

Community Safety Well-Being and Touristic Spots: A Model Proposal to Correlate Safe Places and Touristic Spots Based in Open Data

Dineu Assis, Miguel de Castro Neto & Marcel Motta

NOVA Information Management School, Universidade Nova de Lisboa, Lisbon, Portugal

This is the accepted version of the conference paper published by IEEE at *2020 4th International Conference on Smart Grid and Smart Cities (ICSGSC)*:

How to cite: Assis, D., Castro Neto, M. D., & Motta, M. (2020). Community Safety Well-Being and Touristic Spots: A Model Proposal to Correlate Safe Places and Touristic Spots Based in Open Data. In *2020 4th International Conference on Smart Grid and Smart Cities (ICSGSC)* (pp. 136-144). [9248556] (4th International Conference on Smart Grid and Smart Cities, ICSGSC 2020). Institute of Electrical and Electronics Engineers Inc.. <https://doi.org/10.1109/ICSGSC50906.2020.9248556>

(©2021 IEEE)

Community Safety Well-Being and Touristic Spots: A Model Proposal to Correlate Safe Places and Touristic Spots Based in Open Data

Dineu Assis
NOVA Information Management School
Universidade Nova de Lisboa
Lisbon, Portugal
e-mail: m20170164@novaims.unl.pt

Miguel de Castro Neto
NOVA Information Management School
Universidade Nova de Lisboa
Lisbon, Portugal
e-mail: mneto@novaims.unl.pt

Marcel Motta
NOVA Information Management School
Universidade Nova de Lisboa
Lisbon, Portugal
e-mail: mmotta@novaims.unl.pt

Abstract—When choosing a tourism destiny, besides knowing what are the attractions that must be visited, it is also common to look for some information regarding the safety conditions of the supposed place, neighborhood, city or in a more specific case, the destined country itself. This kind of information is usually found in a structured way at official Open Data sources (although is a government decision to accept or refuse to offer this data) and in a less structured form through the public collection of web sites. However, the information exists – most in a textual form - there is a lack of a common standard to define safe from unsafe places. In this document, the proposed model combines open data, social networks and other web sites contents that after passing through a classification process will allow the definition of a score. When it becomes ready, the consoled will also be offered as an open source, giving a chance to not only help tourists in their traveling decisions but also, in the other hand, give government enough information to better deal with safety and accessibility issues.

Keywords-community safety; well-being; tourism; smart cities; urban analytics; data mining

I. INTRODUCTION

Portugal's tourism sector has grown steadily in recent years. It is a fact that since 2010, the volume of international arrivals worldwide has increased by around 5%, every single year [8] [9] [10]. Europe concentrates more than half of this result and when the benefits of this growth are analyzed in the country aspect of Portugal they become relevant both in the financial context – through positive impacts on revenue [11] [12] – as well as in the planning of public policies, which need to be developed to accompany sustainable growth of tourism and its impact on the well-being of locals.

In addition to the rise of the public revenues, the development of the tourism sector brings with it concerns about crime rates in cities and tourist sites. A tourist can be described as is a sum of relationships and phenomena that result in travel and experiences as non-residents [4]. When

choosing a destination, the tourist often has two options: also seek safety conditions, logistics and comfort or just worry when in the chosen location. Usually travelers go for the second option or, when they decide to inform themselves in advance, find it difficult to interpret indexes and recommend travelers or local news sources.

Similarly, for the local citizen of the selected destination, it is relevant to know how good are the safety, logistics and comfort indexes, at least known so that public policies can be planned and implemented to reduce crime and provide a social good for the population.

In this context, comfort and safety become special while defining what region will be visited by the tourist. More than that, in the perspective of the local resident, it is also interesting to know if the public safety policies adopted for their city (or district) follow any kind of standards or have been monitored in order to improve the citizen life qualify and traveler experience. In this context open data initiatives, considered as defining elements of emerging smart cities as they provide citizens with the tools necessary to create new, innovative services or applications [17] play a critical role.

II. LITERATURE REVIEW

The meaning of “safe places” is described in the literature as a broad concept of community. The earliest publication referring to the term “Safe Community” still refers to the situation analyzed in the context of England in 1986 [14]. At that time, the British government underwent a change of concept. from “crime prevention” to “safety in communities”. The purpose of this change was to expand responsibility for crime prevention beyond the police and thus to consider social aspects of crime that are affected by perceived risk, organizations, families and individuals. It was then that the government management of England and Wales described the concept of “Safe Communities” as follows: “Safe Community is generally one or more community actions to inhibit and remedy the causes and consequences of criminal, intimidating and related antisocial behavior. Its

goal is to ensure sustainable reductions in crime as well as the perception of crime in local communities. Its approach is based on forming multi-agency partnerships between the public, private and voluntary sectors to formulate and introduce community crime measures”.

Complementarily to this definition, the “well-being” state is achieved when all the psychological, social and physical resources required by the community are used to meet a particular psychological, social and/or physical challenge [7]. When individuals have more challenges than resources, the balance falls apart and the welfare state changes.

However, in order to evaluate Community Safety and Well-Being (CSWB) – even though its measurement is abstract, as already mentioned – there is a need to create an index that fulfils the role of evaluating the contribution of social in multisector to create conditions for global improvement. These indicators should aggregate the results of related sectors based on their results that should be shared [5] [6] [14]. When aggregated, there is an index that determines the CSWB level. By doing this, it will be possible to achieve scalability, complexity and systemic perspectives [13] – aiming to derive a result that can be associated to the whole community.

Typically, the results described during the process come from Economic, Health, Safety and Environmental perspectives, that are going to be detailed in the next sections.

A. Safety Perspective: Crime Modelling

Articles usually explore the occurrence of crime in events and unique occasions around the world (i.e. a concert, a sporting event, a conference), which allows the definition of crime and its related entities (author, victim, types of crimes, scene) [1]. In a specific case, where criminal occurrences were found in Auckland (New Zealand), it is possible to see variations in the volume of certain types of crime during the late 1990s and early 2000s. Additionally, it is also suggested a way to follow the evolution of the occurrences over time, corroborating the definition of a score at an opportune moment [1].

Following the same reference, it is clear that there is a concern to distinguish the place where the crime occurred. In Fig. 1, it is possible to check the correlation between the volume of assets lost for every type of crime scene. The data portrayed Auckland during a sporting event in the year 2000.

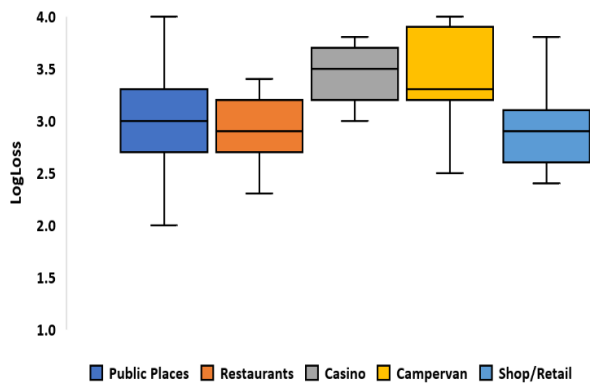


Figure 1. Loss of assets in relation to the crime scene.

The proposed segmentation meets the classification criteria that may compose a crime index score. But one should still explore not only the sites, but also the regions of the city and, to an even greater extent, its district.

As with the Social quantity defined in the previous topic, the information related to the occurrence of the crimes (place where it was committed and the type of crime) will be responsible for composing the Safety Perspective.

B. Economic Perspective: Hotels and the Public Transport

Public transport networks are very critical components of large cities. It is difficult to estimate the capacity of these networks, especially in places where tourism directly is a great influencer of people who need to be served. Previous works suggests real-time analysis to anticipate actions and ensure greater resilience in public transport networks [18]. There is also a specific study on subway networks (which usually require greater investment) that considers adaptations to this transport system to ensure their availability given a planned schedule, but also with consequent action analysis of the facts in real time [19].

The concept of the term “public transport” is still being transformed over the years. In fact, after the emergence of “Smart Cities”, transport is no longer just a service offering that provides mobility to the people, but also considers the availability of information. It is no longer just a question of what means are available for transport from one place to another. People need to know when the transport will be available, what is the best available route, what is the best mode of transport to that destination, and the capacity of the available mode. In this context, previous works discussed solutions adopted in Budapest to improve the available vehicle traffic system as well as future implementations being made [15].

Based on the above sources, it is essential to use public data – preferably in real time – on the conditions and availability of the flow that buses, subways and any other public means of transport for a correct assessment of citizens’ perception of a locality. city specific.

Still in the Economic context, in addition to public transport and its availability, there is also the figure of tourism and its derivatives: hotels. The literature explores several possibilities of defining tourism and its relations with other sectors. There is also the concern to define the main actor that moves tourism, that is, the tourist himself.

Moving forward, the Economic perspective mentioned in this study will be described measuring transport and hotels availability, as well as the proximity of the sights, which will eventually match the ability of the tourist spot to attract new tourists.

C. Environment Perspective

Other indicators that represents the citizen's quality of life and are essential for the creation of a safety index score correspond to the environmental quality indices. The availability of parks and green areas notoriously makes up the environmental picture of urban sub regions and measures

the quality of life of these places. Miscellaneous occurrences or requests for public spaces that require government action (e.g. sanitation, maintenance of pavements, various situations related to public lighting or high noise levels, maintenance of green areas) define the role of the state regarding its obligations in managing the public environment and, consequently, its impact in people's daily life.

Health and Environment should be analyzed separately when observed in the context of the community [14]. Although they are intrinsically related within the community (the reduction of green areas – defined in the "Environment" perspective, for example, leads to variations in sanitation indicators – which belong to the "Health" perspective), the indicators will be treated independently.

To define what should be observed and thus guarantee the citizen a good quality of life in the Smart Cities, a survey was conducted in 2018 [20]. The term "Urban Sustainability" encompasses the perception of the population from the perspective of pollution indicators (sound/visual/climate), government actions, sustainable development and economy. All indicators cited in the study are linked to the perception of the quality of life of citizens living in the areas where the survey was conducted.

Similarly, the different forms of requests made by the population to governmental bodies to maintain the quality of life and to implement the concept of well-being in safe communities will be observed.

D. Health Perspective

It has been previously shown that one of the ways to model health quality in a community occurs is by assessing the availability of different public health modalities (e.g. health centers, hospitals, gyms) [16]. In this regard, the offer of Health Clinics near the observed points indicates that the quality of life of that place is relevant.

Comparing this with the "Environment" perspective, the health indicators will measure the availability of health services in areas close to tourist locations. Requests made by the population to governmental bodies for health maintenance and urban hygiene related to well-being in safe communities will also be observed.

III. CONCEPTUAL MODEL

In order to aggregate and evaluate all sources of information that define the quantities described above, we chose to use a crawler-based model, defined as "a set of programs that are able to download pages interactively or automatically by extracting HTML content from predefined URLs [2]. A web crawler, for example, can be fed with a URL and then download all the content of the pages that are related to it in the form of hyperlinks.

In web crawler programs, it is possible to add calculations during searches to identify content that is considered most relevant to the search, or reject it (if it turns out to be duplicate or already visited content, for example). An important role for crawlers is to support search engines (e.g. Microsoft Bing, Google Search) in setting indexes after retrieving web page content. Regarding data mining, it is

possible to build a crawler in a separate application and then perform the analysis of the data that is captured.

Using crawlers, all the essential information for defining community safety indices will be captured from a variety of sources related to previous topics available on the Internet. After collecting, treating and classifying the data, a model will be created to compute a score for each environment classified as "Tourist Attraction", assigning it a specific rating according to the evaluation of the previous assumptions.

In summary, the captured data will be associated with a category that will compose the score index in the 5 main quantities:

- Security: Crime and Police Action
- Social: Sanitation Activities, Availability of Public Buildings
- Health: Availability of Health Clinics and Urban Hygiene Conditions
- Economy: Tourism and Public Transportation
- Environment: Availability of Green Areas and Parks

As representative of the Economy perspective, also responsible for guiding the studies regarding the tourist aspect, we will define indicators related to the economy in Tourism and Public Transport. In the Social perspective, we will define indicators that evaluate the availability of public services and people's quality of life (i.e. sanitation conditions). For the Health perspective, the availability of health clinics and issues to be resolved within urban hygiene will be observed. Regarding the Security perspective, we will observe the availability of police stations and indicators of urban crime. Finally, the Environment perspective will be defined by indicators related to visual and noise pollution, as well as the availability of green areas.

A. Planned Steps

First, it will be needed to define all data sources that will initially compose the proposed model. Then, for each source, the available subsections will be derived, which will relate to the final category that will compose the score index, with a positive or negative value.

The final result will be a weighted average of the relative universe of the positive items that were found. The normalized score will be between 0 (zero) and 5 (five), where 0 (zero) will represent the minimum value and 5 (five) will represent the maximum value of the indicator.

In the end, the higher the value is, the better is also the result for that tourist attraction.

B. Data Classification

For every perspective described, a rating must be assigned to tell whether the occurrence found has a positive or negative influence on each mapped context.

- Security: Availability of police stations are positive items. Enhancement requests in any categories that match public safety are negative. For evaluation purposes, only enhancement requests that are running at the time of data processing will be considered.

- **Social:** The proximity of public buildings of the observed locality will have a positive representation. Requests for improvement regarding the sanitation conditions will negatively influence the social result of that area.
- **Health:** The offer of Health Clinics in the observed area will have a positive representation. On the other hand, requests for improvement regarding Urban Hygiene conditions will negatively influence the social outcome of that area.
- **Economy:** Availability of hotels, bus and metro stops (public transport) will represent this perspective positively. Any requests regarding the poor quality of housing near the observed location will be negatively represented.
- **Environment:** The availability of green areas and parks near the observed area will have a positive representation. However, any requests for improvements in infrastructure conditions such as paving, road signs, street lighting as well as the maintenance of the green areas found will be assessed negatively.

In all cases where the population requests for the improvement of some service or public facility, only those requests that are being executed/under analysis at the time of data processing will be considered.

C. Process Flow

The following steps define the how data will be processed on each part or the pipeline.

Once the data sources have been defined, data will be extracted by crawling raw information from each source available. The crawler will parse HTML/Javascript and store the data in a more structured format.

In the second step the raw information obtained in the first step will be cleaned. At this time, the goal is to identify and treat anomalies found in the data set, such as erroneous values, missing data, among other data quality issues.

Considering that the data is structured and all relevant information was parsed, it will be needed to apply some filters in the result from previous steps. These filters will help to identify patterns in the remaining texts and propose their classification. Items that could not find a proper classification at this time, will be stored in a staging area that might be used in an eventual manual classification later.

Now that all data is ready to be read, data will be loaded into a Data Warehouse, allowing the definition of result indices and results comparison over time.

The final step is to extract valuable information from the Data Warehouse, leading into a KPI that will provide a business perspective about the CSWB rates.

Fig. 2 illustrates the workflow of the conceptual model previously described:



Figure 2. Process Flow.

D. KPI Definition

As formulated in the literature [14], the Safe Community and Welfare indicator requires the aggregation of different results observed by the analyzed quantities, but no method is defined to be used for such aggregation. Thus, it was decided to aggregate these results by means of a weighted average [3]. The weights attributed to each of the quantities will be responsible for quantifying their relevance in the final result.

Given each perspective (n), the average “SafetyScore” that will rank the results found will be defined in:

$$\frac{POS_n}{TOT_n} * W_n \quad (1)$$

Where POS represents all positive occurrences for that perspective, TOT represents all the occurrences for that same perspective, and W corresponds to the weight that will be used to evaluate the perspective relevance.

We have the final formula that represents the “ScoreIndex” by:

$$\frac{\sum_{Perspective\ 1}^{Perspective\ n} \frac{POS_n}{TOT_n} * W_n}{\sum_{Perspective\ 1}^{Perspective\ n} W} \quad (2)$$

By using the “ScoreIndex” calculation each perspective can be evaluated individually and each result can be aggregated and weighted according to how relevant that information is characterized in the context of the Safe Community and Welfare indicator.

Each equation represents a proportion of the occurrences found and classified in the observed quantities. These ratios are defined by the total positive occurrences (given by the factors prefixed with the label “POS”) over the total occurrences observed for this same quantity over the period analyzed (factors prefixed with the label “TOT”).

As already mentioned in previous topics, there is also the element that defines the “weight.” This is composed of the factors prefixed with the label “W” and is responsible for giving due relevance to the analyzed quantity.

IV. DEVELOPMENT

This section will detail the resources used, as well as the data structures that manage the results of the proposed model. Thus, the auxiliary interpretation of the data set can be performed from the sources described below, by adopting a classic Business Intelligence framework.

A. Extract-Transform-Load (ETL)

In order to store and manage data sources, database tables were used to store different data types for each source. Whenever the database is refreshed, incremental data is loaded and added for future analysis through the Data Warehouse.

The complete workflow consists of using data with extract, transform, and load (ETL - Extract, Transform,

Load) procedures, and then a final structure known as Data Warehouse will store critical information that will allow future analysis and support customer’s business decisions.

B. Data Sources

The following sources describe each one of the twelve data sources defined:

TABLE I. DATA SOURCES

Data Sources Name	Source
BaseNationalMonument	Lisbon City Hall
BasePublicMonument	Lisbon City Hall
BaseGreenArea	Lisbon City Hall
BasePark	Lisbon City Hall
BaseNeighborhood	Dados Abertos Website
BaseSubway	Lisbon City Hall
BaseBusRoute	Dados Abertos Website
BaseTourismDevelopment	Portugal Tourism Website
BasePoliceDepartment	Lisbon City Hall
BaseCrime	Portugal National Statistics Institute
BaseHealthCenter	Lisbon City Hall
BaseAppLx	Mobile Application (Minha Rua LX)

Each source is considered as an “Open Data Source”, and it is refreshed dynamically by its own providers or users (considering the Mobile Application that reflects population opinions around certain perspectives). Added to this, each source is structured in one of the following formats: JSON or pure HTML.

The JSON data sources are all available following the syntax in Fig. 3:

```

"type":"FeatureCollection",
"features":[
{
"type":"Feature",
"properties":{
"OBJECTID":1,
"COD_SIG":"2105705006001021",
"IDTIPO":"999",
"COD_SIG_EDIF":"2105705006001",
"NOME":"38ª Esquadra",
"MORADA":"Rua Ricardo Ornelas Lote 378, R/C-A",
"TIPO_UNIDADE":"Esquadra de Policia",
"TELEFONE":"-",
"EMAIL":"-",
"CODPOSTAL":"-",
"FUNTE":"PSP",
"MORADA_RMOG":"Rua Ricardo Ornelas ",
"GlobalID":"c9475c99-47fa-458a-a458-7a523d689846"
},
"geometry":{
"type":"Point",
"coordinates":[
-9.124781943050156,
38.752016513493594
]
}}

```

Figure 3. Example of a JSON Data Source (BasePoliceDepartment).

The “BaseCrime” database was taken from a raw website, and its format is an HTML table, as it can be seen in Fig. 4:

Geographic localization (NUTS - 2013) (1)	Registered crimes (No.) by the police authorities by Geographic localization (NUTS - 2013) and Category of crime; Annual (3)
	Data reference period (2)
	2018
	Category of crime
	Total
	No.
Portugal	333 223
Continente	314 312
Região Autónoma dos Açores	8 956
Região Autónoma da Madeira	6 311

Figure 4. Example of a HTML Data Source (BaseCrime).

C. Data Warehouse

The Data Warehouse, designed to manage the Safety Score Index and related data, is structured into dimensions (defined by the presented data sources) and fact tables (defined by each perspective already presented – Economy, Health, Social, Safety and Environment), all consolidated in a Star Schema.

For every data source described in the previous section, it is created a corresponding dimension (except the “DimTime”, which represents a time frame and it is not derived from any data source) with relevant data from source tables that will assist in classifying information further ahead.

The list of entities created to represent the processed information can be found below:

TABLE II. DATA WAREHOUSE DIMENSIONS

Dimension Name	Source
DimAppLx	BaseAppLx
DimBusRoutes	BaseBusRoute
DimCrimes	BaseCrime
DimHealthCenters	BaseHealthCenter
DimHotels	BaseTourismDevelopment
DimNeighborhoods	BaseNeighborhood
DimParks	BasePark
DimPoliceDepartments	BasePoliceDepartment
DimSubways	BaseSubway
DimTouristSpots	BaseNationalMonument
DimTime	Not Applied

The relationship and summarization between the dimensions are defined for every fact table (where each fact Table represents one of the perspectives previously presented), as listed below:

- “FactTouristSpotsEconomy”: Summarizes Public Transportation and Hotels availability, and the AppLx Mobile Results related to economic events.
- “FactTouristSpotsHealth”: This summarizes Health Clinics availability and AppLx Mobile Results for health events.
- “FactTouristSpotsSocial”: This summarizes social indicators from AppLX, and the availability of Public Monuments.
- “FactTouristSpotsSafety”: This summarizes crime events, availability of police departments and AppLx safety events.
- “FactTouristSpotsEnvironment”: This summarizes environmental indicators from AppLx, and also the availability of green areas such as gardens and parks.

D. Key Performance Indicators (KPIs)

For each perspective described in the previous sections, there were defined KPIs to allow the possibility to measure and compare results between the touristic spots. These KPIs are represented in the table below and their evaluation is defined such as an “up arrow” (meaning that greater values represent best indicators) or a “down arrow” (meaning the opposite):

TABLE III. KEY PERFORMANCE INDICATORS

Perspective	Measure Name	Evaluation
Security	# Crime Events	↑
	# Police Departments	↑
	# AppLx Noisy Public Events	↑
Social	# Public Monuments	↑
	# AppLx Sanitation Events	↑
Health	# Health Clinics	↑
	# AppLx Urban Hygiene Events	↑
Economy	# Bus Stops	↑
	# Subway Stations	↑
	# Hotels	↑
	# AppLx Housing Events	↑
Environment	# Green Areas	↑
	# Gardens and Parks	↑
	# AppLx Green Areas Events	↑
	# AppLx Sidewalk and Accessibility Events	↑
	# AppLx Municipal Assets Events	↑
	# AppLx Road Events	↑
	# AppLx Street Lights	↑

Perspective	Measure Name	Evaluation
	Events	

The eighteen indicators used to describe each perspective around its CSWB context are balanced between dynamic values – such as the results provided by the Mobile Application – and other typical static values – such as the annual reports provided by the Instituto Nacional de Estatística – Portugal’s official authority for statistical studies, responsible for carrying out public enquiries and collecting/providing data at country level.

V. RESULTS & DISCUSSION

Every national monument defined in the “DimTouristSpots” (63 national monuments extracted from the original “BaseNationalMonument”) was classified by all the indicators defined in the previous section and after all these values were evaluated, the final result was represented by a scale from 0 to 5, graphically defined by stars with 0.5 units range.

In order to give a better view about the final result, the following chart represents the consolidated result for each touristic spot following the “proximity axis” [14]. That compares each touristic spot starting from an “ideal entity” – an ideal touristic spot – compounded by the best results from each spot around every perspective defined previously, followed by the real results found during the data analysis.

As it can be seen in Fig. 5, the data series are split into quartiles (in order to make it easier to visualize what are the spots that need more or less attention).

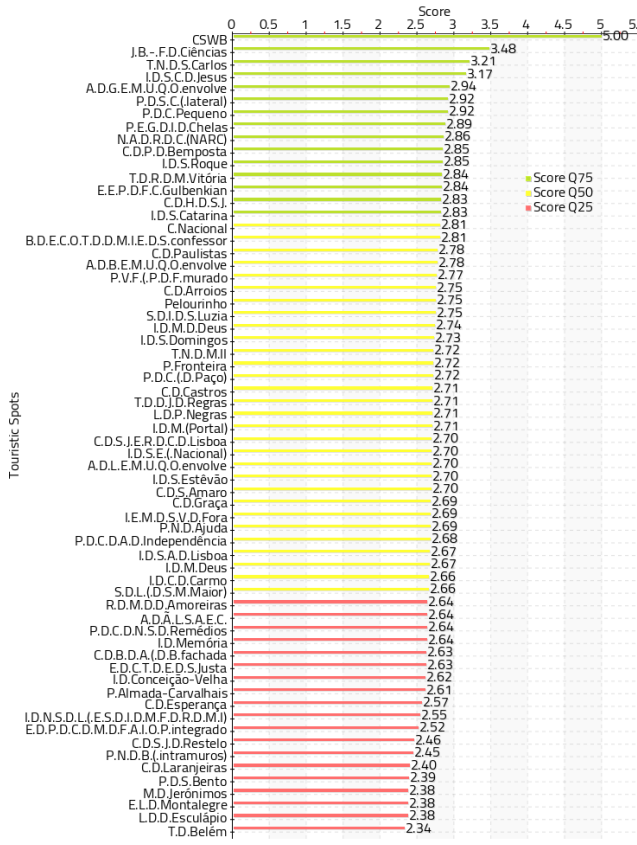


Figure 5. Proximity axis.

The “Score Q25” series groups the results 25% lower. These results are part of the first quartile of the database. The “Score Q50” series represents the results from Quartile 2 (those that are higher than the limit defined by Quartile 1 and lower than the one defined by Quartile 3). Finally, the results grouped in the “Score Q75” series are 25% higher.

A first approach shows which tourist sites need immediate response from the public sector: all those found in “Score Q25”.

Added to this, it is possible to evaluate eventual opportunities to improve the quality of life around the spots, superimposing the results of the equivalent perspective and comparing the index evaluated in each one of them. Fig. 6 shows how two spots can be compared over the five perspectives already defined:

Safety Spots - Lisbon - KPI

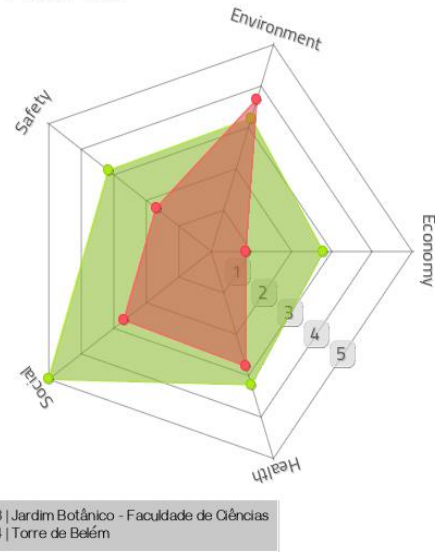


Figure 6. Comparison between first and last positions.

Considering the data evaluated, the best spot found was the “Botanical Garden”, which when converting the final result to a scale from zero to five, obtained the score of 3.48. At the other end is the iconic “Belém Tower”, with a score of 2.34.

The chart area shown in the figure above allows us to evaluate that the result of the “Botanical Garden” is superior in almost all quantities, being inferior only when it comes to the Environmental indicators.

The analysis around the Environment perspective shows the importance of dynamic indicator terms, in this case, provided by the application “Minha RuaLX”. It is easy to see that although both regions have a large number of green spaces and parks, it was in the evaluation of the entries found in the application that resulted in a better rating of the “Belém Tower”.

Apparently, residents around the “Botanical Garden” are more concerned with reporting environmental problems and filing requests to the Lisbon City Hall than those residing near the “Belém Tower”.

In all other measures, the “Botanical Garden” always shows a better result, which justifies its position as the best rated touristic spot in Lisbon, as we can see below:

- **Economy:** High number of bus stops, subway stations, rooms and accommodation. In addition, a small number of open requests from residents in the AppLx application. The difference in the availability of public transport near “Belém Tower” is evident, clearly shown by the data found in the KPI.
- **Health:** Almost equal to each other, both have many entries in the AppLx application, but the “Botanical Garden” still has a larger number of Health Clinics.
- **Social:** Because it is closer to other public monuments and has fewer entries in the AppLx application, it is easy to see why the “Botanical Garden” was also highly ranked.

However, the granularity of the indicators available in the open data presented a challenge to overcome: for the “crime” indicator for example, the greatest detail available is only at the municipality level – which individually does not make sense, since all touristic spots are in the city of Lisbon. And, in order to overcome this granularity, it was decided to evaluate the crime conditions in the neighboring municipalities and to influence the results considering the geographical proximity of the points.

Another item still to be solved refers to the behavior of the “Score Index” throughout the seasons. Current data do not yet allow us to say whether climate change can interfere with the interpretation of the information obtained here or the definition of the result throughout the year – which was also one of the objectives of this work. This is because current data sampling has not made it possible to compare them across other seasons. However, once the model is already defined, this assessment becomes possible: just look at the behavior of the Score Indexes over time.

The evaluation of the touristic spots from the perspectives defined here allows us to evaluate the opportunity for improvements in the tourist areas and thus provide a better quality of life not only for the citizen residing in the evaluated region, as well as for the tourists who come to know Lisbon and Portugal. its surroundings. The comparison here between the first and last ranked in the rule to which it was applied shows that there are opportunities to direct investments and improve the *status quo* of touristic spots. Regarding the continuity of the model, it is noteworthy that the adaptation to new scenarios and challenges is feasible and may assist in decision making in other situations through the evaluation of indicators.

VI. CONCLUSION

The definition and measurement of community safety indicators [14], namely the “Community Safety Well Being Index” (or as labelled here: “Score Index”), allows to create a competitive ecosystem among the public entities responsible for maintaining touristic spots. Assessing and maintaining a “Safe Place” can no longer be viewed as reactive initiatives by a government or any other form of organization.

As shown, the definition of “Welfare” in cities is the responsibility of entities from various sectors who will need to work together for a favorable environment to be established. This ecosystem creates opportunities for improving social, economic, health, safety and environmental conditions around the points analyzed – each of which will be responsible for proposing actions to improve the individual results of the element in question.

The main objective of this article – the definition of an interactive model that evaluates tourist attractions from different perspectives responsible for characterizing the safety and welfare issues of the communities – was achieved. Even so, the model is flexible enough to be adapted to every new scenario. Regarding the definition of indicators in each of the proposed perspectives, the current solution cannot be considered as an exact science: indicators were selected from literature recommendations [16]. However, it should be

noted that the selected indicators allow for certain adaptations – as long as the scope is not changed – following the framework defined in the literature [14].

The main question that this model seeks to solve (“how safe is this place?”) can be answered by looking at the tourism data, since the analysis focused on each of the touristic spots defined by the Lisbon City Hall – and created a result for each of these. Therefore, depending on the area in which the tourist will want to visit, the model will support the decision making process by presenting an index of the most favorable places in all aspects: economic, social, health, safety and environment.

By capturing data and structuring it in the form of a Data Warehouse we were able not only to map the areas of interest (i.e. touristic spots with their geographic location and other information that is available from open data sources) and the characteristics that make up the surroundings of the sights, but also create new information from the joint evaluation of these data.

The challenge to map and identify problems faced by tourists, was made possible by reading the open data provided by the Lisbon City Hall (such as the availability of public transport and the volume of hotels and rooms available, among others) but also by reading the open/dynamic data available through the “Minha Rua LX” application.

Suggestions for future work associated with the proposed model:

- Adaptation of the model to other entities such as parishes, districts, shopping centers and restaurants.
- Adapting the model to other countries, notably developing countries, to assist in classifying localities and identifying points of improvement.
- Creating a mobile application that enables information to become widely available.

Finally, the indicators evaluated along this work could help to gradually increase the number of visits to these sites, as they will become more attractive in all aspects. To achieve this, the indicators must be used to allow actions that could improve the current results for each touristic spot and hopefully change the tourist experience around the city of Lisbon.

ACKNOWLEDGMENT

We thank Universidade Nova de Lisboa for the support and opportunity to research on this Smart Cities subject and allow the possibility to improve life quality around the world through research. We thank Camara Municipal de Lisboa (Lisbon City Hall) for providing relevant sets of data through its “Open Data Website” and “Minha AppLX” mobile application.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- [1] M. Barker, S. Page, and D. Meyer, "Modeling tourism crime: The 2000 America's Cup," *Annals of Tourism Research*, vol. 29(3), January 2002, pp. 762-782.
- [2] M. Thelwall, "A web crawler design for data mining," *Journal of Information Science*, vol. 27(5), October 2001, pp. 319-325.
- [3] A. Sharpe, "Literature Review of Frameworks for Macro-Indicators," Centre for the Study of Living Standards, Ottawa, Canada, 2004. [Online]. Available: <http://www.csls.ca/reports/LitRevMacro-indicators.pdf>.
- [4] N. Vanhove, *The economics of tourism destinations*, Oxford, UK: Butterworth-Heinemann, 2005, p. 2.
- [5] A. Gilchrist, *The Well-Connected Community*, 2nd ed., Bristol, UK: The Policy Press, 2009.
- [6] J. Kania, and M. Kramer, "Collective Impact," *Stanford Social Innovation Review*, Winter 2011, pp. 36-41.
- [7] R. Dodge, A. Daly, J. Huyton, and L. Sanders, "The Challenge of Well-Being," *International Journal of Well-Being*, vol. 2(3), August 2012, pp. 222-235.
- [8] "Estatísticas do Turismo 2014," Instituto Nacional de Estatística, Lisbon, Portugal, 2015.
- [9] "Estatísticas do Turismo 2015," Instituto Nacional de Estatística, Lisbon, Portugal, 2016.
- [10] "Estatísticas do Turismo 2018," Instituto Nacional de Estatística, Lisbon, Portugal, 2019.
- [11] "Portugal Recebeu mais de 60 milhões de turistas em 2016." *Diário de Notícias*. <https://www.dn.pt/dinheiro/interior/portugal-recebeu-mais-de-60-milhoes-de-turistas-em-2016-5624832.html> (accessed May 20, 2018).
- [12] "Turismo é o setor com maior crescimento no mundo." *Dinheiro Vivo*. <https://www.dinheirovivo.pt/economia/turismo-e-o-setor-com-maior-crescimento-no-mundo> (accessed May 25, 2018).
- [13] S. Huddart. "Seven Years on and Seven Years Out: Revisiting Patterns, Principles, and Patterns in Social Innovation." *The Philanthropist*. <https://thephilanthropist.ca/2017/04/seven-years-on-and-seven-years-out-revisiting-patterns-principles-and-practices-in-social-innovation> (accessed December 12, 2018).
- [14] C. Nilson, "Community Safety and Well-Being: Concept, Practice, and Alignment," *Journal of Community Safety & Well-Being*, vol. 3(3), October 2018, pp. 96-104.
- [15] J. Mezei and K. Lazányi, "Are We Ready for Smart Transport? Analysis of Attitude Towards Public Transport in Budapest," *Interdisciplinary Description of Complex Systems*, vol. 16(3-A), September 2018, pp. 369-375.
- [16] H. Sung, and R. G. Phillips, "Indicators and community well-being: Exploring a relational framework," *International Journal of Community Well-Being*, vol. 1, October 2018, pp. 63-79.
- [17] M. Neto, "Urban intelligence," in *18th Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI)*, Santarém, Portugal, October 2018, pp. 12-13.
- [18] P. Noursalehi and H. N. Koutsopoulos, "Real-time Predictive Analytics for Improving Public Transportation Systems' Resilience," in *Bloomberg Data For Good Exchange Conference*, New York, USA, September 2016.
- [19] H. N. Koutsopoulos, P. Noursalehi, Y. Zhu and N. Wilson, "Automated data in transit: Recent developments and applications," in *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, Naples, Italy, June 2017, pp. 604-609.
- [20] D. Petrova-Antonova and S. Ilieva, "Smart Cities Evaluation – A Survey of Performance and Sustainability Indicators," in *2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEEA)*, Prague, Czech Republic, August 2018, pp. 486-493.