

Article

Passive Mobile Data for Studying Seasonal Tourism Mobilities: An Application in a Mediterranean Coastal Destination

Benito Zaragoza¹ , Sergio Trilles²  and Aaron Gutiérrez^{1*} 

¹ Departament de Geografia, Universitat Rovira i Virgili, c/ Joanot Martorell, 43480 Vila-seca, Spain; benito.zaragozi@urv.cat

² Institute of New Imaging Technologies (INIT), Universitat Jaume I, Av. Vicente Sos Baynat s/n, 12071 Castelló de la Plana, Spain; strilles@uji.es

* Correspondence: aaron.gutierrez@urv.cat

Abstract: The article uses passive mobile data to analyse the complex mobilities that occur in a coastal region characterised by seasonal patterns of tourism activity. A large volume of data generated by mobile phone users has been selected and processed to subsequently display the information in the form of visualisations that are useful for transport and tourism research, policy, and practice. More specifically, the analysis consisted of four steps: (1) a dataset containing records for four days—two on summer days and two in winter—was selected, (2) these were aggregated spatially, temporally, and differentiating trips by local residents, national tourists, and international tourists, (3) origin-destination matrices were built, and (4) graph-based visualisations were created to provide evidence on the nature of the mobilities affecting the study area. The results of our work provide new evidence of how the analysis of passive mobile data can be useful to study the effects of tourism seasonality in local mobility patterns.

Keywords: passive mobile data; tourism mobilities; tourism seasonality; urban mobilities; data visualisation



Citation: Zaragoza, B.; Trilles, S.; Gutiérrez, A. Passive Mobile Data for Studying Seasonal Tourism Mobilities: An Application in a Mediterranean Coastal Destination. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 98. <https://doi.org/10.3390/ijgi10020098>

Academic Editor: Wolfgang Kainz

Received: 12 December 2020

Accepted: 19 February 2021

Published: 22 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

There is a growing body of literature on the multiple applications of big data for tourism studies [1]. Studies at different scales can be distinguished, ranging from those that refer to the study of tourists individually to other studies that analyse tourist destinations as a whole [2,3]. Obviously, each type of study needs data with different characteristics. For example, in order to study the *tourist experience*, it is already possible to gather a large volume of “active” information from a significant number of tourists through different platforms, which turns out to be a valuable complement to the studies based on surveys and other traditional techniques. While this type of analyses is known to be useful, other data sources, such as activity records for credit card transactions or data collected from Mobile Network Operators (MNO), have another type of value, as they collect “passive” data from *virtually* all tourists.

Passive Mobile Data (PMD) allows the creation of high-resolution spatio-temporal empirical models of human mobility [4]. Moreover, the characteristics of these data make it possible to differentiate the visitor flows from the local population in any area under study. These data provide a more comprehensive image of tourists’ mobility than traditional individual surveys or GPS tracking. For these reasons, PMD is increasingly being used to analyse tourists’ mobility patterns and their spatial behaviour at a destination [5]. The main limitation of this data source, however, is the lack of individual sociodemographic information about the travellers.

A key potentiality of PMD is the possibility of reproducing analyses on different temporal and spatial scales. In the case of tourism studies, it allows measurement of the effect of specific events on tourist flows. One of these possibilities is the study of the effect

of the seasonality of tourism on the overall mobility in a city or a region. The seasonal nature of tourism activity has deep implications for destinations that need to plan the provisions of services and resources, such as public transportation, for both the peak and the low season [6]. Detailed information on the spatial and temporal variability of flows is of special interest for both the transportation and the tourism sectors.

Taking this context into account, the main objective of our article is to propose and implement a methodology to measure the seasonality of tourism mobility in a mature coastal destination using passive mobile phone data. More specifically, the methodological objectives of the article are the following:

1. To confirm the potential value of mobile phone data for the study of the multiple geographical and temporal scales of seasonal events, such as tourism mobility in cities and regions.
2. To implement an approach that allows the visualisation of these multiple geographical and temporal changes using graph-based representations.

In addition to these methodological objectives, our study aims to achieve the following specific research objectives:

1. To measure the effect of tourism seasonality on the reconfiguration of overall intra-urban and inter-urban mobility.
2. To identify the differentiated mobility patterns of tourists and locals.
3. To verify how diverse intensities in tourism activity within the different areas of the city involve different mobility patterns and timings.

A case study conducted on the Costa Daurada (Catalonia, Spain) is used as a typical example of a destination shaped by tourism seasonality. More specifically, the municipality of Salou, the main destination on the Costa Daurada, was taken to analyse the inter-urban and intra-urban movements. To the best of the authors' knowledge, this is the first work that explicitly addresses the analysis of tourism seasonality by creating origin-destination visualisations from PMD (see Section 2 for a literature review and Section 3 for more details on the methodological approach). Another consideration in this work is to provide results of a big data analysis that could be interpreted by a greater number of people. For this purpose, our approach is based on producing clear data visualisations. This will help in the transfer of information to stakeholders, enrich decision-making processes and generate more value from the data. The second novelty of the work is the multi-scale approach implemented. It combines the representation of trips made within the city under study (intra-urban) and the flows that take place in surrounding cities (inter-urban).

The rest of this article is structured as follows. Section 2 outlines the relevant research work related to analysing tourists' mobilities using PMD. Section 3 introduces the selected case study and presents the methodology defined to carry out the PMD analysis. Section 4 shows the outcomes using two different kinds of visualisation. The paper culminates in Section 5 with the discussion and final concluding remarks.

2. Background

2.1. Mobile Phone Data

Over the past two decades, mobile phones have achieved a high rate of penetration in society [7]. This increased presence has led to millions of phones each producing a fingerprint in the shape of its own particular list of events based on the user interaction, such as calls, Short Message Service (SMS), or 3G/4G connections. Each event is associated with the location of the closest cell tower and is stored automatically by each mobile network operator. This process goes completely unnoticed by the user and is commonly called "passive mobile data" or Call Detail Records (CDR). The main advantage of this type of data compared to other ways of monitoring human mobility, such as social networks or direct surveys, is that mobile data have an intrinsic geospatial component, only limited by the density of the network of cell towers. The analysis of mobility data from social networks and mobile phone data can give similar results in some urban contexts and in

certain types of analysis, such as the analysis of densities or the temporal distribution of the population. However, it has been found that using social network data can present serious problems if, instead of analysing recurrent mobility, the focus is on shorter term mobility [8].

CDR data are sampled unevenly over time so that logs can be very sparse for users with a lower level of activity. This can lead to biased estimates when applied across the board [9]. One solution to this problem is the use of Mobile Signalling Data (MSD) [10]. This kind of data can offer a more detailed view of the footprints of human mobility, especially as regards the temporal aspect. MSD allows the activity of the mobile phone to be collected more constantly thanks to events launched by the telecommunications systems (e.g., cellhandovers, SMSs, or opening of sporadic data connections to perform some background activity, among others) [11]. This feature means that MSD can add more detail about the behaviour of users longitudinally in time and obtain greater granularity than CDR. MSD have been seen by the scientific community as more attractive for carrying out mobility studies.

In the same way as the CDR data, a problem associated with MSD is the “ping-pong” effect, which occurs when there is an oscillation from one cell tower to another. This effect is produced by the transfer of the user’s phone to nearby cell towers due to operations of the telecommunications systems or load balancing. Incorrect movements are produced, often continuously (bounces), without the user having moved. Wu et al. [12] proposed a possible solution to this problem.

As noted above, mobile data are collected at a cell towers level (Figure 1 will be introduced in the next section). Consequently, cell tower density has a significant effect on positioning accuracy. At the same time, the distribution of cell towers is not uniform throughout the territory, which gives rise to a difference in terms of accuracy depending on the area to be analysed. For example, in urban areas, the density of cell towers is much higher than in rural areas, and this causes a considerable loss of accuracy in location. The distribution of these towers defines a grid of cells and these cells are used to account for the activity within them. The activity is recorded with the unique identity of each mobile phone detected by the cell tower belonging to the cell. For example, during the day, users will have their activity linked to the cell tower overlapping their work place, and, at night, the activity will be linked to the cell tower overlapping their home.

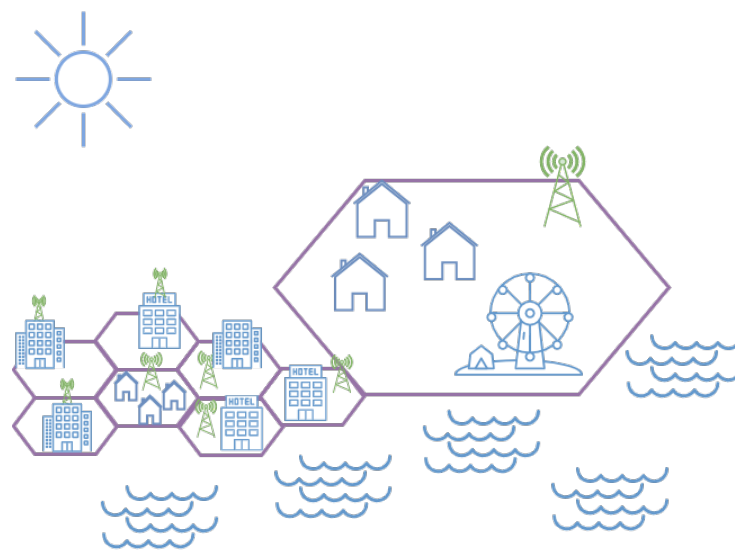


Figure 1. Urban, rural and isolated cells in a mobile phone network. Voronoi polygons are commonly used to build these cells, but those geometries are replaced by hexagons to emphasise the concept of cell.

2.2. Tourism Mobility Studies Based on Mobile Phone Data

Although the use and analysis of this type of data are still relatively novel and involve a high economic cost to acquire these data [13], they have been used in a large number of mobility studies [14]. It should be noted that, apart from tourism studies, some attempts have been made to use PMD to investigate other geographical issues, such as generational differences in spatial mobility [15], internal migration [16], estimation of literacy rates [17], measuring ethnic segregation [18,19], mapping changes in residence [20], tracking population movements after disasters [21], or, more recently, how PMD can be used to inform different aspects of COVID-19 response [22].

More specifically, there are also studies that use PMD to research tourist mobility. One of the pioneering studies in this respect was presented by Ahas et al. [23]. In that work, the authors analysed the main patterns within the seasonal movement of tourists at the regional level in Estonia. The results yielded a strong seasonality in coastal areas during the summer and inland areas during the winter season. In Reference [24], the same authors analysed the “first call” made by tourists in Estonia to detect the main point of entry into the country, and this enabled them to detect some divergence between the patterns of Latvian and Russian tourists.

Raun et al. [25] analysed a three-year dataset of foreign visitors in Estonia. They focused on five dimensions (spatial, temporal, compositional, social, and dynamic) and analysed the inter- and intra-destination movements at the national level. The results showed that smaller destination areas can be differentiated inside the whole country by the geographical, temporal and compositional parameters of the visits. In a more recent work [26], the same authors showed the impact of significant gateways on national tourism flows by using passive mobile positioning and GPS data in Estonia and Israel. In both cases, most of the tourists spent their time visiting areas close to the gateway, and a dramatic decrease was seen in visitation to areas that were some distance away from the gateway location.

In Reference [27], mobile data were used to complement and validate traditional face-to-face surveys about tourism carried out in Saudi Arabia. The results highlighted the popular destinations for domestic tourists and how nearby cities received visitors from the most popular destinations.

Chen et al. [28] used PMD to predict patterns of tourists’ travel behaviour. Using one month’s mobile data from Andorra, the authors applied different prediction algorithms, and, in the best of cases, prediction of future short stays was obtained with a success rate of 94.8% using Long Short Term Memory (LSTM). With the same dataset, in Reference [29], the same authors analysed the Spanish and French tourist patterns in Andorra. They detected that, depending on the country of residence (Spain or France), tourists travelled more frequently to specific towns in Andorra. Using the same dataset, another study evaluated different marketing strategies in tourism, by including tourists’ experiences and evaluating the impact of touristic events [30].

In Reference [31], the author developed a platform to process CDR from mobile users in Hainan (China) in order to identify the different tourist behaviours. Similar to previous studies, in Reference [32], the authors designed a framework to discover tourist groups using the same dataset from Hainan. Chen et al. [33] presented an analysis framework for discovering valuable travel patterns of tourists. They defined a three-layer architecture (data, algorithm and application layer). A testbed in Hainan (China) has been proposed to validate the defined framework. A new study using mobile data (CDR) from Beijing showed that big data analysis methods can determine where the main tourist attractions are and draw tourist routes in a city [34]. The same work stated that data from mobile phones can provide real-time information about tourist behaviours in a timely and effective manner.

PMD have been used to study different aspects of domestic tourism trips in France using a dataset of 18 million users over 154 consecutive days [35]. The analysis extracted 18,380 domestic tourism trips in 32 of the biggest cities in France. A classification method using mobile data was defined to determine different tourism-related profiles in Italy [36].

The authors indicated that residents and long-term commuters are difficult to define with a dataset covering only one month. Tourist profiles, however, can be classified easily in such a short period, especially in the case of tourists who are continuously on the move.

2.3. Main Contributions of This Work

The use of PMD to analyse the mobility of people is not a new topic in itself, but it is true that these data present a series of limitations that may be important according to each case study. In this sense, it is important to propose new applications and determine the pros and cons of PMD in each context.

If we compare the studies reviewed in this section with the main goals of the current research, we can observe that some of them [23–26,33] analysed seasonality in a general context. Regarding inter- and intra-urban movements, only Reference [25,26] included these types of movements at a national level. The current work determines these inter- and intra-urban movements at a local scale, taking into account the mobility implications caused by seasonal reconfiguration, the differences between mobility according to the origin of the visitors, and how the density of tourist resources in an area conditions the flows in its surroundings.

Finally, concerning the way to visualise the results, most of the studies have used simple graphs or thematic maps to describe spatial relationships or population densities. Only Reference [34,35] added more innovative visualisations—some of them specific for PMD—that help to represent the flows of movement and the distribution of tourists in a more understandable way. However, Xu et al. [37] suggested that network science and its outputs enable tourism researchers to identify better insights of tourist flow. Following this idea, we propose the use of different graph-based visualisations to analyse seasonality effects on tourist destination mobilities. These representations should be more readable to researchers, but also to important stakeholders (e.g., local governments or tourism organisations).

3. Methodology

3.1. Case Study

According to the premise of this article, PMD could be analysed using big data techniques and data visualisations to study very particular mobility patterns that take place in some tourist destinations. In this case, we are referring to the strong irregularities in mobility that occur between seasons and time-slots, but which also translate into a greater spatial concentration of visitors around the main tourist services and attractions. To evaluate such patterns, we focused on the tourism mobilities in Salou (see Figure 2).

Salou is a relatively small town in the coastal area of the province of Tarragona (Catalonia, Spain). According to official data for 2019, Salou has a population of 27,476 inhabitants. Together with Cambrils and Vila-seca (33,898 and 22,187 inhabitants, respectively), it makes up the central Costa Daurada, one of the major tourist destinations in Spain. Salou is the main tourist zone in the area, both in number of hotel rooms and number of national and international visitors per year. During 2018, Cambrils, Salou, and Vila-seca received a total of 2.8 million tourists, who spent 11.9 million overnight. These figures account for 51.2% of the total number of tourists and 59% of the total number of overnight stays on the whole Costa Daurada. Tarragona and Reus, the two main cities in the region (134,515 and 104,373 inhabitants, respectively) are located less than ten kilometres from Salou.

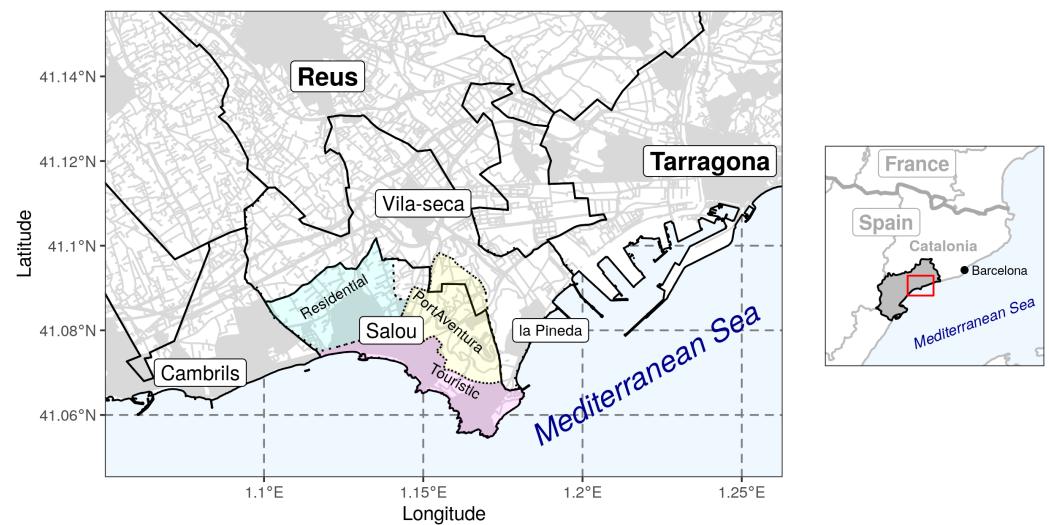


Figure 2. Location map of the study area. The main map shows the principal municipalities in the area and the proposed zoning of Salou (residential, PortAventura, and tourist areas). The reference map shows the study area in context within the province of Tarragona, Catalonia, and Spain.

The main attraction of the central Costa Daurada lies in its sun-and-beach tourism. Except in the case of Reus, the other three towns and Tarragona are on the coast and boast good quality beaches. In the case of Vila-seca, the main urban area is situated somewhat further away from the coast, but it is in the secondary nucleus of la Pineda where tourist activity is mainly concentrated. Another important tourist resource is PortAventura World, a theme park that exceeded 5.2 million visits in 2019. It should be noted that in 2019 PortAventura also set a record for visits in the Halloween (more than 900,000 visits) and Christmas (more than 500,000) campaigns. Finally, Tarragona and Reus also attract urban, cultural, and leisure tourism.

Regarding the intensity of the presence of the tourism sector, Salou can be separated into three very different areas: (1) the PortAventura area and the associated tourist resort, (2) the tourist area next to the beach, where the hotels and most of the accommodation, restaurants and leisure offer are located, and (3) the residential area, where these activities have an almost anecdotal presence.

The mobility of tourists in this area has previously been studied from different perspectives using other data sources and offering a very specific point of view. There are several studies that used public transport data collected by an automated fare collection system to understand topics, such as the impact of tourism seasonality on the bus transport network [38,39], to determine the spatial coverage of public transport networks and their spatial dependence with tourism accommodation establishments [40] and to identify groups of public transport users based on their trip patterns [6]. Another study combined the use of public transport data and surveys in order to determine the tourist profiles that preferred the use of public transport over private cars [41]. Other studies collected data by using GPS loggers to track cruise tourists in order to understand their spatial behaviours in the city of Tarragona [42,43]. In all these studies, there were tourism mobility behaviours and patterns that were known to be of an important magnitude but can vary very quickly in time and space. In such a volatile context, having quality information and being able to analyse it with agility can be essential in many circumstances.

3.2. Data

In this work, we used PMD, which combined the two record typologies described in the background section (CDR and MSD). In this way, each record provided spatio-temporal information about the phone each time it interacted with the network, both for

active events from the CDR (e.g., calls, SMS, data connections) and for certain passive events from the MSD (e.g., changes in the coverage area, network updates, among others). These data have a high spatio-temporal granularity, which is normally higher in more populated urban areas than in sparser or rural areas. The data provided by some operators also included sociodemographic information linked to users, such as age and gender, but we only accessed anonymised and aggregated records. All the mobile data records from Spanish residents came from the Orange network—Orange Espagne S.A.U., which is the Spanish subsidiary of the French multinational company Orange—while the data for foreigners came from mobile operators in different countries (Orange partners). The pre-processing of these data has to be performed by Nommon Solutions and Technologies, which is a technological partner of Orange Espagne that has a collaboration agreement, allowing them to access data directly from the Orange network.

Table 1 shows the number of Orange users in Spain for different age groups and the percentage of the total population that they represent. The structure of the sample of Orange users is similar to the structure of the population in Spain, with a clear predominance of the intermediate and high age groups, so this can be considered a good sample of the total population. These data can be used to study two of the profiles that we want to analyse, the residents in a particular area and the trips by national visitors to that area.

Table 1. Percentage of sample by age group. Sources: “Padrón Municipal de Habitantes” (2019) and Orange Spain (2019).

Age Group	Total Population	Orange Users	Percentage
10–19	4,813,817	683,995	14.2%
20–44	15,238,220	5,489,663	36.0%
45–65	13,522,821	4,045,965	29.9%
Over 65	9,057,193	1,282,015	14.2%

Similarly, only a sample of international visitors is directly available. Table 2 shows the comparison between the total number of tourists and visitors in Spain during November 2019 and the size of the effective sample of foreign tourists and visitors [44]. These percentages are also considered an adequate sample of the total number of visitors by country of origin.

Table 2. Percentage of foreign tourists and visitors by nationality in the sample (November 2019). Sources: Orange Spain and National Statistics Institute of Spain.

Origin	Total Visitors	Roaming Users	Percentage
France	2,060,689	383,927	18.63%
United Kingdom	982,514	94,568	9.63%
Germany	641,242	30,678	4.78%
Portugal	738,530	97,696	13.23%
Rest of the world	3,166,237	333,437	10.53%

Considering the proposed research questions and the already mentioned cost of accessing and processing these data, the population movements selected for the study area were those that occurred between the municipalities of Tarragona, Reus, Cambrils, Salou, and Vila-seca, with a higher level of disaggregation for Salou, where the three main areas of this important tourist destination were distinguished (residential, tourist areas, and PortAventura). Origin-Destination (OD) matrices were required on four representative days, distinguishing between mobility in summer and winter and whether it was a weekday or the weekend. The days chosen to represent the typical activity in summer were Wednesday, 8 August 2018, and Saturday, 11 August 2018. Additionally, Wednesday, 23 January 2019, and Saturday, 26 January 2019 were the representative days of the winter period. Four

different time ranges were considered: early morning (06.00–10.59), midday (11.00–15.59), afternoon (16.00–21.59), and night (22.00–05.59). Finally, it was important to distinguish whether the movements were carried out by (1) users residing in the study area and its vicinity (Tarragona province; national and living in the area for at least three weeks), (2) national visitors (Orange clients or extrapolated estimations), or (3) users residing abroad (roaming users).

Today, the number of PMD records generated for Spain is about 1.5 billion a day. It should be noted that, for any study, the entire dataset has to be analysed, even if only a few indicators for a specific group of users are to be calculated. For example, a minimum of 3 weeks of records are usually analysed to observe longitudinal patterns and infer attributes, such as the place of residence or type of activity performed (i.e., business or tourism). Therefore, it can be said that the volume of data used in the study is around 315,000 or 350,000 million records, before filtering and selecting the necessary information. For processing this volume of data, Nommon Solutions use servers from Amazon Web Services (AWS), in which characteristics depend on the needs of a particular project. In their analyses, they use proprietary software developed mainly in Python and Cython programming languages.

Nommon Solutions ensure PMD data quality and usefulness by processing raw data to obtain an enriched dataset tailored to analyse the case study population's behaviours (see Figure 3). This process can be summarised in the following stages:

1. Pre-processing and cleaning. Based on the mobile operator's anonymised data (Table 3), a pre-processing stage is carried out to facilitate the management, ordering and grouping of the data for later stages. In this same stage, data cleaning and quality control tasks are also performed. These tasks are applied to anonymised mobile data records and user profile data.

Figure 1 shows a representation of a possible scenario for analysing PMD. In the vast majority of studies, the location of the mobile phone is based on the locations of the cell towers of the mobile phone network used (see Table 3). A Voronoi tessellation is constructed based on the cell tower locations and is used to determine the position of each user in a specific moment; in this way, each polygon defines the maximum level of granularity. The Voronoi polygon layer is used to zone and extrapolate the user activity in each cell, depending on other datasets, such as sociodemographic data or land use and cover from the Spanish Land Occupation Information System (SIOSE). We did not have access to the location of the antennas nor the Voronoi tessellation, only to the trips already aggregated at the municipal or district level.

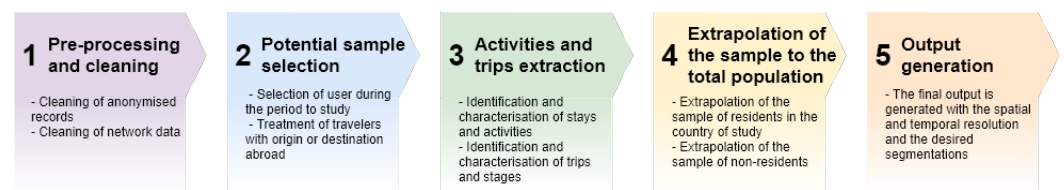


Figure 3. Workflow used by Nommon Solutions to process mobile data.

Table 3. Examples of mobile operator's anonymised data.

Device ID	Cell Tower ID	Date	Time
125976C2BB0F73BDD81888E	2140307102	20190126	161438
322405D3AC0F22ABB72352B	2150335289	20190126	181436
228852A3DB0D33CAE15688E	2140307102	20190811	192361
522345B6BD1A83BBB31932B	2150335289	20190811	202513
...

2. Potential sample selection. The next step is to select the users whose mobile activity provides correct information in the study area. The main objective of this stage is to build a sample that involves a compromise between quantity and quality. Some indicators to consider are the number of records generated and the time lapses between them. Special treatment is required in the case of foreign visitors, since their period of activity is limited to a few specific days, and their activity starts and ends mainly at airports.
3. Activity and trip extraction. At this stage, activity and travel indicators are generated from mobile data records. This stage is divided into two sub-tasks: (1) the identification and characterisation of stays and activities, and (2) the identification and characterisation of travel and stages. The first is responsible for analysing consecutive records in the same mobile cell, and they are identified as stays when they remain for longer than a particular threshold (e.g., an average of 30 min, but it depends on the land use and other variable considerations). The stays that correspond to activities are then identified. The second task is responsible for defining the trips made using the activities and stages detected in the previous task. For each trip, a list of features are defined: a destination, start and end time, a mode of transport and a route. A distinction is drawn between medium- and long-distance trips and short-distance movements in urban environments. For long-distance trips map-matching and the average velocity of the trip is enough to determine the transport mode. For shorter trips, data fusion techniques are also necessary to specify a mode of transport (e.g., transport networks, airports, bus stations, surveys, land use).
4. Extrapolation of the sample to the total population. The fourth stage is responsible for extrapolating the selected sample to the total population. Two different types are considered: residents in the country under study and foreigners. For the first of them, different factors are used, such as the census section, and, for foreigners, the total number of official visitors segmented by nationality, type of visitor, and entry point are used. The mobility of international tourists and other visitors is analysed from the data on users who are roaming on the Orange network. The analysis of the roaming data offers the following peculiarities: (1) Since they are not customers of Orange Spain, the only sociodemographic information available is the nationality of the mobile operator of each user, so it is taken as a proxy of their nationality; (2) roaming users can connect to different networks during their stay, so the percentage of users present in the sample over the total number of foreign visitors is higher than the market share of the operator; and (3) many of these users generate an insufficient number of records to analyse their mobility, so these data are usually excluded from the analysis.
5. Output generation. The last stage is responsible for generating the final dataset, adding different spatial and temporal resolutions and the segmentation needed to visualise the behaviour of the selected population. The results of this stage are a collection of OD matrices with different naming conventions that we merged and adapted for the proposed visual analysis (see Table 4).

Table 4. Extract from the already pre-processed OD matrix.

ID	From	To	Time_range	Residence	Date	Trips
1	1	1	06:00–11:00	Locals	2018-08-08	11,308
2	1	1	06:00–11:00	Locals	2018-08-11	9132
3	1	1	06:00–11:00	Locals	2019-01-23	9004
4	1	1	06:00–11:00	Locals	2019-01-26	7158
5
6272	7	7	22:00–06:00	Tourists-Spaniards	2019-01-26	87

3.3. Data Visualisation

The pre-processed OD matrix is a relatively small table that could be parsed using standard tools. However, there are still too many relationships to extract interesting patterns from the data without the help of special techniques, such as proper data visualisation. For example, it was necessary to distinguish the moments and/or places of maximum congestion, when inter-urban mobility or intra-urban mobility predominated, or how the main poles of attraction of PortAventura and the tourist area (near the beach) acted in different time-slots and seasons, among other interesting patterns of mobility.

In Section 2, we have seen many examples of research that have used PMD to study tourism mobilities. Among these studies, very few paid more attention to the visualisation phase, and usually common charts and maps were drawn. In this paper, we consider data visualisation an important stage. The seasonality patterns on different scales (intra-urban and inter-urban) can be shown with a variety of graph-based visualisations that make good use of different visual variables (e.g., colour, size, direction). After testing with different types of charts, we think that two visualisations were the most appropriate for this case study: flow maps and alluvial diagrams. While different types of flow maps are more common, alluvial diagrams are a specific type of parallel axis plots that have not been used in the literature reviewed. However, we consider that these visualisations are useful for showing relationships of directional and qualitative data [6,45]. In this paper, we have used the R platform and several R packages for data management and data visualisation purposes. The description of these methods is detailed in the R scripts shared in their own Git repository (<https://github.com/gratet/pmd-seasonality-visualisations>-accessed on 21 February 2021). Together with Python, the R platform is a powerful framework for data science and data visualisation [46], which provides access to all the necessary resources for this stage of the research.

4. Results

The main OD matrix extracted from the data pool still had 6272 rows, which is too much information to assimilate without finding a good way to present it. The generation of suitable visualisations is usually of great interest in these cases. However, not all charts or maps work in the same way, and each question must be solved in a particular manner. In this section, we first explain a few details of the selected visualisations and then we use these figures to address the proposed research questions.

4.1. Visualising Tourism Mobilities Using Mobility Data

Faceted visualisations can be especially useful to compare different dates and time-slots. We have developed several R scripts to create a variety of representations. Among all the figures created, the clearest and most useful ones have been included here to answer the proposed study objectives. As mentioned above, figures of two types, flow maps (spatial graphs) and alluvial diagrams, have been used.

Figures 4–6 are flow maps showing the mobility in the three zones of Salou—PortAventura, residential, and tourist—and between these areas and the most important municipalities nearby (Tarragona, Reus, Cambrils and Vila-seca). These figures are spatial graphs. In other words, they are graphs in which nodes are located in a geographical space of coordinates. Nodes, represented as circles, are located at the centroid of the main urban area in each municipality or zone. In almost all cases, this location coincides with the main urban area. However, in the case of Vila-seca, in summer, the bulk of the population is concentrated mainly in *la Pineda* [47], which is the part of the municipality closest to the sea (see Figure 2).

These nodes are sized according to the resident population. The edges show the number of trips, and their thickness (size) is scaled so that they do not become overlapped. Their directionality, represented by an arrowhead, indicates whether it is a round trip, so that they are all represented at the same time. In general, lines running in opposite directions have been represented in parallel, whenever possible, so as not to interfere with

readability. In other cases, lines with sufficient curvature have been used to show the most distant displacements more clearly. Self-loop edges have been used to show the internal mobility of each of the areas of Salou. In both cases, edges and nodes, the same colour has been used to show the municipality or area where each set of trips begin. Finally, in order to be able to display the graphs correctly on a cartographic basis, we decided to make facets by dates and not to disaggregate by time-slots. The resulting figures shed a considerable amount of light on the mobility patterns of each population group. Figure 7 shows a synthesis of all the previous flow maps where the three selected profiles (locals, national visitors, and international tourists) can be compared in relative terms. This figure measures the absolute difference of trips between summer and winter days (summer minus winter trips).

Figures 8 and 9 are faceted alluvial diagrams that show the information using almost the same visual variables. Unlike the previous flow maps, in this case, the alluvial diagrams are focused on the internal trips between the three zones of Salou and trips from sub-areas of Salou to the other municipalities. This means that the trips from other municipalities to Salou are not shown, since the main advantages of alluvial diagrams, such as the clustering of attributes, would be lost. In an alluvial diagram, blocks represent clusters of nodes, and stream fields between the blocks represent changes in the composition of these clusters over time. The height of a block represents the size of the cluster, and the height of a stream field represents the size of the components contained in the two blocks connected by the stream field.

Some basic vocabulary is necessary to describe these plots. Essentially, alluvial diagrams are arranged in axis, alluvium, stratum, lode, and flow. In the alluvial diagrams in this work, two axes showing the origin and destination of trips are used, and alluvium represents the groups of clustered trips. Each colour shows trips by place of residence (local, Spaniards, and foreigners).

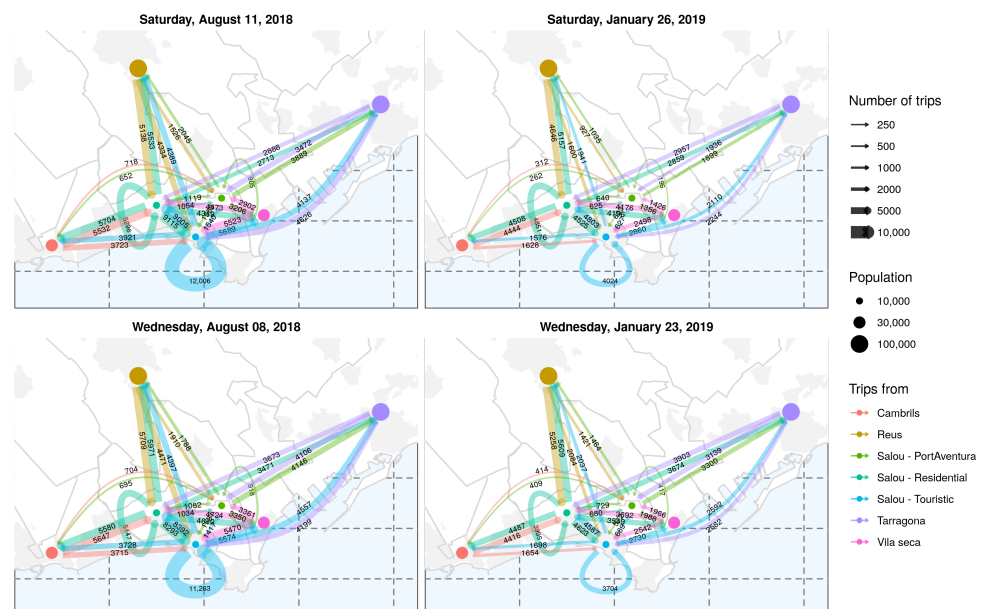


Figure 4. Flow maps of the intra-urban and inter-urban mobility in Salou of the residents in the province of Tarragona.

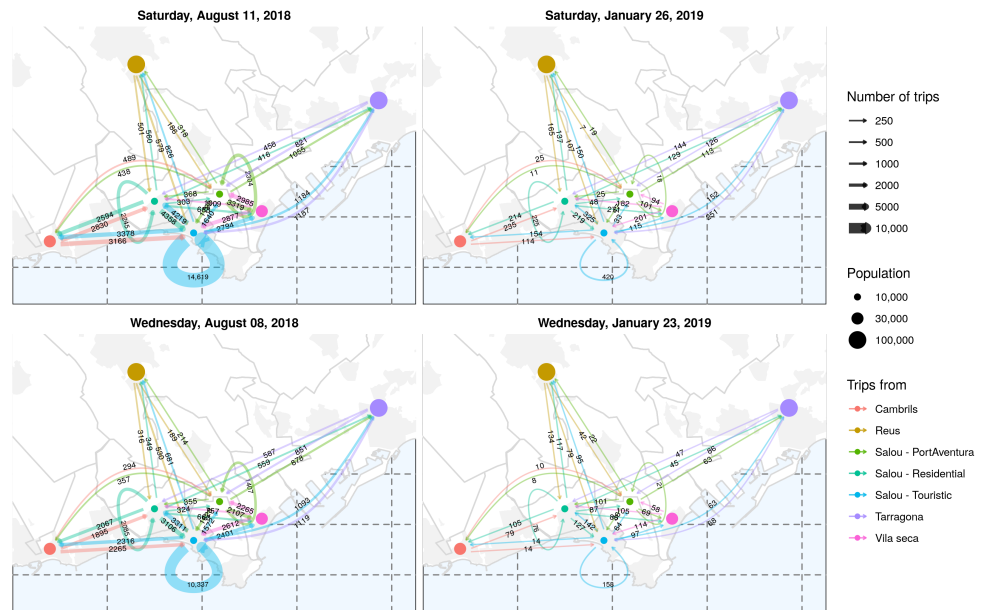


Figure 5. Flow maps of the intra-urban and inter-urban mobility in Salou of the Spanish tourists (from outside the province of Tarragona).

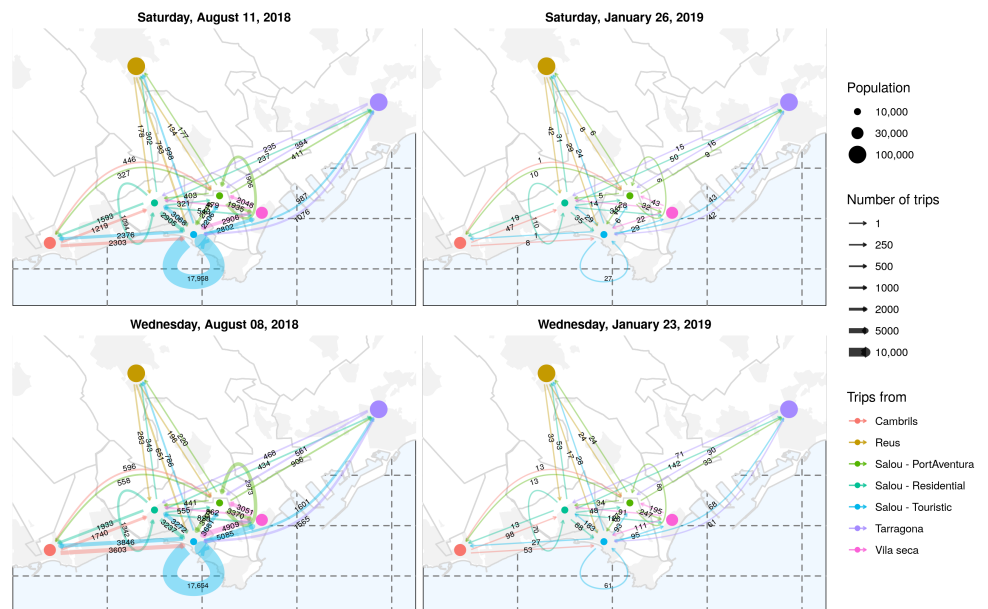


Figure 6. Flow maps of the intra-urban and inter-urban mobility in Salou of the international tourists.

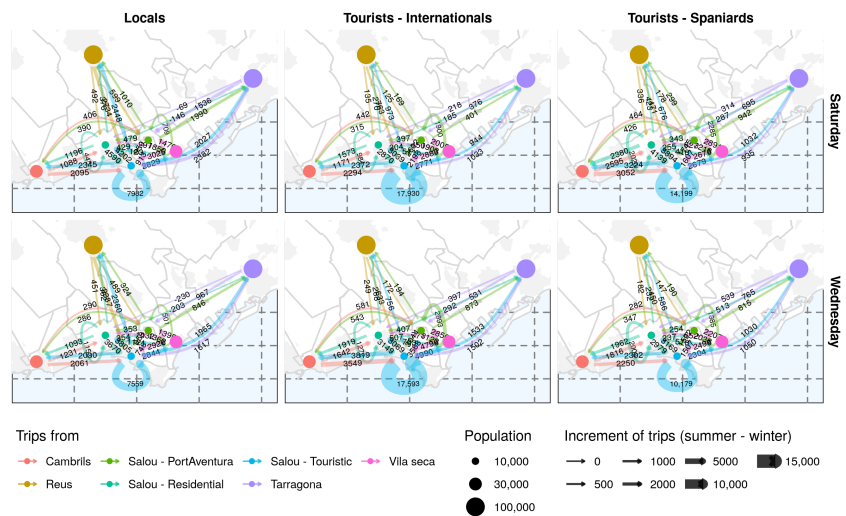


Figure 7. Seasonal increment in the number of trips for local residents, Spanish, and international tourists.

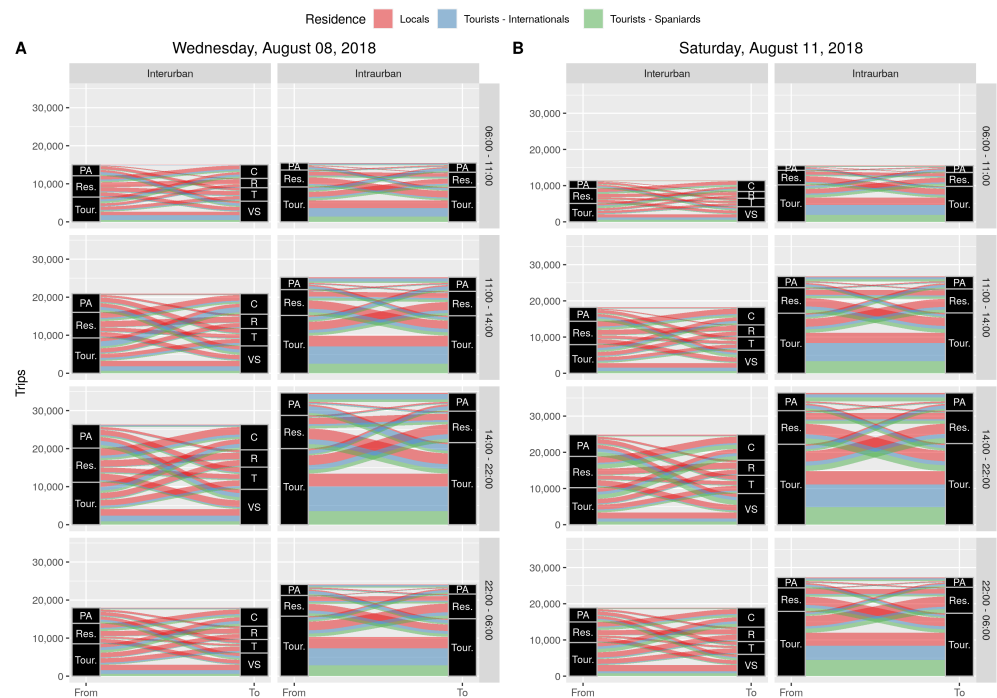


Figure 8. Alluvial diagrams showing the summer mobility of local residents, national tourists, and international tourists in Salou. Intra-urban mobility shows the movements between the different areas of Salou—PortAventura (PA), Residential (Res), Tourist (Tour), while inter-urban mobility shows the movements from these areas to the most important municipalities in the area—Cambrils (C), Reus (R), Tarragona (T), and Vila-seca (VS).

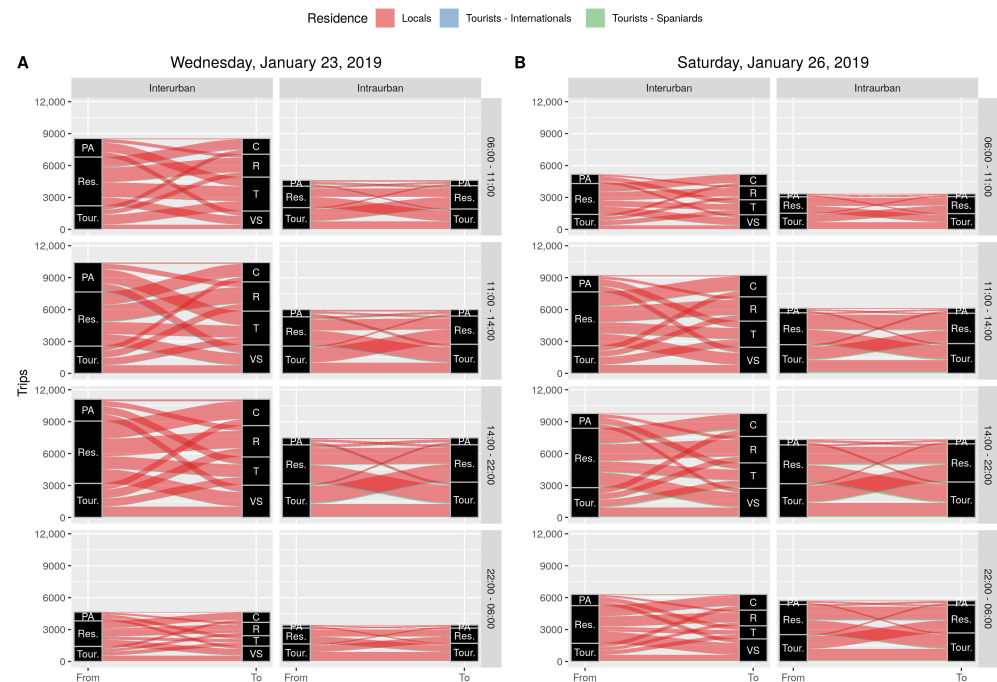


Figure 9. Alluvial diagrams showing the winter mobility of local residents, national tourists and international tourists in Salou. Intra-urban mobility shows the movements between the different areas of Salou—PortAventura (PA), Residential (Res), Tourist (Tour), while inter-urban mobility shows the movements from these areas and to the most important municipalities in the area—Cambrils (C), Reus (R), Tarragona (T), and Vila-seca (VS).

4.2. Seasonality at the Intra-Urban and Inter-Urban Scale

To address the research questions proposed in the introduction, we first want to highlight the main patterns that we can see thanks to our analysis. Later, it will be possible to extract considerations of a more general nature.

The first alluvial diagram (Figure 8) shows the trips made on summer days, at both the intra-urban and the inter-urban levels, of all the population groups at the time (local residents, national visitors, and foreign visitors). Being able to see the trips sorted by place of residence makes it quite clear that *locals* (red alluvium) predominate in terms of inter-urban mobility. At the same time, visitors, both national and international, tend to perform short-distance mobility in the same urban area (intra-urban mobility). These patterns of each profile have deep implications for the overall mobility of the region. The predominant trips in winter are inter-urban (and mainly carried out by *locals*), whereas, in summer, intra-urban trips are clearly higher due to the predominance of short-distance urban tourism mobilities.

In Figure 9, the importance of the local residents in winter is more evident than in the previous flow maps. The comparison between the two alluvial diagrams (Figures 8 and 9) allows measurement of the tourism seasonality effect in the overall mobility in the area. It must be said that the vertical scale, where the number of trips is measured, is almost three times smaller than the one in the faceted summer alluvial diagrams. On the other hand, the number of trips by time-slots is in line with expectations, that is, in the morning the number of trips is lower at weekends than on weekdays, while the number of trips at night is higher at weekends.

Figure 4 shows both the local population resident in Salou and that of the other municipalities in the province of Tarragona (see Figure 2). In general, there is a greater number of trips in summer than in winter. This finding helps to demonstrate that the seasonality of tourism activity in the area also implies higher mobility of the local population in summer. It could be related to the growth of economic activity in summer but also to the greater

mobility of locals for leisure during the summer months. Previous research conducted in the study area has also highlighted this situation using data from public transport passengers throughout the year [39,41]. These studies concluded that the seasonality of tourism activity implies a redefinition of regional functional hierarchies during the year, the coastal tourist cities being the area with the higher concentration of flows. The current study allows us to evidence this situation not only related to public transport flows, as happens in overall flows in the area. Moreover, we can now demonstrate that these differences are also significant at the intra-urban level and are concentrated in the most touristic area of Salou. For example, taking the trips from Reus to the three areas of Salou as a reference (e.g., focusing on Wednesday), trips to the PortAventura area increase by a few hundreds (from 1421 to 1910 trips), approximately the same as to the residential area of Salou (from 5609 to 5971 trips). On the other hand, trips to the tourist area clearly increase twofold (from 2084 to 4471). The internal mobility of the three zones defined in Salou also experiences a contrast between summer and winter, which is more significant in the tourist area with almost three times the number of trips. Among the locals, in winter it seems that mobility is slightly higher on weekdays than at weekends. Internal mobility in the PortAventura area also shows less activity at the weekend (about half of the trips).

In Figure 5, all trips made by residents in Spain from outside the province of Tarragona are considered. In this case, the contrast between the number of trips in summer and winter is much higher. This is logical, as there is clearly a higher concentration of Spanish visitors in summer. The figure could also be focused at first on the trips between Reus and the three areas of Salou, but the contrast is enormous in the trips from Tarragona (10 times bigger in summer), Cambrils, and Vila-seca (20 times bigger in summer, in both cases). The detail of the OD matrices is such that in winter they are able to show edges or flows of less than ten trips between the municipalities in the area. If we consider the internal mobility of the areas of Salou, the differences are considerable (more than 50 or 60 times bigger in summer). The attraction of the tourist area for Spanish tourists is quite clear. Once at the destination, most of their trips and activities are concentrated in the same area, leaving the residential area and PortAventura in a secondary position. Contrary to what happens with the *locals*, the total number of trips made by Spanish tourists is higher at the weekend than during the week. This value is explained by the short-time weekend national visitors that this destination receives [48].

Figure 6 shows the international tourists' (roaming users) trip patterns. These patterns are somewhat similar to those of the Spanish visitors, but the contrasts with the locals are much bigger. As an example, the internal mobility of the most touristic area of Salou is huge during the summer (up to 300 times bigger than in winter). The lack of short-time or weekend visitors in the case of international tourism explains why, in that case, no growth in global trips is observed during weekends. On the other hand, the tendency to concentrate trips in a small area is also greater than in the case of Spanish visitors. In other words, our work reveals that the international tourists tend to make shorter distance trips during their stay than the national tourists. This result is consistent with the literature on tourist mobility in the city [49]. Previous studies have also highlighted that repeat visitors (that could be represented by the Spaniards in the study area) tend to use a bigger area in the city than first-time visitors [50]. Finally, it was observed that international tourists have a greater weight of mobility linked to PortAventura. Moreover, in contrast to what happened with Spanish tourists, their mobility in the Theme Park area is higher during weekdays.

Finally, a global comparison can be made without losing the geographic context. In Figure 7, the map shows the differences between summer and winter, during the week and at weekends. In this figure, the differences between summer and winter are exaggerated, even more so for those trips that are very small or even null for some time slots in winter. Given the size of this figure, labels cannot be added to all arrows, but some significant differences can still be read well that could not be seen on previous maps. International tourists and Spaniards have quite similar travel patterns in the two seasons, but it is striking

that internal displacement in the theme park area varies between weekends and midweek. It seems that the Spanish have a certain preference for visiting the park at the weekend, while international tourists visit the park more during the week. This type of pattern is of interest, since it could be contrasted with the park capacity data in an attempt to validate the results obtained with PMD.

5. Discussion and Conclusions

5.1. Main Findings and Contributions

Our study has enabled us to demonstrate the overall mobility changes due to tourism seasonality in a coastal destination. Our approach combines the study of inter-urban and intra-urban mobility together with the different profiles of population (locals, national tourists, and international tourists). Through this methodology, a deeper knowledge of mobility in the region is achieved, along with new evidence that is added to that of previous survey-based studies of the resident population or public transport data [39,41,48]. On this occasion, having a more complete record made it more feasible to differentiate between national and international tourists. This is useful for the study of the spatial behaviour of different visitors and is of interest for tourism planning and management.

Seasonality affects especially Salou and its more touristic areas, but it also contributes to redefining the overall mobility patterns in the region. The mobility of the local population is also higher during summer. These findings provide new evidence on the utility of PMD for the study of mobility in territories in which visitors play a key role in the overall mobility. It also provides evidence on the utility of PMD for the analysis of seasonal episodes, such as coastal tourism.

Regarding the methodological issues raised in this study, we were interested in knowing to what extent PMD is useful for the study of seasonal tourist mobility at different scales and, as other authors have suggested [37], whether the proposed graph visualisations really help to analyse these data. This research work shows how PMD can be analysed to obtain valuable information and knowledge for decision-making in tourism. More specifically, we have conducted a case study showing how PMD can be aggregated and reduced in complexity in order to extract useful information about the mobilities of locals and visitors associated with tourism seasonality. Considering the nature of this information, it is necessary to elaborate clear and innovative visualisations to better understand and share the extracted knowledge.

PMD provides a good sample of the mobility of the total population staying in an area (both residents and visitors). For example, it is a more inclusive data source that does not require modern devices with GPS or other internet-based APPs—such as smartphones, tablets, wearables—, but, rather, any mobile phone can generate data once connected to the network. As a consequence, PMD allows the characterisation of different population profiles; in this study this is carried out by dividing the population into three different groups (local residents, Spanish tourists, and international tourists). In addition to this advantage, PMD provides quite homogeneous and multi-scale spatial resolutions, especially in very populated urban areas where the mobile phone network is denser. This feature facilitates the study of tourism seasonality, since different time-slots can be selected and the planning of the data acquisition can be carried out at any time. In this research, we selected four representative days—two during the week and two at the weekend—in summer and winter, but many other time periods and time-slots could be analysed using the same techniques. This is essential in a highly changeable sector, like tourism.

Finally, our work provides innovative visualisations for the study of tourism mobility flows. After conducting the literature review presented in Section 2, it has been detected that most of the analyses of tourists' mobilities performed using PMD do not provide detailed visualisations that could help to understand the knowledge extracted from those large amounts of data. This may depend on the type of problem being studied and the different needs to summarise the data. Most of the works reviewed presented their results in tabular form (tables and OD matrices) or using simple line or bar plots. Only

a few articles provided any kind of map-based visualisations [27,32,34–36] or proposed alternative ways of representing mobility flows in tourist areas (e.g., the use of cartograms in Reference [26]). However, on a broader scope, there are interesting proposals for tools and new types of visualisations applied to urban mobility that could be explored in order to find better ways to analyse problems, such as tourist seasonality and flow changes, at different spatio-temporal scales. As an example, one interesting study proposed visualisation of spatio-temporal profiles (urban *beats*) for urban areas [51]. In addition, with the aim of integrating non-experts in the query creation process, another study proposed a graph-based visual query method for massive human trajectory data [52]. This research work does not provide any new visualisations, but it does apply two types of visualisations that have not previously been adopted in tourism mobility analyses. Flow maps and alluvial diagrams have proved to be useful in representing the flows, distribution, and proportion of the mobilities of the population groups defined in the proposed case study. On the one hand, flow maps add a geographical perspective that helps to contextualise in which directions these flows of movements occur. On the other hand, alluvial diagrams offer an adequate vision to better quantify the proportions of these flows. Alluvial diagrams facilitate the drawing of inter- and intra-urban movements, since the flow map should generate visualisations at different scales to cover all cities (inter-urban) and their sub-areas (intra-urban). However, these types of visualisations were better adapted to some specific purpose (e.g., alluvial diagrams were not suitable for showing bidirectionality of flows and flow maps were confusing when trying to draw different groups on the same map).

5.2. Limitations

The analysis presented here offers a reasonably detailed view of the structure of tourism mobilities. However, there are some limitations that should be taken into account in future studies. For example, the availability of data, influenced by the density of the mobile phone network, varies with the MNO and this data heterogeneity can influence the type of analysis that can be carried out. Another point to note is that it cannot be guaranteed that the owner of the phone is the real user (e.g., the user could be an employee or a relative, among other possibilities). This point can add uncertainty to the results due to the fact that, in many cases, demographic data are used. Furthermore, data privacy issues are important aspects of PMD. There are distinct regulations at different levels on how all data collection must be carried out [53]. These regulations guarantee that neither persons nor their location can be identified. MNO are responsible for anonymising data so that users cannot be identified. For instance, when the mobility of a user between two areas is lower than a threshold (i.e., less than one hundred), it is removed from the analysis.

Although the spatial granularity of the data allows for a high level of detail, there are shortcomings in specific scenarios. An example of this problem occurs in the area of PortAventura. Since it is a theme park with a considerably large area (up to 825.7 ha), it is not possible to guarantee the correct recording of movements within the park. This is due to the proximity to the tourist-residential area of la Pineda and the distribution of cell towers in the area, meaning that some of the movements within the park could be counted as inter-urban mobility with Vila-seca.

Another important limitation of PMD comes from the cost of accessing and processing the data. This issue depends on the network provider and the system designed for exploiting these data. In this work, we demonstrated that seasonality can be shown correctly by selecting a few representative days. This restriction reduces the cost of this type of analysis. However, failing to determine an average day could affect the results (e.g., selecting days with special events, music festivals, etc). In addition, depending on the operator, accessing this type of data can be driven by an intermediary partner. In this context, research reproducibility can be compromised for two reasons: (1) the use of anonymised and proprietary data, and (2) the pre-processing of the data using proprietary software in a closed environment. These problems do not affect research conducted using other data sources, software, and practices [46].

One last limitation is related to following the requirements (static and space) that must be met to include these visualisations in a journal article. For example, using faceted plots has many advantages, but not all combinations of variables could be shown correctly if there are space restrictions. For example, creating more detailed faceted flow maps makes labels and arrows difficult to read. In this sense, new forms of interactive visualisations for the web can help users to understand the research outcomes [54] or even to interactively arrange the graphs [55]. However, representing too much information in the same plot can affect the readability, and some degree of generalisation is always necessary.

5.3. Future Work

Immediate directions for future work include other data sources to enrich and try to validate the PMD used in the current research work. In the results section, we already mentioned that the PortAventura entrance data could be used for this purpose. Another kind of data could be the records from smart card validations of the regional transport agency in the study area, where the authors of the current manuscript have previous experience [6]. The records of credit card transactions could also be useful, as these data reveal the origin of the cardholder, which is useful to identify whether a visitor is a foreigner. Another possible extension is the characterisation of the PMD with the type of transport (car, bus, bicycle, etc.) used to carry out the movements between the different areas of the case study, and it is now clear that the length of the stay could be an important element to analyse. The results of this analysis would be of benefit to potential stakeholders, such as policymakers, in order to make mobility or urban planning decisions. On the other hand, it would be of interest to replicate this approach in other destinations that also face the challenge of managing the seasonality of tourism demand.

Finally, an interesting aspect to explore and validate in future studies is the possibility of replicating the same study during the COVID-19 pandemic. The current global health crisis and the restrictions applied by different governments to combat the spread of the virus have had a profound influence on human mobility, in general, and tourism activity, in particular. A new study using telephone data during the different levels of restrictions would be of great interest to analyse the impact of COVID-19 on tourism, especially by analysing the influence on each of the profiles defined in this work (local residents, Spanish tourists, and international tourists).

Author Contributions: Conceptualisation, Aaron Gutiérrez; Data curation, Benito Zaragoza; Formal analysis, Benito Zaragoza and Aaron Gutiérrez; Funding acquisition, Aaron Gutiérrez; Investigation, Benito Zaragoza; Methodology, Sergio Trilles; Project administration, Aaron Gutiérrez; Resources, Aaron Gutiérrez; Software, Sergio Trilles; Supervision, Aaron Gutiérrez; Validation, Sergio Trilles; Visualisation, Benito Zaragoza; Writing—original draft, Benito Zaragoza, Sergio Trilles and Aaron Gutiérrez. All authors have read and agreed to the published version of the manuscript.

Funding: Research funded by the Spanish Ministerio de Ciencia e Innovación (grant number CSO2017-82156-R), the AEI/FEDER, UE, the Departament d'Innovació, Universitats i Empresa, Generalitat de Catalunya (grant number 2017SGR22), the Escola d'Administració Pública de Catalunya, Generalitat de Catalunya (grant number 2018 EAPC 00002) and the Generalitat Valenciana through the Subvenciones para la realización de proyectos de I+D+i desarrollados por grupos de investigación emergentes programme (GV/2020/035). Sergio Trilles has been funded by the postdoctoral Juan de la Cierva fellowship programme of the Spanish Ministry for Science and Innovation (IJC2018-035017-I).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Li, J.; Xu, L.; Tang, L.; Wang, S.; Li, L. Big data in tourism research: A literature review. *Tour. Manag.* **2018**, *68*, 301–323. [\[CrossRef\]](#)
- Alaei, A.R.; Becken, S.; Stantic, B. Sentiment analysis in tourism: Capitalizing on big data. *J. Travel Res.* **2019**, *58*, 175–191. [\[CrossRef\]](#)
- e Silva, F.B.; Herrera, M.A.M.; Rosina, K.; Barranco, R.R.; Freire, S.; Schiavina, M. Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tour. Manag.* **2018**, *68*, 101–115. [\[CrossRef\]](#)
- Chen, C.; Ma, J.; Susilo, Y.; Liu, Y.; Wang, M. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transp. Res. Part Emerg. Technol.* **2016**, *68*, 285–299. [\[CrossRef\]](#)
- Reif, J.; Schmücker, D. Exploring new ways of visitor tracking using big data sources: Opportunities and limits of passive mobile data for tourism. *J. Destin. Mark. Manag.* **2020**, *18*, 100481. [\[CrossRef\]](#)
- Gutiérrez, A.; Domènech, A.; Zaragoza, B.; Miravet, D. Profiling tourists' use of public transport through smart travel card data. *J. Transp. Geogr.* **2020**, *88*, 13. [\[CrossRef\]](#)
- Howard, P.N.; Mazaheri, N. Telecommunications reform, Internet use and mobile phone adoption in the developing world. *World Dev.* **2009**, *37*, 1159–1169. [\[CrossRef\]](#)
- Lenormand, M.; Picornell, M.; Cantú-Ros, O.G.; Tugores, A.; Louail, T.; Herranz, R.; Barthelemy, M.; Frías-Martínez, E.; Ramasco, J.J. Cross-Checking Different Sources of Mobility Information. *PLoS ONE* **2014**, *9*, 1–10. [\[CrossRef\]](#) [\[PubMed\]](#)
- Blondel, V.D.; Decuyper, A.; Krings, G. A survey of results on mobile phone datasets analysis. *EPJ Data Sci.* **2015**, *4*, 10. [\[CrossRef\]](#)
- Xu, Y.; Li, J.; Xue, J.; Park, S.; Li, Q. Tourism Geography through the Lens of Time Use: A Computational Framework Using Fine-Grained Mobile Phone Data. *Ann. Am. Assoc. Geogr.* **2020**, 1–25. [\[CrossRef\]](#)
- Janecek, A.; Valerio, D.; Hummel, K.A.; Ricciato, F.; Hlavacs, H. The cellular network as a sensor: From mobile phone data to real-time road traffic monitoring. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2551–2572. [\[CrossRef\]](#)
- Wu, W.; Wang, Y.; Gomes, J.B.; Anh, D.T.; Antonatos, S.; Xue, M.; Yang, P.; Yap, G.E.; Li, X.; Krishnaswamy, S.; et al. Oscillation resolution for mobile phone cellular tower data to enable mobility modelling. In Proceedings of the 2014 IEEE 15th International Conference on Mobile Data Management, Brisbane, Australia, 14–18 July 2014; Volume 1, pp. 321–328.
- Sørensen, A.Ø.; Bjelland, J.; Bull-Berg, H.; Landmark, A.D.; Akhtar, M.M.; Olsson, N.O. Use of mobile phone data for analysis of number of train travellers. *J. Rail Transp. Plan. Manag.* **2018**, *8*, 123–144.
- Jiang, S.; Ferreira, J.; Gonzalez, M.C. Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore. *IEEE Trans. Big Data* **2017**, *3*, 208–219. [\[CrossRef\]](#)
- Masso, A.; Silm, S.; Ahas, R. Generational differences in spatial mobility: A study with mobile phone data. *Popul. Space Place* **2019**, *25*, e2210. [\[CrossRef\]](#)
- Blumenstock, J.E. Inferring patterns of internal migration from mobile phone call records: Evidence from Rwanda. *Inf. Technol. Dev.* **2012**, *18*, 107–125. [\[CrossRef\]](#)
- Schmid, T.; Bruckschen, F.; Salvati, N.; Zbiranski, T. Constructing sociodemographic indicators for national statistical institutes by using mobile phone data: Estimating literacy rates in Senegal. *J. R. Stat. Soc. Ser. A (Stat. Soc.)* **2017**, *180*, 1163–1190. [\[CrossRef\]](#)
- Silm, S.; Ahas, R.; Mooses, V. Are younger age groups less segregated? Measuring ethnic segregation in activity spaces using mobile phone data. *J. Ethn. Migr. Stud.* **2018**, *44*, 1797–1817. [\[CrossRef\]](#)
- Yip, N.M.; Forrest, R.; Xian, S. Exploring segregation and mobilities: Application of an activity tracking app on mobile phone. *Cities* **2016**, *59*, 156–163. [\[CrossRef\]](#)
- Kamenjuk, P.; Aasa, A.; Sellin, J. Mapping changes of residence with passive mobile positioning data: The case of Estonia. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1425–1447. [\[CrossRef\]](#)
- Bengtsson, L.; Lu, X.; Thorson, A.; Garfield, R.; Von Schreeb, J. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: A post-earthquake geospatial study in Haiti. *PLoS Med.* **2011**, *8*, e1001083. [\[CrossRef\]](#)
- Grantz, K.H.; Meredith, H.R.; Cummings, D.A.; Metcalf, C.J.E.; Grenfell, B.T.; Giles, J.R.; Mehta, S.; Solomon, S.; Labrique, A.; Kishore, N.; et al. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nat. Commun.* **2020**, *11*, 1–8. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ahas, R.; Aasa, A.; Mark, Ü.; Pae, T.; Kull, A. Seasonal tourism spaces in Estonia: Case study with mobile positioning data. *Tour. Manag.* **2007**, *28*, 898–910. [\[CrossRef\]](#)
- Ahas, R.; Aasa, A.; Roose, A.; Mark, Ü.; Silm, S. Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. *Tour. Manag.* **2008**, *29*, 469–486. [\[CrossRef\]](#)
- Raun, J.; Ahas, R.; Tiru, M. Measuring tourism destinations using mobile tracking data. *Tour. Manag.* **2016**, *57*, 202–212. [\[CrossRef\]](#)
- Raun, J.; Shoal, N.; Tiru, M. Gateways for intra-national tourism flows: Measured using two types of tracking technologies. *Int. J. Tour. Cities* **2020**. [\[CrossRef\]](#)
- Alawwad, Z.; Alkowiileet, W.; Alsubaiee, S.; Alfaris, A.; Alansari, M.M.; Ragab, A.M. Analyzing call detail records to support tourism statistics in Saudi Arabia: An exploratory study. In Proceedings of the 14th Global Forum on Tourism Statistics, Venice, Italy, 23–25 November 2016.
- Chen, N.C.; Xie, W.; Welsch, R.E.; Larson, K.; Xie, J. Comprehensive Predictions of Tourists' Next Visit Location Based on Call Detail Records Using Machine Learning and Deep Learning Methods. In Proceedings of the 2017 IEEE International Congress on Big Data (BigData Congress), Honolulu, HI, USA, 25–30 June 2017; pp. 1–6.

29. Chen, N.C.; Xie, J.; Tinn, P.; Alonso, L.; Nagakura, T.; Larson, K. Data Mining Tourism Patterns. Available online: http://papers.cumincad.org/data/works/att/caadria2017_070.pdf (accessed on 12 December 2020).
30. Leng, Y.; Noriega, A.; Pentland, A.; Winder, I.; Lutz, N.; Alonso, L. Analysis of tourism dynamics and special events through mobile phone metadata. *arXiv* **2016**, arXiv:1610.08342.
31. Ling, F.; Sun, T.; Zhu, X.; Chen, Q.; Tang, X.; Ke, X. Mining travel behaviors of tourists with mobile phone data: A case study in Hainan. In Proceedings of the 2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 14–17 October 2016; pp. 1524–1529.
32. Zhu, X.; Sun, T.; Yuan, H.; Hu, Z.; Miao, J. Exploring Group Movement Pattern through Cellular Data: A Case Study of Tourists in Hainan. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 74. [[CrossRef](#)]
33. Chen, Q.; Hu, Z.; Su, H.; Tang, X.; Yu, K. Understanding travel patterns of tourists from mobile phone data: A case study in Hainan. In Proceedings of the 2018 IEEE International Conference on Big Data and Smart Computing (BigComp), Shanghai, China, 15–17 January 2018; pp. 45–51.
34. Qin, S.; Man, J.; Wang, X.; Li, C.; Dong, H.; Ge, X. Applying big data analytics to monitor tourist flow for the scenic area operation management. *Discret. Dyn. Nat. Soc.* **2019**, *2019*, 8239047. [[CrossRef](#)]
35. Vanhoof, M.; Hendrickx, L.; Puussaar, A.; Verstraeten, G.; Ploetz, T.; Smoreda, Z. Exploring the Use of Mobile Phone Data for Domestic Tourism Trip Analysis. Available online: <https://journals.openedition.org/netcom/2742> (accessed on 12 December 2020).
36. Mamei, M.; Colonna, M. Analysis of tourist classification from cellular network data. *J. Locat. Based Serv.* **2018**, *12*, 19–39. [[CrossRef](#)]
37. Xu, Y.; Li, J.; Belyi, A.; Park, S. Characterizing destination networks through mobility traces of international tourists—A case study using a nationwide mobile positioning dataset. *Tour. Manag.* **2021**, *82*, 104195. [[CrossRef](#)]
38. Gutiérrez, A.; Miravet, D. The determinants of tourist use of public transport at the destination. *Sustainability* **2016**, *8*, 908. [[CrossRef](#)]
39. Domènech, A.; Miravet, D.; Gutiérrez, A. Mining bus travel card data for analysing mobilities in tourist regions. *J. Maps* **2020**, *16*, 40–49. [[CrossRef](#)]
40. Domènech, A.; Gutiérrez, A. A GIS-Based Evaluation of the Effectiveness and Spatial Coverage of Public Transport Networks in Tourist Destinations. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 83. [[CrossRef](#)]
41. Gutiérrez, A.; Miravet, D. Estacionalidad turística y dinámicas metropolitanas: Un análisis a partir de la movilidad en transporte público en el Camp de Tarragona. *Rev. Geogr. Norte Gd.* **2016**, *89*, 65–89. [[CrossRef](#)]
42. Domènech, A.; Gutiérrez, A.; Clavé, S.A. Cruise Passengers' Spatial Behaviour and Expenditure Levels at Destination. *Tour. Plan. Dev.* **2020**, *17*, 17–36. [[CrossRef](#)]
43. Domènech, A.; Gutiérrez, A.; Anton Clavé, S. Built environment and urban cruise tourists' mobility. *Ann. Tour. Res.* **2020**, *81*, 102889. [[CrossRef](#)]
44. INE. Estadística de Movimientos Turísticos en Fronteras (FRONTUR) Noviembre 2019. Datos Provisionales. Available online: <https://www.ine.es/daco/daco42/frontur/frontur1119.pdf> (accessed on 12 December 2020).
45. Zaragoza, B.; Giménez-Font, P. A visual data analysis for determining the geographical extent of the cabreves. *J. Cult. Herit.* **2021**. [[CrossRef](#)]
46. Zaragoza, B.M.; Trilles, S.; Navarro-Carrión, J.T. Leveraging Container Technologies in a GIScience Project: A Perspective from Open Reproducible Research. *ISPRS Int. J. Geo-Inf.* **2020**, *9*. [[CrossRef](#)]
47. Pujol-Perdices, M.; Gutiérrez, A. Segregación residencial de la población extranjera en un destino turístico maduro: El caso de La Pineda (Costa Dorada). *Lurralde Investigación y Espacio* **2017**, *40*, 167–192.
48. Gutiérrez, A.; Miravet, D.; Saladié, Ò.; Anton Clavé, S. Transport mode choice by tourists transferring from a peripheral high-Speed Rail station to their destinations: Empirical evidence from Costa Daurada. *Sustainability* **2019**, *11*, 3200. [[CrossRef](#)]
49. Shoal, N.; McKercher, B.; Ng, E.; Birenboim, A. Hotel location and tourist activity in cities. *Ann. Tour. Res.* **2011**, *38*, 1594–1612. [[CrossRef](#)]
50. McKercher, B.; Shoal, N.; Ng, E.; Birenboim, A. First and repeat visitor behaviour: GPS tracking and GIS analysis in Hong Kong. *Tour. Geogr.* **2012**, *14*, 147–161. [[CrossRef](#)]
51. Miranda, F.; Doraiswamy, H.; Lage, M.; Zhao, K.; Gonçalves, B.; Wilson, L.; Hsieh, M.; Silva, C.T. Urban Pulse: Capturing the Rhythm of Cities. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 791–800. [[CrossRef](#)] [[PubMed](#)]
52. Wang, Z.; Yuan, Y.; Chang, L.; Sun, X.; Luo, X. A Graph-Based Visual Query Method for Massive Human Trajectory Data. *IEEE Access* **2019**, *7*, 160879–160888. [[CrossRef](#)]
53. Parliament, E. Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation). *Off. J. Eur. Union L* **2016**, *119*, 1–88.
54. Trilles, S.; Granell, C.; Degbelo, A.; Bhattacharya, D. Interactive guidelines: Public communication of data-based research in cities. *PLoS ONE* **2020**, *15*, e0228008. [[CrossRef](#)]
55. Zaragoza, B.; Giménez-Font, P.; Belda-Antolí, A.; Ramón-Morte, A. A graph-based analysis for generating geographical context from a historical cadastre in Spain (17th and 18th centuries). *Hist. Methods J. Quant. Interdiscip. Hist.* **2019**, *52*, 228–243. [[CrossRef](#)]