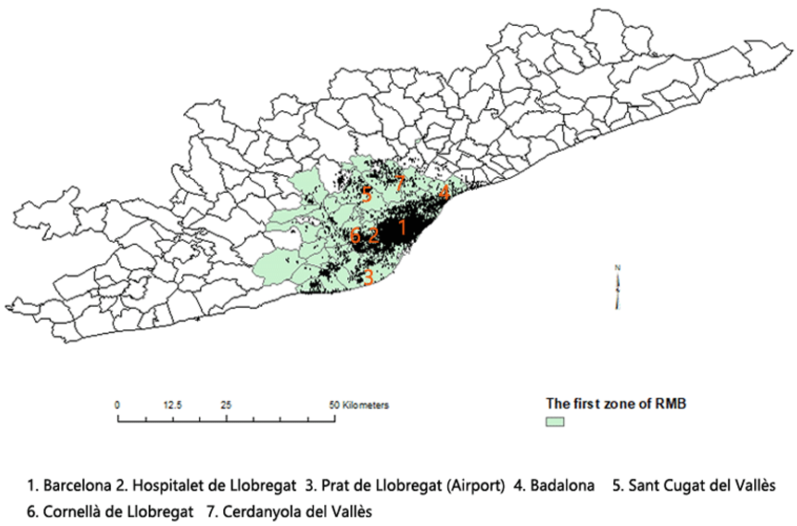
Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

ISBN: 978-84-8157-661-0

2014, the total number of tourists reached more than 7.5 million, being the 20th of the most visited cities in the world[2].

The studied area includes the first zone of Barcelona Metropolitan Region, due to the monitoring range of Foursquare data (see Figure 1). In fact, according to the new urban planning 2011, this first zone is also called Metropolitan Area of Barcelona (AMB), consists of 36 municipalities. The population was 3,239,337 in 2014[3]. In other words, the tourists doubled the number of residents in that year. The actual range of Foursquare data is a little wider than AMB, because the monitoring range is not strictly matched.

Figure 1. **Distribution of check-ins in Barcelona Metropolitan Region**



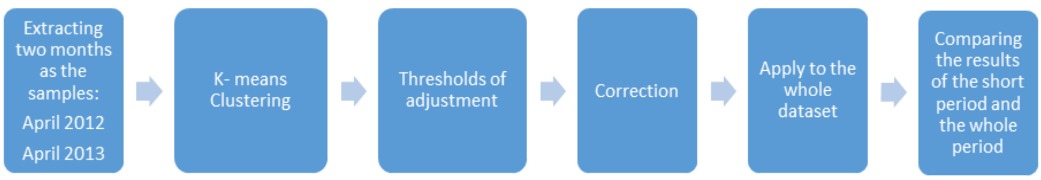
Source: Own elaboration

4. Methodology of classification

4.1 Structure of the methodology

The identification of locals and tourists on Foursquare is based on two assumptions: first, the number of tourists' check-ins, travel distance, and duration of stay are lower than residents on average; second, tourists have different characteristics of urban usages from residents, which can reflect on Foursquare POIs. The following is the structure of methodology.

Figure 2. **Structure of methodology**



Source: Own elaboration

[2] Source: https://ajuntament.barcelona.cat/turisme/sites/default/files/documents/150514_the_tourism_sector_eng_0.pdf

[3] Source: Institut d'Estadística de Catalunya

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

For identifying the touristic and local users through Foursquare behaviors, this paper selects two months as samples to make tests at first. This paper describes user's behaviors though three indicators: total duration of stay, total travel distance and numbers of check-ins.

$$\text{Total Duration} = \sum_{i}^{n} (T_{i+1} - T_i) \qquad (1)$$

where T is the timestamp, *i* is the ordinal number of timestamp which starts from the first timestamp of a user.

$$\text{Total Travel Distance} = \sum_{i}^{n} (D_{i+1} - D_i) \qquad (2)$$

where D is the location of a user at timestamp *i*.

Before classification, it uses Z-score to standardize the three indicators:

$$z = \frac{(x - \mu)}{\sigma} \qquad (3)$$

where z is the standardized score of indicators, x is the value of indicator, μ is the mean of x, σ is the standard deviation.

Next, K-means clustering is applied to classify users. For discrete data, algorithms of grouping data are classification and clustering. Classification requires a training dataset which contains samples whose category is known. As the characteristics of tourist behavior are unknown in our case, clustering is the better approach to divide users. Hierarchical clustering and K-means clustering belongs to the most common algorithms of clustering. Hierarchical clustering does not need to set the number of categories initially. It will produce all possible results of categories. On contrary, K-means clustering requires to set the number of cluster at first. It is widely applied due to its simplicity. Moreover, the quality K-means algorithms performances very good in huge dataset (Abbas, O.A., 2008).

After manual examination, it detects some local users in the tourist group because they are not very active in the analyzing period. Therefore, it needs to improve results through artificial interference. Those users whose indicators are above the threshold will be grouped into locals. This paper tests four different thresholds to improve the initial result and select the optimum one as the final outcome.

Next, the same process of identification is going to apply to the whole dataset. Finally, the paper examines the consistency of the results between the sample months and the whole dataset. Based on the result of classification, it compares the usages of Foursquare POIs between tourists and locals.

$$\text{Difference percentage} = PT_i - PL_i \qquad (4)$$

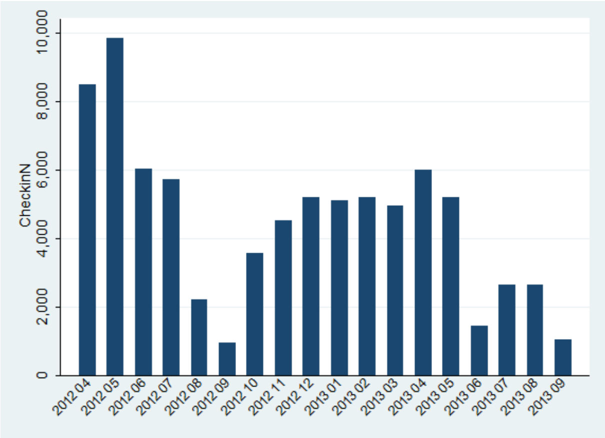
where PT_i is the percentage that tourists made check-ins at the i^{th} type of POIs, similarly, PL_i is the local's.

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

4.2 Preliminary statistics and thresholds of improvement

This paper extracts data from a global Foursquare check-ins dataset[4]. The period of the dataset is from 2012-04-03 to 2013-09-16. After filtration, total check-ins in Barcelona is 80936 items. The volume of check-ins declined significantly after June of 2013. It only has 6931 check-ins from June to September of 2013 (see Figure 3). It may be caused by the reducing of active users of Foursquare, or some changes of privacy policy, or unknown technic problems. It exists loss of data in two periods: from 25 of August to 03 of September, and 25 of September to 16 of October in 2012.

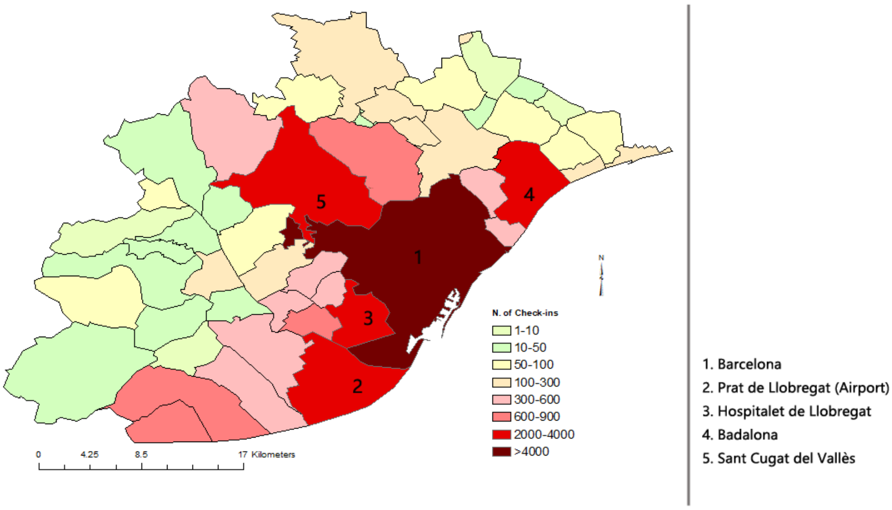
Figure 3. **Monthly Check-ins of Foursquare in Barcelona**



Source: Own elaboration

Most check-ins are located in Barcelona city, which is 57764 items (see Figure 4). Except for Barcelona, only four cities have check-ins over 1000: Hospitalet de Llobregat, Prat de Llobregat, Badalona, Sant Cugat del Vallés and Conellá de Llobregat. They are nearby cities of Barcelona city.

Figure 4. **Distribution of Foursquare check-ins**



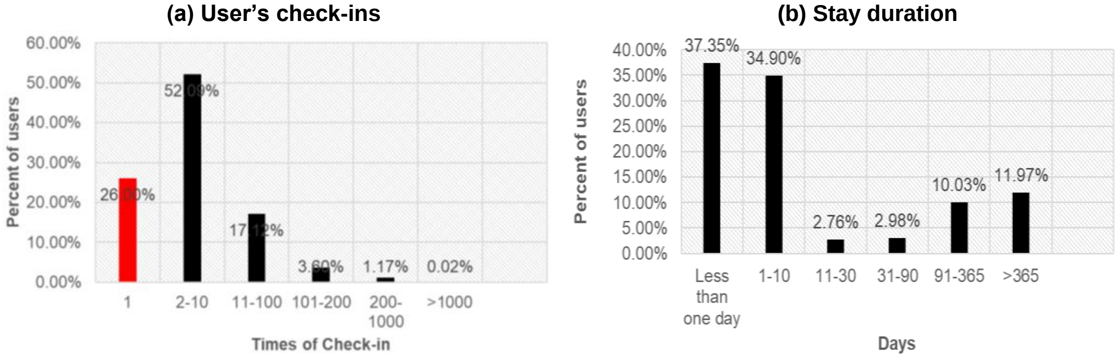
Source: Own elaboration

4 Data source: Dingqi Yang, Daqing Zhang, Bingqing Qu. Participatory Cultural Mapping Based on Collective Behavior Data in Location Based Social Networks. *ACM Trans*. on *Intelligent Systems and Technology (TIST)*, 2015

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

According to the unique ID, 4527 users appeared in the whole period. 1177 users only have one check-in. More than half of users' check-ins number are between 2 and 10 (see Figure 5). With Regard to the duration of stay, 3390 users stay less than 30 days in Barcelona. Over 20% of users' duration is more than 365 days. Nevertheless, it is possible that some of them are returned visitors in the second year, then the stay duration will be very long according to the method of calculation. Therefore, the check-in of users is another important indicator of identification.

Figure 5. **Foursquare users' check-ins and stay duration**



Source: Own elaboration

As a result, the adjusting thresholds consist of two conditions: check-ins and stay duration. Three of thresholds are based on the mean values of the dataset, one is from the empirical study (see Table 1).

Table 1. **Description of four adjusting thresholds**

Threshold	Total duration (days)	Check-ins	Description
1	84	18	Mean value of total 4527 users
2	90	24	Empirical threshold
3	98	21	Mean value of threshold 1 and 4
4	113	24	Mean value of total valid 3350 users

Source: Own elaboration

For reducing data noise, only valid users will be involved in the statistical analysis. It defines a "valid user" as a user who made check-in at least 2 times in the whole period. Thus, the total valid users are 3350 (see Table 2). The average of check-ins are about 24 times per user.

Table 2. **Summary of Valid Users**

Total Users		Mean	Std. Dev.	Min	Max
3350	Check-ins	23.8203	54.14576	2	1182
3350	Duration (days)	112.5728	179.0711	0	531

Source: Own elaboration

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

4.3 POIs of Foursquare

There are 13887 unique POIs of Foursquare in Barcelona that are labeled by 385 categories. Restaurants take a large portion of all types of POIs. This paper assembles these categories into 21 types for analyzing purpose. The similar items are grouped into one type: for example, all kinds of restaurants will be grouped as "restaurant". The table below lists the new classification with a brief description.

Table 3. **New Category of POIs**

Types of POIs	Number of POIs	Description
Restaurant	3318	Mediterranean Restaurant, Japanese Restaurant, Food, Diner, etc.
Transport	668	Train Station, Subway, Airport, Boat, Airport Terminal, Light Rail, etc.
Café	540	Café, Tea Room,Cafeteria, etc.
Education places	634	University, College, Elementary School, Student Center, etc.
Hotel	478	Motel, Hotel, etc.
Market	74	Fair， Farmers Market， Flea Market， Fish Market, etc.
Bar	1138	Bar, Beer Garden, Cocktail Bar ,Jazz Club, Nightclub, etc.
Sports center	308	Athletic & Sport, Baseball Field, Basketball Court ,Football Stadium, Golf Course, etc.
Shop	940	Bike Shop, Dessert Shop, Frozen Yogurt, Gift Shop, etc.
Store	839	Kids Store, Pet Store, Paper / Office Supplies Store, Video Store, etc.
Museum, Art, Historical place	206	Public Art, Performing Arts Venue, Museum, Historic Site, Castle, etc.
gym	148	Gym Pool, Gym, Yoga Studio, etc.
Opera, concert, cinema	192	Indie Movie Theater, Concert Hall, Movie Theater, Opera House, etc.
Work place	1174	Building, Campaign Office, Co-working Space, Design Studio, etc.
Services	1321	Medical, Finance, Post Office, Bakery, Salon, Barbershop, Spa, Tattoo, etc.
Outdoor Resorts	1122	Rest Area, Park, Plaza, Scenic Lookout, etc.
residential place	506	Neighborhood, Residential Building (Apartment / Condo), etc.
infrastructure	131	Bridge, Harbor / Marina, River, etc.
conference center	88	Conference Room, Meeting Room, Convention Center, etc.
Touristic Info center	5	Tourist Information Center
Others	57	Stables, Track ,Planetarium, etc.

Source: Own elaboration

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona*. En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

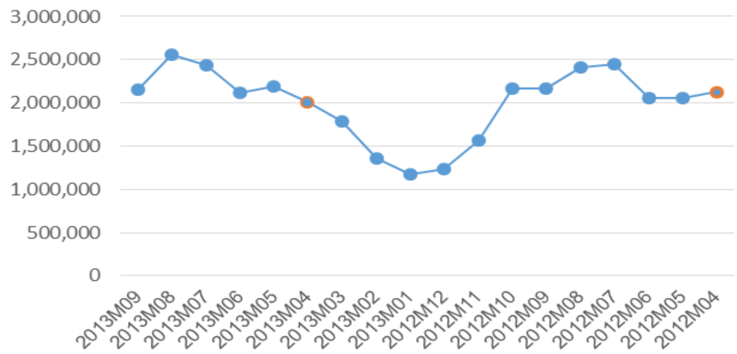
ISBN: 978-84-8157-661-0

5. Results

5.1 Selection of samples

This paper chooses two months as samples to distinguish locals and tourists: from 03-Apr-2012 to 03-May-2012 and the same period in 2013. Both of them have higher check-ins data of Foursquare and higher number of travelers (see Figure 6).

Figure 6. **Travelers and overnight stays by tourist sites in Barcelona**



Source: I.N.E. Available at : http://ine.es/jaxiT3/Tabla.htm?t=2078&L=0

What's more, those users who only make one check-in in the analyzing months but have more check-ins in other months are also removed from the short-term datasets. After data cleaning, the summary of samples is the following:

Table 4. **Summary of the analyzing months**

Time window	Number of users	Number of check-ins	Average check-ins per user
April 2012	679	9136	13.455
April 2013	625	5902	9.443

Source: Own elaboration

5.2 Identification of locals and tourists

After standardization, it checks the correlation among three indicators. The indicators of two selected months appear the similar correlation. It indicates that these indicators are applicable to different periods (see Table 5).

Table 5. **Pearson correlation of three standardize Indicators**

April 2012	**S_Check-ins**	**S_Distance**	**S_Duration**
S_Check-ins	1.0000		
S_Distance	0.7195	1.0000	
S_Duration	0.5265	0.4924	1.0000
April 2013	S_Check-ins	S_Distance	S_Duration
S_Check-ins	1		
S_Distance	0.7287	1	
S_Duration	0.4629	0.4486	1

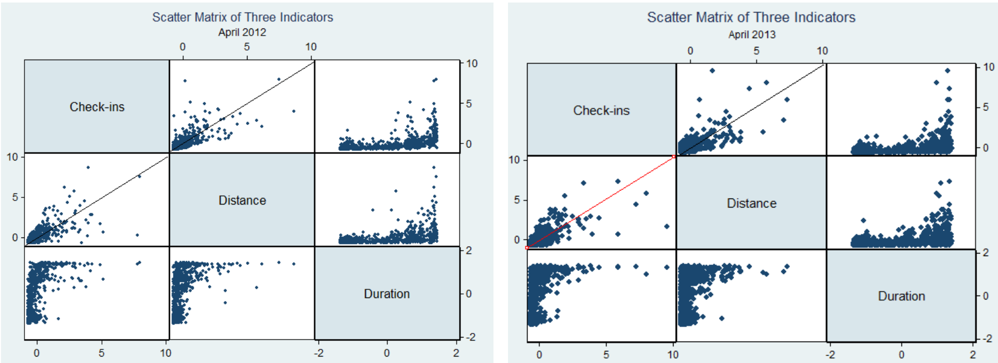
Source: Own elaboration

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona*. En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

ISBN: 978-84-8157-661-0

The correlation between check-ins and travel distance is higher. However, the scatter plot (see Figure 7) shows that the increase of check-ins is divergent with the increment of distance. It implies that both of them should be considered as factors of clustering.

Figure 7. **Scatter matrix of three indicators**



Source: Own elaboration

Next, it adopts k-means cluster to divide the users into locals and tourists by the three indicators. The initial result is the following:

Table 6. **Initial Identification of locals and Tourists**

	Number	Mean	Std. Dev.	Min	Max
Apr-2012					
Local users	**284 (41.8%)**				
Check-ins		23.9331	19.48105	5	139
Total Duration(days)		25.2204	3.727705	10.10531	30.10678
Travel Distance(m)		77223.49	71527.62	0	529345.6
Touristic users	**395(52.2%)**				
Check-ins		5.921519	4.485896	2	40
Total Duration		7.06041	7.095543	0.000116	24.06613
Travel Distance		15960.65	14330.25	0	73888.91
Apr-2013					
Local users	**267(42.7%)**				
Check-ins		16.07116	15.33849	2	122
Total Duration		25.44819	3.778656	2.971343	30.05297
Travel Distance		51861.31	49748.36	0	312638.8
Touristic users	**358(51.3%)**				
Check-ins		4.5	3.083116	2	26
Total Duration		6.571311	6.631968	0.000324	24.93288
Travel Distance		12134.43	11311.89	0	51337.16

Source: Own elaboration

In general, the touristic group's values are lower than the residents'. It accords with our first assumption. Moreover, it is worth to notice that the minimum value of travel distance is zero, which means that all check-ins of that user are at the same place, but in different time. It is possibly caused by users' interests or living habits. For example, we discover that one user only checked-in at his office every morning during one month. Next, four thresholds are applied to improve the initial results as we discussed above. The new results are the following (Table 7).

Table 7. **Identification of locals and tourists after improvement**

April 2012	Locals	Tourists
Before Correction	284	395
Threshold 1	442	237
Threshold 2	431	248
Threshold 3	437	242
Threshold 4	428	251
Maximum difference among 4 thresholds	**14**	
April 2013	**Locals**	**Tourists**
Before Correction	267	358
Threshold 1	431	194
Threshold 2	422	203
Threshold 3	427	198
Threshold 4	422	203
Maximum difference among thresholds	**9**	

Source: Own elaboration

The result is that both months have the similar proportion of locals and tourists. A large number of users transfer from touristic group to local group. Over 85% of these changed users have less than 10 check-ins during 30 days. It proves that inactive local users are easy to be identified as tourists by clustering. Moreover, results of the four thresholds are also similar. It indicates that the result of classification tends to stable when time span is larger than 84 days.

The mean values of indicators are lower than the initial classification. The local groups' values decrease evidently, especially the travel distance. It is caused by those local users who were less active in that month (See Table8). After correction, the behaviors of tourist's group tend to be more concentrated than locals: their cumulative travel distance are less than 15km and check-ins are below 6 times.

Table 8. **Mean Values of Different Thresholds' Result**

2012 April		Before Correction	Threshold 1	Threshold 2	Threshold 3	Threshold 4
Check-in	Local	23.9331	17.73077	18.02552	17.86728	18.10514
Duration(days)	Local	25.2204	20.77791	21.09073	20.95605	21.09725
Travel Distance(m)	Local	77223.49	56158.12	57215.25	56670.52	57546.13
Check-in	Tourists	5.921519	5.481013	5.512097	5.487603	5.525896
Duration	Tourists	7.06041	3.238897	3.473183	3.279598	3.672632
Travel Distance	Tourists	15960.65	14405.23	14419.98	14342.6	14367.26
2013 April		**Before Correction**	**Threshold 1**	**Threshold 2**	**Threshold 3**	**Threshold 4**
Check-in	Local	25.44819	20.10384	20.369	20.2278	20.369
Duration	Local	16.07116	11.60557	11.76066	11.66745	11.76066
Travel Distance	Local	51861.31	36873.26	37238.53	36984.68	37238.53
Check-in	Tourists	4.5	4.639175	4.625616	4.646465	4.625616
Duration	Tourists	6.571311	2.486797	2.716645	2.575371	2.716645
Travel Distance	Tourists	12134.43	11849.07	12199.18	12114.31	12199.18

Source: Own elaboration

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

As the results of four thresholds are very close to each other, we select two thresholds which have the largest difference to make further comparison. Thus, threshold 1 and 4 are involved in the following analysis of usages of Foursquare POIs.

From the view of volume of check-ins, local users contribute to a large portion of check-ins (see Table 9). On a year-on-year basis, the proportion of locals and tourists is at the same level, though the total check-ins in 2013 declined significantly.

Table 9. **Summary of Check-ins**

Time window	Threshold	Total check-ins	Check-ins of local users	%	Check-ins of touristic users	%
April 2012	1	9136	7,837	85.78	1,299	14.22
	4	9136	7,749	84.82	1,387	15.18
	1	5902	5002	84.75	900	15.25
April 2013	4	5902	4,963	84.09	939	15.91

Source: Own elaboration

According to the category of POIs that we set before, it summarizes the features of POI usages of the two groups. It calculates the number of check-ins and the corresponding percentage of each sub-category based on the POIs where users checked in. It lists the top 10 items of usages of Foursquare POIs according to threshold 1 and 4 (see Table 10).

Table 10. **POI Usages based on Threshold 1**

A. Locals

Rank	Locals 2012	%	Description	Locals 2013	%	Description
Total	**7837**	**100.00%**		**5002**	**100.00%**	
1	1297	16.50%	Restaurant	793	15.85%	Restaurant
2	838	10.69%	Outdoor Resorts	625	12.50%	Work place
3	829	10.58%	Work place	478	9.56%	Outdoor Resorts
4	700	8.93%	Transport	454	9.08%	Education places
5	622	7.94%	Education places	341	6.82%	Store
6	582	7.43%	Store	340	6.80%	Services
7	517	6.60%	Residential place	324	6.48%	Transport
8	504	6.43%	Services	286	5.72%	Gym
9	463	5.91%	Bar	272	5.44%	Shop
10	309	3.94%	Shop	259	5.18%	Residential place

B. Tourists

Rank	Locals 2012	%	Description	Locals 2013	%	Description
Total	**1299**	**100.00%**		**900**	**100.00%**	
1	217	16.71%	Transport	155	17.22%	Outdoor Resorts
2	188	14.47%	Restaurant	143	15.89%	Restaurant
3	185	14.24%	Outdoor Resorts	122	13.56%	Transport
4	140	10.78%	Hotel	88	9.78%	Hotel
5	106	8.16%	Museum, Art, Historical Place	73	8.11%	Museum, Art, Historical Place
6	83	6.39%	Bar	55	6.11%	Store
7	60	4.62%	Sports Center	51	5.67%	Bar
8	57	4.39%	Store	44	4.89%	Shop
9	48	3.70%	Shop	35	3.89%	Services
10	34	2.62%	Infrastructure	28	3.11%	Infrastructure

Source: Own elaboration

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

After comparison, threshold 4 has better performance than threshold 1(see Table 11) because the two years' ranks of locals obtain the consistency. Both ranks of tourists of two thresholds are the same in 2013 and only presents slightly differences in 2012.Therefore, this paper decides to use threshold 4 as the final adjusting criteria.

Table 11. **POI Usages based on Threshold 4**

A. Locals

Rank	Locals 2012	%	Description	Locals 2013	%	Description
Total	**7749**	**100.00%**		**4963**	**100.00%**	
1	1285	16.58%	Restaurant	787	15.86%	Restaurant
2	827	10.67%	Work place	620	12.49%	Work Place
3	814	10.50%	Outdoor Resorts	473	9.53%	Outdoor Resorts
4	687	8.87%	Transport	452	9.11%	Education Places
5	618	7.98%	Education places	339	6.83%	Store
6	578	7.46%	Store	339	6.83%	Services
7	516	6.66%	Residential Place	314	6.33%	Transport
8	502	6.48%	Services	286	5.76%	Gym
9	306	3.95%	Shop	271	5.46%	Shop
10	245	3.16%	Gym	258	5.20%	Residential Place

B. Tourists

Rank	Locals 2012	%	Description	Locals 2013	%	Description
Total	**1387**	**100.00%**		**939**	**100.00%**	
1	230	16.58%	Transport	160	17.04%	Outdoor Resorts
2	209	15.07%	Outdoor Resorts	149	15.87%	Restaurant
3	200	14.42%	Restaurant	132	14.06%	Transport
4	140	10.09%	Hotel	90	9.58%	Hotel
5	108	7.79%	Museum, Art, Historical Place	74	7.88%	Museum, Art, Historical Place
6	89	6.42%	Bar	57	6.07%	Store
7	62	4.47%	Sports Center	52	5.54%	Bar
8	61	4.40%	Store	45	4.79%	Shop
9	51	3.68%	Shop	36	3.83%	Services
10	37	2.67%	Infrastructure	28	2.98%	Infrastructure

Source: Own elaboration

According to threshold 4, the top five usages of locals are: restaurant, work place and outdoor resorts, education and store. The top five usages of tourists are: outdoor resorts, restaurant, transport, hotel and museum\art\historical place.

The ranks show an evident correlation with the two groups' typical activities in the city. Moreover, such correlation is consistent in two different years. It means that the method of classification is effective.

Next, we use the percentage of each item of tourists group to subtract the corresponding percentage of the local group. The values of difference reveal how they use the city and their characteristics of urban usages.

Hotel, Transport, and Touristic attractions are positive, which means that these places take more important roles in tourists' activities than local activities. On contrary, workplace, educational places are more checked-in by locals.

Table 12. **Difference of POI usages of two sample months**

April 2012				April 2013			
Locals	Tourists	Description	Difference of usage	Locals	Tourists	Description	Difference of usage
1.07%	10.09%	Hotel	9.02%	1.17%	9.58%	Hotel	8.42%
8.87%	16.58%	Transport	7.72%	6.33%	14.06%	Transport	7.73%
1.30%	7.79%	Museum, Art, Historical Place	6.48%	9.53%	17.04%	Outdoor Resorts	7.51%
10.50%	15.07%	Outdoor Resorts	4.56%	1.55%	7.88%	Museum, Art, Historical Place	6.33%
2.85%	4.47%	Sports Center	1.62%	0.62%	2.98%	Infrastructure	2.36%
1.29%	2.67%	Infrastructure	1.38%	4.35%	5.54%	Bar	1.19%
0.18%	1.01%	Market	0.83%	2.20%	2.88%	Sports Center	0.68%
0.40%	1.01%	Conference Center	0.61%	0.38%	0.64%	Market	0.26%
5.90%	6.42%	Bar	0.52%	0.73%	0.96%	Conference Center	0.23%
0.00%	0.14%	Touristic Info Center	0.14%	0.24%	0.43%	Others	0.18%
2.39%	2.31%	Cafe	-0.08%	15.86%	15.87%	Restaurant	0.01%
0.35%	0.14%	Others	-0.20%	0.02%	0.00%	Touristic Info Center	-0.02%
3.95%	3.68%	Shop	-0.27%	5.46%	4.79%	Shop	-0.67%
1.96%	0.94%	Opera, Concert, Cinema	-1.02%	6.83%	6.07%	Store	-0.76%
16.58%	14.42%	Restaurant	-2.16%	3.28%	1.81%	Café	-1.47%
3.16%	0.50%	Gym	-2.66%	2.06%	0.43%	Opera, Concert, Cinema	-1.63%
7.46%	4.40%	Store	-3.06%	6.83%	3.83%	Services	-3.00%
6.48%	2.24%	Services	-4.24%	5.20%	1.28%	Residential Place	-3.92%
6.66%	1.87%	Residential Place	-4.78%	5.76%	0.32%	Gym	-5.44%
7.98%	2.02%	Education Places	-5.96%	9.11%	1.17%	Education Places	-7.94%
10.67%	2.24%	Work Place	-8.44%	12.49%	2.45%	Work Place	-10.04%

Source: Own elaboration

5.3 Results of whole dataset

After the sample test, the same method is applied to classify the visitors and locals in the whole dataset. The total valid users are 3350 (see Table 13), total check-ins are 79798 items. After examination, the classification results of samples are accord with the result of whole dataset. Regarding the frequency of check-ins, 580 residents created 76% of check-ins. Touristic users checked-in less than 20,000 times during 18 months.

Table 13. **The classification of residents and tourists of whole dataset**

	Number	Check-ins	Mean of check-ins
All Users	3350	79798	23.8203
Residents	580	60618	104.5138
Tourists	2770	19180	6.924188

Source: Own elaboration

In terms of urban usages, the top five of locals are restaurants, workplaces, outdoor resorts, educational places, and transport. It matches with the results from Marmolejo, C. & Cerda, J. (2012). They pointed out that centers of leisure and education also take a large portion of people's timeline. The corresponding rank of tourists is outdoor resorts, transport, restaurants, hotel, and store. Although it is slightly different from the samples' results, all of them are also belong to the typical activities of tourists (see Table 14).

The difference of POIs usages illustrates the difference between tourists and locals in urban activities. For an instance, it is reasonable that hotel and workplaces have the biggest differences between locals and tourists. It is worth to notice that the conference center is also positive for tourists because Barcelona holds many international conferences every year, such as World Mobile Congress and so on.

Table 14. **The difference of usages of POIs**

Description	Locals	%	Tourists	%	Diff.
Hotel	698	1.15%	1576	8.22%	7.07%
Transport	4724	7.79%	2792	14.56%	6.76%
Outdoor Resorts	5768	9.52%	3079	16.05%	6.54%
Museum, Art, Historical place	907	1.50%	1156	6.03%	4.53%
conference center	333	0.55%	412	2.15%	1.60%
Infrastructure(port)	691	1.14%	443	2.31%	1.17%
Sports center	1358	2.24%	581	3.03%	0.79%
Market	238	0.39%	177	0.92%	0.53%
Touristic Info center	3	0.00%	4	0.02%	0.02%
Store	4428	7.30%	1398	7.29%	-0.02%
others	157	0.26%	24	0.13%	-0.13%
bar	3398	5.61%	1031	5.38%	-0.23%
Shop	2804	4.63%	789	4.11%	-0.51%
café	1763	2.91%	330	1.72%	-1.19%
Opera, concert, cinema	1367	2.26%	186	0.97%	-1.29%
Restaurant	9533	15.73	2726	14.21%	-1.51%
residential place	3237	5.34%	563	2.94%	-2.40%
Services	4234	6.98%	746	3.89%	-3.10%
gym	3161	5.21%	154	0.80%	-4.41%
Education places	4982	8.22%	428	2.23%	-5.99%
work place	6834	11.27%	585	3.05%	-8.22%
Total	**60618**		**19180**		

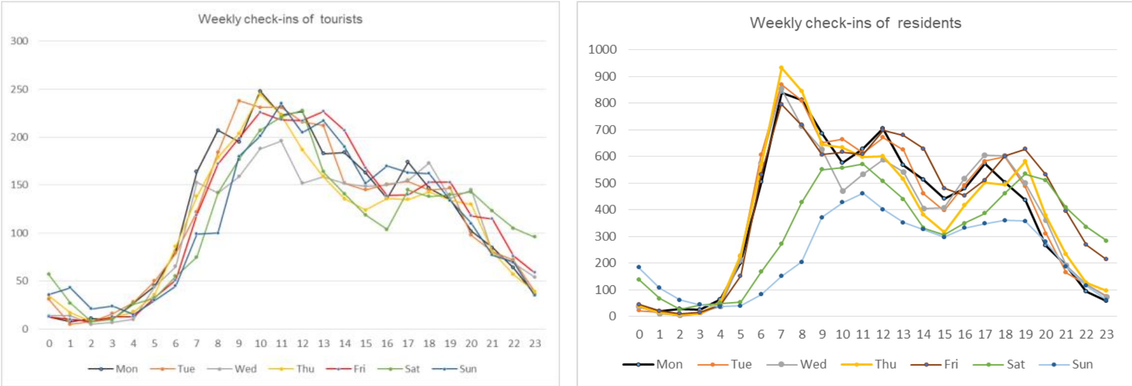
Source: Own elaboration

With regard to the active time (see Figure 8), the daily active degree of tourists is similar from Monday to Sunday. The resident group appears a periodic variation – the active degree of weekdays is higher than the weekends.

Moreover, during weekdays, the rush hour of locals' daily activities starts from 6 o'clock, is earlier than tourists. The rush hours of tourists are from 9:00 to 14:00, then from 17:00 to 20:00. During weekends, the rush hours of locals delay until 9:00.

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

Figure 8. **Temporal difference of Foursquare activities**

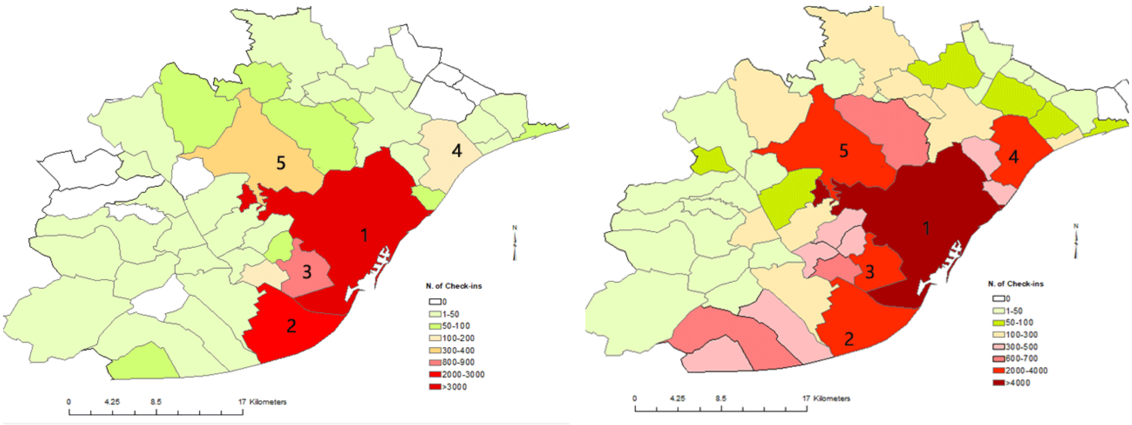


Source: Own elaboration

In terms of the geo-spatial distribution of activities, both groups mainly concentrate in Barcelona city (see Figure 9). Barcelona city and the airport account for 86.6% of all check-ins of tourists. The range of resident activities is larger than tourists. Locals are active in several municipalities nearby in which many companies are located, such as Hospitalet de Llobregat, Prat de Llobregat, Badalona, Sant Cugat del Vallés

Figure 9. **Geo-spatial distribution of check-ins**

(a) Tourists **(b) Locals**



1. Barcelona 2. Prat de Llobregat (Airport) 3. Hospitalet de Llobregat 4. Badalona 5. Sant Cugat del Vallès

Source: self-elaboration

It also depicts the distribution of check-in points by kernel density, which can uncover areas of intensive activities. The locals' active areas actually contain the active area of tourists (see Figure 10). Due to the compact urban model of Barcelona, the central area undertakes multi-functions, such as employment, recreation, and tourism etc. About half of residents of Metropolitan Area of Barcelona live in Barcelona city.

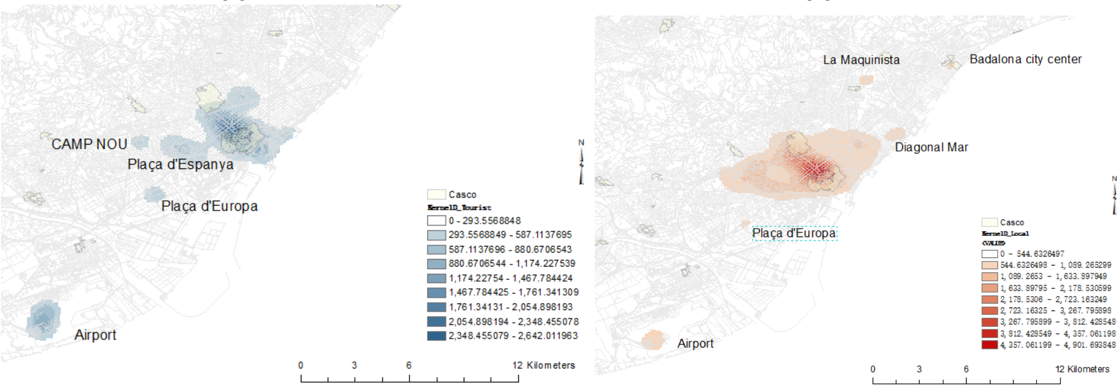
Therefore, both locals and tourists share the city center, though their reasons may be different. Tourists gather in the central area because the historical center and famous tourist attractions are located in this area. Meanwhile, the city center provides employment positions, dwellings and food places for locals.

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

As the density map shows, besides the central area, the dense areas of tourists include CAMP NOU (Barcelona Football Club), Plaça d'Espanya which is one of the most important plazas in Barcelona, and Plaça d'Europa in which many hotels, several commercial centers and the Exhibition and Trade Center are located.

Figure 10. **The Foursquare activities centers of Barcelona**

(a) Tourists **(b) Locals**



Note: Casco means historical places
Source: Own elaboration

The dense area of residents almost covers the whole expanded area of Barcelona, several commercial centers (Plaça d'Europa, La maquinista and Diagonal Mar), and the central area of Badalona where is third largest city in Catalonia.

6. Discussions and conclusions

This paper classifies Foursquare users into tourists and locals, then compares their usages of Foursquare POIs. It shows that the two groups present different characteristics on Foursquare behaviors, regardless of the length of analyzing period. The unsupervised method (k-means clustering) actually can identify users who have the extreme attributes. However, for those users without typical characteristics, the human intervention is unavoidable.

The difference of POI usages of two groups verifies the identification is valid. It reflects the typical land uses of tourists and locals in a city. What's more, the geospatial distribution and active time also reflect that the two groups are different.

The activities of tourists concentrate in the airport and the center of Barcelona city. Locals' activities spread to the nearby cities. The active time of locals is earlier than tourists. Tourists' active period is similar every day. On the contrary, locals show an evident periodic variation daily and weekly.

It is undeniable that there are some limitations in this work. The bias of Foursquare data itself causes that the check-ins concentrate on the category of restaurants because the function of Foursquare is to provide practical information about places for users. What's more, with the decline of popularity degree, Foursquare data tends to shrink in Barcelona. The lack of background information of users also limits the further exploration of the study.

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.

Nevertheless, this study demonstrates that it is possible to identify locals and tourists through Foursquare data, though the uncertainty of data is recognized. How to improve the accuracy of the unsupervised identification and cooperate with other dataset will be the further investigation. Furthermore, whether the identification model can be universally applied is another issue that is worth to test in the future.

Bibliography

ABBAS, O. A. *Comparisons between data Clustering Algorithms*. In: The International Arab Journal of Information Technology [on line] 5 (3): 320-325, 2008. Available in: http://iajit.org/PDF/vol.5,no.3/15-191.pdf

AGRYZKOV, T.; MARTÍ CIRIAQUIÁN, P.; TORTOSA, L. & VICENT, J. F. *Measuring urban activities using Foursquare data and network analysis: a case study of Murcia (Spain).* In: International Journal of Geographical Information Science [on line] 31 (1): 100-121, 2017. DOI: https://doi.org/10.1080/13658816.2016.1188931

CHO, E., MYERS, Seth A. & LESKOVEC, J. *Friendship and mobility: user movement in location-based social networks*. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining [on line] p. 1082-1090, 2011. Available in: https://cs.stanford.edu/people/jure/pubs/mobile-kdd11.pdf

CÓCOLA GANT, A. & PALOU RUBIO, S. *Tourism promotion and urban space in Barcelona: historic perspective and critical review, 1900-1936.* In: Documents d'anàlisi geogràfica [on line] 61 (3): 461-482. 2015. Available in: https://dag.revista.uab.es/article/view/v61-n3-cocola-palou DOI: https://doi.org/10.5565/rev/dag.266

CRANSHAW, J.; SCHWARTZ, R.; HONG, J. I. & SADEH, N. *The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City*. In: Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media [on line] p. 58-65, 2012. Available in: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/download/4682/4967

DA RUGNA, J.; CHAREYRON, G. & BRANCHET, B. *Tourist behavior analysis through geotagged photographies: a method to identify the country of origin.* In: 2012 IEEE 13th International Symposium ON Computational Intelligence and Informatics (CINTI) [on line] p. 347-351. 2012. DOI: https://doi.org/10.1109/CINTI.2012.6496788 Available in: https://ieeexplore.ieee.org/document/6496788

DUARTE, C. M. & TRONCOSO, J. C. *La densidad-tiempo: otra perspectiva de análisis de la estructura metropolitana.* In: Scripta Nova: revista electrónica de geografía y ciencias sociales [on line] XVI, 402. 2012. Available in: http://www.ub.edu/geocrit/sn/sn-402.htm

GAO, Q.; Abel, F.; Houben, G.-J. & Yu, Y. *A comparative study of users' microblogging behavior on Sina Weibo and Twitter.* In: International Conference on User Modeling, Adaptation, and Personalization, 2012, pp. 88-101. Masthoff J., Mobasher B., Desmarais M.C., Nkambou R. (eds) User Modeling, Adaptation, and Personalization. UMAP 2012. Lecture Notes in Computer

Science, v. 7379. Available in: https://link.springer.com/chapter/10.1007/978-3-642-31454-4_8 DOI: https://doi.org/10.1007/978-3-642-31454-4_8

GARCÍA-PALOMARES, J. C.; GUTIÉRREZ, J. & MÍNGUEZ, C. *Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS*. In: Applied Geography [on line] 63, 408-417. September 2015. Available in: https://www.sciencedirect.com/science/article/abs/pii/S0143622815001952 DOI: https://doi.org/10.1016/j.apgeog.2015.08.002

GIRARDIN, F.; FIORE, F. D.; RATTI, C. & BLAT, J. *Leveraging explicitly disclosed location information to understand tourist dynamics: a case study*. In: Journal of Location Based Services [on line] 2 (1): 41-56. 2008. DOI: https://doi.org/10.1080/17489720802261138 Available in: https://www.tandfonline.com/action/showCitFormats?doi=10.1080%2F17489720802261138

GONZALEZ, M. C.; HIDALGO, C. A. & BARABASI, A. L. *Understanding individual human mobility patterns*. In: Nature International Journal of science, *453* (7196): 779-782. 2008. Available in: https://www.nature.com/articles/nature06958

HASAN, S.; ZHAN, X. & UKKUSURI, S. V. *Understanding urban human activity and mobility patterns using large-scale location-based data from online social media*. In: UrbComp '13 Proceedings of the 2nd ACM SIGKDD international workshop on urban computing, 2013. Art. 6. DOI: https://doi.org/10.1145/2505821.2505823

HAYLLAR, B.; GRIFFIN, T. & EDWARDS, D. *City Spaces-Tourist Places*. London, Routledge, 2010. 400 p.

JUE, J. & XIAOLU, G. *Identifying the scope of daily life in urban areas based on residents' travel behaviors*. In: Progress in Geography [on line] 31 (2): 248-254. 2012. Available in: http://www.progressingeography.com/EN/article/downloadArticleFile.do?attachType=PDF&id=13281 DOI: https://doi.org/10.11820/dlkxjz.2012.02.014

KÁDÁR, B. *Measuring tourist activities in cities using geotagged photography*. In: Tourism Geographies, An International Journal of Tourism Space, Place and Environment [on line] 16 (1): 88-104, 2014. DOI: https://doi.org/10.1080/14616688.2013.868029 Available in: https://www.tandfonline.com/doi/abs/10.1080/14616688.2013.868029

LUO, F.; CAO, G.; MULLIGAN, K. & LI, X. *Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago*. In: Applied Geography, 70, 11-25, May 2016. DOI: https://doi.org/10.1016/j.apgeog.2016.03.001 Available in: https://www.sciencedirect.com/science/article/abs/pii/S0143622816300194

MCKERCHER, B. & LAU, G. *Movement patterns of tourists within a destination.* In: Tourism geographies, An International Journal of Tourism Space, Place and Environment [on line] 10 (3): 355-374, July 2008. DOI: https://doi.org/10.1080/14616680802236352 Available in: https://www.tandfonline.com/doi/abs/10.1080/14616680802236352

NOULAS, A.; SCELLATO, S.; MASCOLO, C. & PONTIL, M. *Exploiting Semantic Annotations for Clustering Geographic Areas and Users in Location-based Social Networks*. In: The social mobile web, Association for the Advancement of Artificial Intelligence. [on line] 11 (2), 2011. Available in: https://www.cl.cam.ac.uk/~cm542/papers/SMW11.pdf

PONTES, T.; VASCONCELOS, M.; ALMEIDA, J.; KUMARAGURU, P. & ALMEIDA, V. *We know where you live: privacy characterization of foursquare behavior*. In: Proceedings of the 2012 ACM conference on ubiquitous computing, UbiComp '12, Sep 5-Sep 8, 2012, Pittsburgh, USA2012 [on line] 898-905. DOI: https://doi.org/10.1145/2370216.2370419 Available in: http://lbsn2012.cmuchimps.org/papers/Paper12_Pontes.pdf

SILVA, T. H.; VAZ DE MELO, P. O.; ALMEIDA, J. M.; SALLES, J. & LOUREIRO, A. A. *A comparison of foursquare and instagram to the study of city dynamics and urban social behavior.* In: UrbComp '13 Proceedings of the 2nd ACM SIGKDD international workshop on urban computing, 2013. Art. 4. DOI: https://doi.org/10.1145/2505821.2505836

SUN, Y. *Investigating "Locality" of Intra-Urban Spatial Interactions in New York City Using Foursquare Data*. In: ISPRS International Journal of Geo-Information, 2016, 5 (4): 43. DOI: https://doi.org/10.3390/ijgi5040043

VU, H. Q.; LI, G.; LAW, R. & YE, B. H. *Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos*. In: Tourism Management [on line] 46, 222-232. February 2015. Available in: https://www.sciencedirect.com/science/article/pii/S0261517714001356 DOI: https://doi.org/10.1016/j.tourman.2014.07.003

Citación: YANG, L. *et al. Identifying the urban space for locals and tourists through "Foursquare" data in Barcelona.* En: Libro de proceedings, CTV 2018. XII Congreso Internacional Ciudad y Territorio Virtual. "Ciudades y Territorios Inteligentes". UNCuyo, Mendoza, 5-7 septiembre 2018. Barcelona: CPSV, 2018, p. 143-164.