Points of Interest in Smart Cities and Visitor Behavior

Luís B. Elvas[1,2], Miguel Nunes[1], Bruno Francisco[1], Frederica Gonçalves[3], Ana Lucia Martins[4], and Joao Carlos Ferreira[1,2,4](✉)

[1] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisbon, Portugal
jcafa@iscte.pt
[2] Department of Logistics, Molde University College, 6410 Molde, Norway
[3] Universidade da Madeira, ESTG/ITI LARSyS, Funchal, Portugal
[4] Inov Inesc Inovação—Instituto de Novas Tecnologias, 1000-029 Lisbon, Portugal

Abstract. Smart cities leverage technology and data to enhance the quality of urban life, including the management of points of interest (POIs) and visitor experiences. This paper explores the relationship between POIs and visitor behavior in smart cities, examining the impact of technology-driven solutions on understanding, analyzing, and optimizing visitor experiences. It highlights the importance of data-driven approaches in identifying and managing POIs, enhancing visitor satisfaction, and driving economic growth. The paper reviews existing literature, discusses key concepts, and presents case studies to illustrate the role of POIs in smart cities and their influence on visitor behavior. Our major contribution is a data driven approach to extract useful information from real data to municipality decisions and understand the problem. It concludes with recommendations for future research and practical implications for city planners, policymakers, and tourism authorities.

Keywords: tourist behaviour · location data · data analytics · mobile phone sensing · Internet of Things · smart cities

1 Introduction

The significance of points of interest (POIs) in attracting visitors and driving economic growth cannot be overstated. POIs are key attractions and destinations within a city that draw the attention and interest of visitors. These can include tourist attractions, cultural landmarks, historical sites, entertainment venues, parks, shopping centres, and dining establishments. Here are some of the key reasons why POIs play a vital role in attracting visitors and fostering economic growth:

- Tourist Attraction: POIs are often the primary reason why visitors choose to travel to a particular city or destination. Iconic landmarks, natural wonders, and cultural sites have a unique appeal that entices tourists from around the world. These attractions create a sense of place and identity for the city, making it a must-visit destination.

A. L. Martins et al. (Eds.): INTSYS 2023, LNICST 540, pp. 69–91, 2024.
https://doi.org/10.1007/978-3-031-49379-9_4

- Increased Tourism Revenue: Visitors attracted to POIs contribute significantly to the local economy through spending on accommodation, transportation, dining, shopping, and entertainment. The revenue generated from tourism activities helps support local businesses, creates job opportunities, and drives economic growth in the hospitality, retail, and service sectors.
- Destination Competitiveness: Cities with well-developed and diverse POIs have a competitive advantage in the tourism industry. The presence of unique and compelling attractions differentiates the city from others, attracting a larger number of tourists and extending their length of stay. This, in turn, boosts the city's reputation as a desirable destination and enhances its overall competitiveness in the global tourism market.
- Urban Regeneration: Developing and promoting POIs can contribute to urban regeneration and revitalization efforts. Neglected or underutilized areas can be transformed into vibrant and attractive districts by focusing on the development of key POIs. This rejuvenation can attract new investments, businesses, and residents, leading to increased property values and overall urban improvement.
- Cultural Preservation and Heritage Conservation: Many POIs are of historical, cultural, or architectural significance. By attracting visitors to these sites, cities can raise awareness about their cultural heritage and promote preserving and conserving historical landmarks. This enhances the city's cultural identity and contributes to the sustainable development of the destination.
- Social and Community Benefits: Vibrant POIs often serve as gathering places and community hubs, fostering residents' pride, and belonging. These attractions can provide spaces for cultural events, festivals, and community activities, creating opportunities for social interaction, cultural exchange, and community cohesion.
- Supporting Ancillary Industries: The presence of POIs stimulates the growth of supporting industries, such as transportation, accommodation, retail, and food services. These industries benefit from the increased demand generated by visitors to the POIs. For example, hotels near popular attractions experience higher occupancy rates, and local businesses around POIs thrive due to increased foot traffic and customer engagement.

In summary, points of interest play a critical role in attracting visitors to a city and driving economic growth. They create unique experiences, contribute to tourism revenue, enhance destination competitiveness, promote cultural preservation, support urban regeneration, and provide social and community benefits. Strategic planning and management of POIs are essential for cities seeking to leverage their attractions as catalysts for sustainable economic development and overall urban well-being.

Municipalities must make data-driven decisions to effectively leverage the potential of Points of Interest (POIs). Here are some key reasons why data-driven decision-making is crucial: 1) Understanding Visitor Behaviour: Data collection and analysis provide insights into visitor behaviour, preferences, and patterns. By analysing data on visitor demographics, visitation patterns, and interaction with POIs, municipalities can gain a deeper understanding of what attracts visitors and how they engage with different attractions. This information helps in making informed decisions about marketing strategies, resource allocation, and infrastructure planning; 2) Optimizing Resource

Allocation: Data-driven insights enable municipalities to allocate resources effectively. By analysing visitor data, authorities can identify peak periods of activity, understand visitor flow between different POIs, and allocate resources such as transportation, staff, and amenities accordingly. This helps in optimizing the utilization of resources and improving the overall visitor experience; 3) Enhancing Visitor Experiences: Data-driven decision-making allows municipalities to personalize and enhance visitor experiences. By analysing data on visitor preferences, feedback, and behaviour, authorities can tailor their offerings and services to meet visitor expectations. This includes providing personalized recommendations, customized itineraries, and targeted promotions to enhance visitor satisfaction and engagement; 4) Planning Infrastructure and Services: Data analysis helps in planning and developing infrastructure and services around POIs. By examining data on visitor demand, transportation patterns, and accessibility, municipalities can make informed decisions about infrastructure improvements, public transportation routes, parking facilities, and the provision of amenities and services near POIs. This ensures a seamless and enjoyable experience for visitors; 5) Measuring Impact and ROI: Data-driven decision-making enables municipalities to measure the impact of their efforts and evaluate the return on investment (ROI). By analysing data on visitor numbers, spending patterns, and economic indicators, authorities can assess the economic impact of POIs on the local economy. This information helps justify investments, secure funding, and demonstrate the value of POIs to stakeholders and the community; and 6) Responding to Emerging Trends and Challenges: Data analysis allows municipalities to monitor emerging trends and challenges related to visitor behaviour and POIs. By continuously analysing data, authorities can identify changing visitor preferences, emerging attractions, and potential issues such as overcrowding or safety concerns. This helps in adapting strategies, implementing measures to address challenges, and staying responsive to the evolving needs of visitors.

By analysing visitor data, authorities can gain insights into visitor behaviour, optimize resource allocation, enhance visitor experiences, plan infrastructure, measure impact, and respond to emerging trends and challenges. This approach leads to more informed and effective decision-making, ultimately contributing to the success and sustainability of POIs and the city's overall development.

2 Literature Review

Tourism is one of the economic activities of strong relevance worldwide. According to the World Travel and Tourism Council [1], before the COVID-19 pandemic, its impact on the world GDP was 10.3%, and 1 in 4 new jobs were related to its development. According to the same source, it is also estimated that the revenue generated from international visitors was US$1.8 billion.

In Portugal, official data indicate that the wealth generated by tourism is equivalent to 8.8% of the GDP, and the sector represents 7.5% in total employment [2], with the Municipality of Lisbon being one of the main destinations for visitors, whether national or foreign. Tourism is, therefore, one of the main branches of the Portuguese economy.

The current scenario of competitiveness among organizations is characterized as a time when they need to be ahead of others in time to gain advantage in a continuous series of periods, so it is necessary to invest in constant innovation of their products and services [3].

Moreover, balancing the needs of residents with the expectations of visitors to tourist sites by providing them with high-quality experiences requires this population to be well characterized and managed [4].

The application of Big Data knowledge to the development of Business Intelligence systems applied to tourism stands out, in this sense, as a useful tool for the State and organizations to know this phenomenon and to better manage its economic and social performance in a timely manner.

In the specific case of this work, in which the goal is to analyze the movement of visitors in points of interest in the Municipality of Lisbon, several studies [3] show the importance and usefulness of geospatial analysis of large volumes of data for a better experience by the tourist, and for decision making by managers. In the case of Lisbon, some others have demonstrated how such decision support systems, based on analysis can become important information on the decision making [5–8].

Points of interest are, in the words of Gil et al. (2020), "regions of influence where citizens concentrate because of attractions or facilities and (…) places where the energy of the city is focused, and understanding the shape and size of POIs provides insights into how people experience the city" [9].

From a tourism perspective, it should be noted that the term tourist encompasses all visitors, which is why this term will be widely used in this study, including residents, day-trippers, as well as foreign visitors.

According to Harris & Howard (1996), as cited in Spangenberg (2013), "the focus here is on attractions, defined as the physical or cultural feature of a particular place that travellers or tourists perceive capable of satisfying one or more of their specific leisure-related demands. Such features may be environmental in nature (e.g., climate, culture, vegetation, or landscape), or they may be site-specific, such as a theatre performance, a museum, or a waterfall" [10].

Studies investigating human movement have a strong appeal for understanding this phenomenon, especially for the motivations that lie in its background [11]. They indicate that these movements are related to different activities (e.g., education, work, catering, etc.) in different urban environments.

Nearly a decade ago, in a study advocating the use of mobile device data for analysing people's movement [12], Zhang (2014) examined mobility patterns of users of such devices. Having noted the increasing use of devices such as smartphones, the author indicated that their data could be managed for application development and predictive modelling.

More recently, in a paper that indicated it was one of the first to adopt the Dictionary Learning method for characterizing human mobility patterns [13], Wu et al. (2017) indicated a step forward in this process, advocating that the large volume of data generated by mobile devices be handled in multiple collection centres due to the fact that we live in what they consider the era of data explosion.

It can be seen, therefore, that the use of large volumes of data from mobile devices, namely cell phones, has been asserting itself as a necessary tool for the analysis of people's movement behaviour with a focus on the most varied fields, such as urban planning, epidemiology, telecommunications, etc.

When discussing the importance of Big Data for the tourism industry, Shafiee & Ghatari (2017) state that big data leads to more efficient travel experiences in line with visitors' expectations, and also improves the levels of innovation, so necessary for this sector of the economy [14].

From this interest in innovation, studies have been developed around the analysis of visitor movement in tourist points of interest based on mobile devices. These points of interest (POI) correspond to larger areas or contiguous to tourist sites, such as monuments, museums, parks, buildings, etc.

Some recent studies deal with the development of intelligent systems for recommending POI to be visited based on the user's location and on characteristics of the site to be visited, such as distance, time spent travelling, evaluation given by other users, the desired category (e.g., museum, historical building, beach), etc. [15, 16].

In others, of particular interest for this work, human mobility in POIs is analysed from data visualization and modelling, in order to assist for the definition of urban planning strategies and investments for better exploitation of these territories [9–11, 17].

3 CRISP-DM

In developing this research work, the CRISP-DM (Cross Industry Standard Process for Data Mining) methodology [18] was followed to structure data to extract statistical and predictive information.

The efficiency of the method's application consists of the development in six stages: business understanding, data understanding, data preparation, modelling, evaluation, and deployment.

3.1 Business Understanding

POIs are places of agglomeration of people in each region characterized by their attractiveness and for being focuses of a great flow of visitors. How such mobility occurs is relevant for stakeholders to manage tourism-related problems.

Using big data from mobile devices used by visitors, the analysis of this problem becomes an indispensable tool for urban strategic planning in the tourism area.

Based on data from Vodafone mobile devices made available by CML, and aiming to provide an analysis tool, this work aims to evaluate the movement of visitors to certain areas of the city of Lisbon, differentiating their behaviour according to the winter and summer periods the day of the week the hours/periods of higher concentration in the POI the rainy days of the nationality.

This data is representative through an extrapolation of the market share of the other mobile operators and will be used for the behavioural study of visitors and their movement in the different POIs of the city.

This study will use a Python algorithm with a Jupiter Notebook interface for data analysis and transformation. The final statistics will be visualised using the Microsoft PowerBI tool. The predictive models will be developed using the Prophet algorithm.

3.2 Data Understanding

The base element of this work comes from a dataset made available by CML. This is composed of 12 monthly files, from September 2021 to August 2022, containing information about active cell phone terminals. The database sections the Lisbon Municipality in 200 m $\times$ 200 m^2 (grids) and indicates the number of active terminals in that zone, grouped every 5 terminals and in 5-min intervals, comprising the 24 h of each day. It contains all the data concerning the location of the POIs of local and roaming users, the latter with their nationality.

All files are composed of the same variables, and each dataset has the number of instances referenced. Each file contains a large volume of data that, to be processed together, would prevent processing on personal computers. Thus, the processing will be done on each monthly file for later grouping of the data to be treated.

To enrich the original information, data from the Portuguese Institute of Sea and Atmosphere (IPMA) will be used to collect the rainfall conditions in the studied period. These data contain the precipitation value at each hour of the day. To fill possible gaps in the IPMA data, the research will be complemented with data from the Time and Date website [20].

The available data present some gaps in the information:

- The IPMA database comprises a period of 5 months, from September 2021 to January 2022;
- The Vodafone dataset has a gap in the information for the month of February 2022 due to a "deliberate and malicious cyberattack" [21] suffered by the company on February 7 of that year.

Thus, due to lack or insufficiency of information, it was decided to limit the analysis to 6 months, divided into two 3-month periods to delimit two seasons: winter, from November 2021 to January 2022; and summer, from June to August 2022. The IPMA data complements the winter information with on-the-day precipitation. Of the total, 14,313,762 instances will be analysed, corresponding to approximately 58% of the available Vodafone data, with summer comprising 52%.

The data was also be augmented with another dataset made available by the faculty as part of this work: Wktcomplete. This dataset contains specific and relevant information about the city division grids, namely a name for the location, the parish and its georeferencing.

The data was also complemented with another dataset made available by the teaching team as part of this work: Wktcomplete. This dataset contains specific and relevant information about the city division grids, namely a name for the location, the parish and its georeferencing. This is geographic information.

To make the analysis more objective and targeted, we resorted to the study of visit recommendation websites to delimit the zones of the city to be analysed. In line with the recommendation of the websites Visit Lisbon [22], Turismo de Lisboa [23], O Guia da Cidade [24], Direção Geral do Património Cultural [25], LisbonLisboaPortugal [26] and Civitatis Lisbon [27] 4 zones of the city were selected for the analysis: Belém, Baixa/Chiado, Alfama/Castelo and Parque das Nações.

In the final dataset machine learning and forecasting algorithms will be applied for pattern detection and learning, developing a predictive model of visitors and their behavior.

3.3 Data Preparation

To optimize computational resources, data preparation began. We chose to keep the most relevant variables for the work: Grid_ID, Datetime, extract_month_3, extract_day_4, C1, C2, C9, C10 and D1.

Aiming to facilitate the analysis, it was decided to change the name of the columns to an easily identifiable nomenclature, namely:

- C1 → TN (National Terminals): active national terminals.
- C2 → TR (Terminals in Roaming): active roaming terminals.
- C9 → DN (National Data): national terminals with active data.
- C10 → DR (Roaming Data): roaming terminals with active data.

The datasets were analysed individually, month by month, and do not present duplicate lines. As for missing data, these were detected only in variable D1, which will be described in the item regarding the treatment of nationality.

After cleaning the dataset, the feature engineering process was started to add information for a better classification and understanding of the data. The dataset was enriched by creating new columns.

Day of the Week - To allow segmented analysis by day of the week, the date was used to create two nominal variables:

Weekday - Based on the date, the variable "Weekday" indicates the day of the week, Monday through Sunday.

Weekdays - Using the variable "Weekday" a new variable was created that classifies the days of the week into working days ("workday") or weekend days ("weekend").

3.3.1 Totalizers of the Number of Terminals

Given that the variables TN, TR, DN and DR comprise the number of active terminals and terminals with active data, national and roaming respectively, we created two variables that indicate the total of these variables, with the following information:

- "TT" - Total Terminals - indicates the total number of active terminals, obtained by adding the active national terminals and active roaming terminals (TN + TR);
- "DT" - Total Data - indicates the total number of terminals with active data, obtained by adding up the national terminals with active data and roaming terminals with active data (DN + DR)

Period - Aiming at an analysis that identifies the concentration of people at certain times of the day, it was decided to group the continuous variable of time into a categorical variable that comprises specific periods, as follows:

- Morning - comprised between 7 am and 11 am,
- Lunch - from 11 am to 3 pm,

- Evening - from 3 pm to 7 pm,
- Dinner - from 7 pm to 10 pm,
- Night - from 10 pm to midnight,
- Down - from midnight to 7 am.

Public Holiday - To analyse the impact of holidays on the affluence to POIs, the variable “Holiday” was introduced. The observations were identified with the respective holiday (national or municipal) according to the date. The others were identified with the nomenclature “na” to indicate “not applicable”.

Parish and Place - To add easily recognizable information, the external database Wktcomplete.csv was used to create the variable “Parish”. The relationship was made through the grid identifier and the name of the parish in which the respective grid cell is located is referred to.

Through the grid it was possible to identify, in the same source, the name/location of the grid (for example: Av. Brasília - Belém, Área ribeirinha Pedrouços or Feira Popular). This information was also added in a “Local” variable.

Precipitation - To identify behavior due to weather conditions, a binary variable was introduced to indicate the rainfall on the date.

This column was added using IPMA data. The values “0” and “1” are taken to indicate whether there was (1) or no (0) precipitation on that day.

The IPMA data was cleaned. For the imputation of missing data, the historical meteorological data for the Lisbon municipality available at the timeandate website [28] were used. In the total dataset considered there are 154837 instances with precipitation, which corresponds to 2.4% of the total for the winter period, comprising 13 days of a total of 92 days in the considered winter.

Data Selection - To maximize the efficiency of the statistics and presentations, it was decided to create a dataset limited to the zones object of the analysis, comprising Belém, Baixa/Chiado, Alfama/Castelo and Parque das Nações.

Once the information was limited, a detailed analysis of the variables targeted for analysis was carried out and possibilities for improvement were observed.

Location - With the goal of easy and immediate identification of the location, the nomenclatures identified in the “Location” variable were changed.

Zone - To provide a more comprehensive geographic visualization, a column was created with the information of the city zone where the “Location” is located. Thus, the categorical variable “Zone” was created, which now has the following nomenclatures: Belém, Alfama/Castelo, ‘Baixa/Chiado and Pq Nações.

Station - Aiming at a quick sectioning of the data, it was decided to create the variable “Season” to identify the season of the year to which the data refers. The instances of the months of November, December, and January were identified as “winter” and the months of June, July, and August were identified as “summer”.

nac1 (nationality) -Variable D1, representing the list of nationalities of the terminals, was disaggregated. It originated the variable “NAT1” to identify the country of origin of the respective terminal.

In this variable, nomenclature inconsistencies were identified, so a standardization was performed.

After analysing the missing data in this variable, it was found that they exist in only 59.6% of the total observations. Thus, it was decided to keep the whole dataset to obtain complete statistics on the number of active terminals, using this variable only for the analysis on the nationality of the visitors.

Country/Region - To facilitate the visualization of the geographic region, the variable "Country/Region" was created to group the terminals by region. The regions were selected according to the continental geographic location. In the case of Europe, it was divided between Europe-Shengen [29] and other European countries, thus allowing the identification of countries whose entry into Portugal does not require a visa.

4 Analysis of Results

To evaluate the behaviour of visitors in the selected zones, we started by visualizing the total number of active cell phones.

As shown in Fig. 1, the zones of the city with the highest number of terminals are, respectively, Baixa/Chiado, Parque das Nações, Alfama/Castelo and Belém. Referring to roaming visitors, one notices a slight difference in this trend. The areas with more active cell phones are Baixa/Chiado and Alfama/Castelo, followed by Parque das Nações and Belém. This difference can be explained by the business activity in the city, with the Baixa/Chiado and Parque das Nações areas indicating a higher concentration of offices.

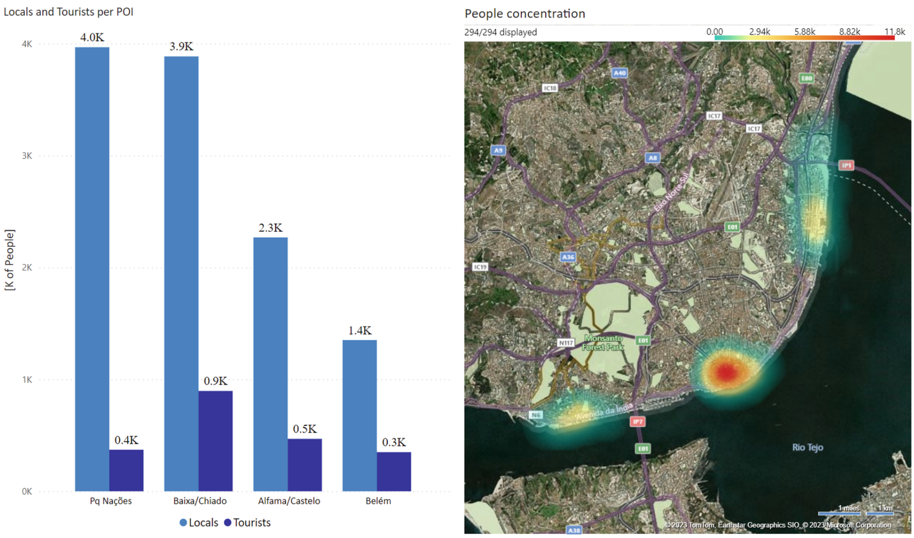


Fig. 1. Total active terminals per city zone

4.1 Concentration of Visitors by Nationality

The analysis of cell phones with active roaming allows us to identify the list of countries with more visitors.

The selection of the 10 countries with the highest number of terminals shows us that 8 belong to the European area, the others being from North and South America (Fig. 2).

Overall, the global trend of most visited zones is maintained in these 10 countries, with Baixa/Chiado and Alfama/Castelo showing the highest concentration of visitors (Fig. 3). For 7 of the 10 countries (France, United Kingdom, Brazil, Belgium, United States of America, Switzerland, and Ireland visit Parque das Nações more) the third most visited city area is Parque das Nações, while the Spanish, Italians and Germans have the highest concentration of visitors in Belém.

The analysis of the roaming terminals shows a higher prevalence of visitors in the summer period, with the Baixa/Chiado and Parque das Nações areas being more popular in the summer period, while Belém and Alfama/Castelo have a slightly higher demand in the winter period (Fig. 4 and Fig. 5).

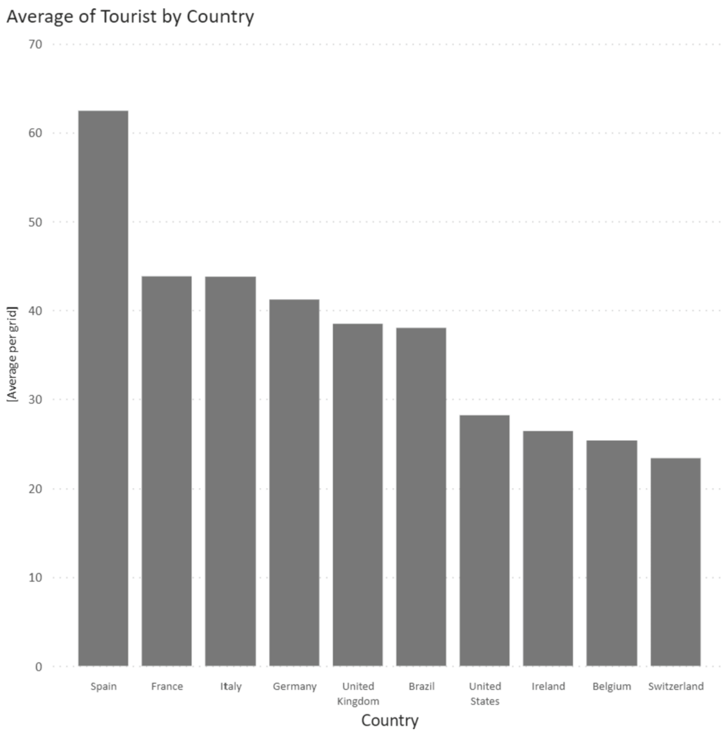


Fig. 2. Top 10 roaming terminals by nationality

The top 10 countries by season show that visitors from Italy and Switzerland have similar affluence of visits in summer and winter. Visitors from France, the United Kingdom, Brazil and Ireland are concentrated in the summer period. The United States and Belgium bring visitors mainly in the summer, and these two countries drop out of the top 10 list in number of visitors in the winter, see Fig. 6. Germany is the only country that brings more visitors in winter than in summer, increasing the influx by roughly 1.3%.

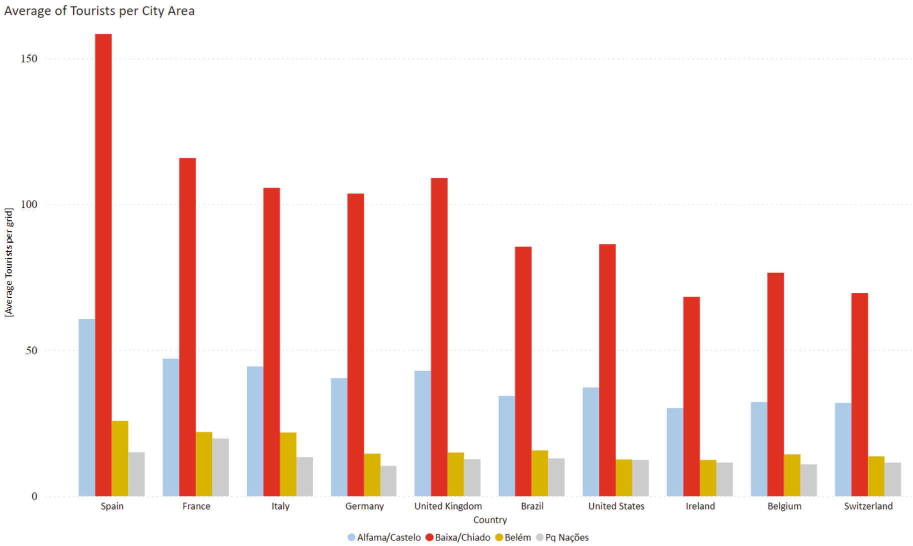


Fig. 3. Top 10 roaming terminals by nationality and city zone

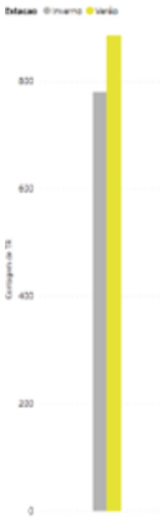


Fig. 4. Overral Roaming Terminals

As shown in Figs. 7, 8, 9 and 10, visitors tend to concentrate in certain locations in the specific area.

In Belém, during the summer, there is a higher concentration of visitors in the area of the Jerónimos Monastery and the whole extension of the Praça do Império, followed by the garden area of the Belém Tower. In winter it is noted that on rainless days there

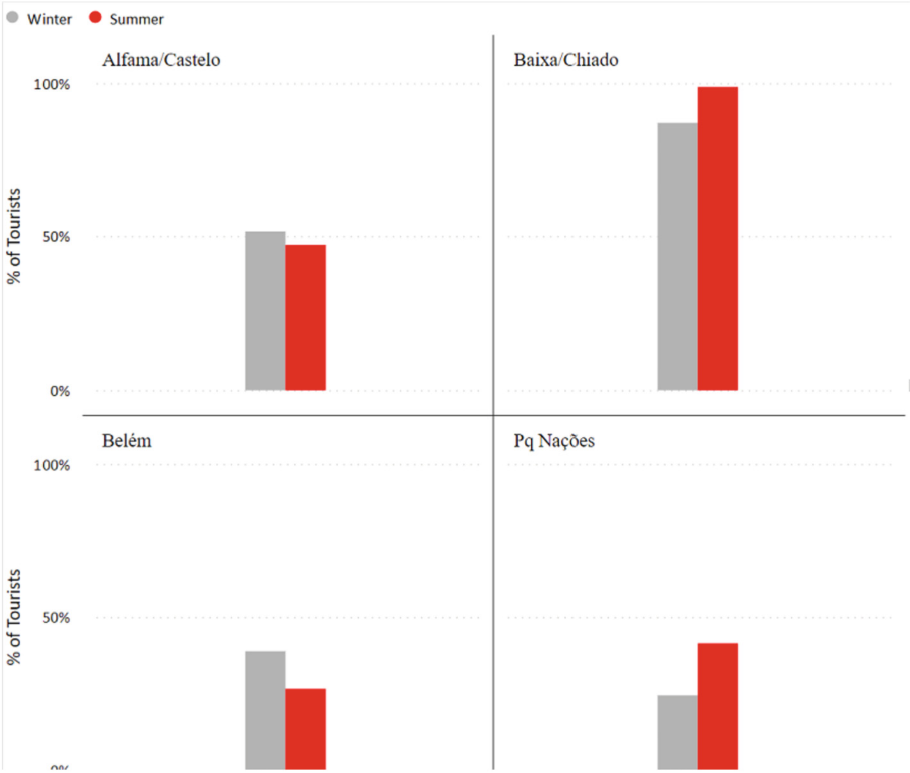


Fig. 5. Overall Roaming Terminal by year season and City area

is a greater concentration of visitors at the Monastery of Jerónimos and the Tower of Belém, while on rainy days there is an evident concentration at the Monastery.

In the Baixa/Chiado area, Praça Luís de Camões, Rossio and Rua Augusta are the areas that concentrate the highest number of visitors. On winter days when there is no rainfall the behavior of visitors is very similar to summer, and on rainy days there is less preference for the squares and a higher concentration near the local commerce is evident.

Baixa/Chiado – Summer Winter (no rain) Winter (rain)

In the summer visitors tend to concentrate near the Sé Patriarcal cathedral, coming from downtown. The second most visited place in this area is the Miradouro de Santa Luzia, followed by the Castelo de S. Jorge.

On winter days is also greater concentration of visitors in the vicinity of the Cathedral, and the Castle of St. George has more visitors on rainy days of winter than on days of the cold season when there is no precipitation.

When there is precipitation, the circulation of visitors in this area reduces by 70% to 80%, with Ireland and Denmark giving way to Austria and Finland in the top 10 list of visitors. In summer, visitors to Parque das Nações tend to concentrate on the Pavilion of Knowledge and the Altice Arena (Pavilhão Atlântico). On winter days when there is no

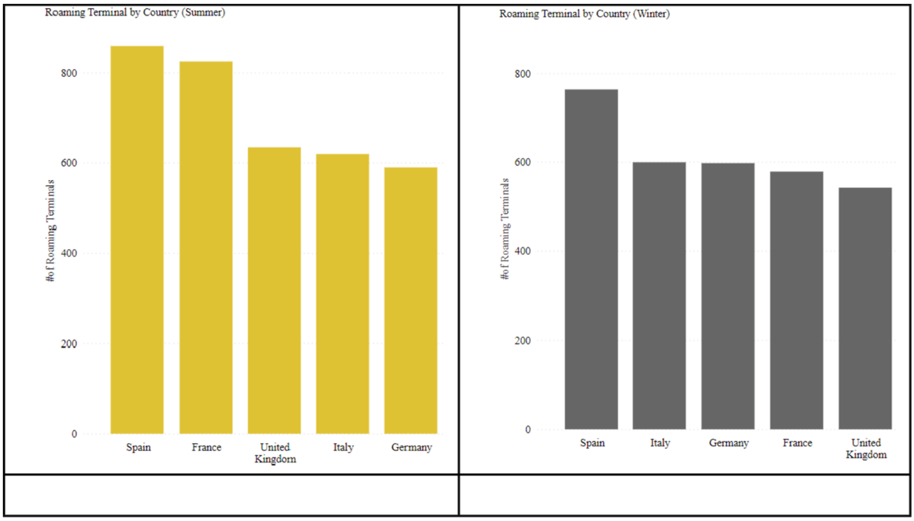


Fig. 6. Top 5 visitors by nationality/season

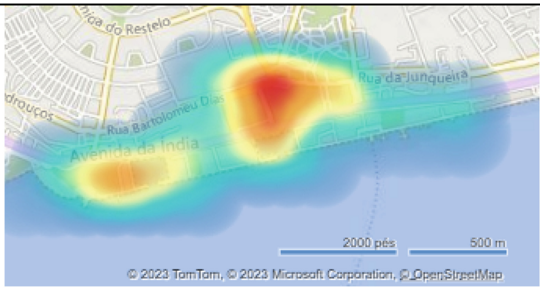
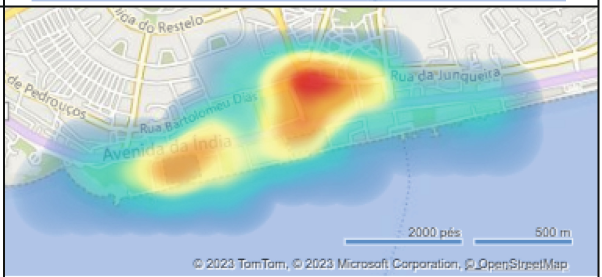
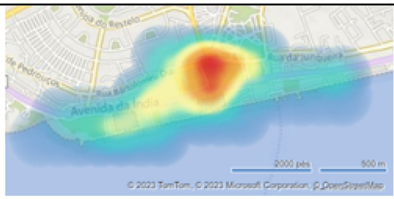
Belém	
Summer	
Winter (no rain)	
Winter (rain)	

Fig. 7. Visualizing the concentration of visitors according to the season and rainfall conditions in the Belém.

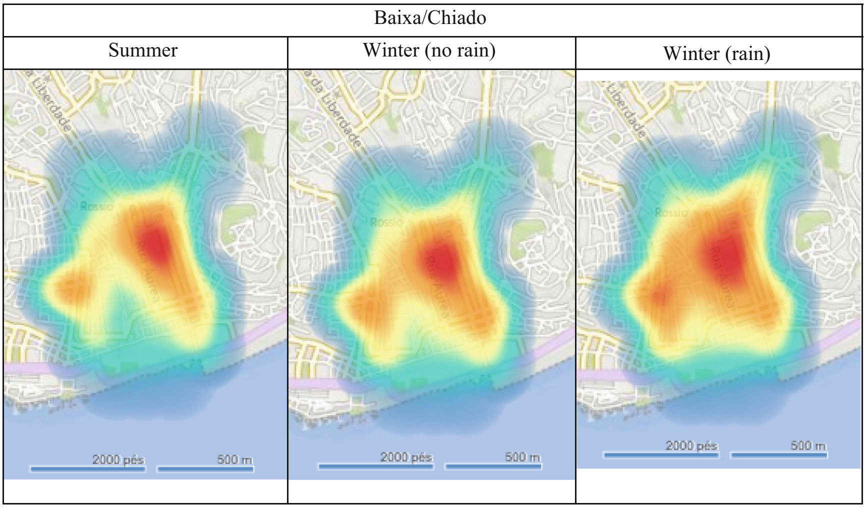


Fig. 8. Visualization of visitor concentration according to the season and rainfall conditions in the Baixa/Chiado area.

rainfall, visitors prefer the area around the Pavilion of Knowledge, while on rainy days, the concentration is higher in the Vasco da Gama Shopping Center area.

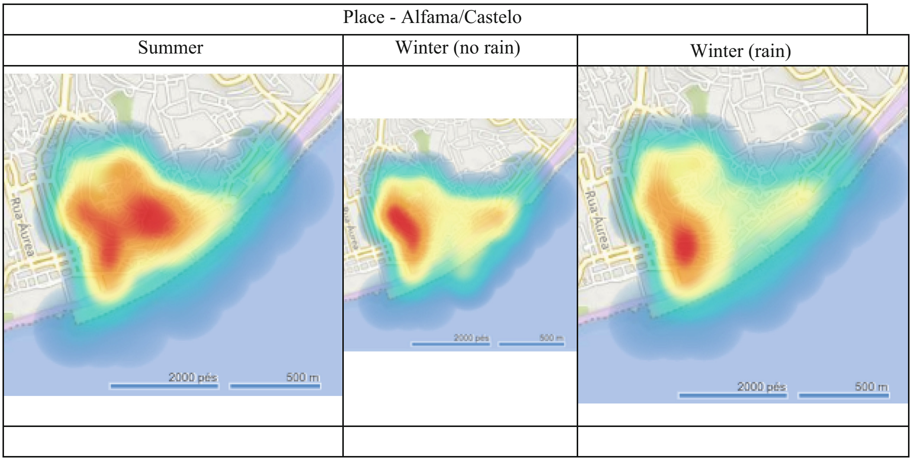


Fig. 9. Visualization of visitor concentration according to the season, season, rainfall conditions in the Alfama/Castelo

4.1.1 Difference in Behaviour by Nationality (Top 5)

The analysis by nationality shows that on rainy days (precipitation variable = 1), the top 5 foreign visitors tend to concentrate more in the Parque das Nações and Baixa/Chiado areas, to the detriment of the others (Fig. 11). The zone that suffers the largest decrease in affluence is Belém, with Alfama/Castelo losing slightly. This trend allows us to observe that the behaviour of all groups seems to be identical, regardless of nationality.

Place – Parque das Nações – Summer Winter (no rain) Winter (rain)

Place – Parque das Nações – Summer Winter (no rain) Winter (rain)

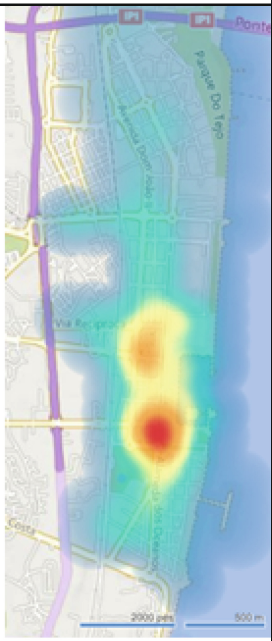
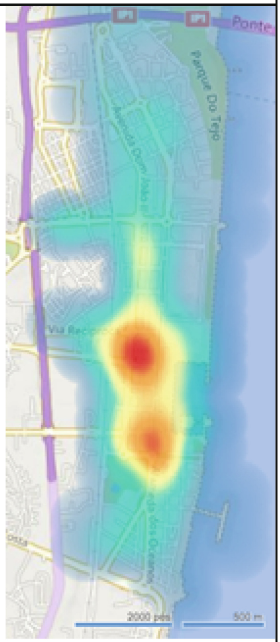
Place – Parque das Nações		
Summer	Winter (no rain)	Winter (rain)
		

Fig. 10. Visualization of visitor concentration according to season and precipitation season and rainfall conditions in the Parque das Nações area

4.2 Day of the Week

Figure 12 shows the turnout of roaming terminals on the different days of the week. It is important to reinforce that the percentage of visitors is significantly higher during weekends.

Regardless of the day of the week there are always more active roaming terminals in the Baixa/Chiado zone, followed by Alfama/Castelo. And as for Belém, only on weekends this zone is not the least visited, with Parque das Nações.

The peak of visitors in Belém is on Sundays, in Baixa/Chiado and Alfama/Castelo on Saturdays and in Parque das Nações on Mondays.

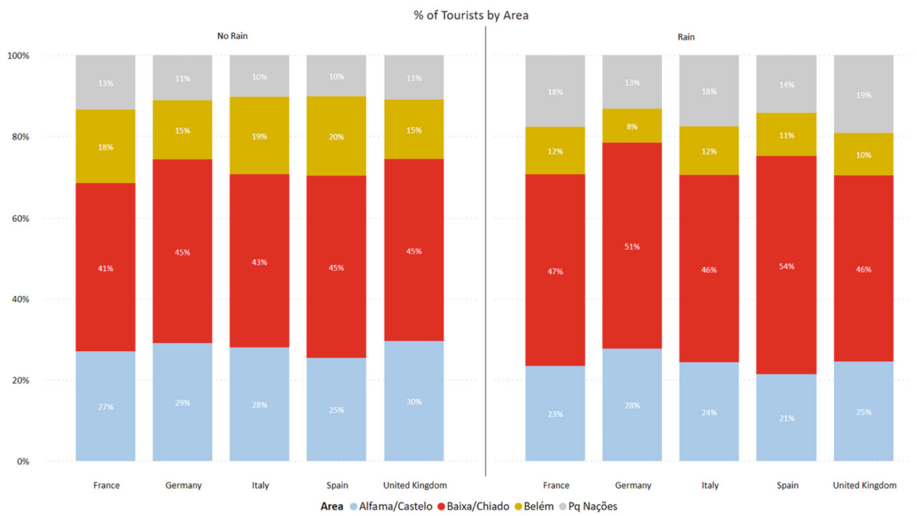


Fig. 11. Top 5 by nationality: comparing visitor behavior visitors behavior on rainy days

In the opposite direction, the lowest affluence is detected in Parque das Nações and Baixa/Chiado on Fridays, Belém on Mondays, and Alfama/Castelo on Sundays.

The area that shows less variation in the affluence of visitors is Alfama/Castelo.

In relation to the nationalities that visit Lisbon on the weekend vs. the whole week, practically no changes were noted in the Top 10. Only visitors from the United States of America appear in seventh when considering the whole week and Switzerland in eighth, reversing the positions when it is only the weekend. The top three, in both cases, are Spain in first position, followed by France and Italy.

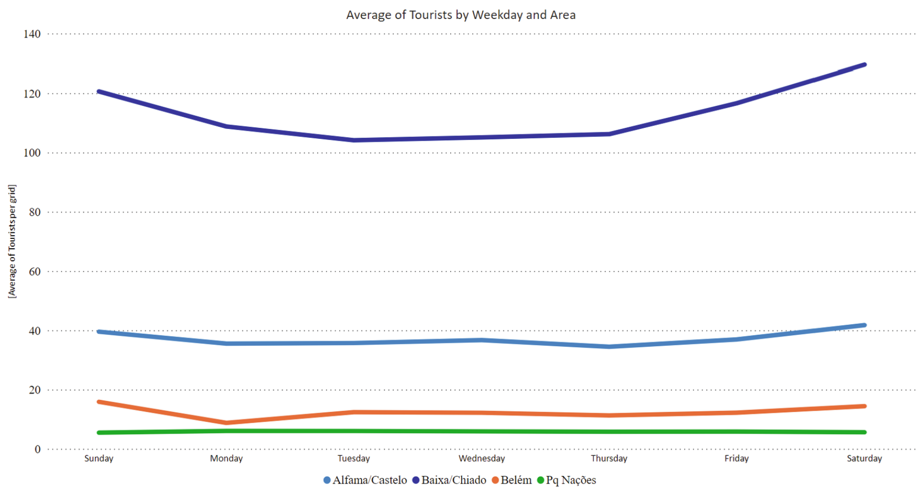


Fig. 12. Roaming terminal affluence in the city zones on the different days of the week

The analysis per daily period (Fig. 13) and per hour (Fig. 14) allows us to identify some trends:

- Visitors arrive in Bethlehem close to lunch time (from 11 am to 3 pm) and leave from there before dinner time (7 pm). It is the area with the lowest frequency of roaming terminals in the early morning period. Factors that contribute to this displacement may be related to the hotel supply and restaurant diversity in the area.
- Parque das Nações also has its peak of visitors at lunch time, decreasing this affluence as the end of the night approaches (1 h). It is an area of the city with a low frequency of roaming terminals during the early morning hours.
- Alfama/Castelo has a certain constancy in the volume of visitors. There is a balance in the affluence from lunch time, which starts around 11 am, until the end of the night, around 1 am. The period with the lowest incidence of roaming terminals is the early morning, even though it does not differ much from the other periods. It is also the area of the city whose affluence starts earlier in the morning.
- Baixa/Chiado, being the most frequented area by roaming visitors, has an increasing affluence since the morning, dropping only slightly at dawn. The affluence at dawn can be explained by the offer of entertainment establishments in the Chiado/Bairro Alto area.

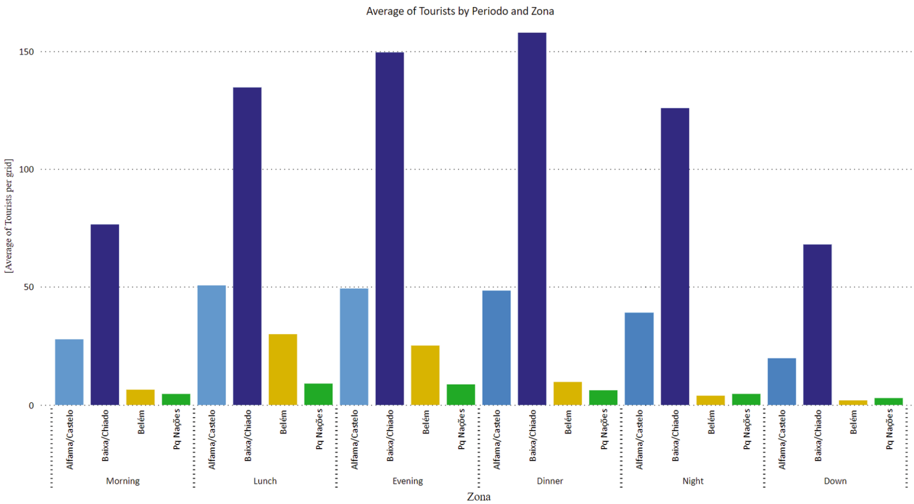


Fig. 13. Affluence of roaming terminals in the city zones the different periods of the day

Thus, the following visitor behaviors are evident:

- in Belém they flow fundamentally by lunch time and early afternoon.
- the movement in Parque das Nações starts late in the morning and gradually decreases until late in the evening.
- They arrive at Baixa/Chiado and Alfama/Castelo in the morning and keep the affluence high until the early morning.

The behaviour of visitors as a function of time was also analysed. In all the zones considered the value drops to less than half of the maximum value reached during the day in the period between midnight and 8 am. This factor may be related to visitors not sleeping in any of the zones considered.

Also, through these graphs we can consider that the number of roaming terminals always reaches its maximum value between 1 pm and 2 pm, except in the Baixa/Chiado area, which happens at 7 pm. The maximum number reached in the period between 1 pm and 2 pm may be related with the arrival time of several cruise ships in Lisbon, also with people who choose to make stopovers in Lisbon and with visitors who are staying in the surrounding area of the city like Cascais, Sintra and Oeiras.

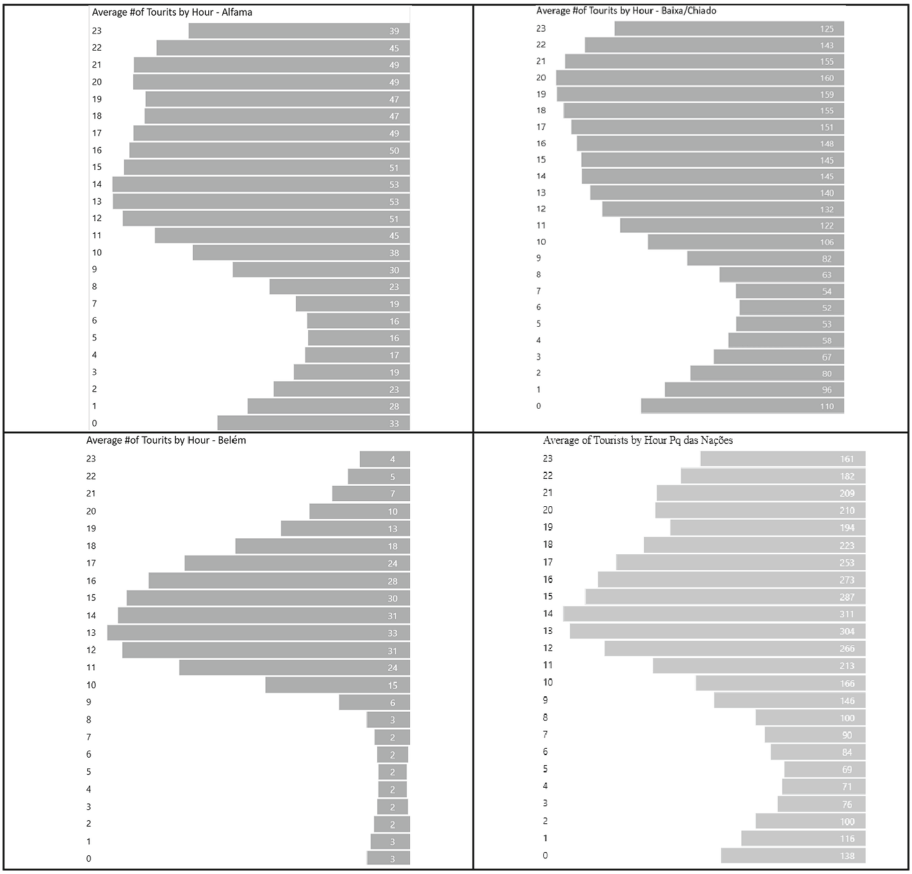


Fig. 14. **R**oaming terminals in the city zones city zones at different times of the day

Comparing the hours per zone on the days of the week, it can be concluded that:

– Alfama/Castelo shows practically no differences in the number of roaming terminals depending on the day of the week.

– Baixa/Chiado there are mostly differences during the afternoon to the end of the night, with on average approximately 10% more terminals in active roaming at the weekend during these hours.
– Belém, where the differences seem to be more significant after 10 am, increasing on average 25%.
– Parque das Nações has an opposite trend, with more data roaming during the week compared to the weekend. The maximum difference observed is at 2 PM, where during the week there are 309 roaming visitors compared to 225 with active roaming. It is also worth noting that at 2 pm the maximum number of visitors is reached in this zone.

5 Modelling and Evaluation

The generation of the prediction model was developed using Prophet, an open source software produced by Facebook [19]. This option is justified by the simpler and faster way of programming and for being visually more intuitive to reach conclusions or insights.

After re-analyzing the database, the DN, DR, and DT columns were eliminated because they are like the TN, TR, and TT columns, which represented optimization of memory space.

All observations with null values and less than five in the TT column were also eliminated, to take into consideration only information from instances with more than five active devices in total.

After verifying the almost irrelevant impact of the number of roaming devices on the total number of devices, it was decided to treat only the TT column as the prediction variable for the forecast model.

To facilitate visualization and to implement the prediction algorithm, the database was divided into two: “winter”, which includes all the information from 11/01/2021 to 01/31/2022; and “summer”, with observations dated between 06/01/2022 and 08/31/2022.

In the winter period, there is a tendency for a decrease in visitors in all POIs, except for the Baixa/Chiado POI, which registers an increase in visitors between the end of November and the beginning of December, probably due to Christmas shopping.

In the summer, there is a tendency of stability in the number of visitors in the POIs. However, one notices a downward movement in the week following the holidays of June 10th and 13th at the Baixa/Chiado POI, that is, between June 18th and 24th. This oscillation is not due to the lower amount of data.

The Prophet algorithm’s prediction for the days of the week in the winter period shows that: in Belém, the number of visitors increases as the weekend approaches; in Baixa/Chiado, the number of visitors increases sharply between Thursday and Saturday, falling sharply on Sunday; in Alfama/Castelo, the largest number of visitors is found on Saturday, Tuesday and Wednesday; and in Parque das Nações, the largest flow of people is on weekdays, especially on Tuesday, which is justified by the fact that many workers who don’t live in the area are allocated there.

For the summer period, and following the same forecasting approach, it is expected that: in Belém, the same trend as in the summer is observed, except for the decrease in visits on Wednesdays; differently from the winter, the movement of people in Baixa/Chiado is higher on weekdays and Saturdays, falling sharply on Sundays; in Alfama/Castelo, visitors prefer Friday and Saturday, with the lowest movement being observed on Wednesdays and Thursdays; finally, in Parque das Nações, the same trend predicted for the winter is maintained.

The prediction model did not have the intended performance evaluation metrics, but the results of predicting values for the summer period are better than the winter, especially for the MAPE values that are between 0.07 and 0.18. This means an accuracy between 93% and 88%, respectively, in predicting the next 30 observations.

In Figs. 15, 16, 17 and 18 below, the graphs with the lines of observed versus predicted values in the summer season for the month of August can be observed.

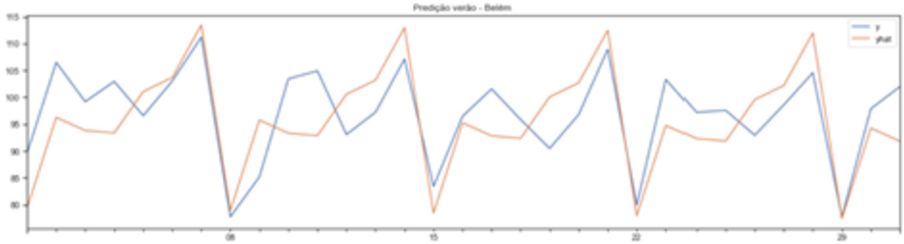


Fig. 15. Forecasting with observed and forecasted values for Belém POI in summer season

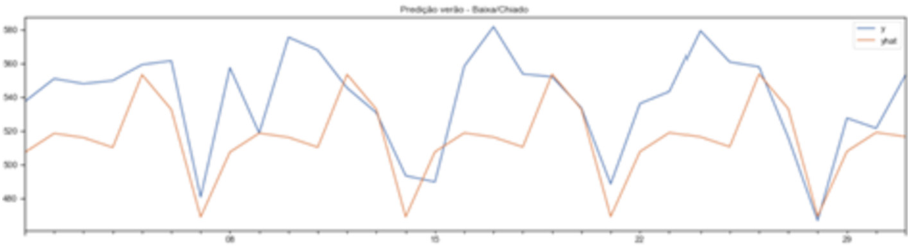


Fig. 16. Forecasting with observed and forecasted values for the POI of Baixa/Chiado in the summer season.

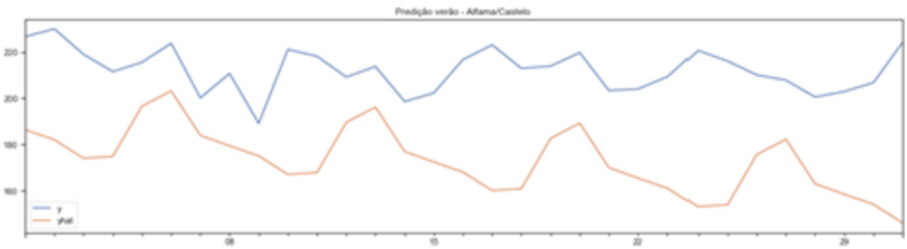


Fig. 17. Forecasting with observed and predicted values for the Alfama/Castelo POI in the summer season

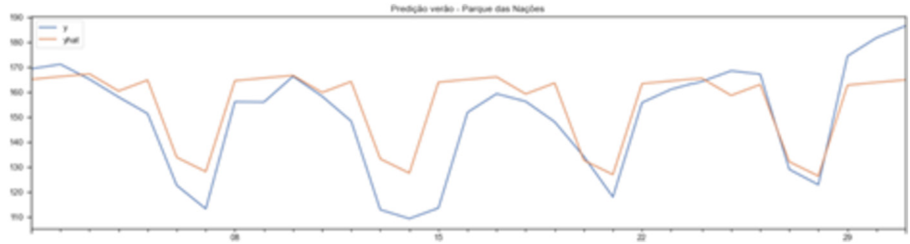


Fig. 18. Forecasting with observed and predicted values for the Parque das Nações POI in the summer season

We conclude that the results obtained with the data modelling using the Prophet software can help forecast and understand the evolution of the number of visitors at the analysed POI and its trend on the days of the week.

6 Conclusion

The paper concludes by summarizing the key findings and emphasizing the importance of understanding visitor behaviour within smart cities. It highlights the potential of utilizing visitor behaviour insights to optimize urban planning, enhance visitor experiences, and create sustainable and inclusive smart city environments.

In summary, this paper explores the concept of POIs in smart cities and investigates visitor behaviour patterns. By understanding visitor preferences, routing patterns, and interactions with POIs, policymakers, urban planners, and tourism stakeholders can make informed decisions to improve urban environments, optimize resource allocation, and deliver personalized visitor experiences in smart cities.

Summary of key findings and contributions.

Reiteration of the importance of POIs in smart cities and visitor behaviour.

Call for continued research and collaboration to advance POI management practices in smart cities.

Acknowledgment. We thank the Lisbon City Council for providing us with the data necessary for this study, namely Mr. António Costa (Lisbon City Council), Mrs. Helena Martins (Lisbon City Council) and Mrs. Paula Melicias (Lisbon City Council). We also thank IPMA for providing us with the data necessary to complement our study.

Funding. This work was supported by EEA Grants Blue Growth Programme (Call #5). Project PT-INNOVATION-0069 – Fish2Fork. This research also received funding from ERAMUS+ project NEMM with grant 101083048.

References

1. Travel & Tourism Economic Impact | World Travel & Tourism Council (WTTC). https://wttc.org/research/economic-impact. Accessed 22 May 2023

2. LCG. TravelBI by Turismo de Portugal. https://travelbi.turismodeportugal.pt/. Accessed 22 May 2023
3. Mariani, M., Baggio, R., Fuchs, M., Höepken, W.: Business intelligence and big data in hospitality and tourism: a systematic literature review. Int. J. Contemp. Hosp. Manag. **30**(12), 3514–3554 (2018). https://doi.org/10.1108/IJCHM-07-2017-0461
4. Supak, S., Brothers, G., Bohnenstiehl, D., Devine, H.: Geospatial analytics for federally managed tourism destinations and their demand markets. J. Destin. Mark. Manag. **4**(3), 173–186 (2015). https://doi.org/10.1016/j.jdmm.2015.05.002
5. Mesquitela, J., Elvas, L.B., Ferreira, J.C., Nunes, L.: Data analytics process over road accidents data—a case study of Lisbon city. ISPRS Int. J. Geo-Inf. **11**(2), Article no. 2 (2022). https://doi.org/10.3390/ijgi11020143
6. Elvas, L.B., Gonçalves, S.P., Ferreira, J.C., Madureira, A.: Data fusion and visualization towards city disaster management: Lisbon case study. EAI Endorsed Trans. Smart Cities **6**(18), e3–e3 (2022). https://doi.org/10.4108/eetsc.v6i18.1374
7. Elvas, L.B., Mataloto, B.M., Martins, A.L., Ferreira, J.C.: Disaster management in smart cities. Smart Cities **4**(2), Article no. 2 (2021). https://doi.org/10.3390/smartcities4020042
8. Elvas, L.B., Marreiros, C.F., Dinis, J.M., Pereira, M.C., Martins, A.L., Ferreira, J.C.: Data-driven approach for incident management in a smart city. Appl. Sci. **10**(22), Article no. 22 (2020). https://doi.org/10.3390/app10228281
9. Gil, E., Ahn, Y., Kwon, Y.: Tourist attraction and points of interest (POIs) using search engine data: case of Seoul. Sustainability **12**(17), Article no. 17 (2020). https://doi.org/10.3390/su12177060
10. Spangenberg, T.: Standardization, modeling and implementation of points of interest – a touristic perspective. Int. J. U- E- Serv. Sci. Technol. **6**, 59–70 (2013). https://doi.org/10.14257/ijunesst.2013.6.6.07
11. Zeng, W., Fu, C.-W., Müller Arisona, S., Schubiger, S., Burkhard, R., Ma, K.-L.: Visualizing the relationship between human mobility and points of interest. IEEE Trans. Intell. Transp. Syst. **18**(8), 2271–2284 (2017). https://doi.org/10.1109/TITS.2016.2639320
12. Zhang, Y.: User mobility from the view of cellular data networks. In: IEEE Conference on Computer Communications, IEEE INFOCOM 2014, pp. 1348–1356 (2014). https://doi.org/10.1109/INFOCOM.2014.6848068
13. Wu, T., Rustamov, R.M., Goodall, C.: Distributed learning of human mobility patterns from cellular network data. In: 2017 51st Annual Conference on Information Sciences and Systems (CISS), pp. 1–6 (2017). https://doi.org/10.1109/CISS.2017.7926085
14. Shafiee, S., Ghatari, A.R.: Big data in tourism industry. In: 2016 10th International Conference on e-Commerce in Developing Countries: with focus on e-Tourism (ECDC), pp. 1–7 (2016). https://doi.org/10.1109/ECDC.2016.7492979
15. Suakanto, S., Andreswari, R., Albasori, E.P.: SEMPIR: sequence multiple point of interest recommender system for overland tourism. In: 2021 International Conference on ICT for Smart Society (ICISS), pp. 1–7 (2021). https://doi.org/10.1109/ICISS53185.2021.9533194
16. Ajantha, D., Vijay, J., Sridhar, R.: A user-location vector based approach for personalised tourism and travel recommendation. In: 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), pp. 440–446 (2017). https://doi.org/10.1109/ICBDACI.2017.8070880
17. Giglio, S., Bertacchini, F., Bilotta, E., Pantano, P.: Machine learning and points of interest: typical tourist Italian cities. Curr. Issues Tour. **23**(13), 1646–1658 (2020). https://doi.org/10.1080/13683500.2019.1637827
18. Wirth, R., Hipp, J.: CRISP-DM: towards a standard process model for data mining. In: Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining (2000)

19. Prophet. http://facebook.github.io/prophet/. Accessed 22 May 2023
20. Current Local Time in Lisbon, Portugal (Lisboa). Consultado em 10 de abril de 20223. Disponível em https://www.timeanddate.com/worldclock/portugal/lisbon
21. Vodafone Portugal alvo de ciberataque. Press Release, 8 de fevereiro de 2022. Consultado em 03 de maio de 2023 (2022). Disponível em https://www.vodafone.pt/press-releases/2022/2/vodafone-portugal-alvo-de-ciberataque.html
22. 10 things to do and see in Lisbon. Visit Portugal. Consultado em 01 de abril de 2023. Disponível em https://www.visitportugal.com/en/destinos/lisboa-regiao/73773
23. Visit Lisboa. Turismo de Lisboa. Consultado em 01 de abril de 2023. Disponível em https://www.visitlisboa.com/pt-pt
24. O que visitar em Lisboa. O Guia da Cidade. Consultado em 01 de abril de 2023. Disponível em https://www.guiadacidade.pt/pt/discover-pois-lisboa-11
25. Indicadores TOP 100 - Registos de Património + Vistos. Direção Geral do Património Cultural. Consultado em 01 de abril de 2023. Disponível em http://www.monumentos.gov.pt/site/APP_PagesUser/SIPATop100.aspx?it=1
26. Lisbon Portugal Guide. LisbonLisboaPortugal.com. Consultado em 01 de abril de 2023. Disponível em https://lisbonlisboaportugal.com/index.html
27. Lisbon Travel Guide. Civitatis Lisbon. Consultado em 01 de abril de 2023. Disponível em https://www.lisbon.net/
28. Past Weather in Lisbon, Portugal. Consultado em 12 de março de 2023. Disponível em https://www.timeanddate.com/weather/portugal/lisbon/historic?month=12&year=2021
29. Europa sem fronteiras: O Espaço Schengen. Publicação da Comissão Europeia – Migração e Assuntos Internos. Consultado em 13 de abril de 2023 em (2019). https://doi.org/10.2837/71557, https://home-affairs.ec.europa.eu/system/files_en?file=2020-09